

# HISTORY, ONTOLOGY, SCIENCE STUDIES

## HOW TO STUDY OPEN SCIENCE AND SCIENTIFIC DATA

Judit Gárdos

**BASED ON THE TRACKS “OPEN SCIENCE” AND “LIVES AND DEATH OF DATA” AT THE YEARLY EASST/4S CONFERENCE IN BARCELONA I MAKE SOME REMARKS ON HOW SCIENCE STUDIES ANALYSE THE SO-CALLED “OPEN SCIENCE” AND THE “LIFECYCLE OF DATA”. I PROPOSE THAT INCLUDING ONTOLOGICAL AND HISTORICAL ASPECTS WHEN STUDYING THESE TOPICS MIGHT BENEFIT OUR UNDERSTANDING OF THE METHODOLOGICAL, POLITICAL, SCIENTIFIC OR CULTURAL DETERMINANTS OF THE EMERGENCE OF THESE CATEGORIES, AND ALSO HELP UNVEIL WHAT OUR OWN ROLES AS SCIENTISTS ARE IN SHAPING THE THINGS WE THEN ANALYSE.**

An event as the EASST/4S annual meeting can be a key object to study the ever-forming discipline of *science studies*. I will try to use my scattered experiences of two different tracks to draw a few draft conclusions about some methodological features of our discipline, and propose some more critical research questions that could shape *science studies*.

In the track Open Science we heard presentations, among them many case studies on how and which actors should be and are included in science (for example local communities affected by macro-level political decisions about the environment, the general public concerning GMO, Wikipedia-contributors, lay biologists, urban communities, etc.). There were lively discussions about who thinks what is important about open science. We also heard recommendations how to achieve the goal of “open science”. A considerable part of the scholars working on open science (and some of them presenting in Barcelona) are also working in policy-related committees or other organizations to foster open science.

Often, open science is contrasted with the issue of whom the data belongs to. Making such a contrast – we have to be aware – constructs a space where property issues are opposed to a concept of an open science where openness is a normatively positive entity (which should be achieved, with the help of scholars/experts in committees). If we understand science in this manner and in this context of contrast, then propriety issues will tend to attain a negative connotation. However, in my incomplete perception of this track and its presenters, not very often were such ontological questions about the category “open science” asked: it was not in the main focus to address why and with the help of which people or groups this category emerged and was shaped during history; under what political, cultural, scientific contexts it operates; what functions it has or had in shaping society, business, culture or science. Rather, most of the time, the discussions covered the different semantics of open science (of course, not always – for a detailed and thorough analysis of the track see Mayer and Aibar’s review in this issue, analysing the different semantics of openness in the presentations as well), about the different perceptions (for example of stakeholders, policy-makers) on what open science is, or about how open science is performed (on Wikipedia, in journals, in participatory science projects, etc.). We also heard many presentations on the question of how to implement open science; this latter question can be characterized as presupposing a *normative* understanding of open science.

The notion of openness that was so frequently used has not really been critically analysed in the majority of the contributions – apart from few, but notable exceptions. It seems for sure that “openness” is positively connoted. Such a connotation has been part of Western scientific tradition since modernity (see for example Merton’s

scientific norm of communism, Merton 1942). In his paper on magic and science, talking about historiography, Láng (2015:127) rightly points out that “(r)esearchers simply accepted the view that openness is a positive value that supports academic research, and that secrecy, which is more characteristic of the history of technology, was fortunately abandoned by modern science.” But Láng (2015: 125) also observes that “many scholars have shown how secrecy in science became not only a tool of protecting knowledge from intellectual competitors, but also a dynamic social practice, a force that creates and organizes groups, and influences the mechanisms of exclusion-inclusion”. The analysis of these and similar questions in relation to the many practices around the definition and practice of so-called “open science” might produce valuable knowledge for science studies.

In a way Western democracy seems to be the normative backdrop of the dialogue on open science; but let us play a little bit: What could be the antonym of open science? It could be many things: closed science, science for the few, science for the privileged ones, etc. All of the antonyms shed light on one or another aspect of open science that could be studied by science studies scholars, ever more so if we wanted to accept the normatively positive notion of open science, as it is widely accepted nowadays. Some questions for future empirical analyses of scientific practices could include: What are the normative, scientific or political stakes for different disciplines in performing the movement of open science? Which groups are leading the discussion in this field? Why and how do disciplines, scholars, policy-makers focus on activities regarded as fostering open science? What are the performances in this field? What is regarded as closed science? How does this narrative of bad closed and good open science shape scientific activities? How did this opposition come into being in the first place? These possible questions would shed light on open science from a meta-level: they would show us the processes how the *concept* of open science is shaped culturally, socially or scientifically, and those cultural, scientific or social entities and their networks that emerge from these processes. Such an approach would *not* focus on – as did many of the excellent presentation we heard in the track – how a *pre-defined* “open science” is made, manufactured, constructed, performed or used. It would rather study how the thing we now call “open science” came into place, what its ontological status is, what its changing roles and relationships are in the cultural, political, scientific landscapes of other entities. Steven Shapin (2008: 222-223), in a similar approach, describes for example how what we now define as openness was a normality in the 1970s among biologists in the academia, most of them living scarcely off their salaries; but when the industry became aware of the profitable nature of genetics, scientists were tempted to change their workplace and work in less open circumstances. The opposition in the narratives about dirty secretive industry as opposed to pure and virtuous open science emerged because of these developments: “Since there was no money, a sense of sainthood was required in the situation”, said a student about research in academia (Philip J. Hilts: Scientific Temperaments: Three Lives in Contemporary Science, quoted by Shapin 2008: 223).

