Open Science and the Academic Profession

Rosanna De Rosa,[1*] Biagio Aragona[2]

[1] ORCID Nr: 0000-0002-3114-4184
Vico Monte di Pietà 1, Department of Social Sciences, Università degli studi di Napoli Federico II, rderosa@unina.it

[2] ORCID Nr: 0000-0001-8697-2932
Vico Monte di Pietà 1, Department of Social Sciences, Università degli studi di Napoli Federico II, aragona@unina.it

Abstract: Open science is considered a new science paradigm to make research accessible, accountable, and effective. Open science is already changing the academic profession starting from micro-practices to professional relations with epistemic communities and stakeholders, with implications that we are not yet able to predict. The article delves first into literature and official documentation to unfold the discursive regimes which sustain the spread of open science. A specific focus is then devoted to the professional transition, highlighting the role of funding organizations in setting the new science environment and the subjective experience of academics. The article is completed by a case study in the field of Research Data Management where the misalignment among incumbent/changing processes can be more apparent. Finally, a research agenda that focuses on how academic micro-practices are affecting organizations and science structures is proposed. This article aims at beginning to plow the ground for new research directions to emerge.

Keywords: Open science, research data management, academic profession, open access, citizen science

1. Introduction

The academic profession has always been the process of change, but the factors constituting the change have evolved over time. New aspects which were not present previously or which were not considered as scientific or purely academic have started to play a role. One of the most challenging factors is the increasing claim for open access to scientific knowledge production and processes (Kogan, Teichler, 2007), namely open science. This happens for example when researchers feel obliged to incorporate openness in peer judgment to remain credible, for example, by giving priority to the social relevance or impact of research outcomes. Indeed, open science represents: «an approach to research that is collaborative, transparent and accessible» (O'Carrol et al., 2017: 8).

The article aims at investigating some implications that open science has on the academic profession shedding light on the process of "industrialization of academic activities" (Kogan, Teichler, 2007), namely the passage from craft production to (organized) production processes based on specialization of tasks, and standardization. Our study is inspired by Science and Technology Studies and theoretically underpinned by the work of the pragmatist philosopher of science John Dewey (1958) on the knitting between science and the social, political, and economic apparatuses of society. Open science is here considered as a socio-technical construct, which combines the social and the technical and it is, at the same time, shaped by academic professionals, as well as shaping the academic profession.

Firstly, we retrace the guiding discursive frameworks which sustain in literature the spread of open science and the adoption of the open paradigm within research funding schemes and institutions. Furthermore, we give quantitative evidence and figures drawn from secondary data analysis highlighting actual difficulties in practicing open science. By searching the Scopus catalogue, we detected the amount of reflection about open science through the number of publications and the spread of open access practices among the different disciplines. We also present the results of several international surveys conducted on academic researchers to analyze their behaviors and opinions on open science. The article is completed by a case study, the OBERRED project, which aimed to construct an ecosystem for training and certification in Research Data Management (RDM). Finally, by combining the theoretical framework with the quantitative data on open science spread, and the findings from our case study, we contribute to identifying critical points and implications in the long run to be further investigated.

A paramount critical point for open science is the development of appropriate skills and dedicated professional profiles. In this respect, the open science paradigm requires that the hybrid nature of researcher's profiles is recognized and assessed. In addition, more consistent and dedicated research aiming to bind together the space of policy, the space of debate and the space of practices should be developed. Model publishing, RDM practices, as well as governance approaches (top-down -vs – bottom-up) and claims for reproducibility are the research streams of the proposed research agenda.

2. How open science is depicted: three narratives and an ontology

Open Science is a prismatic notion based on a series of assumptions about the future of scientific knowledge and its dissemination. Every single concept variation is sustained by a system of norms, values and narratives which act as a 'discursive regime', counter-narratives capable of changing, re-orienting, determining, modeling new epistemologies of science and research (Foucault, 1977).

A first, multi-layered set of narratives is about the democratization of scientific results and processes (Feyerabend, 1975). The social and educational benefits to making research outputs accessible without financial, legal and technical barriers (Van der Zee, Reich, 2018) have been highlighted, and open science has been said also to enable the rights for all people, because scientific results have "a powerful effect on people's capacity to speak up as active citizens on matters that are shaping their

city and their world" (Appadurai, 2013: 281). Open access, open data practices, and policy frameworks are those that may have the biggest impact on the academic profession, contributing to accelerating scientific progress; as the possibility to access a consistent amount of data on Covid-19 has largely demonstrated. The pandemic spread and emphasized the narrative that opening access to knowledge and data can help tremendously in identifying effective therapies and vaccines. In 2020 alone, more than 85 thousand publications about Covid were indexed by Scopus, of which 80% are open access. The Coronavirus Disease Research Community - COVID-19 on Zenodo.org is a concrete example of informal sharing micro-practices combined with community collecting research outputs and datasets that may be relevant to mitigate Covid-19 effects and to increase pressure for research findings accountability and democratization, according to the Findable, Accessible, Interoperable and Reusable (FAIR) principles of Open Science.

From a societal outlook, the European Commission is strongly fostering the development of initiatives about citizen science, which means, on one side, a science that is responsive to citizens' concerns and needs and, on the other side, that citizens themselves could produce reliable scientific knowledge (Irwin, 1995).

Citizen science can play a transformative role on science-society relations widening the social relevance of academic research and contributing to spread concepts, methods, and procedures of scientific reasoning to unconventional publics. Examples are developing in natural and social sciences. In natural science amateur bird watchers voluntarily contributed scientific data, in humanities, Zooniverse is a transcription tool to help volunteers participate in more than 50 scientific projects and to help to transcribe ancient papyri. Citizen science aims at increasing participation and transparency, and this may have a positive impact on democracy.

