Roadmap on Distributed Computing Infrastructure for e-Science and Beyond in Europe
Disclaimer

This document has been produced with the funding of the European Commission. The content of this publication is the sole responsibility of the SIENA Consortium and its experts and cannot be considered to reflect the views of the European Commission.
Index

Foreword ................................................................................................................................................................................................... 3
1 Executive Summary ........................................................................................................................................................................... 5
2 Introduction .......................................................................................................................................................................................... 6
3 Facing up to the Challenges.......................................................................................................................................................... 9
4 Call for Action – Short and Medium Term Priorities ............................................................................................................. 15
5 Capitalising on e-infrastructure investments in Europe ......................................................................................................... 19
6 Future directions - driving positive change for Europe in a global context ................................................................. 34

ON-LINE ANNEXES TO THE ROADMAP
www.sienainitiative.eu/SIENA_Roadmap_Annexes.aspx

Terms of reference
Annex 1 – The European e-infrastructure landscape – How Europe can capitalize on expertise and knowledge gained through past and current investments, and pinpoint opportunities to reinforce links with research infrastructures.
Annex 2 – European cloud initiatives - How Europe can gain from greater awareness of cloud initiatives, identify common requirements for standards implementation and strengthen its overall strategic positioning.
Annex 3 – Sustainability strategies – How European initiatives can define tangible and intangible assets and plan their reuse.
Annex 4 - The global standards landscape – How Europe can identify relevant standards and studies to address interoperability, portability and legal impediments in a global landscape
Annex 5 – Legal brief – An investigation into current legal regulations in Europe.
Annex 6 - Glossary
Scope of the SIENA roadmap

SIENA is funded European Commission under the 7th Framework Programme (FP7). The goal of SIENA, Standards and Interoperability for eInfrastructure impleMeNtation initiAtive, is to accelerate and coordinate the adoption and evolution of interoperable distributed computing infrastructures through engagement with industry, standards development organisations (SDOs) and major stakeholders to forge community agreements on best practices and standards for distributed computing.

SIENA is the first initiative to bring to the same table standardization bodies to support the analysis of open standards-based interoperable grid and cloud computing infrastructures. The SDOs that have contributed include: the Open Grid Forum (OGF), the Institute for Electrical and Electronics Engineers Standards Association (IEEE SA), the European Telecommunications Standards Institute (ETSI), the Distributed Management Task Force (DMTF), the Organization for the Adoption of Structured Information Standards (OASIS), the Storage Networking Industry Association (SNIA), and International Telecommunication Union Telecommunication Standardization Sector (ITU-T).

The primary focus of the SIENA initiative is European electronic infrastructure for research. We therefore, take into consideration the points of view of both the people responsible for infrastructure creation and operation, such as service and application developers, and the “end users”, for example, researchers in science, the arts and humanities. The focus on standards and interoperability and the offer of services through European grid and cloud infrastructures will benefit the education sector, and will also be relevant for e-government and commercial enterprises that can take advantage of the substantial research background that underlies most e-infrastructures.
Foreword

As science is increasingly driven by the processing of big data, researchers need access to science clouds and other e-infrastructure that satisfy their requirements. Interoperability of services and applications is a key concern because it broadens choice and ensures a level playing field for both service providers and users, driving competition and innovation. I am a fervent defender of interoperability and it is a priority of the Digital Agenda for Europe.

Standards are the key tools to achieve interoperability. However, standards sometimes take too long to develop in comparison to the speed at which the ICT world functions. The aim of the SIENA roadmap is to reduce the required time to reach consensus in distributed computing standardisation initiatives by focusing attention on the main challenges and charting the possible paths to follow.

I therefore welcome the SIENA roadmap and I invite all stakeholders to use it as a reference.

Neelie Kroes
Vice-President of the European Commission
1. Executive Summary

This roadmap assesses the situation, identifies issues, and makes recommendations regarding the adoption and evolution of open standards-based interoperable grid and cloud computing infrastructure (e-infrastructure) to support research in Europe. The vision for such a European e-infrastructure is to empower productivity of research communities through ubiquitous, trusted, and easy trans-national access to services for data, computation, communication and collaborative work. Some considerations in this roadmap apply also to computing in industry and the public sector.

Projects to develop technology to support research computing in Europe have typically worked independently. The drive towards an open standards-based interoperable e-infrastructure highlights the important need for cooperation and coordination to communicate requirements and participate in the standardization process. A coordinating advisory body with representation from all projects developing e-infrastructure technology is one possible vehicle to these ends\(^1\). This approach has been trialed with the Distributed Computing Infrastructure (DCI) projects funded in the 7th call from the e-Infrastructure and GÉANT Unit through a collaborative roadmap.

The substantive issues of the need for enhancements to appropriate existing standards, or development of new standards if necessary, must be addressed by the providers of e-infrastructure technology. Providers need to commit time to participating in such work. An approach of joint collaboration between current and future DCIs, commercial vendors, and standard development organizations will help find a way through the standards challenges mentioned in this report.

An important conclusion that emerges from our analysis is that there are important areas in cloud computing where Europe can establish leadership without going into head-to-head competition with leading commercial cloud service vendors. One could be the development of a large scale private cloud, jointly financed and operated by European research funding agencies, available to researchers, and optimized for research. Another is to continue the efforts begun in the EGI-InSPIRE\(^2\) project to build the standards and software to support clouds formed by federating multiple smaller cloud providers, with possible extension to commercial public clouds. These two initiatives are complementary and can be integrated through the adoption of standards such as those discussed in this roadmap.

The structure of the document is as follows. An introduction sketches the origins of grid and cloud computing, setting the stage for e-infrastructure standardization. This is followed by a description of a number of challenges to the adoption of cloud computing, not only in research, but also in the public sector and industry. The third section contains the actual roadmap, in the form of a Call for Action containing recommendations with short- and medium-term priorities. The following two sections then expand on the challenges, and on the calls for action. Further supporting material, as described in the table of contents, is provided in on-line annexes available at www.sienainitiative.eu/SIENA_Roadmap_Annexes.aspx.

---

1. See Call for Action number 8.
2. Introduction

Prior to the mid-20\textsuperscript{th} century, computers were people doing calculations by hand. The first electronic digital computers were built for the decipherment of secret codes and for scientific computing: numerical simulation in science and engineering. These applications have strongly influenced the evolution of what we now call information technology. Originally, all computing was centralized frequently accessed by individual users sharing time on a single mainframe.

With the invention of the internet, it became possible to distribute computing over separate machines. Over the past decades, distributed computing has become increasingly prevalent. More recently, the evolution has been towards a utility model of computing. Grid and cloud computing represent the current state in this evolution towards a utility model. Grid computing enables sharing of distributed computing resources; cloud computing offers computing as a service.

A major motivation for the development of grid computing in Europe was to share and process data generated by the Large Hadron Collider (LHC) at CERN\textsuperscript{3}, the world’s largest scientific experiment. That requirement, together with those of research communities in other fields, has ultimately led to the world’s largest production grid computing infrastructure\textsuperscript{4}, which now benefits wide areas of scientific and engineering research beyond physics, and beyond science to the social sciences and humanities. The LHC computing grid has been a great success\textsuperscript{5}.

The development of grid computing to accelerate scientific discovery has been an essential stage in the evolutionary development of electronic infrastructure that will contribute to supporting the research and innovation flagship initiative of the Europe 2020 Strategy\textsuperscript{6}.

The first meeting of an international organization devoted to the promotion and standardization of grid computing, the Global Grid Forum (GGF; now the Open Grid Forum – OGF)\textsuperscript{7}, was held in Amsterdam in March, 2001. The ongoing effort to build the European grid infrastructure (EGI)\textsuperscript{8} has been closely tied to efforts to standardize grid computing, to a large extent under the umbrella of GGF/OGF. As a complementary effort, the Distributed European Infrastructure for Supercomputing Applications (DEISA) has helped ensure that GGF/OGF standards are also relevant to high-performance computing environments. Other international standards development organizations (SDOs) have become involved as technologies from other areas of information technology became incorporated into grid middleware.

The standardization effort in grid computing has been largely driven by the requirements of scientific research, which, unlike technologies for most enterprise information technology, requires federation of resources that span administrative, ownership, and national boundaries.

\textsuperscript{3} http://lhc.web.cern.ch/lhc/.
\textsuperscript{4} http://www.egi.eu/.
\textsuperscript{7} http://www.ogf.org.
\textsuperscript{8} http://www.egi.eu/.
By about 2008, it had become apparent that grid computing would not be the universal utility computing solution. While grids have been successful in scientific computing, grid computing has had little adoption in industry or commerce. One of the reasons is the lack of a general business model for grid computing. A few years ago, cloud computing (easily available computing on demand) burst onto the scene with a business model (pay-as-you-go) that is broadly applicable. Very quickly, a correspondingly wide spectrum of actors has become involved in adopting cloud computing, building businesses around cloud computing, and standardizing cloud computing.

We are now at a cusp in the development of electronic infrastructure to support research in Europe. It seems clear that, with the exception of high-end capability computing (application of the largest and most powerful computers to the most challenging problems), variations on the cloud computing model will ultimately dominate computing of most kinds in the future. Therefore, it behooves us to adapt our e-infrastructure to this inevitability. To a significant degree, the current European distributed computing initiatives (DCIs) discussed in this roadmap are exploring such an adaptation.

It is important to recognize that classical scientific computing is far from being the driving application for cloud computing. We use cloud services every day for internet search, mobile applications, photo sharing, messaging, and social networking. The internet of things is expected to drive many new applications in the cloud.

The fact that e-infrastructures continue to evolve in response to both new user requirements and the arrival of new technologies is a sign of vigorous health. The fact that the European DCIs are now actively exploring the value of the assets they have acquired with a view towards longer term sustainability adds greater strength to the evolution towards fulfilling the Europe 2020 Strategy.

It is important to emphasize that cloud computing for research is work in progress. Cloud computing is not a magic wand for scientific computing.

There are many research applications that can run equally well in the cloud as on a dedicated high-performance computer. There are also applications that are better suited to cloud computing, and others

---

9 Another is the visible complexity of securely sharing resources across administrative and ownership boundaries.
13 High-end capability computing in Europe is under the auspices of PRACE http://www.prace-project.eu.
15 The European Grid Initiative has a formal process to gather requirements. See https://wiki.egi.eu/wiki/Track_User_Support_Requirements.
16 Shane Canon, “Debunking some common misconceptions of science in the cloud” http://datasys.cs.iit.edu/events/ScienceCloud2011/.
that will continue to require dedicated supercomputers\textsuperscript{17}. High-end capability resources\textsuperscript{18} (high-performance computers, or supercomputers) and their corresponding e-infrastructures will continue to use traditional grid computing technologies and tools for the foreseeable future. A great deal of work remains to be done to provide effective support for research in the cloud.

Cloud computing is a fast moving field, and the number of standards development organizations (SDOs) and other collaborative groups working independently on standards is large. These bodies represent diverse interests. The following section describes some of the challenges in providing a roadmap on distributed computing infrastructure for research in Europe.

\textsuperscript{17} Geoffrey Fox, Dennis Gannon, op cit.

\textsuperscript{18} Defined, e.g. by their presence near the top of the list of the top 500 supercomputer sites at http://www.top500.org.
3. Facing up to the Challenges

Powerful economic and environmental forces are driving a major evolution in the way computing resources are provisioned for user communities in research, in government and the public sector, and in industry. Economies of scale are driving consolidation of such resources into a smaller number of ever larger data centers. Considerations of the cost of powering and cooling huge concentrations of electronic equipment, together with environmental concerns, drive the placing of such data centers in geographic locations where power is plentiful and inexpensive. These forces and their consequences both enable and drive the move towards a utility computing model. The latest manifestation of this model is cloud computing. The dynamic flexibility and reduced cost of accessing resources in the cloud are beginning to overwhelm most other considerations on provisioning IT resources and providing services to many user communities.

