EasyChair Preprint

№ 7351

Open Science for an Open Future

Mayukh Bagchi

January 19, 2022

Open Science for an Open Future

Mayukh Bagchi

Institute for Globally Distributed Open Research and Education *(IGDORE)*

mbagchi.research@gmail.com;

mayukh.bagchi@unitn.it

Abstract:

Open Science, in simple terms, refers to a movement that attempts to make the entire research cycle of a research pursuit, freely and openly available (the ingredients for an Open Future). The paper attempts a review of the various component-concepts in Open Science, deliberates on its various schools of thought, portrays its multifaceted character, and explores the various perceptions about the concept itself. It also focuses on the role of libraries in Open Science and provides a set of recommendations to make it a successful endeavour.

Keywords:

Open Science, Open Data, Open Definition, Libraries, FAIR Data, Citizen Science, Open Access.

1. Introduction

In a generic sense of the term, science can be understood as a methodical examination of the multifarious and multifaceted phenomena occurring in the vast world of nature and beyond, enquired primarily through the all-encompassing binary of observation and experimentation. Specifically, in terms of data, science can also be ideally interpreted as the unfettered collection, analyses, dissemination, scrutiny, re-evaluation & re-use of organized data (i.e., information). But, at least until recently, this very notion of science from the perspective of unrestrained data activities was a chimaera in practicability. The concept of Open Science is slowly but steadily changing the perspective.

Open Science (also known as Open Scholarship; reflective of its broad scope) refers to both the concept and the campaign of facilitating equitable and libre societal access to research data, research methods and methodologies, and, dissemination of research findings, i.e., opening up of the scientific research cycle as a whole to all those interested, regardless of societal hierarchy or professional background. It is often used as a hypernym, incorporating emergent practices such as Open Data, Open Educational Resources, Open Peer Review,

Open Access to Information, Open Innovation, Citizen Science and a host of other novel concepts and practices. Importantly, it also involves a change of culture, and considerable change of practices amongst participating entities in different arenas of society.

There are several motivations behind the global emergence of the integrated concept of Open Science. From the sociological viewpoint, Open Science is a harbinger of increased effectiveness, transparency, global participation, cross-disciplinary efforts and output-oriented research. That the findings and subsequent products of publicly funded research can be freely utilised by the public-at-large, is also an important economic marker of Open Science. Further, there are benefits of Open Science vis-à-vis digitisation, institution-society relationship and policy amendments. These, alongside other aspects of Open Science, are critically reviewed and recommended in the following sections.

2. Related Work

Till date, there has been a considerable number of studies which actually examine the various sub-concepts, issues and ramifications associated with the broader concept of Open Science.

The European Commission (2016) has interpreted Open Science as a latest approach to scientific research, highlighting the importance of scientific collaboration and modern ways of knowledge dissemination through digital techniques as its key enablers. The Organisation for Economic Co-operation and Development (OECD) in its landmark report in 2015, “Making Open Science A Reality'', sums up Open Science as the “encounter” between long held notions of openness in the scientific arena, and ICT tools to promote and publicise research in newer ways. Fecher and Friesike (2014) discusses how any change towards futuristic creation and publication of scientific knowledge, will be attributed to the very concept of Open Science, and its various schools of thought. Picarra (2016) stresses the importance of Open Science for policymakers, and how they can affect strategic policy changes towards such an agenda at various levels in public and private polity. The inter-relationship between Open Science and Open Innovation has also been explored, along with parameters like funding, policy incentives and intellectual property arising out of such a relation (Chesbrough, 2015).

Martens (2015) discusses in rich detail, the various benefits arising out of an adoption of an Open Science agenda. Centre National de la Recherche Scientifique (CNRS, France) in a white paper in 2016, hailing Open Science as a renaissance, urged the creation of a “Right to Open Science” and simultaneous balancing - cum - protection of genuine interest in scientific endeavours. Levin *et al*. (2016) attempts to understand and annotate the perception of Open Science amongst researchers (a cohort of 22 biomedical researchers in the U.K.), and provides generalised feedback on how to improve it through increased awareness, content enhancement and structured applications. O’Carroll *et al*., in an European Commission report (2017) introduces the “Open Science Career Assessment Matrix (OS-CAM)”, as a novel evaluation method to evaluate researchers based on Open Science parameters. The