The second narrative is about the effectiveness of research processes through its infrastructure. Indeed, the practices just mentioned could have not become established unless the distribution of knowledge had not been enhanced by means of digital technologies. The open science paradigm gains momentum in the interplay between the humans (the researchers and the other stakeholders involved in research processes) and the technological tools and infrastructures (Decuypere, Simons, 2014). So, the development of frameworks for the sharing of information and data such as the FAIR principles, the Common European Research Information Format (CERIF) and the OpenAIRE guidelines are clear efforts to enable that cultural change required for the development of open science (MacCallum, 2018). Thus, the European Open Science Cloud (EOSC) and the Research Data Alliance (RDA) play the role of organizational nerves that embodied the sharing processes through the building of common data infrastructure and strategies. ICTs facilitate, therefore, the exchange of good practices, the large-scale dissemination of results, and the deployment of referencing management tools, as well as new forms of academic and professional recognition. The web environment has also maximized the fruition of opportunities by pushing cultural institutions to be more responsive in respect of their constituency's needs and expectations. More specifically, the digital tools that served open science are social and collaboration networks for scientists, platforms for sharing frameworks and principles, and data repositories and archives. Nentwich and König (2012) point to social networks for scientists, such as ResearchGate and Mendeley, explaining how they are functional for the establishing of networks, for the creation of thematic expert groups, communication purposes, and

self-promotion. In parallel, alternative impact measurements (such as altmetrics) have been developed starting from the sharing of raw science material as well as experimental designs, nanopublications, self-publishing via blogs, annotations and mentions systems and so on. In this diverse epistemic ecosystem, altmetrics is becoming essential to sift these new forms of academic discourse, since they're not retrieved by the traditional means for the communication of science.

So, we come to the third series of narratives that sustained the discourse about open science: the claim for transparency. Bowman and Keen lay out an 'onion model' of open science that

"*describes increasing levels of transparency and suggests how open science practices can be understood less as a revolutionary concept, but more as a logical extension of some of the historical pillars of scientific norms" (Bowman and Keen, 2018: 363).*
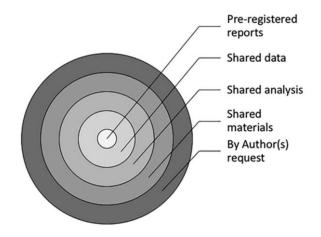
The debate about open science is then, in our opinion a late consequence of the Merton's (1942) discussion of science as the «extension of certified knowledge» (Bowman, Keen, 2018: 117). In other words, over «the reproducibility of established findings» (ivi, 364), which is the pillar that supports the whole scaffolding of scientific communication and results validation. The question represented by Bowman and Keen as an 'onion model' (see figure 1) is, in fact, all over a matter of incremental transparency along the inquiry process, as well as a matter of research integrity. In the 'onion model' the first four layers are almost common practices in the scientific communities, but the highest level of transparency is reached with the pre-registered reports which are

«*broadly understood as a process by which authors publicly document their core study logic, hypotheses, research design, and an a priori plan for their data analysis» (Bowman, Keen, 2018: 368).*

The pre-registered reports allow, in fact, to understand and check the original formulation of research hypotheses by preventing them from altering because of research findings, and to review research material in two stages, at the time of study design and after the research execution[1].

---

[1] Some scientific journals have already experienced the pre-registered model. This is the case of the Journal of Media Psychology, special issue 2017 (Bowman, Keen, 2018).
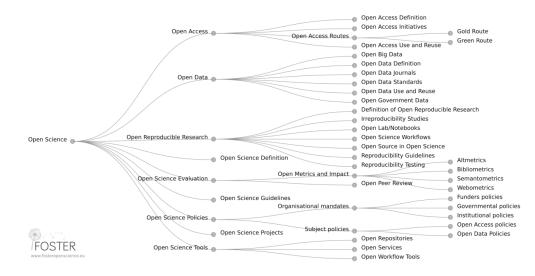
Fig. 1 A conceptual 'onion model' for various open science practices. Inner-most layers represent increased transparency for a given published manuscript. Source: Bowman and Keen (2018).

Open science is therefore said to improve the transparency and validity of research. This perspective has been clearly remarked by the 2017 recommendations of the EU Working Group on Rewards under Open Science. In fact,

"*Open science goes hand in hand with research integrity and requires legal and ethical awareness on the part of researchers. A driver for open science is improving the transparency and validity of research as well as in regard to public ownership of science, particularly that which is publicly funded*" (O'Carrol et al., 2017, p.4).

Fig. 2: Foster Open Science Taxonomy. Source: fosteropenscience.eu

Finally, the concept has been operationalized to shed the light on its multifaceted nature. One of the most comprehensive ontologies was proposed by the Foster Project to support the shift towards open science, by training the young generation of researchers. What figure 2 makes apparent, is the level of the conceptual diversification and the need for formalizing Open Science. We may see this process as the 'constitutionalization' in progress of a new knowledge field, requiring a certain level
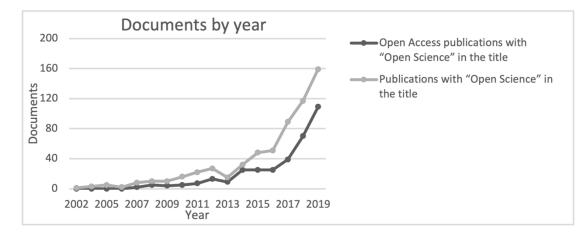
of innovation in the concepts used, processes and tools. There is probably not any other argumentative device in open science with such a normative power.

3. What researchers think about open science

In this section we are going to present figures drawn from Scopus and from international comparative surveys about open science run by academic researchers. The analysis of these different sources aims, on one side, to map the dimensions that form (and the schools that inspire) the open science paradigm and, on the other side, to study the behaviors and the opinion of researchers toward open science by focusing on two of the most crucial dimensions: open access and open data. 'Open Science' has been used as a keyword to search Scopus, the bibliographic search engine created in 2004 by Elsevier. Our query returned 615 documents with 'open science' in the title, within the period 2002-2019. 54,9%, of these 593 publications are open access, distributed on-line, free of cost.

The debate on open science in publications takes on a momentum of its own in 2003, its distribution over time shows an increase in 2007, and an acceleration in 2013 (see figure 3). Among these, open access publications followed the same curve, with a sharp increase in 2013 (see figure 3).

Fig. 3 Number of publications with "Open Science" in the title by year and publishing model (2002-2019). Source: Authors' elaboration on Scopus.

Nevertheless, the diffusion of open access differs sharply among the different disciplinary fields. Social sciences are leading the way, with almost 17% of the total open access publications on open science (see figure 4). Computer Sciences, Medicine and Multidisciplinary follow with 9.5%, all the other disciplines are slightly above, or below, 5% of the total of open access publications on open science.
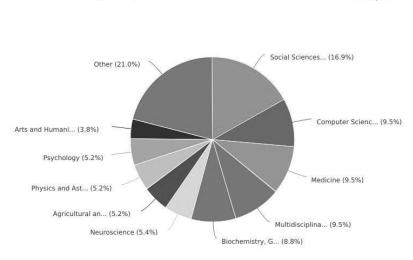
Fig. 4 Open Access publications by subject area (1979-2019). Source: Scopus.