As suggested by analysts such as Gartner, cloud computing represents a service provider-consumer model and is driving a fundamental shift in the consumption model of IT functionality. Such a fundamental shift from asset ownership to service provisioning poses numerous challenges to all players in the transition. Many of the general challenges to the adoption of cloud computing are shared by cloud-based e-infrastructures for research. These include overall coordination of projects and programmes, requirements identification, several issues associated with standardization, security, education, legal impediments, and globalization.

- **Overall coordination of projects and programmes.** A vast array of projects and programmes devoted to cloud computing exists, with new ones appearing almost daily. The EC has a large collection of white papers and communications relevant to cloud computing, some of which are currently open for public consultation and comment, and can therefore be expected to change in the near term. There is so much going on that an active on-going openly accessible “cloud watch” would be very helpful.

- **Conflicting and inappropriate requirements.** The identification and prioritization of cloud computing requirements is an essential part of establishing security, portability and interoperability in cloud computing through relevant standards. A danger in this process is that conflicting or inappropriate requirements be imposed. For example, a requirement that a customer be able to specify the geographical location of data would be fatal for small- and medium-sized businesses proposing cloud services in Europe due to the high cost of reliable storage services.

---

19 Data centers with hundreds of thousands of computational and storage units are not uncommon. A recently announced data center near Dallas, Texas will have 3,600 square meters of raised floor under one roof.

20 Some important applications will continue to require dedicated high-performance computers.

21 Chiemi Hayashi: presentation at the OASIS International Cloud Symposium, op cit.

22 See, e.g. the ESA/CERN proposal for a “science cloud”, http://ur1.ca/9f2jf and the HelixNebula project starting 1 July 2012.


24 Reliable data access would require multiple storage facilities (for redundancy) in each country in which the...
» **Standardization challenges.** As in the case of projects and programmes, the cloud computing standardization landscape is complicated and confusing with many different organizations and interests involved in cloud standardization efforts\(^\text{25}\). The US National Institute of Standards and Technology (NIST)\(^\text{26}\) has compiled an inventory\(^\text{27}\) of standards relevant to cloud computing, and has identified a core suite of emerging high-level cloud-specific standards\(^\text{28}\), although, and as the NIST document points out, standards from other domains such as web services, the internet, and grid computing are available, and can support the functions and requirements of cloud computing.

» **Security challenges\(^\text{29}\).** Industry and government see risk in completely different ways\(^\text{30}\) (accept and manage, balance costs, clear lines of responsibility vs. eliminate risk, zero tolerance, ill-defined lines of responsibility). The vast majority of government data has nothing to do with national security\(^\text{31}\). In most cases for both industry and government, geographical location of data should not be relevant\(^\text{32}\). It has been argued that with appropriate standards in place, moving from proprietary in-house computing to cloud computing improves security\(^\text{33}\).

» **Data Management challenges.** There are a number of significant data management issues relating to the vast amount of data collected and used by research projects. These include ownership and governance of the data, creating trust in the data, sharing and re-use of the data and the long term preservation of the data. All of these issues are addressed in the recent report from the High Level Expert Group on service is offered, making the service prohibitively expensive. Such requirements have been voiced, e.g., http://url.ca/9f2sm.


28 These three standards, OVF (DMTF), OCCI (OGF), and CDMI (SNIA), are described in the first iteration http://goo.gl/ZcF8k of the SIENA roadmap (March 2011). See also the presentation on cloud computing standards work at NIST by Annie Sokol at http://goo.gl/XhcxH, especially slides 17-19.

29 For a general discussion from the government perspective, see, e.g. US Government Cloud Computing Technology Roadmap, Release 1.0 (Draft) Volume 2: Useful Information for Cloud Adopters, section 6, at http://goo.gl/uBxFg.

30 Scott Algeier, OASIS International Cloud Symposium, op cit.

31 Ian McCormack, OASIS International Cloud Symposium, op cit.

32 Megan Richards, in response to a question at the OASIS International Cloud Symposium (op cit.), pointed out that the oft-quoted US Patriot Act is very detailed and restricted in scope, and is “not a fishing license”. Tim Cowen pointed out in his presentation on legal impediments to cloud computing that European law enforcement agencies are also able to obtain access to data, and not necessarily under conditions as strict as in the US. Cloudscape IV, 23-24 February 2012, Brussels, highlighted that the focus on the US Patriot Act often eclipses the larger issue. It is the divergence between European and US data that needs addressing to ensure both parts are compatible and aligned with trade agreements for the open market, http://url.ca/9f2v0.

33 Marnix Dekker, OASIS International Cloud Symposium, op cit.
Scientific Data\textsuperscript{34}. We support the recommendations of this report, and urge immediate attention to their implementation.

» **Education challenge.** Education of decision takers at all levels of business and government, including politicians, is needed to overcome social and cultural obstacles to cloud computing adoption, to overcome inertia, and to ensure that the standardization and security challenges are met appropriately.

» **Challenges in understanding terminology and approaches.** The availability of ever more capable electronic components (e.g. for processing, memory, and networks) at ever lower prices leads to distortion of traditional terminology. The term “high performance computing” traditionally referred to use of systems in the Top 500 and related complex application enabling activities (e.g. parallel programming, achieving application scalability, etc.). In this strict sense, neither a grid nor a cloud is a supercomputer (although a supercomputer can be a node in a grid, or – in principle – be offered as a service in a cloud). Nowadays, one hears about high performance computing (HPC) in the cloud and HPC on the desktop, contributing to confusion for both consumers and providers.

» **Legal impediments.** The current fragmented and inconsistent legal and regulatory environment in Europe is seen as a major inhibitor to the deployment of cloud services. This is a potential problem for sustainability of Europe’s DCIs through uptake of their assets. A dramatic example is the *mimecast* story\textsuperscript{35} in which a rapidly growing UK firm with a cloud-based email service was obliged to seek further growth in North America because of the impracticality of providing the service in Europe due to uncertainties over cross-border data movement. A recent report\textsuperscript{36} from the EC following their public consultation on the subject also highlights the many issues. Annex 5\textsuperscript{37} contains a brief on legal issues in cloud computing adoption.

» **Globalization challenges.** Several major governmental cloud computing programs are underway\textsuperscript{38}. These are being pursued in the absence of an adequate set of agreed and accepted open standards that could guarantee security, portability, and interoperability of cloud services, yet the globalization of business implies that inter-cloud portability, cloud interoperability and cloud federation are global issues\textsuperscript{39}.

» **Entrepreneurship challenge.** Encouragement and support for the creation of new cloud-based businesses could be a motor for wealth creation and employment for young people prepared to take risks and work hard. The knowledge, technical expertise and assets offered by the DCIs can be a basis for the creation of new products and services, either directly through the sustainability efforts described in chapter 5 and Annex 3\textsuperscript{40} or through new start-ups. Creation of an atmosphere conducive to such entrepreneurial activity is still a major challenge for Europe, in spite of many coordinated efforts designed towards

\textsuperscript{34} Riding the wave: how Europe can gain from the rising tide of scientific data, available at: http://url.ca/9f2wt.

\textsuperscript{35} http://url.ca/9f311.


\textsuperscript{37} http://www.sienainitiative.eu/SIENA_Roadmap_Annexes.aspx

\textsuperscript{38} See for example, the Google contribution to the Public Consultation on Cloud Computing at http://url.ca/9f324.

\textsuperscript{39} See, e.g., the Cloudscape IV Executive Summary (Takeaways), highlighting the significant experience in federation gained by European e-infrastructure communities. This experience could be shared with colleagues from NIST, the US National Science Foundation and around the globe (op cit.).

\textsuperscript{40} http://www.sienainitiative.eu/SIENA_Roadmap_Annexes.aspx.
product-oriented research like Eureka/EuroStars\(^{41}\).

» **Rapid Pace of Change.** In cloud computing, the pace of change necessitates agility and flexibility. Efforts have to be made in Europe to support start-ups and facilitate acceleration of progress through appropriate legal, procurement and financial environment.

» **Collaboration efforts.** The open source communities and their open innovation process are paving the way to effective construction of new assets. Today these models and processes are significantly reused by industry and academic players. A significant support and legal framework needs to be provided for such initiatives to be sustainable and even more productive. Cloud computing is a very complex domain where isolated players will not be able to master the whole complexity. For these reasons, now is the time to take advantage of the valuable European open source communities to develop common value that can form foundations for new and innovative businesses and services.

Europe is well positioned to address many of the enumerated challenges, and indeed to take a world leading position in cloud computing, provided key challenges are met, notably to unlock\(^{42}\) and encourage a European cloud computing market for start-ups and small- and medium-sized enterprises. The EC Framework Programme on e-infrastructure\(^{43}\) and research infrastructures\(^{44}\) plays to key European competencies in research. The EGI\(^{45}\) is the largest, most powerful, and most comprehensive distributed computing e-infrastructure supporting research in the world\(^{46}\). Europe can also capitalize on the achievements of DCIs that have placed special emphasis on cloud computing and made concerted efforts to share experiences in the implementation of standards. VENUS-C is among the European DCIs leading the way in user-centric approaches to the cloud, demonstrating its feasibility for the ‘long tail of science’ and supporting a number of small businesses. Two of its pilots are closely associated with EDGI, helping to bring its desktop grid infrastructure into the cloud and thus extend and integrate the e-infrastructure. As the work of the StratusLab project focuses on the IaaS layer, there has been strong support of standards development in this area. StratusLab has established interaction with OGF, following closely the work on OCCI as well as on network management standards. Although not strictly a standards-related group, StratusLab has worked with the HEPiX community to define an Appliance Marketplace metadata format and develop the respective service\(^{47}\). Further, StratusLab and VENUS-C have sought to ensure productive collaboration by sharing experiences on implementations and use of OpenNebula, as well as clarifying interactions between IaaS and PaaS cloud infrastructures and relevant standards\(^{48}\).


\(^{42}\) Justin Pirie, *mimecast* cloud strategist: “Doing business in Europe is frankly a nightmare.”


\(^{46}\) The EGI comprises 346 contributing resource centers in 57 countries, with four non-European operations centers. The infrastructure supports 13,800 users in 219 virtual organizations, running 28 million computational jobs per month on 338 thousand processors, backed up by 107 petabytes of disk and 113 petabytes of tape storage. The infrastructure delivers 75 million processor wall-clock hours of computation per month to its user communities.


\(^{48}\) [http://url.ca/9f91v](http://url.ca/9f91v).
OpenNebula\textsuperscript{49} is an open source project developing solutions for building and managing virtualized data centers and cloud infrastructures. The e-infrastructure users of OpenNebula include EGI, StratusLab, EDGI, and VENUS-C\textsuperscript{50}.

CompatibleOne\textsuperscript{51} is an open source collaborative project providing a model and platform for the description and federation of heterogeneous cloud services.

The EGI-InSPIRE project\textsuperscript{52}, in collaboration with other EC-funded distributed computing initiatives and stakeholders\textsuperscript{53}, is driving the transition through virtualization from a grid-based to a cloud-based e-infrastructure. In particular, the Federated Clouds Task Force\textsuperscript{54} is working on the required standards and methodologies to implement a transition from grid- to cloud- based services for e-infrastructure provision to researchers in the European Research Area by building on current work. This task force has identified requirements from e-infrastructure stakeholders (resource centers offering cloud services, technology

\textsuperscript{49} See also section 5, Capitalizing on e-infrastructure investments in Europe, Boosting entrepreneurial opportunities and industry interaction. See also Ignacio Llorente, Open-source and Standards – Unleashing the Potential for Innovation of Cloud Computing, Position Paper and presentation, Cloudscape IV (op cit.).

\textsuperscript{50} See also Ignacio Llorente, Open-source and Standards – Unleashing the Potential for Innovation of Cloud Computing, Position Paper and presentation, Cloudscape IV (op cit.).

