Benefits and obstacles of openness in science: an analytical framework illustrated with case study evidence from Argentina

August 2016

Valeria Arza & Mariano Fressoli[*]

Abstract

Doing open science is to collaborate openly with others in a scientific endeavor and to share openly the outcomes of the scientific process. Benefits of open science are plenty and diverse, ranging from increasing research productivity, to empowering local population and other participants in the scientific process, to improving the democratization of science. However, there are many meanings and practices of open science and thus when analyzing concrete open science initiatives one finds a full lot of hybrid forms of openness. We identify and discuss the different aspects, elements and meanings of open science and their benefits and obstacles as they were discussed in the literature. Our claim is that both benefits and barriers are somehow related to how openness is achieved. We propose a bi-dimensional framework to characterize openness along research stages. The first dimension accounts for the characteristics of the collaboration, while the second takes into consideration aspects of access and accessibility of shared outputs. Our framework allows us to characterize different open science initiatives in this bi-dimensional space and to anticipate the type of benefits and obstacles to be expected. We illustrate our framework by discussing four Argentinean open science initiatives.

Keywords: open science; Argentina; analytical framework, benefits; obstacles, case-study

[*] Researchers from the National Council for Technical and Scientific Research (CONICET), Research Center for the Transformation (CENIT) associated to *Tres de Febrero* University (UNTREF) and STEPS AMERICA LATINA. Correspondent author varza@fund-cenit.org.ar.

1. Introduction

Open science is meant to be collaborative in processes and freely available in outcomes. Open science practices are inspired and generally based on similar principles as the open source movement. It seeks to share the data, outcomes, tools and problems and also the efforts of producing relevant knowledge. Web-based and electronic tools have created great opportunities to scale up and speed up openness and collaboration. Benefits of open science are plenty. They range from increasing the efficiency in scientific production to fostering collective intelligence for the resolution of intractable problems to empowering local population whose interests get to be better reflected in the research agenda and who could access the latest scientific findings.

These claims have been regarded as the beginning of a new revolution in knowledge production (Bartling and Friesike, 2014) and, unsurprisingly, have attracted a lot of attention and increasing support from scientific institutions, funding organizations and policy makers.

However, as Fecher and Friesike (2014) show, there are plenty of definitions and meanings of open science which in turn inform several approaches. Therefore, it is interesting to note that when analyzing concrete open science initiatives one finds a full lot of hybrid forms of openness. Outcomes may be publicly available but they might be unintelligible for most, or collaboration may be opened to a club of experts, or everyone could be entitled to participate in the collection of data but on individual, rather than interactive, basis. Opening scientific knowledge can also create its own challenges and problems, free riders can take advantage of common knowledge and try to privatize it in some way and neglected scientific agendas can be punished or discredited for using open, participatory approaches.

This working paper aims at identifying different dimensions of openness, how they combine to produce different types of benefits, and the specific obstacles that need to be overcome to increase openness. We propose an analytical framework that could be used as a toolbox to assess different experiences of open science around the world.

The paper is divided in eight sections. Next section 2 presents the discussion on the different dimension of openness. Section 3, in turn, identifies benefits of open science practices that have been mentioned in the literature. Section 4 does the same for obstacles. Section 5 proposes an analytical framework to associate dimension of openness, benefits and obstacles and presents the research methods. Section 6 presents the four case studies. Section 7 illustrates our analytical framework with case study evidence. And finally Section 8 concludes.

2. Dimensions of openness

Following Benkler we identified two main dimensions that characterise open and collaborative knowledge production: firstly, those aspects that characterize how actors collaborate among each other to produce knowledge, and secondly, those other aspects that characterizes access and accessibility to shared outcomes.

The first dimension relates to the process of **collaboration** in peer-production. This in turn comprises aspects of interaction, participation, and diversity.

Actors could participate in producing scientific knowledge on individual basis or collectively and interactively. The level and quality of the interaction affect the characteristics of the knowledge production process. We take notice here of the literature on innovation studies that regards learning as an interactive processes (Lundvall, 1992). In this literature interaction among diverse types of actors renders better quality innovation, which also comes up at a faster rate. Since actors learn by interacting, the whole process is also empowering for participants.