Copyright © 2020 Elsevier B.V. All rights reserved. Scopus® is a registered trademark of Elsevier B.V.

In the second step, we analyzed the keywords associated by the author with the indexed publications, and their relationships. The keywords cloud was developed with VosViewer, a software tool for the construction and visualization of bibliometric networks developed by the Center for Science and Technology Studies of the University of Leiden. VosViewer is able to build conceptual networks starting from the co-occurrences of keywords and displaying them in the form of graphic maps according to emergent clusters[2] (see figure 5). The result is a graph where clusters of co-occurrences highlight the different strands taken by the scientific debate about the open science paradigm.

---

[2] A cluster is a set of elements included in the map. In VosViewer clusters do not work in overlap. An article can belong to only one cluster. Clusters do not have to cover exhaustively all the elements of a map. Therefore, there may be elements that do not belong to a cluster. For this analysis, the minimum number of occurrences of a keyword is 5. On a total of 1016 keywords, 25 meet the threshold. For each of the 25 keywords, the total strength of the occurrence links with other keywords has been calculated by the software. The graph represents only the keywords with the greatest total link strength.
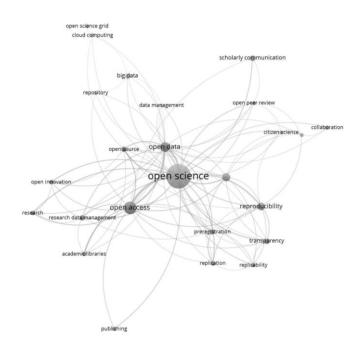
Fig. 5 The network of keywords of the publications on Open Science (1979-2020). Source: Scopus.

In the left down part of the graph, we find the keywords connected with open access, research data management, open data and libraries. This is definitely a cluster related to the availability of research outputs, but two different arguments emerge, one concerned with the accessibility of the research process (open data, research data management, open source) the other with the accessibility of the research products, the publications (open access, academic libraries). The second category intertwines the infrastructural elements of open science, the majority of the non-human actors that structure open science as socio-material construct. In the top left section of the graph, we find cloud computing, repositories, big data, data management, all the material actants (Latour, 1985) that interact with researchers and sometimes resist them. The top right corner of the graph is the space of collaborative research. Here we find peer review, communication between scholars, and participation of citizens to science, the network of human actors that can play a role in the making of open science. In the cluster at the right bottom of the graph we find all the values that inspire the debate on open science for the sake of science. Reproducibility, replicability, transparency are all keywords linked to the idea that open science means more accountable and recursive protocols. Finally, it is worth noting that publishing creates a very consistent link with open science and open access, as it is the very basic material act of academic work. It is not accidental that the future of scholarly publishing was one of the eight pillars established by the European Commission in 2018 together with FAIR data, the European Open Science Cloud, education and skills, rewards and incentives, next-generation metrics ('Altmetrics'), research integrity and citizen science (Ayris et al., 2018). Also, publication policies have lots of implications on researcher careers and peer assessment as demonstrated by the growing discontent towards current evaluation systems.

Based on a literature review, Fecher and Friesike (2014) retraced five open science schools of thought – 'iterative motives and patterns of argumentation that form more or less distinct streams' (p.18) - which perfectly fit the categories (access, data storage, collaboration, reproducibility, publishing) highlighted in the VosViewer graph:

- The public school, concerning public science accessibility and comprehensibility.
- The democratic school, concerning the access to knowledge production.
- The pragmatic school, concerning the efficiency of research through collaboration.
- The infrastructuring school, concerns the technological architectures including social networks and distributed computing.
- The measurement school, concerning the seeking for alternative impact assessment processes.

These schools of thought, though deeply intertwined, also evoke different understandings of open science, according to the diversity of stakeholders and their needs (researchers from all fields, policy makers, platform programmers and operators, publishers, and the interested public).

The systematization work by Fecher and Friesike (2014) focused mainly on two paradigm shifts, concerning science-society relationship: from closed to open and from individual to collective inspired by the search for efficiency and collaboration. However, in light of progress made by the academic reflection and policy making (Owen et al., 2012; Owen, von Schomberg, & Macnaghten, 2021; Novitzky et al. 2020), and taking into account the pressing demand for transparency, gender equality, sustainability and responsiveness, Fecher and Friesike's table needs to be further integrated. In the last ten years, in fact, several collective actors (policy institutions, advocacy organizations etc.), started to foster anticipative, reactive and reflective science actions to foresee (and adjust when necessary) the side effects of technology innovation processes and to address societal challenges (i.e SDGs).

Thus, two new directions clearly emerged in EC official documentations and in STS literature: an ethical and a deliberative school (tab. 1) aiming at moving open science a step forward in Responsive Research and Innovation. RII is a policy-driven framework defined as "the on-going process of aligning research and innovation to the values, needs and expectations of society" requiring that all stakeholders are mutually responsive and responsible to each other for the processes and outcomes of research and innovation (Rome Declaration, 2014)[3].

Again, with these two new schools, we are witnessing another paradigm shift, from the efficiency of research to its responsiveness to societal challenges[4]. Both are likely to have interesting implications for academic and research work (Owen and Pansera, 2019).
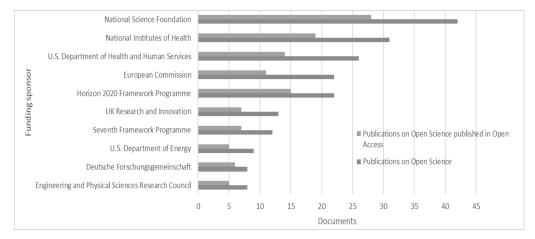
---

[3]   https://digital-strategy.ec.europa.eu/en/library/rome-declaration-responsible-research-and-innovation-europe

[4]   https://ec.europa.eu/programmes/horizon2020/en/h2020-section/societal-challenges

Table 1 Open Science schools of thought. Source: Authors' elaboration on (Fecher and Friesike 2014)

