Integration to European Research and Education Networks and EaPConnect Project

Vugar Musayev[1], Maria Minaricova[2]
[1]Institute of Information Technology of ANAS, Baku, Azerbaijan
[2]EaPConnect Project, GEANT Limited, Cambridge, United Kingdom
[1]*vuqarmusa@gmail.com*, [2]*Maria.Minaricova@geant.org*

***Abstract* — The paper introduces the open science concept in the context of a recent initiative, namely European Open Science Cloud and the e-infrastructures as building blocks of it. An overview of the framework, policy, requirements, challenges, the present and potential actors in establishing the European Open Science Cloud are briefly summarized. The natural role of GEANT community together with the national research and education networks is emphasized. EaPConnect Project is considered as a successful and promising platform for the Eastern Partnership Countries, specifically for Azerbaijan to integrate to the European e-infrastructures, Open Science Cloud and European Research Area.**

Keywords — Open Science Cloud; networks; HPC; GEANT; NREN; e-infrastructures; EaPConnect Project; AzScienceNet

I. INTRODUCTION

Open Science has emerged as a new level of research policy and framework of scientific collaboration recently. On 4 April 2016 in Amsterdam, a EU Conference focused on transition towards open access to scientific publications and the re-use of research data. Open Science transforms the approach to research, research cooperation and knowledge sharing by using the latest digital technology. Open Science means more transparency, well-organized research, more collaboration, and quality of science in terms of efficiency and coverage [1].

II. HISTORICAL OVERVIEW OF OPEN SCIENCE

The concept of open access to scientific data was institutionally established in 1957-1958, with the formation of the World Data System [2]. The International Council for Science building World Data Centers considering the risk of data loss and data accessibility [3]. In 2004, the Science Ministers of the nations of OECD signed a declaration stating that all publicly funded data should be made publicly available [4]. Since 2011, a new Linked Open Science approach developed and LinkedScience.org initiated as a platform for sharing and interconnecting scientific datasets, methods, tools and vocabularies [5]. In 2013, the G8 Science Ministers approved a "Statement on open research data" [6], which proposes new areas for collaboration, and agreement on global research infrastructure, open scientific research data, and increasing access to the scientific publications. The statement proposes that "open scientific research data should be easily discoverable, accessible, assessable, intelligible, useable, and wherever possible interoperable to specific quality standards". In 2015, the World Data System of the International Council for Science adopted a new set of Data Sharing Principles [7] of open science.

III. EUROPEAN OPEN SCIENCE CLOUD

In European research area, open access is meant to provide researchers, businesses and citizens with improved and free of charge online access to public funded research results, publications and data [8]. Another important issue, particularly in Horizon2020 program, is the concept of "Excellent science" which means not only excellent research but also excellent research infrastructures, which allows interoperable data exchange, storage and processing.

In 2015, European Commission adopted the Digital Single Markets strategy, and announced the "Open Science Cloud". The European Open Science Cloud (EOSC) will serve 1.7 million European researchers and 70 million science and technology professionals [9]. European Open Science Cloud offers researchers open access to a wide range of advanced research technology, resources and expertise. The main players of the European Open Science Cloud are GÉANT, EGI, OpenAIRE, PRACE, EUDAT, and LIBER, which have been strongly involved in building the separate but fundamental infrastructures and services for open science. These organizations have been collaborating on establishing a joint vision and roles of e-Infrastructures in several recent workshops and conferences considering the realization of Digital Single Market [10].

It should be noted that the European Open Science Cloud initiative is important not only for scientific development, but also for economic and societal challenges. Data, specifically research data should be reconsidered in terms of management, storage, security, sharing and reuse for the most possible wider community. Europe intends to invest on new research facilities, research data centers, high performance computing, cloud technology, advanced networking infrastructure, and special purpose services on the created infrastructure. European Open Science Cloud is a virtual environment where researchers can store, manage, analyze and re-use scientific data. It is clear that the cloud will be an umbrella for the existing research infrastructures and collected data. The new environment will enable the researchers to extend their work to more data driven research with a greater scale. New horizons of interdisciplinary research will certainly emerge in this complex.

There are several challenges on the way to a successful science cloud. The first issue is the common understanding on the value of data and sharing data. Setting widely accepted standards for inter-operability of data is another important issue.