\textsuperscript{51} See also section 5, Capitalizing on e-infrastructure investments in Europe, Boosting entrepreneurial opportunities and industry interaction.

\textsuperscript{52} See also Ignacio Llorente, Open-source and Standards – Unleashing the Potential for Innovation of Cloud Computing, Position Paper and presentation, Cloudscape IV (op cit.).

\textsuperscript{53} European Desktop Grid Initiative (EDGI), European Middleware Initiative (EMI), Initiative for Globus Europe (IGE), StratusLab, and VENUS-C (Virtual Multidisciplinary Environments Using Cloud Infrastructure).

\textsuperscript{54} The outcomes of the first 6 months activities were presented at the EGI Community Forum, 26-30 March 2012, Munich, Germany, http://url.ca/9f3ft. StratusLab and VENUS-C are the two DCI projects that are currently most involved in the Task Force.
providers tendering solutions, and user communities consuming cloud services) for eight key federated cloud capabilities. An ambitious standards and profile development process for these six key capabilities within the OGF is planned, which will deliver recommendations for open standards by the end of 2012 through the DCI Federation Working Group (DCIFed-WG). Interaction with other SDOs with projects in this domain will benefit the further development of appropriate standards leading to interoperability of future cloud computing infrastructures.

Imminent future efforts such as the EC-funded two-year pilot project Helix Nebula will explore a European Cloud infrastructure for science with a commercial basis. The project will map demand from the European Space Agency (ESA) and CERN and other large capacity use cases with IT integrators and telecom companies on the supply side. A pilot testing phase will precede a 2014 rollout. Partners in the Helix Nebula project include OpenNebula and EGI.eu in addition to thirteen European industrial partners.

4. Call for Action – Short and Medium Term Priorities

The following recommendations constitute a “Call for Action” on all stakeholders involved in the development of European e-infrastructure. These recommendations have been formulated and divided into immediate short term priority actions, and actions with a medium-term timeframe. By short-term we mean within a maximum of twelve months, and by medium-term we mean a maximum of three years. These timeframes are the maximum amount of time that we feel it should take to complete each call for action, although we anticipate that each could be completed well within that timeframe. They take into account the life expectancies of the current DCI projects, and the probability of new calls for new projects by the European Commission.

Within this overall plan there is a role for governments where the trend towards procurement of commercial cloud services by the public sector (including research) will generate interest in standards. There is also an important role for industry to play in the international standards dialog, implementation and certification processes, as well as in continued investments aimed at boosting European innovation.

Immediate Priority Actions

We offer a list of actions, building on work already begun across Europe and complementing efforts globally, which merit attention within the next twelve months:

1. Determine optimum deployments of cloud computing for research. An urgent initiative is required to review current deployments of cloud computing for e-research and the broader scientific user and infrastructure communities at various service levels, (infrastructure-, platform-, software-, science-as-a-service) using a representative range of applications and determine the optimum deployment(s) for future use. This will require the collection of case studies and careful requirements analysis, and interaction with infrastructure, standards, software and application stakeholders. The latter would help to understand where emphasis should be placed, such as at which cloud service level a particular application is focused. Taking into account the impact of ICT on the environment and the use of green technology, the criteria must not only be about pure performance but on overall efficiency as well.

2. Strengthen, collaborative international dialogue for achieving interoperability and portability. It is imperative that the current standardization dialog be strengthened and widened to involve all stakeholders, including SDOs, cloud collaborative working groups, open source and commercial infrastructure providers, and other global and regional public sector institutions, in order to guarantee adequate representation of all the needs and to leverage the inherent strengths of each group to the benefit of the cloud computing environment as a whole in the international arena. Continuous dialog with

60 Infrastructure, platform, or software.
all key stakeholders internationally, on priority action plans is a foundation upon which to strengthen respective use cases, best practices and achieving true interoperability and portability. To this end, the EC and European member state governments should include direct support for funded activities to participate in relevant international activities. Furthermore, common ‘synchronized’ calls for proposals between the EC and the US National Science Foundation (NSF) and other national funding organizations, could strengthen international collaborations.

3. **Strive for a common approach for contributing to the European Digital Market.** A common approach for European Member States public administrations is necessary to help citizens and businesses profit fully from the European Union’s digital single market. The e-research communities’ contribution to this goal should be through the EC putting in place plans for the long-term sustainability and reuse of the assets from existing DCI projects. A condition of approval for future DCI projects should be the creation of reusable assets wherever possible for potential uptake by European government, industry and e-research and more widely wherever feasible. As part of plans for sustainability and asset reuse, attention needs to be paid to avoiding multiple independent parallel developments offering almost identical functionalities, and to recognize that these developments can also span diverse areas such as transport, energy or health.

4. **Expand support for DCI efforts to provide mechanisms to federate across multiple cloud suppliers.** A major near-term initiative, supported by the EC, is required to enable federation across multiple globally located cloud providers, including those in the SME bracket, to support European participation in global research enterprises. This federation must build upon the participation of European communities to build the necessary standards, and implement the software profiles and a common administrative framework. The EGI project will be exploring such participation during 2012 since there is a need for a compendium of virtual machines (platforms) to be available for user communities to deploy and link together on European e-infrastructures.

5. **Introduce measures to provide open access to all relevant SDO standard documentation.** The SDOs that can be regarded as producing open standards should be actively encouraged by the EC to share information and documentation about their standards work to avoid duplication of effort and provide better understanding to the many communities seeking to use their standards. Open standards that are readily available to whoever wishes to use them are necessary to support the procurement process by reducing the risk of becoming dependent on a single vendor. Vendors will implement standards if they are listed as technical specifications in calls for tender by public sector customers. One way of achieving this action would be to build on the current plans of the EC’s ADMS project on sharing repository metadata which are very relevant to this need.

6. **Introduce business models for use of clouds by research.** A set of common business models for the use of cloud computing by research communities is required. In order for cloud computing to be used effectively for research, certain special cloud services and properties will be required which are not driven by e-government or business applications. In some cases no good solution may be available, e.g. long-term data curation. There is general agreement that much research data must reside in the cloud, but there is currently no business model to support long-term sustainability, and government funding alone is not expected to be sufficient.

7. **Review public sector procurement regulations.** It is very important to allow the DCIs to liaise with governments to support large research-optimized “private” clouds. To this end a close examination of the existing European and national public sector procurement regulations is necessary to identify and remove any barriers to the procurement and re-use of DCI assets by governments and other public sector organizations before these projects close. Aspects that need to be addressed include: should there be an open tender process to purchase any of the assets; could an overarching framework contract be put in place to handle the disposal of assets; does the OJEU\(^{62}\) limit need to be changed; will the increasing trend for public sector procurements demanding solutions that use open standards be an obstacle for those DCI projects that have used proprietary or non-standard approaches.

8. **Re-use tangible and intangible assets produced by DCIs.** The DCIs should produce tangible and intangible assets together with associated information packages that can be made available to other parts of the public sector, SMEs and industry. The urgency of this action is necessary to protect the assets of those DCI projects that are scheduled to terminate in the short- to medium-term. The awareness of such information packages should be supported not only by DCI projects themselves, but also by the respective units of the funding organizations who should promote re-use of results in projects of other units where such results may be needed. As part of such packages, the DCIs should evaluate using a transparent procedure, explain the maturity of their offerings and include their sustainability plan for each asset. This will assist potential customers in making possible procurement and reuse/participation decisions. The production of this material should be a pre-requisite of future projects approved and funded by the EC. This action should include necessary publically published reviews of the deliverables from the current DCIs such that confidence is raised in the products that have been generated.

9. **Establish an e-infrastructure Implementation Advisory Group.** An e-infrastructure Implementation Advisory Group (eIIAG) should be established under the auspices of the EC to provide communication across all relevant stakeholders, i.e. the e-infrastructure community, the SDOs producing relevant standards, and those organizations and projects consuming their standards. The main roles of this group are to maintain a “cloud watch” on cloud computing projects, programs and service offerings; agree requirements and the standardization route for those requirements; share policy, regulatory and implementation advice; develop standards’ profiles; promote interoperability testing and plug fests/tests; and establish a compliance testing regime, all of which are essential components to achieving true interoperability. Finally, the Group should maintain close liaison with the e-Infrastructure Reflection Group, eIRG\(^{63}\). EC-funded projects should be required to take recommendations of the Group into account in their work. The Group must have representation from major e-infrastructures like EGI and Partnership for Advanced Computing in Europe (PRACE)\(^{64}\) in order to provide a single point of contact for e-infrastructure across Europe. To ensure that research communities have faith in its continued sustainability the Group must also have longer term funding (more than five years and possibly more than ten).

---

62. [Official Journal of the European Community](http://www.ojec.com/)


64. [http://www.prace-ri.eu](http://www.prace-ri.eu)
Medium Term Priority Actions

Within a medium term timeframe the following actions merit attention:

10. **Co-ordinate global regulations for transnational data flows.** An action that allows Europe to take the lead in coordinating the harmonization of regulations on the flows of data across borders is required. Many governments place restrictions on allowing data to be held outside their jurisdictions, primarily because of concerns about the adequacy of data protection legislation in other countries. The longer this situation persists, the slower the take up of cloud computing by the public sector will be. Therefore, this problem needs to be rectified.

11. **Create an inventory of grid and cloud standards for research.** A shared inventory of grid and cloud standards should be developed and maintained for the e-research stakeholder community covering both domain specific and e-infrastructure capabilities, with pointers to proven interoperable implementations and warnings of defunct or unused standards. This inventory should result from an EC-commissioned independent review and validation of current and proposed open standards and be aligned with the NIST Cloud Standards Inventory\(^{65}\) and the outputs of other relevant grid and cloud collaboration groups. The definition of a perhaps roughly defined but guiding reference model and associated standard-based reference architecture(s) should be explored as part of creating the inventory, in order to provide a frame of reference for standard adoption in a wide area of potentially different ICT sectors.

\(^{65}\) See NIST Inventory, op cit.
5. Capitalizing on e-infrastructure investments in Europe

Research is re-organizing itself in response to new data, communication and computing possibilities enabled by the cloud and demands for cost savings and greater efficiency. Aggressive investments are taking place in many parts of the world to support research, government and commercial organizations. Europe must seize the opportunity to build its innovative advantage through reinforced e-infrastructures by evaluating cloud computing in terms of its usability, performance and cost, whether as complementary to existing large-scale supercomputing and grid computing systems boosting scientific productivity, or as an accelerator for small research groups with new ideas as a motor for new specialized research services and entrepreneurial endeavors.

European e-infrastructures are widely recognized as critical infrastructures for e-science and open access to scientific data, as well as knowledge and innovation enablers. E-infrastructures (cyberinfrastructures in the US) are complex systems enabling seamless access to heterogeneous, independently managed resources, supporting e-science and e-research that exploit these advanced computational resources, data collections and scientific instruments. Europe, through ambitious EC programs, national research and infrastructure programs, has played a leading role in building multi-national, multi-disciplinary e-infrastructures. Significant investments by the EC over the past decade have thus resulted in capacity building in DCIs, supercomputing infrastructures, simulation software, and scientific data infrastructures, together with pan-European research infrastructures identified in the roadmap of the European Strategy Forum for Research Infrastructures (ESFRI). However, past achievements are no guarantee of future success. Without a long-term commitment to e-infrastructure, research communities in Europe will not benefit from the significant advantages of a new ecosystem and the current leadership in e-infrastructures within the European Research Area will not be guaranteed to continue in the future as in the past.

The European research and developer communities bring important insights into the implementation of distributed computing techniques (e.g. grid, HPC, cluster, supercomputing) construction that yield the highest performance research platform. In bringing the cloud model to the DCI arena, best practices around architecture should be maintained such that in the converged future, all the lessons learned from the spectrum of optimizations and special purpose systems are well understood and characterized. Leveraging the DCI experience, explicitly understanding the somewhat different cloud model and bringing a spectrum of services to the marketplace would place Europe in a much more sophisticated position than the countries for which cloud has been a one-dimensional “race to the bottom”.