School of thought	Central assumption	Involved groups	Central Aim	Tools & Methods
De-mocratic	The access to knowledge is unequally distributed.	Scientists, politicians, citizens	Making knowledge freely available for everyone.	Open Access, intellectual property rights, Open data, Open code
Pragma-tic	Knowledge-creation could be more efficient if scientists worked together.	Scientists	Opening up the process of knowledge creation.	Wisdom of the crowds, network effects, Open Data, Open Code
Infra-structure	Efficient research depends on the available tools and applications.	Scientists & platform providers	Creating openly available platforms, tools and services for scientists.	Collaboration platforms and tools
Public	Science needs to be made accessible to the public.	Scientists & citizens Making science accessible for citizens.	Citizen Science, Science PR,	Science Blogging
Measure-ment	Scientific contributions today need alternative impact measurements.	Scientists & politicians	Developing an alternative metric system for scientific impact.	Altmetrics, peer review, citation, impact factors
Ethical	Science needs to be aligned with the needs, values and expectations of society. Be responsible, ethically bound and able to take into account the real-world complexity	Scientists, public authorities, citizens, innovators & civil society	Ensuring accountable, responsible, ethical, and socially desirable outcomes	Research Integrity Framework
Delibera-tive	Science should engage, support, and share responsibility with institutions and territories. Implement sustainable institutional changes and innovation	Academia, industry, civil society, and public authorities	Maximizing science-technology social impact by supporting institutional change and innovation governance	New forms of governance and participatory structures

Fig. 6 Number of publications on Open Science by funding sponsors and publishing model (2002-2020).
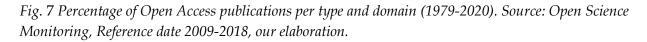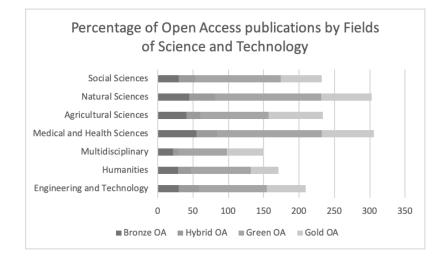*Source: Authors' elaboration on Scopus.*

The role of funding organizations has also been pivotal in dictating the rules for open access publication, also fostering institutional mandates and policies. Figure 6 shows results of open science publications by funding organizations and the model of publishing. By comparing the figures, the role exercised on openness by research sponsors emerged. The compliance to funding rules about open access has been, in fact, one of the key factors enabling a change in research and academic institutions. Organizations such as the US National Science Foundation and the National Institutes of Health have had the strongest role in publishing strategies together with the European Commission and its funding programs. In 2007, the EU Commission adopted a Communication (COM/2007/0056) on scientific information in the digital age for the first time, followed by Council conclusions inviting the Commission to experiment with open access to scientific publications resulting from projects funded by EU research framework program. Later on, in 2012, the EU Commission has recommended (C(2012) 4890) to define clear policies for the dissemination of results and open access to scientific publications resulting from publicly funded research, providing: a) concrete objectives and indicators to measure progress; b) implementation plans, including the allocation of responsibilities; c) associated financial planning.

The European Commission has also created a platform to monitor the process (the Open Science Monitor), obtaining quantitative and qualitative insights on the ongoing development of open science practices. One of the duties of the Open Science Monitor is to highlight emerging trends and specific patterns in the publishing choices of academic researchers. For example, during the period 2009-2018 the general trend of so called 'Green' publications[5] has a positive trend passing from 20,4% in 2009 to 26,9% in 2017, but in 2018 it comes as a surprise the downfall of green publication to 15,3%. The rank clearly highlights the primacy of some countries such as the United Kingdom (43,9 %) and Switzerland (42,8%) against Italy, where Green publications reached only 30,3 %. Furthermore, Green publication in the multidisciplinary domain reaches 68,7%, against the medical and health

---

[5]   Publications are classified in Gold, Green, Hybrid, Bronze according to the degree of openness allowed by journals. Green publications are those journal articles that are also available in an open access repository. Data from Open Science Monitoring have been aggregated per domain according to the OECD classification 2007, Frascati Manual.

science where open journal publication (Gold) and use of open repository (Green) have a longstand-ing history. Publications in social and political sciences step behind with only 13% of publications using Green repositories.

Fig. 7 Percentage of Open Access publications per type and domain (1979-2020). Source: Open Science Monitoring, Reference date 2009-2018, our elaboration.

These differences in open access between disciplinary fields shed the light on somehow divergent practices (see figure 7), questioning the open paradigm as a comprehensive framework, that instead may reproduce the traditional divide among data-driven sciences and the social sciences and hu-manities (Sidler 2014). In fact, as detected by the EOSC-Pillar National Initiative Survey published in 2020[6], research infrastructures at a national level provide Social Sciences (25%), Humanities (18%) and Agricultural Sciences (14%) less frequently, with services while Natural Sciences communities (66%), Medical and Health Sciences (36%), Engineering and Technology (33%) are better served. However, the development of e-infrastructures seems to entail balancing the distributions of ser-vices among the different disciplines, with Humanities (45%) following Natural Sciences (69%) it is likely, thanks to the strong effort spent in organizing digital archives and virtual libraries. Social Sciences (38%) close the list together with Agricultural Sciences (30%). Figures are, of course, coun-try-specific, with Italy and Austria observing for Humanities percentages above the mean.

In the last 10 years an international agenda around open data has also emerged. According to the Open Data Report[7] run on 1200 researchers in 2016 and 2018, open data and open access are an emerging attitude of academic researchers. Almost 76% of the 1200 sampled academics declare to share data in some form. 66% believe that sharing data increases the possibilities for collaboration, the reproducibility of research (57%), encourages other researchers to do the same and, moreover, related publications are more likely to be cited. The open data culture is constantly growing, together with the awareness that open data represents an opportunity to enlarge and enhance academic net-works. The report highlights that the role of funding policies is pivotal in requiring open access and

---

[6]   https://www.eosc-pillar.eu/news/national-initiatives-survey-importance
[7]   A year-long, co-conducted study between Elsevier and the Centre for Science and Technology Studies.

open data, although the effort in research data management seems mainly assigned to researchers individually, since the 52% of respondents said that their institution does not provide funds to cover the costs of managing or archiving research data. Dissemination of data is primarily contained within the current publishing system. Indeed, the preferred method of research data dissemination is publication, though a large amount of data might remain unshared. When asked which Creative Commons license they would make their data available under, 62% answered that they did not know. Where researchers did provide an answer, they tended to favor more restrictive licenses. Overall, this suggests that researchers might have a lack of knowledge regarding the sharing and reuse of data, which may affect their willingness to do so. Indeed, according to the EOSC-Pillar National Initiative Survey (2020), only 50% of the e-Infrastructures in the countries covered by the survey have an access policy for services and data, while a large amount of them are still planning to introduce it in the next two years. There are also meaningful differences among countries; for example, in Italy access to data is usually granted to the internal organization only.

