Do You Speak Open Science? Resources and Tips to Learn the Language.

Paola Masuzzo[1, 2] - ORCID: 0000-0003-3699-1195, Lennart Martens[1,2] - ORCID: 0000-0003-4277-658X

Author Affiliation
[1] Medical Biotechnology Center, VIB, Ghent, Belgium
[2] Department of Biochemistry, Ghent University, Ghent, Belgium

Abstract

The internet era, large-scale computing and storage resources, mobile devices, social media, and their high uptake among different groups of people, have all deeply changed the way knowledge is created, communicated, and further deployed. These advances have enabled a radical transformation of the practice of science, which is now more open, more global and collaborative, and closer to society than ever. Open science has therefore become an increasingly important topic. Moreover, as open science is actively pursued by several high-profile funders and institutions, it has fast become a crucial matter to all researchers. However, because this widespread interest in open science has emerged relatively recently, its definition and implementation are constantly shifting and evolving, sometimes leaving researchers in doubt about how to adopt open science, and which are the best practices to follow.

This article therefore aims to be a field guide for scientists who want to perform science in the open, offering resources and tips to make open science happen in the four key areas of data, code, publications and peer-review.

The Rationale for Open Science: Standing on the Shoulders of Giants

One of the most widely used definitions of open science originates from Michael Nielsen [1]: "Open science is the idea that scientific knowledge of all kinds should be openly shared as early as is practical in the discovery process". With this in mind, the overall goal of open science is to accelerate scientific progress and discoveries and to turn these discoveries into benefits for all. An essential part of this process is therefore to guarantee that all sorts of scientific outputs are publicly available, easily accessible, and discoverable for others to use, re-use, and build upon.

As Mick Watson has recently wondered, "[...] isn't that just science?" [2]. One of the basic premises of science is that it should be based on a global, collaborative effort, building on open communication of published methods, data, and results. In fact, the concept of discovering truth by building on previous findings can be traced back to at least the 12th century in the metaphor of dwarfs standing on the shoulders of giants: "*Nanos gigantum humeris insidentes*"[1].

While creativity and intuition are contributed to science by individuals, validation and confirmation of scientific findings can only be reached through collaborative efforts, notably peer-driven quality control and cross-validation. Through open inspection and critical, collective analysis, models can be refined, improved, or rejected. As such, conclusions formulated and validated by the efforts of many take prominence over personal opinions and statements, and this

[1] Metaphor attributed to Bernard of Chartres, and better known in its English form as found in a 1676 letter of Isaac Newton: "If I have seen further, it is by standing on the shoulders of giants"

is, in the end, what science is about. While science has been based for centuries on an open process of creating and sharing knowledge, the quantity, quality, and speed of scientific output have dramatically changed over time. The beginning of scholarly publication as we intend it today can be traced back to the 17th century with the foundation of the 'Philosophical Transactions'. Before that, it was not at all unusual for a new discovery to be announced in an encrypted message (*e.g.*, as an anagram) that was usually indecipherable for anyone but the discoverer: both Isaac Newton and Leibniz used this approach. However, since the 17th century, the increasing complexity of research efforts led to more (indirect) collaborations between scientists. This in turn led to the creation of scientific societies, and to the emergence of scientific journals dedicated to the diffusion of scientific research. Paradoxically however, knowledge diffusion has dramatically slowed down over the same time. In his review of Michael Neilsen's book "Reinventing Discovery" [3], Timo Hannay describes science as "self-serving" and "uncooperative", "replete with examples of secrecy and resistance to change", and furthermore defines the natural state of researchers as "one of extreme possessiveness" [4]. Hannay might have a point: the majority of research papers are behind a paywall [5], researchers still fail at making data and metadata available [6], reproducibility is hampered by the lack of appropriate reporting of methodologies [7], software is often not released [8], and peer-review is anonymous and slow [9].

As a reaction, the open science movement was born, almost as a counterculture to the too-closed system that re-emerged over the past few decades. More and more academic and research institutions are currently opening up the science they produce, making the scientific research, produced data and associated papers accessible to all levels of an ever more inquiring society, amateur or professional. And increasingly, major funding agencies are mandating the same. For example, the European Commission requires participants of the H2020 funding framework to adhere to the Open Access mandate and the Open Research Data Pilot. Furthermore, both the National Institutes of Health (NIH) and the Wellcome Trust have developed specific mandates to enforce more open and reproducible research. As a result, practicing open science is no longer only a moral matter, but has become a crucial requirement for the funding, publication, and evaluation of research.

Because the many benefits of open science have already been extensively studied and reported [10–16], this article instead intends to be a user guide for open science. The next sections of this article therefore provide an overview of the key pillars of open science, along with resources and tips to make open science happen in everyday research practices. This collection of resources can then serve as an open science guidebook for early-career researchers, research laboratories, and the scientific community at large.