---

Standards development and implementation

The cloud computing standards landscape is dynamic and complex with a plethora of organizations involved in some type of standardization effort. European research and development for distributed computing has played a leading role in standardization efforts, bringing knowledge and experience of value to other stakeholders in the public and private sector.

A core suite of early standards relevant to cloud initiatives has been developed by the Distributed Management Task Force (DMTF), the Storage Networking Industry Association (SNIA) and the Open Grid Forum (OGF):

» Open Cloud Computing Interface (OCCI) – developed by OGF. OCCI describes application programming interfaces (APIs) that enable cloud providers to expose their services. It focuses on IaaS based clouds and allows the deployment, monitoring and management of virtual workloads (like virtual machines), but is applicable to any interaction with a virtual cloud resource through defined http(s) header fields and extensions. While there are several open-source implementations, OCCI has not yet been widely adopted in commercial platforms.

» Open Virtualization Format (OVF) – developed by DMTF. OVF is a packaging standard designed to address the portability and deployment of virtual appliances. This is recognized as a DMTF, ANSI standard categorized under IaaS. There are firms who provide tools for conversion between various appliance formats, including OVF format to Amazon Machine Image (AMI) format.

» Cloud Data Management Interface (CDMI) – developed by SNIA. CDMI defines the functional interface that applications use to create, retrieve, update and delete data elements from the Cloud.

Complementary Cloud Plugfests jointly hosted by OGF and SNIA offer a forum for organizations to take their implementations of cloud standards based products to test, identify and fix bugs in a collaborative setting and thus develop interoperable products. Participants of Plugfests events sign a confidentiality agreement so that unannounced and/or unreleased products can be tested as part of the process while reducing efforts on legal overheads. The Plugfests also offer participants the opportunity to work with and gain knowledge from experts developing CDMI and OCCI, as well as to draw on the support of other developers. Participants at the last two Plugfests include CompatibleOne and VENUS-C. Future Plugfests are extending geographical outreach to China, demonstrating global demand for interoperability testing. Concerted efforts are also underway to broaden the synergistic approach through the support of other standards groups. Active participation in the Plugfests is a best practice that should continue in the future.

---

68 See, e.g. Markus Pleier, The Cloud Data Management Interface, Position Paper, Cloudscape IV (op cit.). On SNIA’s role in the European VISION Cloud project, see Dimosthenis Kyriazis, The VISION Cloud Data Mobility Approach, Position Paper, Cloudscape IV (op cit.).

69 Cloud Plugfests, 22 September 2011, Düsseldorf (Germany) and Santa Clara (US) and 28 February-1 March 2012 Düsseldorf.
Another potentially fruitful direction could be the organization of pooled Cloud Plugfests that implement the assets & user requirements analysis of funded initiatives. The ultimate goal would be to streamline the considerably high number of standards efforts, by bringing together industry forums and groups focused on end-user perspectives on common goals and providing insights and knowledge from developers. Further, application services and smart devices are already driving the market. Standards organizations that aim to cater for the community-expressed needs for portability, interoperability may prove effective in discouraging companies in continuing to build interoperable products and ecosystems by discarding relevant standards.

 Frameworks and ongoing work in other domains

Examples supporting functions and requirements of cloud computing relevant to e-infrastructures include OASIS Security Services (SAML) – providing the Security Assertion Markup Language, developed by OASIS. SAML is an XML-based framework for communicating user authentication, entitlement, and attribute information.70

 Established protocols

Over the past decade European research and development has focused on building large-scale e-infrastructures with attention to end-user facing systems with services such as authentication, authorization, accounting monitoring and service discovery, contributing to the development of several established protocols that are proving to be relevant to cloud infrastructures (see Annex 4)71:

- Basic Execution Services (BES72) and Job Submission Description Language (JSDL)73 used as default protocols for job submission and job life-cycle management for different types of execution (batch processing, workflows and MapReduce74).
- The Usage Record (UR)75 specification is designed to exchange usage and resource consumption records within the infrastructure. The Resource Usage Service (RUS)76 specification, based on the UR format, recommends an interface intended to accommodate requirements on usage auditing and accounting. Both are relevant to cloud infrastructures.

Other standards may emerge that enable interoperability between clouds and grids. For example, the GLUE77 standard from OGF provides an information model for describing grid and cloud entities while the

70 http://url.ca/9f92g.
74 http://www.mapreduce.org/.
Common Information Model (CIM)\textsuperscript{78} model from DMTF provides an alternative model used frequently in industry. The GLUE standards are being used within EGI to describe resources and derivatives of the Usage Record specification to aggregate accounting records on a European wide basis. Much of this information flow is now being supported by messaging technologies implemented through the Java Message Service (JMS) specification\textsuperscript{79}. Several agreed standard interfaces/schemas for interoperability between established middleware technologies are adopted in the European Middleware Initiative (EMI)\textsuperscript{80} and are continuously tested for compliance: Storage Resource Management (SRM)\textsuperscript{81} and GLUE. In addition to OGF standards, EMI also adopts several OASIS specifications such as SAML, XACML, and WS-Trust. Further activities include the addition of storage usage tracking to OGF UR. As a complementary effort, EMI has started to survey the needs of its key user stakeholders to identify requirements towards the use of new emerging cloud infrastructures in conjunction with EMI components and beyond (see Annex 4)\textsuperscript{82}.

However, to ensure standards implementation remains relevant and increasingly user-driven, e-infrastructure providers must set up effective mechanisms to gather requirements and ensure they are clearly defined. Without such definitions and conformance, little can be done to furnish standards-compliant solutions that meet any community requirements. They should also build on current standardization work, facilitate the identification of gaps and not seek to re-invent the wheel by taking a more pragmatic approach that speeds up the whole process.

In this respect the future looks bright. The DCI peer collaborative model that has been established provides a solid foundation for future implementation, testing and certification by leveraging the new knowledge and expertise acquired through current standards adoption\textsuperscript{83}. While it is important to note that project life-cycles inevitably impact on the extent to which standards implementation and interoperability testing takes place, such an approach could help overcome part of the complexity in the current landscape by sharing new knowledge to avoid duplication of effort. Further, an analysis of current and emerging standards implementation by e-infrastructures shows that established protocols for grid computing retain their relevance in cloud infrastructures, demonstrating the transferability of European knowledge and experience from past work, including relevance for supercomputing infrastructures (see Annex 4)\textsuperscript{84}.

A good example of collaboration across e-infrastructure stakeholders (resource centers offering cloud services, technology providers tendering solutions, and user communities consuming cloud services) stems from the Federated Clouds Task Force\textsuperscript{85}, which is coordinated by EGI.eu. The purpose of the Task Force

\textsuperscript{78} CIM Schema v2.30.0, September 2011. See http://www.dmtf.org/standards/cim.

\textsuperscript{79} JSR 343: Java Message Service 2.0 submitted to Java Community Process in March 2011, http://uri.ca/9f92t.

\textsuperscript{80} http://www.eu-emi.eu/.

\textsuperscript{81} http://www.ogf.org/documents/GFD.154.pdf.

\textsuperscript{82} http://www.sienainitiative.eu/SIENA_Roadmap_Annexes.aspx.

\textsuperscript{83} See, e.g., the interoperability layer that has been developed by the Barcelona Supercomputing Center for its COMPS superscalar (COMPSs) programming framework with the aim of making it interoperable with cloud technologies like Amazon, OpenNebula, Emotive and other OCCl-compliant offerings, http://uri.ca/9f8pn.

\textsuperscript{84} See also the presentations at the SIENA Roadmapping session during the EGI Technical Forum, 19-23 September 2011, Lyon, France, http://uri.ca/9f93e.

and related Test Bed\textsuperscript{86} is to support the transition through virtualization from a grid-based to a cloud-based e-infrastructure. Specifically, members of the Federated Cloud Task Force are working to produce a blueprint and profile by re-using existing assets and expertise within EGI and the DCIs involved. Coordinated activities focus on the required standards and methodologies to implement the transition and identifying requirements from e-infrastructure stakeholders (see Annex 4\textsuperscript{87}). As the EGI Federated IaaS platform will provide the foundation for all other platforms built on top of it, full standards coverage of that platform is paramount and a priority\textsuperscript{88}. An ambitious standards and profile development process for eight key capabilities within OGF is planned, which will deliver recommendations for open standards by the end of 2012 through the DCI Federation Working Group (DCIFed-WG)\textsuperscript{89}.

Interaction with the on-going work of other SDOs\textsuperscript{90} will benefit the development of appropriate standards leading to interoperability of future cloud computing infrastructures (see Annex 4\textsuperscript{91}). A number of international collaborations have also been established for interoperability testing, for example, between EGI, PRACE and the US Extreme Science and Engineering Discovery Environment (XSEDE)\textsuperscript{92}. These best practices could pave the ground for cloud computing strategies in regions where uptake is still immature for the benefit of international stakeholders, thus fostering further alignment between global standardization efforts. Further, an investigation into complementary mechanisms seamlessly using clouds in parallel to the well-established BES/JSDL set of standards that currently enables interoperability between EGI, PRACE, Resources linkage for e-science (RENKEI)\textsuperscript{93} and XSEDE would be a valuable contribution towards satisfying the needs of global research.

In order to support these initiatives, a major near-term initiative, supported by the EC, is required to enable federation across multiple globally located cloud providers, including those in the SME bracket, to support European participation in global research enterprises. This federation must leverage the participation of European communities to build the necessary standards, and implement the software profiles and a common administrative framework. The EGI project will be exploring such participation during 2012 since there is a need for a compendium of virtual machines (platforms) to be available for user communities to deploy and link together on European e-infrastructures. The significant experience gained with regard to federation will be a valuable asset that Europe could share with colleagues from NIST, the US National Science Foundation and around the globe.

\textsuperscript{87} http://www.sienainitiative.eu/SIENA_Roadmap_Annexes.aspx.
\textsuperscript{88} See, e.g., Steven Newhouse, EGI – European Grid Infrastructure, Position Paper, Cloudscape IV (op cit.) and presentation, EGI: Building a European Cloud through Federation, Cloudscape IV (op cit.).
\textsuperscript{89} http://www.ogf.org/gf/group_info/view.php?group=dcifed-wg.
\textsuperscript{90} Notably IEEE P2301 and P2302 http://grouper.ieee.org/groups/2301/ and http://grouper.ieee.org/groups/2302/.
\textsuperscript{91} http://www.sienainitiative.eu/SIENA_Roadmap_Annexes.aspx.
\textsuperscript{92} http://www.xsede.org.
\textsuperscript{93} http://www.e-sciren.org/index-e.html.
Leveraging expertise beyond research

The knowledge and experience gained by e-infrastructure providers can serve the whole public sector during the transition to cloud computing. The opportunity for government and cloud is primarily the adoption of standard services offered by the service provider community. It is expected that a significant proportion of government IT needs could be met by standard public cloud offerings for 5-10 years. Generic concerns about data security and governance, and service level agreements are common to both the public and private sector. Key issues for government today include interoperability between service providers and portability of data to allow a change of application service provider. Further, as the market develops, standardized billing systems are likely to become a requirement.

The relevance of DCI expertise for government adoption of cloud computing has been evaluated through an initial mapping between the scenarios and technical issues identified in the US government cloud computing program that has been defined in liaison with NIST and by Fraunhofer FOKUS (see Annex 4). Such mapping exercises are also valuable in aligning priorities and appropriate standards, which need further industry driven consensus processes to ensure wider implementation. Best in class approaches exemplified by e-infrastructure providers could therefore be extended to meet “business” needs.