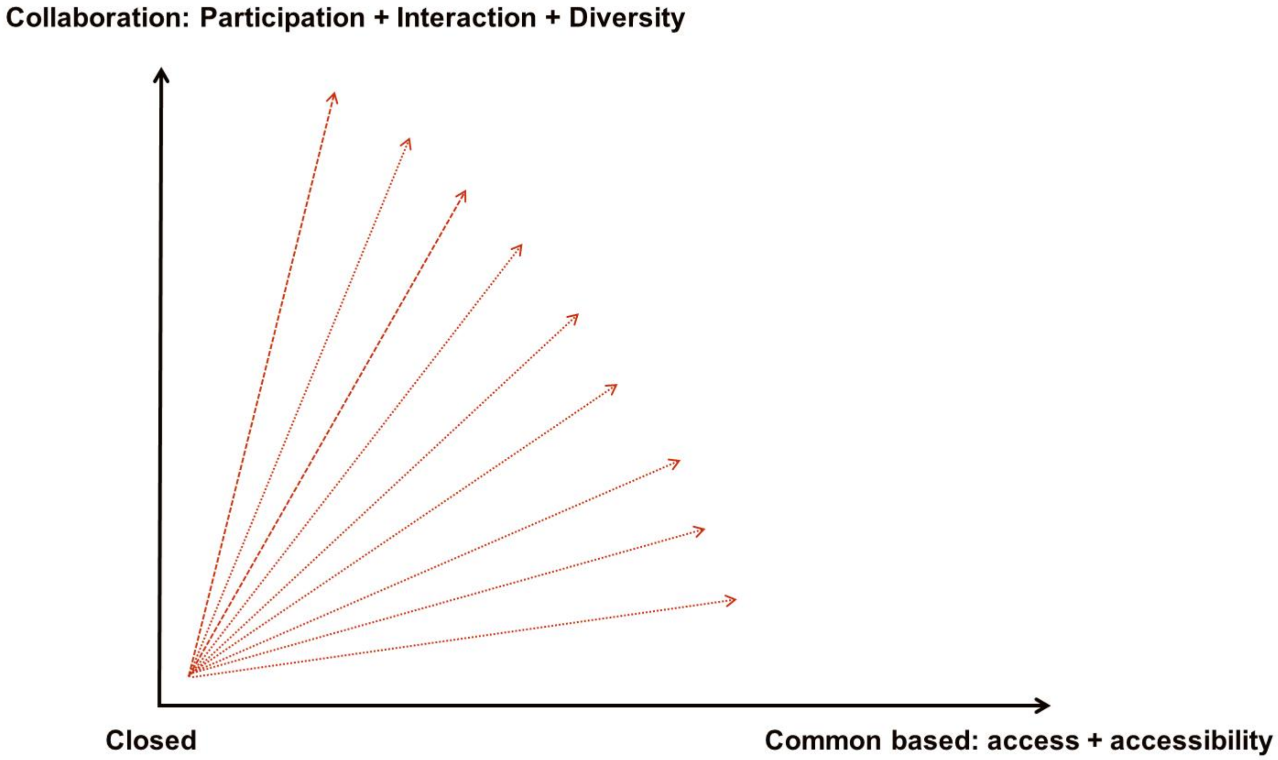
Moreover, actors could be more or less involved or committed to the research. We draw from Arnstein (1969) taxonomy on levels of citizen commitment in policy decisions to analyse participation in open science processes (from sharing information to sharing control).

Finally, this dimension also accounts for *who* participates or better how diverse is the participating group. Researchers are normally more prepared to share their research outputs with other scientists rather than with a diverse audience. Expanding the quantity and diversity of actors involved as users and producers of scientific knowledge is one of the ambitions of open science practices.

