

Comments to Jean-Claude Burgelman's article *Politics and Open Science: How the European Open Science Cloud Became Reality (the Untold Story)*

Dimitris Koureas[†]

Naturalis Biodiversity Center, Museums & Institutions, Leiden, the Netherlands

Citation: Koureas, D.: Comments to Jean-Claude Burgelman's article *Politics and Open Science: How the European Open Science Cloud became reality (the untold story)*. Data Intelligence 3(1), 26-28 (2021). doi: 10.1162/dint_a_00072

It has been genuinely a pleasure reading through Jean-Claude Burgelman's draft. It provided me with clear and straight forward narratives as responses to questions regarding the origins of EOSC.

I want to start formulating a response to this draft by recognising the good place EOSC is today. A strong mission statement and a realisation that EOSC goes above and beyond the technical aspirations its name actually suggests. Today, EOSC concept and implementation actions clearly and powerfully encompass the sum of the socio-technical components that form unique value chains across data-driven science, from ideation to impact. These components are predominantly delivered by a multitude of scientific disciplines, through scientific communities, their respective practices and linked technical solutions supporting data lifecycles. It is probably the involvement of these communities, not just as users of EOSC but as core stakeholders, that helps EOSC transition from an opaque political statement to a tangible socio-technical ecosystem of people, data and services for science.

Looking back at JCB's opening, I cannot but express gratitude and admiration towards JCB and his colleagues for realising early the forthcoming transformation of scientific practice, in response to the unprecedented production of scientific data. Also, I recognise their commitment and perseverance to place this topic high enough on the agenda of science-policy for Europe. JCB refers to a series of workshops that had taken place in 2014 and cites unanimity around "*the key problems of this explosion of data science will be: non-discriminatory access to the data, interoperability across disciplines and making sure these data are managed with respect for European sensitivities*". As such, it is highlighted that the main bottlenecks

[†] Corresponding author: Dimitris Koureas (E-mail: dimitris.koureas@naturalis.nl; ORCID: 0000-0002-4842-6487).

for effectively leveraging European scientific resources in the new big Open Data (OD) era would be mostly socio-technical. Despite early and explicit messages that flagged the need for a more inclusive and broader conceptualisation and implementation strategy for EOSC, the years after its first inception were predominantly focused (through linked investments) in developing the underlying e-infrastructure commons. An attempt to federate existing national and European public infrastructures that provide computing and storage services. The consolidation of the highly fragmented landscape of e-infrastructures assumed a central role in the vision of EOSC.

JCB states that the predominant focus in developing the idea of EOSC was to “get the end-users on board, the scientists”. I would argue that if the investment was geared towards getting the end-users on board the concept of EOSC, that probably did not yield the results needed at that point. By 2017, the European research landscape had already made significant investments in developing their discipline-specific Research Infrastructures (RI). Organisations that were driven by the needs of the communities and supported by member states (MSs). In 2017, we already had tens of those RIs in preparation, construction or operational phase. These RIs managed to address their respective communities' needs, providing valuable end-user services, whilst making use of a multitude of underlying computing, storage, and data resources. As such, they quickly turned into major *de facto* stakeholders of EOSC. In our article back in 2016 (Koureas et al., <https://doi.org/10.3897/rio.2.e9933>) we specifically touched upon the socio-technical distance between end-users and e-infrastructures. We referred to solutions that would allow communities of practice to benefit from a central e-infrastructure service bus, using community-driven intermediate structures (such as research infrastructures or virtual research environments) to bridge the gap, a model that would resemble the value chain of wholesale (e-infrastructure commons) -> retail (community infrastructures) -> user (researchers) approach.

JCB lists, in a remarkably interesting way, the steps taken to bring us in the current realisation of what EOSC can be and offer. EOSC is, in its inception, an innovative concept, which would partly disrupt current established practices and hopefully improve the efficiency of data-driven scientific practices. As such, it was deemed to go through its own hype cycle. Though it is still difficult to assess whether we are still in the troughs of disillusionment or climbing the slope of enlightenment, we can for sure assert that the peak of inflated expectations was primed with a misguided understanding that the success EOSC hinges on the orchestration of distributed technological solutions towards end-users. As a wider community of stakeholders, we have failed to fully realise that the equation from technological capacity to scientific innovation is heavily nuanced between communities. Aspects related to community traditions, established practices and trust mechanisms play a pivotal role in completing each of those equations. In the next steps in completing the EOSC mission, we ought not only to acknowledge these domain-specific aspects but encompass them and work with them to improve the ability of researchers to benefit from the investments in EOSC.

AUTHOR BIOGRAPHY



Dimitris Koureasis, Chief Information Officer at Naturalis Biodiversity Center and executive director of the Distributed System of Scientific Collections (DiSSCo) European Research Infrastructure. He holds a PhD in biosystematics and post-doctoral expertise in biodiversity informatics and in the research data lifecycle. He is an Invited lecturer, teaching biodiversity informatics, across European universities. He is a member of the Technical Advisory Board of the Research Data Alliance (RDA), and has served in various board-level positions across data-related international organisations.

ORCID: 0000-0002-4842-6487