Further, the underlying trend expected to generate interest in standards is the move to the procurement of commercial cloud services by the public sector, including the funding of research. Open standards will be needed to support the procurement process by reducing the risk of becoming dependent on a single vendor. If standards are listed as part of the technical specifications of calls for tender by the public sector, vendors would comply through implementation. To this end, public authorities need to gain a better understanding of existing standardization initiatives and define the level of interoperability required by e-government services in the cloud.

If Europe is to stay ahead of the game and address future challenges stemming from further uptake of commercial cloud services setting de facto standards, the European research community needs to reinforce its engagement in the standardization process and work towards greater awareness. It is vital that the research community fosters wider uptake of relevant standards by commercial organizations. A select industry group appointed by the EC has highlighted the role of industry in driving certification as a form of validation that can help build trust for end-users.

94 On the requirements identified by the UK government as expressed in the Government IT Strategy and Implementation Plans, see Ian Osborne, UK’s Government Cloud – What happened next?, Position Paper, Cloudscape IV (op cit.) and his presentation, Delivering Government Services through the Cloud, Cloudscape IV (op cit.).
95 Fraunhofer Institute for Open Communication Systems, Cloud concepts for the public sector in Germany – Use cases, August 2011, http://url.ca/9f8zm. See also Linda Strick, Cloud Concepts for the public sector in Germany – Use cases, Position Paper, Cloudscape IV and her presentation, Cloud Concepts for the Public Sector – Usage scenarios & use cases, Cloudscape IV (op cit.)
97 Industry recommendations to Vice President Neelie Kroes on the orientation of a European cloud computing strategy, November 2011, http://url.ca/9f8zy.
The use of voluntary internationally recognized standards and certification schemes is expected to avoid multiple, inconsistent and fragmented certification regimes, reducing costs while increasing impact\textsuperscript{98}.

An e-infrastructure Implementation Advisory Group (eIIAG) established under the auspices of the EC would be conducive to capitalizing on DCI expertise in the years to come. Specifically, the group would act as a bridge between all relevant stakeholders, including standards organizations, commercial organizations and projects adopting their standards. The Group should have representation from major e-infrastructures like EGI and PRACE in order to provide a single point of contact across Europe, and should liaise closely with the e-Infrastructure Reflection Group (e-IRG). This group should seek to build consensus on both requirements and solutions, maintain a “cloud watch” on cloud computing projects, programs and service offerings, share policy, regulatory and implementation advice, as well as promote interoperability and portability testing and a compliance testing regime in synergy with relevant policy, industry, and government players. To ensure that research communities have faith in its continued sustainability, the Group should also have longer term funding, that is, more than five years and possibly ten.

Driving e-infrastructure innovation

The European e-infrastructure has continued to evolve in response to both new user requirements and the advent of new technologies, creating a distributed computing infrastructure (DCI) ecosystem that integrates grid infrastructures underpinned by scalable middleware\textsuperscript{99} and more recently cloud computing infrastructures and virtualization planning. Europe must now chart a course for the transition to the next-generation DCI ecosystem in view of harnessing current assets and the objectives of Horizon 2020\textsuperscript{100}. Delivering Europe’s future e-infrastructure vision is critical to the first priority of Horizon 2020 - “excellent science” as the foundation of tomorrow’s technologies, jobs and well-being, developing and retaining research talent as well as on ensuring access to the best infrastructures.

Extending the research user base

Central to the EU2020 strategy is ensuring Europe becomes not only cloud-friendly but also cloud-active by responding to new requirements supported by software development and the creation of new services, which will also help to reinforce e-infrastructures. In the scientific arena, data intensive problems are

\textsuperscript{98} On approaches fostering industry self-certification in the UK, see http://www.cloudindustryforum.org.
\textsuperscript{99} http://www.wikipedia.org/wiki/Middleware.
\textsuperscript{100} Eric Badiqué, presentation Horizon 2020 – the EU framework programme for research and innovation, 30 November 2011.
gaining significant success with cloud computing as they have the closely coupled structure well supported in clouds and are typically “new applications” that are attractive to build from scratch with programming models suited for clouds. However, cloud is not a magic wand for all situations requiring some form of distributed computing. Hence efforts to extend the research user base should draw on insights into the benefits of cloud currently emerging from on-going initiatives as well as the continued need for grid and supercomputing, such as large-scale simulations.

Support for existing EGI user communities continues through existing channels such as the renewed synergy between EGI-InSPIRE and the Worldwide LHC Computing Grid (WLCG). This collaborative agreement ensures resources and expertise can continue to be exploited while providing a framework for managing change as EGI evolves to support researchers from new disciplines, including the humanities and social sciences, where the significant broadening of EGI’s customer base will contribute to EGI evolving into a sustainable, ubiquitous long-term e-infrastructure across Europe. On the other hand, new cloud European partnerships are emerging through flagships uses in fields such as high energy physics, earth observation and genomic assembly, offering a testing ground for new innovations and services. Here, the use of cloud computing is primarily seen as an effective way to provide additional resources rather than as an outright alternative to existing grid-based systems. On-going work with virtualization and cloud technology in the existing publicly managed grid infrastructure is expected to facilitate partial migration of workloads to the cloud.

Current DCI evaluations of cloud computing include case studies from research organizations with earlier experiences in grid and desktop computing (e.g. EGEE, EDGES), particularly in the field of complementary bioinformatics scenarios (EDGI, StratusLab and VENUS-C). Two of these case studies are being further supported by collaborative agreements aimed at delivering best-of-class efforts to reinforce and integrate the European cloud and desktop grid infrastructures, while other examples serve as a testing ground for a wider set of user scenarios relevant to the public and private sector. The evaluations also span fields as diverse as astrophysics, chemistry, civil engineering and architecture, civil protection and emergencies, healthcare, mechanical engineering, aerospace engineering (maritime security) and social media. Use includes seed fund allocation and voluntary experiments with a number of newcomers to distributed computing. More broadly, data intensive problems are gaining momentum because they have the ‘closely coupled structure’ well supported in clouds. They are typically “new applications” that are attractive to build from scratch with programming models suited for clouds. Open Calls have proven a cost-effective tool to extend the user base and could be further pursued in the future through clearer guidelines on sub-contracting procedures.

In summary, these early experiences can serve as a foundation for future assessments across different service levels as the e-infrastructure seeks to build its user base and cater to new requirements.

The next decade will see European e-infrastructure being used as a foundation for establishing multi-national, multidisciplinary research infrastructures such as those described in the ESFRI Roadmap. While the maturity of these individual projects varies, together they have common requirements that if provided consistently will contribute to EU 2020 objectives and provide cost-effective return on investment. For example, fundamental research and advances in areas such as biotechnology could trigger new opportunities.

for industrial leadership participation and private investment. A number of ESFRI projects also have a key role to play with regard to the third priority of Horizon2020 - societal challenges. Advanced computing, storage and data resources have an important role to play in supporting multi-disciplinary collaborative research aimed at driving breakthrough solutions in key areas such as climate change, the environment and energy, and challenges associated with an ageing population, among others.

Europe must seize the opportunity to analyze the requirements of ESFRI projects in order to pinpoint the best possible forms of distributed computing and related applications and services in the years to come. The EGI-InSPIRE project, for example, could leverage early collaborative work with the ESFRI projects, Common Language Resources and Technology Infrastructures (CLARIN)\(^{102}\) and Digital Research Infrastructure for the arts and Humanities (DARIAH)\(^{103}\) to define potential services to other projects, while all relevant DCI assets should be considered for re-use wherever feasible. Another pertinent example comes from European Bioinformatics Institute (EMBL-EBI), which is looking to minimize the movement of data, bringing the compute power as close to the data as possible through ‘virtual machines’. Algorithms are being parallelized, effectively creating a Grid solution, while cloud solutions are also being explored where commoditized data and computing can be utilized (see Annex 1\(^{104}\)). Finally, the EGI-InSPIRE project has an important role to play in evaluating cloud computing adoption strategies and trends at National Grid Initiative (NGI) level so as to harmonize future steps. Such an approach would reduce duplication of effort and fragmentation through bespoke solutions, while offering further opportunities to capitalize on ongoing standardization efforts.

The EC should support an urgent initiative to review current deployments of cloud computing for e-research and infrastructure communities at the various service levels (infrastructure-, platform-, software-, science-as-a-service) as well as large-scale deployments, with a representative range of applications in order to determine the optimum deployments for future uptake. Such an initiative should ensure interaction with infrastructure, standards, software and application stakeholders aimed at understanding where emphasis should be placed (e.g. at service level). Finally, feasibility criteria should include not only pure performance but also overall efficiency.

Boosting entrepreneurial opportunities and industry interaction

It is important to note that the European ICT market has its own particular structure with small and diversified and a general focus towards B2B. The European telecommunication industry is a major exception to specialized services provided by smaller players. This set-up calls for more work on integration, federation and interoperation to build up a European wide cloud ecosystem leveraging this diversification for richer

\(^{102}\) http://www.clarin.eu/external.
\(^{103}\) http://www.dariah.eu.
service provision\textsuperscript{105}. From a provider perspective, Europe currently appears to be lagging behind in terms of dominant players but is making strides in cloud by focusing support on key areas of strength while also ensuring concerns are appropriately addressed\textsuperscript{106}. To this end, future research needs to focus on cloud ecosystems powering business networks and the provisioning of domain specialized services and tools\textsuperscript{107}. Moving forward, there also needs to be a shift in investment focus towards applications and services across devices, that is, on what’s on top of the stack. Both industry and research adopters must better assess their investments with this in mind\textsuperscript{108}.

When it comes to cloud computing usage, cloud-powered open science and innovation paradigms can provide the engine for innovation by building closer links between academic institutions, industry, small- and medium-sized enterprises (SMEs) and technology parks across EU27, to leverage cloud computing through accelerator initiatives and dedicated support mechanisms that are harmonized across the Union. An appropriate legal, procurement and financial environment would strengthen support to startups and facilitate the acceleration of progress. As part of the drive towards making Europe ‘cloud-active’, initiatives like the European Parliament SME Group, the SME Action Plan\textsuperscript{109} and the European Enterprise Network\textsuperscript{110}, among others, aim to foster the uptake of innovative technologies, either SME-driven or through SME uptake of available technologies. However, government and industry also have a role to play in boosting Europe’s entrepreneurial spirit and finding more effective ways to bind local support and investments.

Currently, DCI success stories include the establishment of a new startup by a think tank serving as an early adopter of cloud computing for architecture and civil engineering\textsuperscript{111}. Early cloud usage in this field has led to the development and rollout of new cloud-based services for the design of eco-friendly buildings in a relatively short time-frame. Support for this startup comes from a combination of subsidized facilities at a local technology park and a national initiative for private investments for startups. In synergy with academic institutions, the start-up has created a set of complementary commercial services, demonstrating benefits to both the public and private sectors\textsuperscript{112}. A similar model, coupled with industry support programs, could be fostered across Europe, showcasing successful uptake and the creation of new services, including practical advice to academics and researchers thinking of taking new services to market. People who have already made the transition to the business world would be particularly suited to fulfilling this role. Examples include but are certainly not limited to the Swedish Institute of Computer


\textsuperscript{106} Robert Jenkins, CloudSigma, Research in Future Cloud Computing, 2 May 2012, EC, Brussels, Belgium, http://url.ca/9f9jl. See also Jan Duffy, Aiming High – Is the EU more or less Cloud-friendly and Cloud-active than other Regions?, Position Paper, Cloudscape IV, op cit..

\textsuperscript{107} Wolfgang Theilmann, SAP, Research in Future Cloud Computing (op cit.), http://url.ca/9f9lg.

\textsuperscript{108} Cloudscape IV Takeaways, op cit.

\textsuperscript{109} http://url.ca/9f93y.

\textsuperscript{110} http://portal.enterprise-europe-network.ec.europa.eu/. See also the European SME Envoy network, http://url.ca/9f9c0.