Amsterdam Call for Action on Open Science (2016) mooted a multi-participant integrated approach for the systematic development and implementation of Open Science across Europe, and the world at large. Ayris and Ignat (2018) discusses how libraries can provide necessary leadership in implementing Open Science policies, and proposes a 4-step model to assess and step-up their engagement. Finally, Munshi and Madalli, in their Open Science country note on India, share fascinating insights regarding the existing & upcoming policy frameworks, skills, co-operations and institutional establishments vis-à-vis Open Science, in the Indian national context.

3. The Open Definition

To reiterate once again, the novelty of Open Science is that it opens up the whole scientific research cycle- from its initiation to its completion, and subsequent publication & application- and, to achieve such a comprehensive task requires widespread changes in scientific cultures & practices.

The Open Definition (also known as the six principles of Open Science) serves as the guiding cardinal on which the very concept of Open Science is based. These six principles of the definition represent the opening up of specific components in a research pursuit, the combination & organization of which constitutes the cohesive concept of Open Science. The six principles are as follows:

i) Open Data
ii) Open Source
iii) Open Methodology
iv) Open Peer Review
v) Open Access
vi) Open Educational Resources

Among these, the principles of Open Peer Review and Open Educational Resources were formulated later than the other four principles, which were proposed in a landmark paper which debated Open Science as a case in Technology Enhanced Learning (Kraker *et al*., 2011). The principles are briefly discussed as follows:

Open Data refers to data that is freely accessible to anyone and everyone interested, open for reuse, sharing, redistribution and republication, without any form of restriction or constraint (at most, there might be a requirement of attribution of the initiator). Alongside textual data, it might include varied non-textual technical representations of the likes of scientific graphs, maps, mathematical formulae and chemical diagrams (etc.). Sources of Open Data may be scientific (Human Genome Project; Dataverse Network Project), governmental (data.gov.in; data.gov.uk; data.gov {USA}) or intergovernmental (EU Open Data Portal). When published as linked data, it is referred to as Linked Open Data (LOD). It is of special interest to scientific workers in the library and information domain.

Open Source, in this context, refers to software that has been released under open licence, i.e., its source code is freely available for use, modification or distribution (Feller and Fitzgerald, 2002). Such software are often products of public collaboration, on open web-based platforms such as SourceForge (etc.).

Several significant studies, over the years, have pointed out that research papers, in general, don't represent the exact methodology required to reproduce research results, i.e., there is a mismatch between how it is actually done, and how it is represented in writing. To mitigate such a discrepancy, ***Open Methodology*** was proposed, in which the whole gamut (inclusive of all steps & ancillary activities) of methodologies adopted for a particular research problem is described in detail (implicit knowledge is often not captured, due to its inherent nature). Projects such as Stanford Exploration Project (SEP) and SAHARA Labs are already delivering on this aspect.

Open Peer Review is a term without a universally agreed definition. Ross-Hellauer (2017) analysed around 122 disparate definitions of Open Peer Review, and, came to a conclusion that it is a hypernym which basically aligns the peer review process with the objectives of Open Science, and includes: Open Identities (authors and reviewers know each others' identities) and Open Interaction between them, Open Pre-Review Manuscripts and Open Review Reports, Open Participation from the community-at-large, Open Final-Version Commenting and Decoupled Review (review done by another organization, other than the publisher).

Peter Suber (2015) defines ***Open Access*** to information as literature which is "digital, online, free of charge, and free of most copyright and licensing restrictions". Ordinarily, it refers to outputs of academic research, which are distributed online, irrespective of any form of restraint. There are several routes/models to achieve Open Access (Green, Gold, Hybrid respectively). Prasad and Madalli (2007) emphasizes the importance of the implication that, for "Knowledge for all" to be achieved, "Information for all" has to be ensured. Further, Prasad and Madalli (2007) proposes OPEN MANTRA, involving OPEN SOURCE tools built upon OPEN STANDARDS to realize OPEN ACCESS to information. It is also, like *Open Data,* of special interest to library and information professionals.