The most important point is that, there is a great necessity for much more advanced computing, storage and network capabilities. In the European Open Science Cloud environment, there will also be issues of use policy, coordination, privacy, data protection, copyright, and intellectual property rights.

European Open Science Cloud has to be a collection of open but trusted infrastructure and services.

IV. European e-Infrastructures

This section explores briefly the most active European Open Science Cloud e-infrastructures: EGI, PRACE GÉANT, OpenAIRE, EUDAT, and LIBER.

A. EGI: European Grid Infrastructure

The European Grid Infrastructure (EGI) is a publicly funded e-infrastructure which gives scientists access to more than 530,000 logical CPUs, 200 PB of disk capacity and 300 PB of tape storage to drive research and innovation in Europe. The infrastructure provides High Performance Computing, cloud computing and storage capabilities [11]. The resources are provided by about 350 resource centers who are distributed across 56 countries in Europe, the Asia-Pacific region, Canada and Latin America. The main active project of EGI is the EGI-Engage project (Engaging the Research Community towards an Open Science Commons) started in March 2015, co-funded by the European Commission for 30 months, as a collaborative effort involving more than 70 institutions in over 30 countries. EGI-Engage aims to serve the Open Science initiative by developing the services for computation, storage, data, communication, and knowledge [12].

A. Partnership for Advanced Computing in Europe

PRACE, the Partnership for Advanced Computing in Europe was established in 2010 to create a persistent pan-European Research Infrastructure of world-class supercomputers. PRACE infrastructure consists of six supercomputers which are hosted by France, Germany, Spain and Italy. The machines are accessible via peer-review evaluation [13].

B. European Data Infrastructure

EUDAT enables the European scientists to store, find, access, and process data in a trusted network of collaborating data centers. EUDAT offers common data services and supports several research communities besides individual researchers. EUDAT connects 35 European organizations where research data is stored in powerful supercomputers distributed across 15 European countries [14].

C. Association of European Research Libraries

LIBER (Association of European Research Libraries) is the main network for research libraries in Europe. The association was founded in 1971 and covers more than 400 libraries from over 40 countries. LIBER represents the interests of European research libraries and their universities and their researchers in many directions [15].

D. The OpenAIRE2020 Project

The OpenAIRE2020 Project has started in January 2015 with a total budget of 13 132 500 € and will last for 42 months. 50 partners, from all EU countries, and beyond, will collaborate to work on this large-scale initiative that aims to promote open scholarship and substantially improve the discoverability and reusability of research publications and data.

The initiative brings together professionals from research libraries, national e-Infrastructure and data experts, and researchers. OpenAIRE2020 will assist in monitoring H2020 research outputs and will be a key infrastructure for reporting H2020's scientific publications as it will be loosely coupled to the EC's IT backend systems [16].

E. Helix Nebula Project

Helix Nebula is a new partnership between big science and big business in Europe. The project is active and promising in Science Cloud formation. The partnership joins important IT companies and the main research centers in Europe like European Organization for Nuclear Research (CERN), European Molecular Biology Laboratory (EMBL) and European Space Agency (ESA).

Helix Nebula presents solutions to the rapidly growing computing power demand by big science. Thanks to Helix Nebula, the global cloud industry has opportunity to work closely with unique organizations like CERN or work on molecular biology and earth observations. Helix Nebula, being able to overcome the demand by the big science data, will be an important builder in European Open Science Cloud [17].

F. GÉANT

GÉANT Association serves as a platform for research and education networking community to share, test and collaborate on the leading technology and expertise. GEANT has been leading successive European projects which stimulate the collaborative research, innovation and technology development in more than 40 countries. GÉANT Association owns and manages GEANT network, the cutting edge, advanced high-speed pan-European e-infrastructure. Data intensive research projects use the reliable GÉANT network thanks to the quality of service.

GÉANT network is a specific purpose ultra-high-speed internet designed for the needs of research and education. Each day 1,000 terabytes of data flows on the GÉANT network. GEANT services like IP connectivity, VPN, dedicated lambdas, security, identity federations, clouds and high definition video conferencing are only a subset of the whole range of services and opportunities that users benefit from [18].