\textsuperscript{112} HUB-ENGINEERING (HUB-E), innovation cluster for architects and civil engineers, http://www.hub-e.com/services.
Science (SICS) Start-up Accelerator and SNIC Cloud Project\textsuperscript{113}, as well as the SME focus of EIT ICT Labs Technology Transfer, EIT ICT Labs Entrepreneurs Club, EIT ICT Labs Cloud Computing Research\textsuperscript{114}.

Collaborative paths to increase e-infrastructure and industry interaction are needed on multiple levels so as to unleash the full potential of public and private investments and expertise. One possible approach could be DCI know-how sharing through a consultancy process, as well as private sector guidance on marketing new services, patents, training and dedicated European startup support programs\textsuperscript{115}. Continued private investments at EU level would play an important role in co-funding further assessments of the cloud in addition to providing resources to initiatives like the EGI Federated Cloud Task Force alongside new software development.

Initiated in the context of pioneering EU-funded projects in cloud computing, OpenNebula.org is an open-source project developing an industry standard solution for building and managing virtualized data centers and cloud infrastructures. As an open, interoperable cloud enabler, OpenNebula is providing significant business opportunities to the ICT Industry, including SMEs, helping them play a more active role in adopting cloud computing, developing more innovative cloud products, shaping the development of innovative cloud technology, and bringing to market interoperable cloud solutions and services. In research environments, OpenNebula is being used as an open platform for innovation and interoperability in leading research and infrastructure projects, and as a reference implementation of cloud standard specifications. OpenNebula is widely used in supercomputing centers, research labs, and e-infrastructures to build and manage their own cloud computing systems for hosting virtualized computational environments or for providing users with new “HPC as a service” resource provisioning models\textsuperscript{116}.

CompatibleOne is an open-source collaborative project providing a model, CORDS, and a platform, ACCORDS, for the description and federation of different clouds comprising resources provisioned by heterogeneous cloud service providers. For example, the CompatibleOne platform can automatically provision and distribute a workload on OpenNebula for its application part and on OpenStack for its database. The CompatibleOne model and platform are based on key open standards such as OGF OCCI and, in the coming months, SNIA CDMI. The flexible service architecture of CompatibleOne is independent of any cloud service provider (from OpenStack to OpenNebula, from Azure to Vcloud) and can address all types of cloud services (IaaS, PaaS, SaaS, XaaS, BpaaS, ...). Services provided by the CompatibleOne platform can be used for different cloud service deployments, e.g. public, private, community and hybrid clouds. CompatibleOne provides the core services described in the NIST Reference Architecture definition of a cloud service broker: intermediation, aggregation and arbitrage. CompatibleOne aims at being interoperable with most platforms so as to provide maximum freedom of choice to cloud service consumers and developers. CompatibleOne provides interoperability, portability and reversibility and thus helps break vendor lock-in.

\textsuperscript{113} These initiatives are cited as readily available examples through current DCI interaction, http://www.sics.se/node/8457 and http://www.sics.se/startupaccelerator.


\textsuperscript{115} See for example http://url.ca/9f9cw. For perspectives from SME associations, see http://url.ca/9f9db.

\textsuperscript{116} http://www.opennebula.org.
CompatibleOne is part of the OW2 Open Source Cloudware initiative (OSCi)\textsuperscript{117}.

The European telecom industry, which is facing increasing competition in the global marketplace, can also play a key role in supporting new cloud partnerships\textsuperscript{118}, helping to pinpoint novel ways in which to boost European scientific productivity. Further, the GÉANT 2020 Vision\textsuperscript{119} proposes new roles for the European telecom industry and industry interaction. Future industry-research collaborative work could also focus on new ways to deal with the current imbalance between increasing amounts of data traffic and revenue streams\textsuperscript{120}, addressing issues like data ownership and security, developing more user-friendly services and building stronger links with the public sector in areas as diverse as energy, health and transport. Such an approach would be conducive in overcoming current separate silos of knowledge. Finally, dynamic and commercially supported initiatives like OpenStack\textsuperscript{121} and jclouds\textsuperscript{122} could also impact on future e-infrastructures and should be appropriately monitored.

Moving forward, DCI sustainability plans will also need to ensure that collaboration between public and private infrastructure provision surpasses the state-of-the-art. Pinpointing proof of concepts and collaborative research topics between providers can advance European capabilities across common use cases. Steps are needed to channel investments and sponsorship at European level to help boost the creation of jobs, skills development, scholarships and new services. However, the legal frameworks needs to break down current barriers rather than create new ones and provide incentives. Further collaboration in these areas can help advance Europe collectively in its cloud development. Achieving these objectives will be important for the second priority of Horizon 2020 - industrial leadership - by attracting more private investment in research and innovation and increasing SMEs to create growth and jobs.

Cloud Watch

While the DCIs themselves are a strong starting point for such joint research, often with industry stakeholders integrated in the project from the outset, there is a pressing need to drive dedicated actions, education and awareness-raising programs coupled with an evaluation of the EU27 landscape in synergy with SME, industry

\textsuperscript{117} http://www.compatibleone.org and http://www.ow2.org/view/Cloud/.
\textsuperscript{118} Strategic Plan for a Scientific Cloud Computing Infrastructure for Europe, http://cdsweb.cern.ch/record/1374172/files/CERN-OPEN-2011-036.pdf. See also Maryline Lengert, Cloudscape IV, op cit...
\textsuperscript{119} Knowledge without Borders - GÉANT 2020 as the European Communications Commons, October 2011, http://url.ca/9f9dw.
\textsuperscript{120} Knowledge without Borders, op cit., p. 34, note 21: “Thus, in boardrooms across Europe executives are trying to square the economic circle, that is, how to continue to expand the network while at the same time generating revenue from the relentless demand for mobile data and services”.
\textsuperscript{121} http://openstack.org/.
\textsuperscript{122} http://www.jclouds.org/documentation/quickstart/openstack.
and tech associations.
To this end, an active ongoing openly accessible “cloud watch” would be very helpful to share information on the vast array of projects and programs devoted to cloud computing exists, with new ones appearing almost daily. The EC’s large collection of white papers and communications on cloud computing could form the basis for a “cloud watch” aimed at clarifying current terminology and how it differs from traditional terms of reference. The “cloud watch” could also play an educational role for decision takers at all levels of business and government, including politicians. This educational role would also help overcome social and cultural obstacles to cloud computing adoption and inertia, as well as ensure that the standardization and security challenges are met appropriately with new skills development in areas such as data architects and contract negotiators.

A strategy for reusable DCI assets

The fast-evolving cloud computing landscape strongly suggests that isolated players will not be able to master the whole complexity. Now is the time to take advantage of the valuable European open source communities to develop common values that can form foundations for new and innovative businesses and services, which can be reused by both academic and commercial organizations. A significant support and legal framework would provide a foundation for more sustainable and productive initiatives in the future. Sustainability is defined as the ability to survive beyond a project life-cycle through both external contributions and internal activities, demonstrating that it is a trusted and demand-driven solution. Sustainability is even more important in the absence of established standards and true interoperability, as users may end up locked-in to a platform that has no clear long-term strategy, and unable to move to potential alternatives without high migration costs.

The long-term sustainability of distributed computing infrastructures, both of the structures themselves and the assets they create (like software, knowledge, experience) is very much part of the SIENA vision as a challenge facing all DCIs, current and future ESFRI projects[^123], as well as a number of scientific data infrastructures. A good case in point is the EGI/NGI survey analysis, which highlights the short-term allocation of funding of the National Grid Initiatives (NGIs) on a one to two-year basis, as well as the uncertainty of funding in some instances. It seems that in most cases, a sustainability funding scheme is still not defined yet funding beyond 2014 is critical[^124]. Further pushing the need for sustainability, public administrations (PAs) need a clear long-term vision before committing to the use of a project or infrastructure (taking into consideration the fact that most PAs must provide services for longer time-frames than commercial entities – sometimes 10 years or more). A project critically dependent on a single funding source, or for which a clear roadmap is not apparent may not be qualified to be used as a basis for a long-term IT project by its

[^123]: Regarding the emphasis on ESFRI project implementation and sustainability, see Report on a consultation workshop on the possible content of “Horizon 2020” - Common Strategic Framework for Research and Innovation (2014-2020), op. cit.
public users.
A number of existing models for tangible and intangible assets could be leveraged by individual initiatives, though it is important to take on board any legal constraints for offerings targeted at non-governmental groups or impacting on possible service offerings. Planned sustainability studies could also draw on the outcomes of dedicated workshops\textsuperscript{125}, as well as in the context of standardization and the creation of standardized platforms as an important step in facilitating sustainability. An example may be the European Middleware Initiative (EMI) and its recent activities on creating an open source software initiative\textsuperscript{126} that is largely driven by the demand to explore sustainability options for middleware developments and maintenance in Europe.
A sustainability analysis starts by identifying tangible or intangible “assets” which provide a differentiating value. From a DCI perspective, assets include infrastructure; software/source code; knowledge; non-material assets and services. These assets serve as potential tools for obtaining external contributions (monetary or otherwise) that can support the long-term viability of a platform or reduce the current costs of maintaining it. An organization may decide to employ a mixed model - that is, an offering directly or indirectly monetized based on the combination of assets, for example, a Platform-as-a-Service (PaaS) offering that integrates software, support and infrastructure. Examples of large-scale cooperation in action include the OpenStack project, which has aggregated over 2000 contributors and 140 companies around several sub-projects in a very short time-frame by fostering collaborative work.
Research has showed that sustainability is bound up with the number of participating developers whereas niche size does not seem to impact significantly on sustainability\textsuperscript{127}. Equally important is intrinsic motivation, that is, contribution probability without a direct, immediate return. Research shows that the credibility of the project sponsor is the main driver behind a successful community of contributors\textsuperscript{128}. This is extremely important for communities created by research projects, which can usually count on substantial credibility in terms of status and their scientific research, thus providing an asset in its own right. The DCIs should therefore dedicate effort to defining value propositions around their assets, potential sponsorship while taking steps to create a community at strategic phases of the project. One possible way of addressing DCI challenges in terms of engagement with commercial entities could be the creation of a separate spin-off company that acts as a single intermediary for the acquisition of resources which are then offered to public administrations or other organizations. This separation of roles guarantees that administrative and legal burdens are reduced on the part of the DCIs, while ensuring support and commercial activities required in a traditional monetary based exploitation are transferred to the spin-off.

**Combined models:** DCIs are in a unique position to capitalize on both tangible and intangible assets. Models that help exploit both types of assets should therefore be investigated. Examples might include offering tuned software designed explicitly for the hardware provided as an execution platform or creating

\textsuperscript{125} One example is the EGI User and General Sustainability Workshop, 24-26 January 2012, Amsterdam, Netherlands https://indico.egi.eu/indico/conferenceDisplay.py?confId=709.
\textsuperscript{126} http://sciencesoft.web.cern.ch/sciencesoft/.
\textsuperscript{127} “Sustainability of Free/Libre Open Source Projects: A Longitudinal Study” InduShobha Chengalur-Smith, Anna Sidorova, Shrae Daniel; Journal of the Association for Information Systems, November 2010.
\textsuperscript{128} “How Firms Make Friends: Communities in Private-Collective Innovation ” Matthias Stuermer, Doctoral Dissertation No. 18630, ETH Zürich.
a combined hardware-software platform that can be offered as a higher-order package compared with current asset providers like Amazon. These PaaS offerings are becoming increasingly relevant especially for commercial organizations that are just starting to face issues like manipulating, analyzing and moving big data sets or using specialized algorithms that until a few years ago more typical to the research environments. Amazon Elastic MapReduce is one example of this evolution, providing on-demand Hadoop service. However, opportunities for more complex environments like scientific workflows or language execution environments exist and could offer a differentiating value as an advantage for DCIs to exploit.