The EOSC survey also contributes to shed the light on the most cumbersome issues for the individual researcher. For example, data curation seems the activity gaining the lowest level of attention (38% of repositories perform basic data curation). Thus,

> "across countries, depositors are most concerned about the 'effort of preparing the data for publication' (58%), followed by `intellectual property (e.g. copyright)' (52%), `lack of control over the usage of data' (50%), `data protection' (47%), `benefit of sharing data' (40%) and `competitive disadvantage when sharing' (39%)" (EOSC-Pillar 2020: 105).

What emerges from the literature about open data is a picture of very scattered initiatives across and within disciplinary fields that are happening primarily at the individual level. Academic researchers may be motivated to share data because it can increase their visibility (Patel, 2016), but at the same time they are demotivated because they may face technical problems, or find the effort too great (Childs et al., 2014). However, skill-gaps (Zuiderwijk, 2015) and lack of control over data production hinder researchers from opening their data up.

The development of skills and dedicated profiles is, in our opinion a very critical point for open data. Also for academia, the research data management is a relatively new area of practices that implies new expertise and engagement programmes, and paradoxically, this is also the least frequently regulated aspect of open science (EOSC-Pillar 2020). For example, although, on average, respondents predominantly indicated that they were familiar or very familiar with the concept of FAIR data with little difference between countries, when surveying about the implementation of each FAIR principle the picture becomes quite scattered. Showing that academic institutions preferred not to adopt formal policies, leaving cultural diversity and professional change to be supported and fostered by other means (ivi). What emerges from the opinions of researchers is that RDM is still considered an administrative burden. In Italy, for example, it has been noted that the choices made by some funders (i.e. European Union Framework Program) impact on departmental research units and researchers in respect to the organization and activities of departments, the type of knowledge produced, and the ways of doing research (Primeri, Reale, 2012). We have already noted that open science has primarily developed as a top-down regulatory policy, in the same fashion open research data seems a reality for expert groups supporting the EU policymaking, but it has

not yet become mainstream for the individual researcher. Referring to public sector information, for example, the open data maturity report 2019 (EU28)[8] identified only three countries as 'trend-setters' (Ireland, Spain and France), while the most relevant group is the 'beginners' one. The report measured the data maturity level focusing on four pivotal dimensions: policy, portal, quality, and impact. It clearly shows the gap among the policy level of data governance (average 74%) and the impact level of re-use practices (average 57%), highlighting the relevance of creating a data culture in society as a whole[9].

In other words, it seems that various aspects of how academic works are supposed to change. Answering the request of opening the research data means that academics must learn about licenses, commons, intellectual property rights and must have a deep knowledge of different public's needs. Things that are somehow obscure for most academics. Such a misalignment is quite evident in our case study.

3.1. Open Science and the academic profession: the case of training for Research Data Management

This complex, layered nature of open science makes it both theoretically and empirically difficult to disentangle its various domains. It is therefore, hard to identify an entry-point where all these different elements may be simultaneously observed. We then decided to isolate a specific practice of open science in which the link between these different domains could emerge, and the consequences on the academic profession could be spotted: the Research Data Management (RDM). We focus on one case study about training in RDM, the project Open Badge Ecosystem for the Recognition of Skills in Research Data Management and Sharing (OBERRED). We participated in this project in order to deliver training to academics on RDM through open education means within an open ecosystem and employing an open skill recognition system.

The OBERRED project will be used to underline, first, how the specific practice of RDM intertwines with all the school of thoughts of the theoretical framework on open science described above (Fecher and Friesike, 2014), second how RDM in open science remodels the academic profession, and, finally, how RDM demands new skills. The project[10] aims at identifying the competences mobilized in RDM processes and making them available to the science communities through the development of open educational resources. 'Education and skills' are in fact one of the eight pillars of Open Science identified by the European Commission. By conducting this project, the ways in which RDM practices lay in the intersection of *people*, *things*, and *values* clearly emerged.

A first issue concerns the identification of how *people* define what research data are, and this affects scholars, publishers, stakeholders, and citizens differently. The first notion of data is, in fact,

---

related to where they come from (Research Information Network, 2008), but this may sharply vary between domains, and therefore between scholars, and also between other stakeholders interested in research data (e.g. citizens). To put it simply, the concept of data can change in relation to its recipient. Data may differ also for technological aspects that depend on how they have been collected and stored electronically through software, platforms and from other material actants (e.g. data protocols, FAIR principles, etc.). In other words, the mediation of *things* changes the concept of data identification and definition, playing an agency role. Finally, the definition of data may be sustained by specific *values*. For instance, the link between data and data life cycle (which distinguish between input data (microdata) process data (metadata) and output data (macrodata)) is clearly inspired by the need for reproducibility and research integrity (Merton, 1942). While the idea of the same data being the research outputs or inputs alternative to new hypotheses, policies, and scientific discoveries is clearly inspired by social impact and research responsiveness. Therefore, when practicing RDM concretely, the differences between the school of thoughts that support open science blur, making RDM a very cross-sectional practice that takes into account democratic as well as pragmatic outlooks, infrastructure and measurement stances as well as ethical, deliberative and public ones.

In addition, RDM is critical in order to reach a common definition of research data. For OBERRED participants it meant it was necessary to adapt the theoretical definitions already existing in its own domains with the practical processes followed in the collection, storage, management, and dissemination of research data. Moreover, the lack of controlled vocabularies and ontologies, as well as of semantic techniques to related data, is a sensitive issue in disciplines such as Social Sciences, and an evident barrier to interoperability and reuse of research data as recently stated by the RDA Social Sciences & Humanities Research Data Interest Group (2021)[11].