The second dimension aims, in turn, to take into consideration aspects of access and accessibility to shared resources. This is related to the **common** based characteristics of shared resources. The principle in open science is that scientific resources could be enjoyed by everyone. This requires, on the one hand, open access. By this we mean either that no group could imposes restrictions, such as the imposition of paid subscriptions or formal licenses (to read/use, to distribute, to reproduce, etc.) (Molloy, 2011). On the other hand, accessibility is also required, and by this we mean the lack of other more informal restrictions, such as specific skills, capabilities or capital resources needed for understanding or using open science products. When access and accessibility are high, the more people will be able to fully enjoy shared outcomes.

These bi-dimensional characteristics of openness could be drawn in a Cartesian diagram such as that in Figure 1, and different open science practices could be placed here and there in each of these dimensions of openness. Our claim is that expected benefits and obstacles will be related to the specific route towards openness each initiative decides to follow. We therefore now discuss benefits and obstacles in Section 3.

Figure 1: Two dimensions of openness in science



3. Benefits of open science practices

3.1. Benefits of open science in the literature

The literature claims that there are many different benefits associated to open science practices. We will review them in an attempt to understand how they interplay with the bi-dimensional framework of openness referred to above.

1. Increased research efficiency and quality

To increase productivity in scientific production means to be able to achieve more or better scientific outputs (i.e. findings, publications, trained scientists) using the same amount of scientific inputs (i.e. resources). It is a measure of efficiency of scientific production; it means to be able to make the most of current research resources. Sometimes, some practices or actions can also increase the likelihood of producing more or better scientific outputs in the future (given current research findings, resources or state of the art), and therefore we may also refer to *dynamic* efficiency. New technologies, in particular online collaboration through social networks, peer-production tools and open access are also fostering new forms of interaction and speeding up problem-solving strategies.

One of strong argument for supporting open science practices is that they increase efficiency (Nielsen, 2012). This occurs either because researchers become better researchers (i.e. technical efficiency) or because research activities become less expensive (i.e. cost efficiency). Different mechanisms related to the different dimensions of openness contribute to this increase in efficiency.

a) Public good characteristics of knowledge mean that the most effective way to exchange it is by opening up the process of creation and distribution. Knowledge, data and information are public goods because they are non-rival in use (i.e. it can be used simultaneously from many actors without losing its properties or functions) and it is costly to exclude other from using or possessing them (see Benkler (2006)). These public good characteristics of knowledge means the most effective way to produce reliable additions to the stock of knowledge is to promote collaboration (*our dimension 1*). As David put it: “wide sharing of information puts knowledge into the hands of those who can put it to uses requiring expertise, imagination and material facilities not possessed by original discoverers and inventors. This enlarges the domain of complementarity among additions to the stock of reliable knowledge and promotes beneficial spillovers among distinct research programs” (David, 2003: 22)

b) Open access to scientific final or intermediate outcomes (*dimension 2*), increases the pool of knowledge in common use. This increases research productivity because i) duplication can be more easily avoided, and ii) all researchers can explore new questions and solutions to problems by standing on the shoulders of a taller giant since a greater pool of global knowledge becomes available to all.

c) Duplication is also avoided when sharing information and insights in collaborative research (*dimension 1*). All researchers get to know faster and more accurately by interacting with other researchers in their fields.

d) Open access (*dimension 2*) also increases technical efficiency because new research outcomes can be drawn from *data driven intelligence*. With the aid of software tools, researchers could reuse online available data to arrive to new findings simply by interconnecting everything that is already known (Nielsen, 2012).

e) Technical and dynamic efficiency increase when researchers interact with peers in collaborative research (*dimension 1*). The constant interaction among researchers using web 2.0 tools promotes processes that amplify collective intelligence of the group by the mere fact of being able to share, validate and quickly rule out different ideas, assumptions, hypotheses or avenues of inquiry. When a group of researchers interact and collaborate technical and dynamic efficiency in knowledge production is improved within the group just because ideas come back and forward feeding from the interaction. However, as Nielsen (2012) argues, the amplification of collective intelligence probably works better when interactive actors share at least some cultures of practice or when they are focused on the same problem-solving strategy.