**Below:** some of the most common examples of business model taxonomies using these combined blocks for cloud computing

<table>
<thead>
<tr>
<th>Model</th>
<th>Components</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cloud Builder</strong></td>
<td>Software, knowledge (&quot;create the cloud for a third party using the same tools and expertise that the DCI uses internally&quot;)</td>
</tr>
<tr>
<td><strong>Cloud Tool Vendor</strong></td>
<td>Software (&quot;sell or provide as a service the software the DCI uses internally&quot;). If the software is open source, it implies a model that is based on proprietary aspects like Open Core or Dual Licensing.</td>
</tr>
<tr>
<td><strong>Cloud Infrastructure provider</strong></td>
<td>Software, Infrastructure, Knowledge (&quot;provide an Infrastructure or a Platform as a service&quot;)</td>
</tr>
<tr>
<td><strong>Cloud Broker</strong></td>
<td>Software, Infrastructure, Financial or Contractual assets (&quot;accept requests from cloud users and dispatch them through an internal infrastructure, through external services using our software platform&quot;)</td>
</tr>
</tbody>
</table>

E-infrastructures in general and DCIs in particular should aim to produce tangible assets that can be made available to other parts of the public sector, SMEs and industry. The urgency of this action is necessary to protect the assets of the DCI projects that are scheduled to close in the short to medium term. The awareness of such a package of information should be supported by not only DCI projects, but also by the respective units of the funding organizations by promoting re-use elsewhere as much as possible so as to reap their full potential without duplicating effort. Wherever possible, the package should indicate the scope, transparent re-use procedure and maturity of individual assets as part of their sustainability plan in order to facilitate the decision-making process, possible procurement and re-use. DCIs can also leverage long-standing strengths such as knowledge, a vast array of software tools that are tried and tested in very large configurations to underpin their differentiating value. Statutory limitations pertaining to research institutions could be obviated through the support of participating organizations managing commercialization. Asset packages should be a prerequisite of all future projects funded by the EC, appropriately furnished with practical guides.

Further, a common approach to Member State public administration is needed to ensure citizens and businesses profit fully from the EU’s digital single market. The EC should put in place plans for the long-term sustainability and reuse of assets from existing DCI initiatives. Wherever possible, the creation of reusable assets for potential uptake by government, industry and e-research should be a condition of approval for
future DCI projects. To this end, greater attention needs to be paid to avoiding multiple independent parallel developments offering almost identical functionalities while also recognizing that these developments can also span diverse areas such as transport, energy and health, among others.

6. Future directions - driving positive change for Europe in a global context

As the market for new technologies grows and as Europe moves towards the implementation of an Interoperability Framework for use by the public sector in the Member States, interoperability and portability of data remain pressing issues that need addressing as central to ensuring broader choice and fairer competition. Allied to this is the persistent need to address legal impediments such as issues surrounding data transfer regulations. If Europe is to drive positive change in a fast-evolving global landscape, it must foster a coherent and pragmatic approach to current and emerging challenges through an internationally coordinated dialog with multi-stakeholder engagement for the collective public good.

Open, supplier and user consensus-driven processes are essential to driving the implementation of a core set of internationally recognized standards and avoiding multiple, inconsistent guidelines and bespoke solutions. From a user perspective, open standard interfaces protect from vendor lock-in, thus avoiding significant migration costs wherever open interfaces are not provided. Further, there are persistent concerns around interoperability, including efficient resource utilization. The portability of applications and associated data across cloud services from different providers also needs addressing to ensure multiple suppliers are capable of providing cloud services with similar interfaces. Finally, there are also concerns about reversibility (moving data from cloud to non-cloud environments). While these requirements are relevant to cloud uptake in both the private and public sector, interoperability of public services is particularly complex as it is critical all the way up the interoperability stack, that is, at legal, organizational, semantic and technical levels. The entire interoperability stack is captured in the European Interoperability Framework 2.0\textsuperscript{129} (see Annex 4\textsuperscript{130}), which provides a common vision and policy guidelines for setting up public services in the Member States due for implementation in 2013.

The importance of addressing multiple interoperability levels is now being recognized\textsuperscript{131} with new efforts to chart a course in this direction. The OASIS Technical Committee on Transformational Government\textsuperscript{132} is aimed at providing an overall framework, identifying use cases and providing guidance to governments around the world\textsuperscript{133}.

Common approaches to the development of standards of practice for online operation in the area of trust,

\textsuperscript{129} http://ur1.ca/9f9ee.
\textsuperscript{130} http://www.sienainitiative.eu/SIENA_Roadmap_Annexes.aspx.
\textsuperscript{132} http://ur1.ca/9f9em.
\textsuperscript{133} On standards development and OASIS, see, e.g., James Bryce Clark, What are the Standards Bodies doing & how do they respond to challenges from the community?, Position Paper, Cloudscape IV (op cit.).
security and certification would be valuable. Industry-led codes of practice\textsuperscript{134} would play an important role in increasing the transparency, capability and accountability for suppliers of cloud services, for example, through self-certification processes, enabling potential contractors and users to make informed decisions. However, significant industry acceptance is still missing in current standardization work but is essential to facilitate the coordination of future. Wider consensus is needed on existing and emerging standards, including but not limited to:

» Cloud profiles: Guide for Cloud Portability and Interoperability Profiles (CPIP) – developed by IEEE Standards Association (IEEE SA)\textsuperscript{135} (P2301)\textsuperscript{136}. The guide is aimed at informing cloud vendors, service providers and users with information on standards-based application, portability, management and interoperability interfaces, as well as file formats and operation conventions with definitions (“profiles”) drawn from multiple sources.

» The Standard for Intercloud Interoperability and Federation (SIIF) – developed by IEEE SA (P2302)\textsuperscript{137}. This standard will define topology, functions, and governance for cloud-to-cloud interoperability and federation. Topological elements include clouds, roots, exchanges (which mediate governance between clouds), and gateways (which mediate data exchange between clouds).

» The Identity in the Cloud Technical Committee (TC)\textsuperscript{138} – led by OASIS. Activities are aimed at addressing serious security challenges posed by identity management in cloud computing, identifying gaps in existing identity management standards and investigating the need for profiles to achieve interoperability within current standards. Risks and threat analyses on collected use cases are performed with guidelines for mitigating vulnerabilities.

» The Privacy Management Reference Model (PMRM) Technical Committee – led by OASIS. Activities focus on providing a standards-based framework that will help implement privacy and security policies in cloud operations. A guideline or template for developing operational solutions to privacy issues is foreseen to assist practical implementation.

» The Topology and Orchestration Specification for Cloud Applications (TOSCA) Technical Committee\textsuperscript{139} – led by OASIS. The work aims to enhance the portability of cloud applications and services to enable the interoperable description of application and infrastructure cloud services, the relationships between parts of the service, and the operational behavior of these services (e.g., deploy, patch, shutdown), independent of the supplier creating the service, and any particular cloud provider or hosting technology. OASIS standardization efforts thus cover security, identity privacy and portability.

» Distributed Application Platforms and Services (DAPS) – developed by ISO/IEC (JTC1 SC38)\textsuperscript{140}.

\textsuperscript{134} \url{http://www.cloudindustryforum.org/},

\textsuperscript{135} The development of standards by IEEE SA occurs simultaneously with the Industry Connections Program aimed at involving, stimulating and transferring technologies to industry as a breakthrough approach, \url{http://standards.ieee.org/develop/indconn/index.html}. See also, David Bernstein, The IEEE and Cloud Computing, Position Paper, Cloudscape IV (op cit.).

\textsuperscript{136} \url{http://grouper.ieee.org/groups/2301}. Group established July 2011. A draft standard will be released in 2012.

\textsuperscript{137} \url{http://grouper.ieee.org/groups/2302}. Group established July 2011. A draft standard will be released in 2012.

\textsuperscript{138} \url{http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=security#technical}.

\textsuperscript{139} \url{http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=tosca}.

\textsuperscript{140} \url{http://url.ca/9f9eu}, developed in synergy with DMTF, INLA, ITU, OASIS and SNIA.
Standardization work focuses on interoperable Distributed Application Platforms and Services and is scoped to cover web services, service-oriented architecture (SOA) and cloud computing. The cloud computing working group focuses on cloud computing reference architecture, terms and definitions; investigates requirements for new work in the area of use case analysis methodology and principle of cloud service delivery underpinned by a use case-based approach.

- Technical Committee Cloud – led by ETSI141. The technical scope covers resource and service access, protocols, middleware and security with the aim of complementing current work through collaboration with OGF and Collaboration on Networked Resource Management (OGF SCRM-CG group) and participation of DMTF, OASIS, SNIA, the Tele Management Forum (TMF), Internet Engineering Task Force (IETF), International Telecommunication Union (ITU) and the World Wide Web Consortium (W3C). ETSI is also investigating the interoperability of heterogeneous elements and lawful interception/data retention with validated standards (Plugtests) to help ensure interoperable standards and that interception capability in the cloud is present “in country” when required. Cloud providers could provide a framework enabling them to provide lawful interception capability. This collaborative work is expected to lead to a Roadmap for interoperable ICT cloud standards.

- Cloud Data Governance Working Group – led by Cloud Security Alliance142. Activities focus on a survey phase aimed at understanding the top requirements and needs of different stakeholders on governing and operating data in the Cloud, and a best practices recommendation phase aimed at prioritizing and answering of the key problems and questions identified by Cloud stakeholders in Phase 1.

However, attention must also be paid to de facto industry standards already in place, including a wealth of existing interoperability between industry leaders, Amazon, Microsoft, VMWare and others. This is particularly important as existing grid and cloud standards could be challenged as more use is made of commercial clouds, increasing the importance of open standards implementation by commercial organizations. Further, the impact of dynamic and commercially backed groups such as the Open Data Centre Alliance143, OpenStack and jclouds and any related standards work could have an impact on the e-infrastructures of tomorrow and therefore need to be taken on board.

With regard to issues like security, an important role has been played by the European Network and Information Security Agency (ENISA), such as the report on Opportunities for European Cloud Computing Beyond 2010 – Security and Resilience in Governmental Clouds144. Other relevant documents include a Cloud Computing Risk Assessment145 providing an in-depth and independent analysis that outlines some of the information security benefits and key security risks of cloud computing; an Assurance Framework146 with criteria designed to assess the risk of adopting cloud services, compare different cloud provider offers, obtain assurance from the selected cloud providers and reduce the assurance burden on cloud providers.

142 https://cloudsecurityalliance.org/research/cdg/.
143 http://www.opendatacenteralliance.org/.
in addition to ongoing work. The Cloud Security Alliance (CSA)\textsuperscript{147} focuses on research around security, data protection governance, assurance, compliance, interoperability and portability in cloud computing. The CSA has been swift in organizing security issues into categories, clarifying which issues are directly related to cloud computing and which ones are broader issues. Further, CSA has produced two practical and publicly available documents, one on ‘cloud security recommendations’ and one on ‘cloud security best practices’. The CSA paper, ‘Contribution to the European Commission Strategy on Cloud Computing’ sets out short and medium actions that support the secure adoption of cloud computing across the EU with particular reference to interoperability and portability; trust security and assurance; security innovation in the cloud\textsuperscript{148}.

Internationally, we have witnessed synergies between the US National Institute of Standards and Technology (NIST) with European initiatives such as SIENA and ETSI. Work in the Japan, through the Global Inter-Cloud Technology Forum (GICTF)\textsuperscript{149} and the Electronics Standardization Institute (CESI)\textsuperscript{150} in China are both potential partners in evaluating cloud potential standards relevant for European e-infrastructures. One of the Japanese national research and development projects has been working on inter-cloud computing technologies for the last three years\textsuperscript{151}. Particular emphasis is placed on an architecture to establish standard interfaces and working towards the standardization of network protocols and the interfaces through which cloud systems interwork with each other, fostering international interworking of cloud systems. For example, GICTF is contributing to global standardization through the International Telecommunication Union Standardization Sector (ITU-T) and other standards organizations. Collaborative technology evaluation efforts are also important with regard to developing R&D test beds for these technologies. However, the advent of cloud computing strategies in BRICSS (Brazil, Russia, India, China, South Africa, South Korea), the South Pacific (Malaysia, Taiwan, Thailand) and Australia, Canada, New Zealand and Singapore calls for concerted efforts across the international arena.