In such a (still domain-based) ecosystem, the definition of skills related to RDM in open science is quite challenging. We can refer to the European Charter for Researchers (ECR)[12], a set of 41 general principles and requirements approved in 2005 which specify the roles, responsibilities, and entitlements of researchers. The ECR has been revisited because of open science by introducing new skills (O' Carrol et al. 2017). For example, communication abilities are becoming pivotal when communicating findings in a comprehensible way, to different public audiences (e.g., citizens), and so managerial expertise that could transform research outputs in policy inputs for institutional and commercial practices. Also, political capabilities are required to address research responsibility and responsiveness in a governance context. To define what RDM is, who impacts on, and what skills should mobilize at what level, remains a tangled question. Also, when we compared three skills frameworks: the European Standard Classification of Occupations (ESCO), the Research Data Alliance, and the European Open Science Cloud (EOSC) we noted a sharp cultural mismatch between the space of policies and that of practices.

In response to funder mandates, the academic research community still needs to build a data culture (Aragona & De Rosa, 2018), especially as research support requirements grow, and open science practices develop. And this data culture is something that needs to be constructed also "from

---

[11]   https://www.rd-alliance.org/groups/social-science-research-data-ig
[12]   https://euraxess.ec.europa.eu/jobs/charter/european-charter

below", trying to engage all the stakeholders in the ecosystem, including citizens. The OBERRED project is a step toward the construction of this culture, but there is still much to do to fully understand what kind of consequences open science is generating, in the short, medium, and long term. The impact of open science on the academic profession should then be scrutinized in a more comprehensive way, with an organic research agenda.

4. An open conclusion and a research agenda

In an over simplified research approach two principles of differentiation organized the academic profession: one separating academics according to their main activities (research or teaching); another drawing territory around the different tribes (Becher, 1994), constituted by the disciplines and sub-disciplines. However, as emerged, for example, in the biography of Pasteur by Latour (2001), many academics were involved in several activities (to build what Latour (1987) and Callon (1989) would describe as «socio-technical networks») functional for their careers, and in turn for their scientific reputation. Bourdieu (1978) distinguished between pure scientific careers and careers built on a participation in the management of science, but at his time other activities were necessary, although they were not expected and not explicitly rewarded. The affirmation of open science seems to affect the academic profession in at least two ways, firstly with the diversification of academic activities, furthermore with the introduction of emergent and more hybrid academic roles, a multi-profiles research identity. Several policy frameworks emphasize the hybrid role of scientists equipped with skills which are often acquired in other activities and/or in other curricula or which are built on digital services provided by other professionals (EOSC 2021). Therefore, academic roles are more and more interrelated, and the process of change becomes a systemic question involving all actors and actants.

To be clear, these changes are part of a wider transition of the academic profession that relates to digitization (Decupyere, Simons, 2014) and with progressive rationalization of higher education institutions (Trela, 2012) that is taking shape in standardized performance evaluation (Espeland, Sauder, 2016). Digitization has affected the working lives and identities of academics (Lupton, 2015) and transformed many aspects of higher institutions (Williamson, 2018). The academic profession is also being re-modelled by the global process of rationalization which impacts on academic autonomy and identity. However, what is disruptive in open science is that it can intertwine these two concurring changes in a comprehensive narrative that blends the ethos of digital transformation with that of managerial rationalization into various aspects of at least three main levels: things (infrastructures, devices, techniques, etc.), values (democratization, replicability, efficiency, etc.), and people (researchers, employees, citizens etc.).

So far the Open Science paradigm seems deeply rooted in a co-evolutionary contest, which affects technologies, science, and epistemologies. The open paradigm is born from the availability of technology, but technology development largely depends on the changing academic culture and on the spreading of research results. Talking about the emergence of digital information infrastructures, Borgman, in fact, argued that:

"Some contend that these changes are revolutionary and will change the world; others argue that the changes are evolutionary, and that individuals and organizations will incorporate networked information technologies into their practices just as they incorporated many earlier media and technologies. […] I take the view that these changes are neither revolutionary nor evolutionary but somewhere in between: that they are co-evolutionary" (Borgman, 2007, p. 30).

Due to its co-evolutionary nature and despite all the positive narratives on open science, the promising idea of open science comes with a multitude of challenges, which need to be addressed to help this paradigm unfold and see how it will change academia. The advantages for the system are communicated both clearly and convincingly, but some risks may still be envisaged.

Open science culture has developed with big players setting the agenda. The idea of open science as a unique framework, guided by research funders may diminish the role of the specific forms of disciplinary knowledge and cultural diversity. In addition, the idea of reproducibility by default is intriguing, but not free from risks. Open research data, and open collaboration in the forms of open peer review and open research through collaboration, may sometimes hide an impulse towards the application of a market-oriented vision of science, with results being used only as products, shared with our peers on a convenience basis. A kind of managerialism of science or, in other words, the application of new public management to the scientific career (De Lourdes Machado-Taylor et al., 2017). Mirowski (2018) argues that «the open science movement is an artifact of the current neoliberal regime of science, one that reconfigures both the institutions and the nature of knowledge so as to better conform to market imperatives». Indeed, the emphasis on transparency does not eliminate the opacity of research governance. While research may be perceived as open and transparent, the digital infrastructures, the algorithms that run on digital repositories, and the frameworks adopted, may instead, as every socio-material entity, be rather opaque (Star, 1999), especially for academic researchers who have not been trained and accustomed to open science policies, procedures, and infrastructures.

In other words, open science is mainly characterized as a top-down approach, basically routed by funding institutions which have met some expectations of open access debate. Many steps forward have been made and several open initiatives raised all over the world. Have they really changed the academic work? How deeply? and in what directions? The impression is that much is still needing to be explored; looking at implications and latent effects of a concept that (despite the huge effort in creating a policy rhetorical set) is non-mainstream in every scientific field.

However, as it has happened in other human activities, the development of social media tools is also changing the professional ecosystem from bottom-up by creating a sharing culture. Platforms, such as Research Gate, already introduced a kind of 'sharing factor' in terms of reads, citations, recommendations, and total research interest as alternative metrics impacting on scholar reputation and research effectiveness as well. In other words, the open science paradigm is finding its own way, despite the official policy frameworks. At this point, new research directions concerning the dimension of micro-practices and methods are needed. These directions are all related to the changing of the academic profession that affect daily work, the way it is intended and the challenges it poses, and finally the role of researchers and science in the public sphere.