V. GEANT and European Open Science Cloud

Considering that GÉANT network and more than 40 National Research and Education Networks (NRENs) already

connect 50 million researchers and students, it turns out that an important part of the European Open Science Cloud has been successfully deployed and is ready for integration. On the other hand, increasing demand for connectivity and integration to high performance computing and huge databases, require advances for the GÉANT network and for the NRENs. GÉANT is owned and managed by the national research and education networks, builds, operates and develops state of the art network infrastructures and services. Hence GEANT is already a pan European digital infrastructure with strong, collaborative management and experience.

VI. EAPCONNECT PROJECT

Eastern Partnership Connect (EaPConnect) project was officially inaugurated at the 1st Eastern Partnership Ministerial Meeting on the Digital Economy on 11 June 2015 in Riga, Latvia [19]. Launched in July 2015 for five years, the EaPConnect project with a budget of 13.6 million euro will build an advanced regional network and deliver tailored services for the research and education community in the region. The European Commission's Directorate-General for Neighborhood and Enlargements Negotiations (DG NEAR) is contributing 95% (13 million euro) towards the cost of the EaPConnect project; the remaining 5% will be co-funded by the six beneficiary countries.

As one of the beneficiaries Azerbaijan`s scientific and educational community will be connected to pan-European network, GEANT. This will be an outstanding opportunity for Azerbaijan as well as the whole EaP region to lessen the digital divide and effectively cooperate with European research institutions for data intensive research. EaPConnect project will establish and operate a high-capacity broadband internet network for research and education and integrate the national research and education networks in the region into the pan-European GÉANT network. It will facilitate participation of local scientists, students and academics in global research and education collaborations. By interconnecting the R&E communities across the region and with their European counterparts, EaPConnect will create a gateway for talented individuals in the EaP countries to be truly global players.

CONCLUSION

A brief overview on the initiative for the European Open Science Cloud and the existing e-infrastructures clearly shows that European science is about to put a step into the new area of scientific development. Considering the need and reality of big data, especially research data has recently been highly valued due to the developments and opportunities that computational power, storage and networking put before the science. Row data, processed information, and gained knowledge are very valuable assets which will be benefited maximally if all the partners of scientific community decide to share as much as possible and collaborate on the shared data and knowledge. Open Science initiative is the formulation and platform for such a new paradigm.

At this point there is an increasing demand for advanced data sharing and processing infrastructures, mostly e-infrastructures. European e-infrastructures like EGI, PRACE, EUDAT, OpenAIRE, LIBER, GÉANT, and EaPConnect provide high speed, reliable and secure communication channels, high performance computing, data repository and access opportunities, federated services, and access to centralized data, libraries and knowledge bases.

Azerbaijan research and education community has joined the GEANT Association from the very beginning of research networking via the Institute of Information Technology of Azerbaijan National Academy of Sciences. AzScienceNet, the first network in Azerbaijan has been serving the research community. AzScienceNet participates in the EaPConnect Project and GÉANT. As a result Azerbaijan is establishing digital highways with the European Open Science Cloud. With huge communication, computing and data repositories, researchers and students in the country will have the chance to cope with the scientific developments and educational opportunities.

As a result, it becomes clear that, there is a need to develop cooperation with all e-infrastructures and put forward a long term strategy and planning in this area.

REFERENCES

[1] https://ec.europa.eu/research/openscience/index.cfm
[2] Earth Observations from space: The First 50 Years of Scientific Achievements. The National Academies Press. p. 6.
[3] www.icsu-wds.org/organization
[4] http://acts.oecd.org
[5] T.Kauppinen and G. M. Espindola, Linked open science communicating, sharing and evaluating data, methods and results for executable papers, Int. Conf. Computational Science (ICCS), 2011, vol. 4, pp. 726-731.
[6] www.g8.utoronto.ca/science/130613-science.html
[7] www.icsu-wds.org/services/data-sharing-principles
[8] https://ec.europa.eu/futurium/en/content/about
[9] https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud
[10] www.eudat.eu/events/european-open-science-cloud-for-research-eosc-workshop
[11] www.egi.eu/infrastructure/index.html
[12] www.egi.eu/about/egi-engage/index.html
[13] www.prace-ri.eu/about-prace-ri
[14] www.eudat.eu/what-eudat
[15] http://libereurope.eu/about-liber
[16] www.openaire.eu/project-factsheets
[17] www.helix-nebula.eu/
[18] www.geant.org/About
[19] www.eapconnect.eu/About