Open Educational Resources (OER) refers to educational resources, of any kind, which are openly accessible for use by all stakeholders in education and knowledge communities, without any constraint or any need for payment. Butcher (2015) asserts it as a game changer in transforming the global education scenario. Neither are all OERs e-learning, nor all e-learning OERs.

These are, therefore, the various components of the Open Science Taxonomy. It is also worthy to mention that these are the necessary steps for a researcher to work within the framework of Open Science.

4. Schools of thought

Based upon analysis of differing and multi-dimensional perception of the concept of Open Science, Fecher and Friesike (2014), in their landmark study, proposed five schools of thought within the ambit of Open Science.

The ***Public School*** stands on the belief that there is an increasing need to make science more and more accessible, and understandable for the public-at-large. For it to fructify, there should be enhanced interaction between two of the most important stakeholders in the Open Science process: scientists and citizens. Citizen Science, Science Blogging and Science Public Relations are believed to be key enablers within the objectives of this school.

The ***Democratic School*** argues that there should be unhindered accessibility to knowledge, in terms of free and open availability of research products. Apart from scientists and citizens, this school offers politicians a golden opportunity to really affect a change in science policy. Advocates of this school are involved in research on Open Access to Information, Open Data and Open Code.

The ***Pragmatic School*** views Open Science as a catalyst to make dissemination of knowledge of scientific research more effective, through modular knowledge creation, opening up of the research chain and increased interdisciplinary collaboration, guided by the true notion of 'openness'. It puts the onus on scientists, and calls for more collaborative research in Open Data and Open Code.

The ***Infrastructure School*** perceives Open Science as a concept rooted in technology, and assumes that the efficiency of research depends on the available scientific and technical infrastructure. It calls on scientists and service providers to create, adopt and use open-collaborative scientific infrastructure for research (such as distributed computing, one famous example of which is the *Open Science Grid*).

The ***Measurement School*** explains that, to measure the impact created by any scientific research, there is an urgent need to implement alternative forms of scientific impact measurement. It encourages rating scholarly outcomes through hidden metrics such as sharing, bookmarking, reading (etc.) (known by the collective term: ***Altmetrics***).

5. The Eight Pillars Of Open Science:

The League of European Research Universities (LERU), in its advice paper no. 24 (May, 2018), deliberates on the eight pillars of Open Science, which has been identified as of utmost importance by the European Commission.

i) **'*The future of scholarly publishing*' (1st Pillar):** It deals with how conventions around research are fast changing with the rapid development in digitisation and digital delivery, and how Open Access to information can be integrated into it. It calls for careful deliberation on the definition of Open Access on a discipline-to-discipline basis, and mentions several leading declarations which were issued in support of it. Further, it discusses several possibilities which might advance the concept of Open

Access scholarly communication and publishing, including concepts which are explicitly existent, like Open Access Monographs, ORCIDs (Open Researcher and Contributor ID), MOOCs (Massive Open Online Courses), OERs, BOOC (Books as Open Online Content), and, emerging concepts like megajournals.

ii) **'*FAIR Data*' (2nd Pillar):** In the present research scenario, it is no longer sufficient to open up access to research publications. It is equally important to freely and openly share research data, in an effort to contain duplication and step-up validation and re-distribution, obviously according to established protocols in *Research Data Management (RDM)*. Moreover, there has been a pressing need to share data in such a form that is both machine and human readable and re-usable. The *FAIR Data Principles* emerged out of this need (Wilkinson, 2016). The acronym FAIR stands for Findable, Accessible, Interoperable and Reusable, and these are the very characteristics that research data should be endowed with, on sharing. Metadata plays a pivotal role in the *"FAIRification Process"* of research data before sharing, and this is exactly where librarians and information scientists have a crucial role to play.

iii) **'*The (European) Open Science Cloud (OSC)*' (3rd Pillar):** Being an important component of the European Commission's Open Science strategy, it attempts to accentuate the transition from the present pre-dominant research scenario to efficient Open Science in the digital, cloud-based platform. It comes with benefits like free access to data and services, regardless of socio-geographical or disciplinary boundaries.

iv) **'*Education and Skills*' (4th Pillar):** It is of paramount importance that the various stakeholders in the research process- students, researchers, staff and the leadership- should be aware of the broad concept of Open Science and its numerous benefits. Necessary training should be imparted to all involved in research, so that they acquire the requisite skills for smoothly transforming in favour of Open Science.