To introduce these micro-practices dimension, J. Dewey's work on the nature of the formation and transfer of knowledge in the process of scientific investigation comes to our aid (1958). For the well-known philosopher, the process of knowledge is contingent, that is, inscribed in the social context and bound to the intersubjective agreement between those who participate in various capacities. It is this agreement that gives the process the legitimacy it needs to reiterate it over time. Created to define that group of people involved in the scientific investigation process, the community of inquiry theory provided the opportunity to review the paradigms on the transmission of knowledge in terms of participation in an open community of investigation. The community of inquiry is a bottom-up model. In fact, according to the cited Open Science Monitor, researchers usually share research protocols, material, and data with the communities they belong to. They consider this step as a preliminary and necessary assessment stage. This is normal behavior for researchers who help in gaining peer recognition and social legitimacy. The same holds true for the claim for access to established findings and data that is usually granted at the level of author requests in the 'Onion Model' (Bowman, Keen 2018). However, the claim for reproducibility as an 'extension of certified knowledge' remains a questionable issue. As Chen (2018) has argued:

> "*The reuse of research results as a goal, requires the adoption of new research practices during the data analysis process. Such practices need to be tailored to the needs of each given discipline with its own research environment, culture, and idiosyncrasies. Services and tools should be developed with the idea of meshing seamlessly with existing research procedures, encouraging the pursuit of reusability as a natural part of researchers' daily work. In this way, the generated research products are more likely to be useful when shared openly*" (Chen, 2018, p. 113).

In this respect, openness is not enough. If sharing is the goal that the research should aim at, then it requires the adoption of a proactive approach from the very beginning of the research process, considering the whole scientific community at large as a community of inquiry. Some questions can be raised at this point. The first question that should be investigated concerns the possible emancipation of researchers from the communities they belong to. Open science has a strong normative nature, it exercises a prescriptive pressure at the individual and collective level in somehow bypassing the epistemic communities. For example, the 'platformization' (Wilkinson et al., 2016; Van Dijck, Poell, De Waal, 2018) of research through extended data infrastructure, common research protocols, open access repositories, and social platforms such as ResearchGate or Academia.edu is thus changing scholar micro-practices enlarging (and disintermediating) the environment in which they operate. For example, the growing referencing to altmetric factors in researchers CV, blog, or others informal writings is a clear demonstration of such a process of disintermediation to show scholarly media and public interest, or to emphasize the use of research in policy documents or other official publications. All information which may not appear in the conventional citation databases (Ayris et al., 2018).

In this context, what role are the epistemic communities going to play (Normand, 2016)? Under what conditions can the corroboration of research results be a seamless system of inward/outward processes? How is open science going to disrupt traditional hierarchies and cultures?

The second question concerns research integrity pre-registration to workflow and report of findings. Research integrity is also a matter of trust and reciprocity, definition of research background

and foreground, a system of non-written rules sedimented by, and routed into the academic culture during centuries of scientific dissemination and research praxis. Epistemic communities and editorial staff have played an active role in controlling and contrasting plagiarism. How will open science constitutionalize the new research environment? How might a system of rules arise between human and non-human actors? How do you transform trust and reciprocity from a value system into a set of performative actions?

The third question concerns the publishing practice (Heller et al., 2013). For a long time, publishing has been the main assessment criterion for researcher careers, a parameter and incentive for his/her productivity (O'Carrol, 2017). The open science regime requires researchers to change the focus of their activity, shifting their attention towards the sharing process and its correct management. How much will the open science regime affect the 'publish or perish' approach? Will new and dynamic publication formats emerge as forms of temporary recognition of compliance within the open science regime?

To conclude, open science is all about methods: accessing, sharing, assembling, pre-registering, reproducing and publishing are all the performative verbs identifying a specific set of actions connected to the process of standardization. But what is at the very core of the debate concerning open science is the science-society relationship. Opening the culture of scientific methods, the whole science instrumentation, is perceived in fact as a changing paradigm in knowledge accumulation which acquire a second dimension: its belonging to the public domain. Science is then called to act in the media sphere, to take part in the public debate to gain legitimation, cognition, and social acceptability (Lakomý et al., 2019), as is happening in the case of the Pandemic.. Clearly, the two dimensions (standardization and mediatization) need to be addressed.

While standardization (requiring a formalization language and a strong pressure for conformity) is the only way to make processes comparable and interoperable in nature, mediatization changes the inner logic of scientific organizations called to act on an uncertain terrain, where science is debatable. This is the space where the accountability of the world meets alternative epistemologies of science. How open can this space be?

References

Appadurai, A. (2013). *The future as cultural fact: Essays on the global condition*. Rassegna Italiana di Sociologia, 14, (4), pp. 649-650.

Aragona, B., De Rosa, R. (2018). *Policy making at the time of big data. Datascape, datasphere, data culture*. AIS Journal of Sociology, 11, pp.173-187.

Ayris, P. et al. (2018). *Open Science and its role in universities: A roadmap for cultural change*. League of European Research Universities (LERU) Advice paper no. 24 – may.

Bartling, S., Friesike, S. (2014). *Opening science: The evolving guide on how the internet is changing research, collaboration and scholarly publishing*. Berling, Springer Open.

Becher, T. (1994). *The significance of disciplinary differences*. Studies in Higher education, 19(2), 151-161.

Borgman, C. L. (2007). *Scholarship in the Digital Age: Information, Infrastructure, and the Internet*. Boston, MIT Press.

Bourdieu, P. (1978). *Les catégories de l'entendement professoral*. Actes de la recherche en sciences sociales, 24, 2, pp. 2–24.

Bowman, N. D., Keene, J. R. (2018). *A layered framework for considering open science practices*. Communication Research Reports, 35(4), 363–372. DOI

Callon, M. (1989). *La science et ses réseaux: genèse et circulation des faits scientifiques*. Paris, Éditions La Découverte.

Chen, X., Dallmeier-Tiessen, S., Dasler, R. (2019). *Open is not enough*. Nature Phys, 15, pp. 113-119.

Childs, S., McLeod, J., Lomas, E. e Cook, G. (2014). *Opening research data: Issues and opportunities*. Records Management Journal, 24, 2, pp.142-162.

De Lourdes Machado-Taylor M., Soare, V. M., Teichler U. (eds) (2017). *Challenges and options: The academic profession in Europe (Vol. 18)*. Berlin, Springer.

Decuypere, M., Simons, M. (2014). *On the composition of academic work in digital times.* European Educational Research Journal, 13, (1), pp. 89-106.