f) There is a wide range of new, open source tools that help to improve scientific collaboration increase scientific productivity. These open source tools cover almost every aspect of the research cycle, from sharing protocols of research (Open notebook) to sharing data (Figshare), to collaboration between scientists (Open Science Framework) to collaboration between scientist and citizen and crowdscience (Petridish, Sciencestarter).[2] Open source tools are important because they promote a wider share of research inputs and hypothesis and data. At the same time, since they use open software and open hardware, these tools are available for modification and improvement by any other scientists or member or the public, which might lead to continuous enhancement of the same instruments. Researchers themselves could improve the tools constantly so as to make scientific collaboration more productive (*dimension 1*).

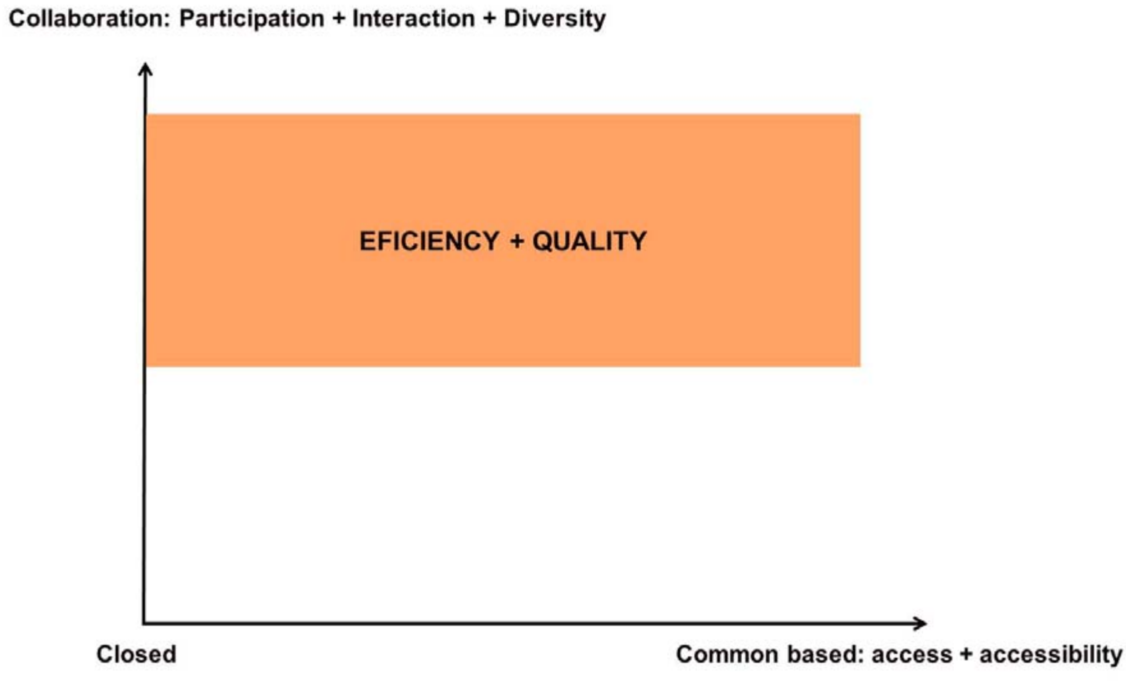
[2] See Tools for Open Science, in http://science.okfn.org/tools-for-open-science/

g) By increasing the quantity of actors participating in data collection (or analysis) (*dimension 1)* new cognitive and manpower resources become available. These are idle resources that would not have been used for scientific purposes in the context of traditional science. They contribute to scientific endeavors responding to intrinsic motivation (pleasure, fun, intellectual interest, etc.) or extrinsic motivation (prestige, recognition, networking, and sometimes paid remuneration etc.) (Lakhani and Wolf, 2005; Shah, 2006)