v) **'*Rewards and Incentives*' (5th Pillar):** Here, rewards and incentives are discussed in the sense of value of Open Science to careers of researchers. The institutional and funding policies are, mostly, still aligned to conventional metrics for the performance evaluation of researchers'. It argues that Next-Generation Metrics should, at least, complement the performance evaluation process, and trigger a change in conventional policies.

vi) **'*Next-Generation Metrics*'' (6th Pillar):** In addition to more cautious use of existing bibliometric measures (journal impact factor, for example), this pillar encourages the supervised use of new measures in altmetrics, which are more aligned towards the objectives of Open Science and which uncovers many of the apparently hidden factors capable of judging the impact of a research contribution. Library and

Information specialists are poised to make lasting contributions to this specific aspect of Open Science.

vii) **'*Research Integrity*' (7th Pillar):** Science and research integrity go hand-in-hand. Practices in research should be based on mutual trust, honesty and transparency. Open Science helps to strengthen such an atmosphere through its novel practices including Open Access, Open Data, Open Methodology, *Open Citations* and *Open Peer Review*.

viii) **'*Citizen Science*' (8th Pillar):** It refers to the involvement of amateur citizens in scientific research. It is a beneficiary as well as a contributor to Open Science. It allows citizens to take part in professionally supervised research projects, contribute (through means like data collection) and in the process develop scientific temper. Open Science practices like *Open Source Software (OSS)*, *Open Data Standards* and Open Access to information helps to encourage such a trend.

6. Perceptions about Open Science

There are mixed perceptions and differing interpretations about the benefits and challenges involved in the realization of Open Science. Several research consortiums, policy advocacy groups and individuals involved in scientific endeavours have voiced their opinions about it. Several of those considerations are reviewed and generalised below (not exhaustive).

In favour of Open Science (inclusive of benefits):

i) Undoubtedly, Open Science actively encourages *transparency, reproducibility and validation* of scientific research.
ii) Freely accessible research data, research methodologies and research reports facilitates strong peer-reviewing.
iii) Public receives benefits out of the research they are funding.
iv) Due to its collaborative and interdisciplinary nature, it is also capable of solving intricate research questions.
v) Most importantly, it magnifies to a large extent the visibility and impact of a specific research exercise (like jump in the number of citations), and brings incentives for the research staff also.

Not in favour of Open Science:

i) There is an apprehension that the various components of Open Science might be subjected to misuse & misinterpretation.
ii) It leads to information overload for the scientific community.
iii) Apprehensions about the quality of research conducted and unmet technical standards.

Challenges:

i) Awareness about copyright assignment and copyright management is still in its infancy among many researchers.
ii) Developing infrastructures for Open Science can be a costly affair.
iii) There are certain data (such as *data concerning national security*) which should, ideally, never be made freely and openly available.
iv) The concept itself is not a universally agreeable one; there are disagreements as regards its models, metrics, infrastructure and genuine openness.
v) There is the huge & overwhelming challenge of *initiating a paradigm shift of culture and practice*, without which, Open Science efforts will remain futile.

7. Libraries and Open Science:

Libraries and librarians, being ageless custodians of data, information and knowledge, can play an engaging and decisive role in the promotion and adoption of Open Science. They are capable of providing leadership to the global movement of Open Science (Ayris and Ignat, 2018) at all levels of the social hierarchy. Nielsen (2013) elucidates the three different shifts that science is undergoing presently:

i) Collaboration leading to scientific knowledge creation
ii) Discover meaning in knowledge
iii) Recalibration of relations between science and society.

Libraries, as leading social institutions of knowledge management, are aptly suited to manage the above listed shifts in science, and subsequently guide the transformation to an era of networked and Open Science. In pursuance of such objectives, they can establish dedicated monitoring groups such as Open Access Team, Open Data Team etc employing specialized professionals like *data analysts, data curators and data librarians*. In fact, University College London (UCL) libraries have such infrastructures in place to take forward Open Science across their institution (Ayris and Ignat, 2018).

