Proceedings of LR4SSHOC: Workshop about Language Resources for the SSH Cloud, pages 37–38
Language Resources and Evaluation Conference (LREC 2020), Marseille, 11–16 May 2020

EOSC as a game-changer in the Social Sciences and Humanities research activities

Donatella Castelli
Istituto di Scienza e Tecnologie dell'Informazione "A. Faedo"
Consiglio Nazionale delle Ricerche
Via Moruzzi,1 – Pisa
donatella.castelli@isti.cnr.it

Abstract

This paper aims to give some insights on how the European Open Science Cloud (EOSC) will be able to influence the Social Sciences and Humanities (SSH) sector, thus paving the way towards innovation. Points of discussion on how the LRs and RIs community can contribute to the revolution in the practice of research areas are provided.

Keywords: European Open Science Cloud, Social Sciences and Humanities

1. Research and technological progress

Research strongly influences, but is also largely influenced by, technological and societal progress. This is true for all the research domains, including the Social Science and Humanities (SSH) ones. So far, research activities in these domains have been primarily based on scattered data collected in long-lasting campaigns and analyzed by the researchers that have collected them. In the last few years, with the advent of big data, often gathered by sensors or through citizen scientist activities, and with the spreading of data innovative analysis techniques, mainly based on artificial intelligence, the situation has started to rapidly change.

2. The European Open Science Cloud

In the near future, many more changes are expected to occur thanks to the European Open Science Cloud (EOSC). With its technological, capacity and governance components, EOSC fits into this already rapidly evolving context, paving the way for a real revolution in the practice of research areas. It is will provide a trusted system giving seamless access to data and interoperable services. Through EOSC the researchers will have access to functionality supporting the whole research data cycle, from discovery and mining to storage, management, analysis and re-use across borders and disciplines.

3. Accessible datasets and services

EOSC is currently being built as a collective effort by relying on existing components, e.g. infrastructures, services, and data resources. Through its system, it will be able to perform functions and carry out purposes that do not reside in any component alone (aka emergent behavior[1]). A large part of these functions will be dedicated to supporting cross-domain and cross-sector activities and, in particular, to make accessible and exploitable datasets and services across borders.
This latter EOSC functionality is likely the one that will produce more innovation in the SSH sector. It will not only allow researchers to access similar complementary datasets collected by others. It will make it possible to enrich these datasets with contextual information in space and time using datasets and services produced in other domains, such as, for example, earth observation, environment, and medicine, just to mention a few. At the same time, researchers in the SSH sector will be able to make their outcomes available to researchers of other domains in a short time and with limited effort. This will widen the diffusion and impact of their research products, not only in the context of the research community, but also to decision-makers and innovators.

4. Removing the language barriers

In order for this vision to fully realize, EOSC will have to necessarily offer functionalities able to remove barriers in the usage of data and services. As emerged in a recent workshop organized by the EOSCSecretariat.eu project[2] dedicated to collect the needs and requirements for future research environments[3], among the barriers "language and communication" are perceived as top ones by many researchers. For example, even searching in the EOSC catalogues of services and data may not be simple. The language used for the descriptions of these resources can be really problematic for those that do not belong to the same domain and sector. For removing these language and communication barriers EOSC should include, at its core, a variety of "translation" services.
The richer these services will be, the better EOSC will be able to reach its cross-border objectives. These core services might include, for example,

(i) human to machine and machine to human translations services,

(ii) horizontal translation services able to convert domain-specific technical terms of research outputs across disciplines or to policymakers,

[1] Crawley, E. F., Cameron, B. G., and Selva, D. (2016) System Architecture: Strategy and Product Development for Complex Systems, Pearson.

[2] http://eoscsecretariat.eu/

[3] Report on the Workshop "Co-creating the EOSC: Needs and requirements for future research environments", DOI 10.5281/zenodo.3701193.

(iii) vertical translation services covering the translation of research specific concepts, as well as the explanations for different career stages, i.e. from specialists to early career researchers, students, and the interested public at large.

5. LRs and RIs contribution

If on the one hand it is indisputable that the availability of these services largely influence the impact of EOSC in revolutionizing the research practices of the SSH community, on the other hand, it is also clear that the community itself, and in particular that part of it dealing with language resources (LRs), can contribute to the realization of these services by sharing their expertise and resources.
This clearly exemplifies the role of the Research Infrastructures (RIs) in EOSC that are called to both exploit but also to contribute to it.