h) Technical and dynamic efficiency may also increase when collaborative research involves the participation of a wider community (*dimension 1*) (Jeppensen and Lakhani, 2010). This is because outsiders may have a fresh look to problems put forward within specific scientific fields; they may contribute by drawing from different knowledge and cognitive tools to the ones well inside the problem field. A fresh look into a problem could be especially relevant when scientific groups have reached a “silo mentality” that refrains to share information or shies away for collaboration with other actors. Social studies of science claimed that major innovation in different fields tend to be put forward by scientists trained in different fields mainly because they are not bound by professional traditions (Ben-David, 1960). A similar phenomenon have been observed in studies about innovation (Bijker, 1997). Jeppensen and Lakhani (2010) go a bit further and claim that it is not just technical marginality but also social-political marginality which may contribute to novel ideas, for similar reasons, these actors are more prone to thinking unconventionally and therefore more creative.

Thus, although both dimensions seem to be relevant to improve scientific research productivity and quality, *dimension 1*, the collaboration in peer-production of knowledge seems to be crucial. Research activity improves as more actors become involved and interact among each other. This is represented in Figure 2.

Figure 2: Research efficiency and quality in the two dimensional space of open science



2. Democratisation

Other proponents of open science emphasize the need to reinforce the public good characteristics of knowledge not to increase scientific productivity but to increase democratization (see Fecher and Friesike (2014))[3]. The use of open tools for scientific research can lower the cost of access which eventually allows a broader access to the public. Moreover, tools that allow public participation through mechanism of crowd science (Wiggins and Crowston, 2011) encourage the public to participate and potentially learn science during the process.

Open access (*dimension 2*) increase the pool of information available to anyone (e.g. people may get to learn about latest treatment of certain diseases, they may get to know about relevant techniques in several application fields, etc.). Digital tools might facilitate the potential for knowledge sharing, rendering negligibly small incremental costs of expanding the quantity of users. However, there are still costs associating to training potential users so they become able to enjoy all functions of shared outputs and make the most of open access. These costs are inversely related to the investment in knowledge translation and communication efforts (*dimension 2*).

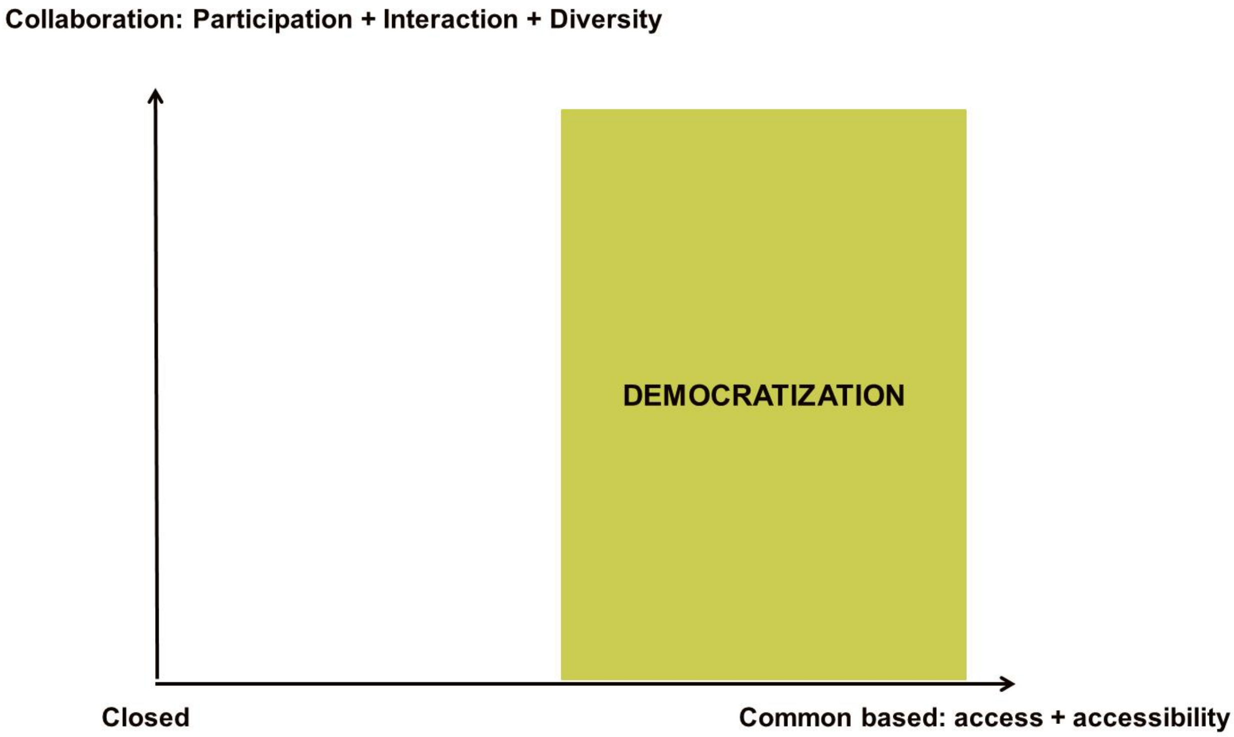
The same kind of electronic tools that encourage increases in productivity and dynamic efficiency might also promote the democratization of scientific knowledge. There are two key aspects of democratization here: first, electronic tools are based in a model of open sharing of knowledge and ideas (*dimension 2*). A wider access to scientific knowledge contributes to a better informed society, and eventually to allow the public to have a say in the construction of the scientific agenda. Thus, the second key element in open science infrastructure is participation (*dimension 1*). Digital tools, especially those oriented to open access (sharing data) and crowdscience (co-production of knowledge) facilitate that a greater quantity and wider diversity of social actors become involved in scientific production (e.g. Galaxy Zoo, Foldit, etc)- see Wiggins and Crowston (2011) for a broad survey of initiatives using electronic tools-, as in citizen science experiences, increasing democratisation not just in the use but also in the production of scientific knowledge.