I will now propose some possible similar research questions based on another field in science studies, dealing with scientific data. Again, I will focus on what could be an *ontological* analysis about the emergence and ever-changing status of the different things we in STS call “data”.

In the track “Lives and Deaths of Data” the focus of many of the talks was on the different ways of the “interpretation” of data, the different stops of their “journey”, the “changes” in the translations of data.<sup>1</sup> The topics were, among others: sensitive health “data” and their context, discussions around and interpretations of astronomical “data”, “data” sharing practices and inequalities, the commodification of “data”, configurations of public and non-public “data”, etc. Among the many possible definitions of data there was one feature that came up quite often: that data is something that can be circulated (implying as well that it can be used several times).

The secondary use of scientific data, that seems to be one defining feature of data in this view, has been a contested issue for decades. The relationship of what is usually called metadata and data or the relationship of data and context are not self-evident. Even these distinctions are under scientific scrutiny (Mauthner-Gárdos 2015). Postmodern theories have questioned the assumption that data are neutral or objective representations of the world. Performative scholars (Barad 2007) have

<sup>1</sup> I quote here the introductory speech of this track by Sabina Leonelli.

<sup>2</sup> One of the notable exceptions in this track was Haider's and Kjellberg's analysis about the relationship of the structure of a big scale experimental facility and the type of data it produces. They stressed that the meaning of data starts before researchers begin their work.

challenged representationalist approaches (many of them postmodernist or constructivist ones); such approaches, these scholars say, still stick to the view that scientific data somehow represent natural or social worlds (even if these approaches do not necessarily narrate around terms like *objectivity* or *subjectivity*). It would be interesting to analyse – in a performative approach – how the notion that data can be circulated itself presupposes a specific notion of data and thus a specific way how data can be analysed; a question that has not quite been in the focus of the presentations in this track. If we understand data as something that can be circulated (and many presenters in this track shared this view), then one of the foci of such science studies analyses about data will be the ways data circulate or data are transmitted, and how different people “interpret” the “same” data. Scientific data will be a well-defined entity without borders that are contested, without ends that may fray. Science on data will be then somewhat less on what cultural, scientific, social, etc. traditions and surroundings influence what counts as data in the first place<sup>2</sup>, on what in our world does not qualify as scientific data, let alone on the ways how we as science studies scholars choose our objects of study; short: on the *ontological* status of data in sciences (in relation for example to other types of data, or other similar entities in sciences that end up not being called data) and, importantly: on data as the *object of scientific enquiry* in STS. In this track, the main focus – of course with exceptions, mostly ethnographic, close-up analyses of processes that result in the production of entities then called data – was less on these latter aspects.

It might be fruitful, if we want to reflect on our own methods as scientists, to look at our ways how we define data, or open science (or anything else as a matter of fact) and at the causes of selection of things that seem worthy of analysing. Also, I propose to analyse to greater extent the *ontology* of data or open science: what is regarded as data or open/closed science, which scientific, methodological or other traditions influence how these notions came into being in a specific scientific discipline at a specific time in history, at a specific place on Earth.

So: the questions that might be valuable to elaborate and that were – in my view – a bit underrepresented in the tracks under review: What is regarded/defined as data or open science and what not? What are the disciplinary, methodological, political etc. factors that play a role in the processes of and the practices resulting in a specific definition? What are the factors that lead to the concept of open (and closed) science and that of scientific data? How is the relationship of the things that then are called “world” and “data” in different methods, sciences and societies? Through which terms, methods and concepts is this distinction conceptualized, made through different practices, and then used in scientific narratives and texts or in the politics and policies of science? It seems to me that science studies might greatly benefit from including approaches and research questions about the ontology and historicity of the objects we choose to study and thus, in and through our actions and choices as researchers, bring into being as scientific objects.

#### LITERATURE

Barad, K. (2007) *Meeting the Universe Halfway*. Duke University Press.

Láng, B (2015) Ciphers in Magic. Techniques of Revelation and Concealment. *Magic, Ritual, and Witchcraft*. Vol 10, No 2: 125-141.

Mauthner, NS. & Gárdos, J. (2015) Archival practices and the making of ‘memories’ *New Review of Information Networking* vol 20, no. 1-2: 155-169.

Merton, R (1942) Science and Technology in a Democratic Order. *Journal of Legal and Political Sociology*: 115–126.

Shapin, Steven (2008) *The Scientific Life: A Moral History of a Late Modern Vocation*. Chicago: University of Chicago Press.

*Judit Gárdos (Hungarian Academy of Sciences, Centre for Social Sciences) is working on the critical analyses of performances around data, the explanatory structures and normative features of sociological practices, especially of quantitative methods. She is also researching how social scientific data archiving is shaped by specific ontological understandings of social sciences.*

[gardos.judit@tk.mta.hu](mailto:gardos.judit@tk.mta.hu)