Dewey, J. (1958). *Experience and nature*. North Chelmsford (MA), Courier Corporation.

Espeland, W. N., Sauder, M. (2016). *Engines of Anxiety: Academic Rankings, Reputation and Accountability*. New York, Sage.

Feyerabend, P. (1975). How to Defend Society against Science, in E.D. Klemke e R. Hollinger, (eds) (1988), *Introductory Readings in the Philosophy of Science*. Buffalo (NY), Prometheus Books, pp.54-65.

Foucault, M. (1977). *Discipline and Punish*. London, Allen Lane.

Foucault, M. (1980). Power/knowledge: Selected interviews and other writings, 1972-1977. Paris, Vintage.

Heller L., The, R., Bartling, S. (2014). Dynamic Publication Formats and Collaborative Authoring. In Bartling, S., Friesike, S. (eds), *Opening Science. The Evolving Guide on How the Internet is Changing Research*, Collaboration and Scholarly Publishing. Berling, Springer Open.

Irwin, A. (1995). *Citizen science: A study of people, expertise and sustainable development*. Hpve, Psychology Press.

Kogan, M., Teichler, U. (eds) (2007). *Key challenges to the academic profession*. UNESCO Forum on Higher Education Research and Knowledge. Paris, INCHER-Kassel.

Lakomý, M., Hlavová, R., Machackova, H. (2019). *Open Science and the Science-Society Relationship*. Society, 56, 3, pp.246-255.

Latour, B. (1987). *Science in action*. Cambridge (MA), Harvard University Press.

Latour, B. (2001). *Pasteur: guerre et paix des microbes; suivi de Irréductions.* Paris, Éditions La Découverte.

Lupton, D. (2015). *Digital Sociology*. Abingdon, Routledge.

MacCallum, C. (2018). *Research Communication: Open Science & the perverse evaluation cycle*. Presentation delivered at the UCL Press – Open Access Megajournal Project Town Hall event, 16 January, London. DOI

Merton, R. K. (1942). *Science and technology in a democratic order*. Journal of legal and political sociology, 1, (1), pp.115-126.

Mirowski, P. (2018). *The future (s) of open science*. Social studies of science, 48, (2), pp. 171-203.

Nentwich, M., Konig, R. (2012). *Cyberscience 2.0: Research in the age of digital social networks*. Koln, Campus Verlag.

Normand, R. (2016). *The Changing Epistemic Governance of European Education: The Fabrication of the Homo Academicus Europeanus?* London, Springer.

Novitzky, P., Bernstein, M.J, Blok, V., Braun, R., Chan, T.T., Lamers, W., Loeber, A., Meijer, I., Linder, R., Griessler, E. (2020). *Improve alignment of research policy and societal values*. Science 369 (6499), 39–41.

O'carroll, C., Rentier, B., Cabello Valdes, C., Esposito, F., Kaunismaa E., Maas, K., Metcalfe, J., Mcallister, D. and Vandevelde K. (2017). *Evaluation of Research Careers fully acknowledging Open Science Practices: Rewards, incentives and/or recognition for researchers practicing Open Science*. Luxembourg, Publications Office of the European Union.

Owen, R., P. Macnaghten, and J. Stilgoe. 2012. *Responsible Research and Innovation: From Science in Society to Science for Society, with Society*. Science and Public Policy, 39 (6), 751–760.

Owen, R., & Pansera, M. (2019). *Responsible Innovation and Responsible Research and Innovation*. In Handbook on Science and Public Policy. Cheltenham, UK: Edward Elgar Publishing. DOI

Owen, R., von Schomberg, R., & Macnaghten, P. (2021). *An unfinished journey? Reflections on a decade of responsible research and innovation*. Journal of Responsible Innovation. DOI

Patel, R., (2016). *Research data management: A conceptual framework*. Library Review, 65, 4/5, pp. 226-241.

Primeri, E., Reale, E. (2012). *How Europe shapes academic research: Insights from participation in European union framework programmes*. European Journal of Education, 47, 1, pp.104-121.

Sidler, M. (2014). Open Science and the Three Cultures: Expanding Open Science to all Domains of Knowledge Creation. In Bartling, S., Friesike, S., *Opening science: The evolving guide on how the internet is changing research, collaboration and scholarly publishing*. Berling, Springer Open.

Star, S. L. (1999). *The Ethnography of Infrastructure*. American Behavioral Scientist, 43, (3), pp. 377-391.

Trela, J. (2012). *The rationalization of academic work*. Presentation at the Second ISA forum "Social Justice and Democratization" Buenos Aires, 1-4 Agosto.

Van der Zee, T., Reich, J. (2018). *Open education science*. AERA Open, 4, (3). DOI

Van Dijck, T., Poell, T., & De Waal, M. (2018). *The Platform Society: Public Values in a Connective World*. New York: Oxford University Press.

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., e Bouwman, J. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific data, 3. DOI

Williamson, B. (2018). *The Hidden Architecture of Higher Education: Building a Big Data Infrastructure for the 'Smarter University'*. International Journal of Educational Technology in Higher Education, 15, (1), p.1-12.

Zuiderwijk, A. (2015). *Open data infrastructures: The design of an infrastructure to enhance the coordination of open data use*. Hertogenbosch, Uitgeverij BOX Press.

About the Author

Rosanna De Rosa

ROSANNA DE ROSA is professor at the University of Naples Federico II where she teaches Political Communication and leads several European projects. Her publications are mainly aimed at analyzing the impact of new technologies on politics - with a focus on the constituency communication, digital democracy and electronic government- and on eResearch and education - with a focus on MOOCs, Learning Analytics, Open Science. Among her publications: *A research agenda on MOOCS: the perspective of social sciences*, Italian Journal of Educational Technolog, Volume 26, Number 2, 2018. With Biagio Aragona, *Big data in Policy Making*, Mathematical Computational Science (MPS), 15 Feb 2018. And *Unpacking Big Data in Education. A Research Framework*. Statistics, Politics and Policy, Vol.8, n. 2.

Biagio Aragona

BIAGIO ARAGONA is Associate professor in Sociology at the University of Naples "Federico II", where he teaches Digital methods and Big Data and Advanced methods for quantitative research. He is managing editor of the AIS Journal of Sociology. His research activity has always been dedicated to the use of secondary data for the analysis of society, to the use of big data and algorithms in policy making, and to the management of open data in social research.