It is vital that international efforts avoid operating in a vacuum with further invention by global standards bodies. What is required is a mapping of standards and their scope, followed by a common sense synthesis of what exists, with some more general guiding principles, fostering information exchange on successful ongoing efforts that do not unduly overlap but that can be combined to be used together in software implementations.

To this end, a shared inventory of grid and cloud standards should be developed and maintained for the wider stakeholder community covering both domain specific and e-infrastructure capabilities, with pointers to proven interoperable implementations, indicating whether versions of the standard have been published, tested, deployed, matured and/or deprecated. This inventory should result from an EC commissioned and independent review and validation of current and proposed open standards, aligned with the NIST Cloud Standards Inventory and the outputs of other relevant grid and cloud collaboration

\textsuperscript{147} https://cloudsecurityalliance.org/.

\textsuperscript{148} http://url.ca/9f8nk. See also Daniele Cattedu, Trust, Security & Legal Issues, Position Paper, Cloudscape IV (op cit.) and his presentation Cloud Security Alliance – what is happening in 2012, Cloudscape IV, op cit.

\textsuperscript{149} http://www.gictf.jp/index_e.html.

\textsuperscript{150} http://www.en.cesi.cn/.

\textsuperscript{151} See, e.g., Atsuhiro Goto, Inter-cloud computing: use cases and requirements for social infrastructures, Position Paper, Cloudscape IV, op cit.
groups. Further, the definition of a roughly defined but guiding reference model and associated standard-based reference model should be explored as part of creating the inventory, so as to provide a ‘frame-of-reference’ for standard adoption in a wide area of potentially different ICT sectors. The inventory should include standards implementation listed in the technical specifications of calls for tender by public sector customers. Finally, the current plans of the EC’s ADMS project on sharing repository metadata could be very relevant to achieving the action.

With regard to the reuse of DCI assets, it is very important to allow the DCIs to liaise with governments to support large research-optimized “private” clouds. To this end, a close examination of the existing European and national public sector procurement regulations is necessary to identify and remove any barriers to the procurement and re-use of DCI assets by governments and other public sector organizations before these projects close. Aspects that need to be addressed include: should there be an open tender process to purchase any of the assets; could an overarching framework contract be put in place to handle the disposal of assets; does the OJEU limit need to be changed; will the increasing trend for public sector procurements demanding solutions that use open standards be an obstacle for those DCI projects that have used proprietary or non-standard approaches.

Legal impediments & data issues

There is a general consensus that data and legal issues represent a real barrier to the widespread adoption of cloud computing (see Annex 5)\(^1\). There are general concerns in the industry impacting both public and private sector users, further complicated by country specific laws at odds with service provider regional investment strategies and lax policies for replicas and data flow management. Specifically, from a vendor perspective, barriers to innovation and uptake are seen to include divergent national regulatory frameworks creating difficulties in achieving required scale when different rules pertaining to data location as well as business model transformation since the cloud model requires all ICT companies to change their business model without creating significant disruption\(^2\). While the situation is improving, the absence of genuine management policies and controls will inhibit widespread adoption of services in the short to medium term.

The EC Select Industry Group has provided an in-depth analysis of the European Data Protection Directive (95/46/EC)\(^3\), highlighting a number of substantial impediments for business caused by the lack of harmonization. These issues are particularly important for services based on cloud computing as they most often rely on cross-border data flows and also derive benefits from economies of scope and scale. In summary, benefits and opportunities for cloud computing are significantly limited by a) how data is defined, maintained, processed secured or deleted and b) variations in filings related to collection processing, use

---

1\(^{1}\) On data challenges in the scientific realm related to cloud computing, see, e.g. Dennis Gannon, Cloud Policy, Big Data and the Long Tail of Science, Position Paper, Cloudscape IV, op cit.

1\(^{2}\) Industry Recommendations to Vice President Neelie Kroes, op cit.

or storage of data, resulting in divergent limitations on collection, use or transfer of data. Harmonization is needed to address the important differences in the way the Directive is implemented across EU27. However, an important conclusion relates to restricting the movement of data as contrary to the well-known advantages of cloud computing.

In the scientific arena, data challenges have been addressed by the High Level Expert Group on Scientific Data, calling for immediate attention while taking on board emerging trends in open access strategies and early sharing of research results through dedicated forums, which is expected to impact on traditional practices. While the Stork project\textsuperscript{155} is addressing the need for a coherent strategy for federation of identity recognized across the European Union, future international cooperative work needs to cover comparable Quality of Service (QoS) measures, vocabularies for Service Level Agreements (SLAs) and dashboardability, as well as all contractual issues (scope, SLAs liability cover, risk and governance)\textsuperscript{156}.

Two actions are recommended to address issues pertaining to long-term data curation and the harmonization of data flow regulations.

Firstly, the provision of a set of common business models for the use of cloud computing by research communities is required. In order for cloud computing to be used effectively for research, certain special cloud services and properties will be required which are not driven by e-government or business applications. In some cases no good solution may be available, e.g. long-term data curation. There is general agreement that much research data must reside in the cloud, but there is currently no business model to support long-term sustainability, and government funding alone is not expected to be sufficient.

Secondly, a mid-term action that allows Europe to take the lead in coordinating the harmonization of regulations on the flows of data across borders is required. The primary objective is to accelerate the uptake of cloud computing by the public sector by overcoming current impediments.

Calling for an international, multi-stakeholder dialog

Several major governmental cloud computing programs are underway (see Annex 4\textsuperscript{157}). These are being pursued in the absence of an adequate set of agreed and accepted open standards that could guarantee security, portability and interoperability of cloud services, yet the globalization of business implies many of the issues are global, calling for concerted international efforts.

Pragmatic international cooperation has already identified both near-term (see Annex 4\textsuperscript{158}) and future priority areas, paving the way for further cooperation to encompass emerging initiatives globally. Future priorities should focus on comparing and strengthening respective use cases in Europe and globally, with the possible implementation of test beds on strategic adoption techniques, creating a bridge between countries with uptake experiences and those countries where uptake is still immature. The mapping of priority plans

\begin{itemize}
  \item \textsuperscript{155} https://www.eid-stork.eu/.
  \item \textsuperscript{156} OASIS International Cloud Symposium, op cit.
  \item \textsuperscript{157} http://www.sienainitiative.eu/SIENA_Roadmap_Annexes.aspx.
  \item \textsuperscript{158} http://www.sienainitiative.eu/SIENA_Roadmap_Annexes.aspx.
\end{itemize}
should encompass frameworks to support the seamless implementation of federated community cloud environments and collaborative parallel strategic ‘future cloud’ initiatives for the collective public good. From a government perspective, emerging hiring trends such as service management, contract experts and data architects provide early experiences of value to peers nationally and internationally. It is therefore important to ensure guidance to architects and engineers by facilitating dialog between government officials and standards organizations in Europe and globally. A conformity assessment system is required to leverage certification and accreditation and reduce the burden on industry, particularly SMEs, and government. Finally, light-weight accreditation should also address different degrees of portability and interoperability. It is imperative that current standardization dialog be strengthened and extended to involve all stakeholders, including standards organizations, cloud cooperative working groups, open-source and commercial providers, as well as global and regional public sector institutions to ensure adequate representation of all needs and leverage inherent strengths. The objective of this action is to benefit the cloud computing arena globally for the collective common good. This continuous open dialogue on priority actions would provide an important forum through which to share respective use cases and best practices in the drive towards achieving true interoperability and portability. The EC and EU27 member state governments should provide direct support for funded activities to participate in relevant international initiatives. Further, common ‘synchronized’ calls for proposals between the EC and the US National Science Foundation (NSF), and possibly other national funding agencies, could strengthen international cooperation.

Multi-stakeholder engagement on a level playing field would ensure greater awareness of reciprocal needs, raise much-needed awareness on proven investments not only in cloud computing but also in standards engagement at all levels and ultimately foster a more conducive environment in which to pursue the actions captured in the SIENA and other relevant roadmaps.
**Benefits of e-research community engagement:**
Ensure user requirements are appropriately addressed, particularly in the light of broadening the user base. Share knowledge and expertise on current implementation and development efforts for the benefit of public authorities, international projects and commercial organizations. Define value propositions on standards-based reusable assets for both private and public organizations and synergize on interoperability testing and conformance regimes. As the user base is extended, organizations with wide stakeholder representation should actively participate in the standardization process.

**Benefits of public authority engagement:**
Increase understanding of standardization efforts so as to accelerate adoption and define the level of interoperability/portability required. Push for standards-based public procurement and educational programs to build new skill sets. Foster trust in cloud computing technologies and create leadership by example. Channel national e-research funding more effectively and enhance policies to support SMEs.

**Benefits of industry engagement:**
Drive increased implementation of relevant, internationally recognized standards and actively support testing and certification. Foster leadership by example, including continued investigations into concepts like accountability and privacy by design. Promote data stewardship and the continued development of overall data governance practices. Play a more active role against illegal access to services and data. Continue investments in research, uptake and support schemes for technology transfer and entrepreneurial initiatives with a focus on standards-based approaches.

**Benefits for standards organizations & collaborative working groups:**
Share information and documentation on standards work and studies, including all relevant information for the proposed inventory. Ensure implementation and testing is as cost-effective as possible, particularly where funding is dependent on public money or provided directly by SMEs. Work towards combining successful ongoing standards that can be used together in software implementations, help pinpoint overlaps and synergize on development, testing and certification. Define value propositions for each relevant stakeholder group and ensure best practices and requirements are appropriately covered.
Editorial Board

The SIENA roadmap has been designed, drafted and compiled by Martin Antony Walker, SIENA Roadmap Editorial Board Chair; Silvana Muscella, SIENA Technical Coordinator; John Borras and Carlo Daffara, Roadmap Editorial Board Co-Chairs; James Ahtes, Nicholas Ferguson, Daniele Lezzi, Mirco Mazzucato, Stephanie Parker, Elisabetta Ronchieri, David Wallom, and Johannes Watzl, SIENA Consortium Members.

Acknowledgements

Numerous individuals including members of three SIENA expert groups – the Industry Expert Group (IEG), the Special Liaison Group (SLG) and the Roadmap Editorial Board (REB) - have provided valuable feedback on earlier draft versions of this document. The SIENA Consortium is grateful to Joe Baguely, David Bernstein, Robert Bohn, Goetz-Philip Brasche, Winston Bumpus, Mark Carlson, Daniele Catteddu, Subrata Chattopadhyay, Guy Coates, Marnix Dekker, Michel Drescher, Ake Edlund, Thomas Edwall, Ad Emmen, Mike Fisher, Vangelis Floros, Geoffrey Fox, Vincent Franceschini, Fabrizio Gagliardi, Dennis Gannon, Yury Gilkman, Atsurhiro Goto, Patrick Guillemin, Gudmund Høst, Gershon Janssen, Bob Jones, Jean-Pierre Laisne, Erwin Laure, Dawn Leaf, Craig Lee, Ignacio Martin Llorente, Andrea Manieri, Bob Marcus, Karl Mayrhofer, Thijs Metsch, Pietro Michiardi, Steven Newhouse, Ian Osborne, Alexander Papaspyrou, Dana Petcu, Morris Riedel, Hiroshi Sakai, Alan Sill, Paul Strong, Etienne Urbah and Rigo Wenning.
SIENA (RI-261575) is funded by the European Commission under Framework Programme 7 (2007-2013) Research infrastructures projects.