OECD (2015) defines the role of libraries in the Open Science movement as that of enablers & infrastructure providers, considering their inherent expertise in collection, curation and dissemination of scientific research through digital means. The different routes through which libraries can fulfil their role as enablers in the Open Science process are:

i) Libraries can play a *pivotal role in campaigning* for the various benefits and incentives that are associated with the adoption of Open Science.
ii) Libraries can provide crucial *scientific infrastructural support* (governance of Open data repositories, metadata management, information retrieval etc.) to promote Open Science.

iii) Librarians can help to develop *Research Data Management infrastructure*, theoretically and empirically.
iv) Most importantly, librarians can guide, train and serve researchers on how to open up access to their research cycle.
v) Librarians, employing their thorough expertise in *meta-knowledge*, can even guide researchers on their quest to find research funding agencies with compulsory Open Science requirements, thus giving a boost to the overall movement.

Libraries can also take inspiration from several organizations and projects (Centre for Open Science; *Open Knowledge Foundation*; Polymath Project; *Public Library Of Science [PLOS]*; Pre-print servers like *ArXiv* etc.) which have done genuine work in advocating Open Science.

8. Recommendations:

Finally, after a thorough review of the spectrum of concepts involved in Open Science, the author has several recommendations to make, primarily based on *observable lacunae or possible conceptual extensions*, in an attempt to make the Open Science framework more robust. The recommendations are listed in points below:

i) Active advocacy, aided by competent management professionals and an objective communication outreach, will go a long way in effecting *cultural changes*, enhancing awareness about the potential benefits of Open Science while being cautious about its challenges.
ii) *Pushing for statutory mandates* in institutions in support for implementation of full Open Science practices.
iii) Compartmentalization of Open Science tasks among related departments in an institution, led by its *'library and information arm'*.
iv) Framing of an institutional Research Data Management policy and a similar policy on *Data Stewardship*, encapsulating the essence of FAIR Data and based on Open Science foundations.
v) Initiating expert training sessions on the various technicalities faced in the implementation of Open Science practices.
vi) Establishment of *Open Discovery services*, which will aid researchers in their search for various open components of Open Science, and guide them on how to re-use them.
vii) Formal establishment of incentives, recognition and continued support (in terms of skill-based training) for those embracing Open Science practices (including for participants in Citizen Science).
viii) Implementation of *Open Science based HR (Human Resources) management*- including recruitment, promotion and performance assessment- thus giving a strong thrust to the adoption of Open Science practices.

ix) Continuous monitoring of *Open Science based research workflows*, ensuring adherence to the best standards of integrity in research, and codification of institutional guidelines for researchers regarding research integrity.
x) Formation of a standard checklist to evaluate research institutions on several counts of the Open Science agenda, through frequent updation.

9. Conclusion & Future Work:

Before concluding the discussion, it would be pertinent to note that Open Science has boundless power to induce a tectonic shift in the way universities and research institutions initiate, handle and disseminate research outputs. It also has the ability to effect a change of culture and practices, thus leading to increased two-way interaction between citizens and science- a perfect recipe for an Open Future (a future where data, information and knowledge will be freely, fairly and openly accessible). Future work in the arena of Open Science would mostly revolve around technological solution-support, policy formulations, country-specific models/plan of actions towards Open Science implementation and more, in-depth research on hitherto unknown concepts, which might strengthen it more in the upcoming future.

References

Del Soldato, Eva (2016). Zalta, Edward N., ed. *The Stanford Encyclopedia of Philosophy* (Fall 2016 ed.). Metaphysics Research Lab, Stanford University

David, P.A. (2007) 'The Historical Origins of 'Open Science': An Essay on Patronage, Reputation and Common Agency Contracting in the Scientific Revolution', Capitalism and Society, Vol. 3, Iss. 2, Article 5.

Eamon, W. (1985) "From the Secrets of Nature to Public Knowledge: The Origins of the Concept of Openness in Science." Minerva, Vol. 23, No. 3, pp. 321-347.

Fecher, B. and Friesike, S. (2013) Open Science: One Term, Five Schools of Thought, RatSWD, Vol. 218, pp. 1-11.

Council of the European Union (2016). "The transition towards an Open Science system", 27 May.

European Commission (2016). Open Innovation, Open Science, Open to the World. A vision for Europe. May. https://doi.org/10.2777/061652; last accessed 05 June 2018.

Wilsdon, J.R. et al. (2017). Next-generation metrics: Responsible metrics and evaluation for open science. Report of the European Commission Expert Group on Altmetrics, European Union, March. https://doi.org/10.2777/337729; last accessed 5 June 2018.

Bartling, Sönke and Friesike, Sascha (2014): Opening Science: The Evolving Guide on How the Internet is Changing Research, Collaboration and Scholarly Publishing, Cham et al.: Springer.

Suber, Peter (2012). Open access (The MIT Press Essential Knowledge Series ed.). Cambridge, Mass.: MIT Press. ISBN 9780262517638. Retrieved 2018-04-28

OECD (2015), "Making Open Science a Reality", OECD Science, Technology and Industry Policy Papers, No. 25, OECD Publishing, Paris. http://dx.doi.org/10.1787/5jrs2f963zs1-en.

Nielsen, M. (2013). Reinventing Discovery: The New Era of Networked Science. Princeton, New Jersey: Princeton University Press.

Ayris, P. and Ignat, T. (2018). Defining the role of libraries in the Open Science landscape: a reflection on current European practice. Open Information Science, 2(1), pp.1-22.

Wilkinson, M.D. et al. (2016). "The FAIR Guiding Principles for scientific data management and stewardship", Scientific Data 3, 15 March. https://doi.org/10.1038/sdata.2016.18; last accessed 05 June 2018.

A.R.D. Prasad and Devika P. Madalli (2007). OPEN MANTRA for Open Access to Information. In ISI 2007 - 10th International Symposium for Information Science May, 30th - June, 1st, 2007. University of Applied Sciences, Cologne Germany

Butcher, N. (2015). *A basic guide to open educational resources (OER)*. Commonwealth of Learning (COL);.

Ross-Hellauer, T. (2017). What is open peer review? A systematic review. *F1000Research*, *6*.

Feller, J., & Fitzgerald, B. (2002). Understanding open source software development (pp. 143-159). London: Addison-Wesley.

Kraker, P., Leony, D., Reinhardt, W., & Beham, G. (2011). The case for an open science in technology enhanced learning. *International Journal of Technology Enhanced Learning*, *3*(6), 643-654.

Usha Munshi and Devika P. Madalli (n.d.). India Open Science country note. Retrieved from www.innovationpolicyplatform.org/content/india-open-science-country-note.

Enserink, M. (14 April 2016). "Dutch push for a quantum leap in open access". *Science*. **352** (6283)

Chesbrough, H. (2015). From open science to open innovation. *Institute for Innovation and Knowledge Management, ESADE*.

Picarra, M. (2016, April 4). Discussion Paper: Researchers and Open Science. Zenodo. http://doi.org/10.5281/zenodo.51858

Further Readings:

Collaboration, Open Science (2012-11-01). "An Open, Large-Scale, Collaborative Effort to Estimate the Reproducibility of Psychological Science". *Perspectives on Psychological Science*. **7** (6)

Belhajjame, Khalid; et al. (2014). "The Research Object Suite of Ontologies: Sharing and Exchanging Research Data and Methods on the Open Web". arXiv:1401.4307.

Bagchi, M. (2020). Conceptualising a Library Chatbot using Open Source Conversational Artificial Intelligence. DESIDOC Journal of Library & Information Technology, 40(6).

Bagchi, M. (2019). A Knowledge Architecture using Knowledge Graphs. MS Dissertation, Indian Statistical Institute.

Bagchi, M. (2021). A Large Scale, Knowledge Intensive Domain Development Methodology. Knowledge Organization, 48(1), 8-23.

Bagchi, M. (2021). Towards Knowledge Organization Ecosystem (KOE). Cataloging & Classification Quarterly, 1-17.

Satija, M. P., Bagchi, M., & Martínez-Ávila, D. (2020). Metadata management and application. LIBRARY HERALD, 58(4), 84-107.

Bagchi, M. (2019). Smart Cities, Smart Libraries and Smart Knowledge Managers: Ushering in the neo-Knowledge Society. 50 Years of LIS Education in North East India.