To exploit this technological opportunity, barriers to participation should be lifted in order to increase the likelihood that everyone could contribute in the scientific endeavor (Wagner, 2009; Wiggins and Crowston, 2011). In data collection this means, for instance, to create protocols to automatically validate data collected by a wide range of participants.

Thus, in this case again both dimensions are relevant for the democratization of scientific outcomes; but dimension 2, the one representing the common based aspect of openness, that allows everyone to access and understand the knowledge production process, seems to be particularly important because it guarantees that no one could exclude others. This is represented in Figure 3.

[3] We refer here to both, the democratization school and the public school in Fecher and Frisike (2014) taxonomy.

Figure 3: Democratization of science in the two dimensional space of open science



3. Empowerment

Open access is potentially empowering because it reduces the costs of using and reusing the worldwide accumulation of knowledge assets. On the one hand, more people could get access to more resources, which they potentially could grab to solve their problems. On the other hand, the dissemination of open access information also allows that problems affecting powerless actors become known much more effectively (*dimension 2*). Open access and open data supporters affirm that wider availability of information can improve efficiency in problem-solving strategies, allow a better accountability from governments and incumbent powers and foster new process of learning (Gregson *et al.*, 2015; UN Independent Expert Advisory Group Secretary, 2014; World Bank, 2015) (*dimension 2*). In a similar fashion open and free access to scientific information could do a lot to improve the political position of marginalized groups in order to engage with other actors like authorities, the press or other potential supporters.

Another key aspect for empowerment is learning and capability building. Many open science projects pursue the goals of fostering scientific education among their participants (Wiggins and Crowston, 2011). Learning in open science projects can take varied forms. It can happen as part of the interaction among participants where the “experts” guide the beginners (*dimension 1*). This explains the interest of open science initiatives in providing i) interactive tools such as online forums, e-mail groups, etc. ii) online training courses such (tutorials, massive online courses (mooc)) and iii) open tools such as open notebooks, open software or open hardware (Baden *et al.*, 2015). Some open science initiatives are starting to introduce open science tools in students curricula as a way to improve learning and research capabilities (Molloy, 2014).

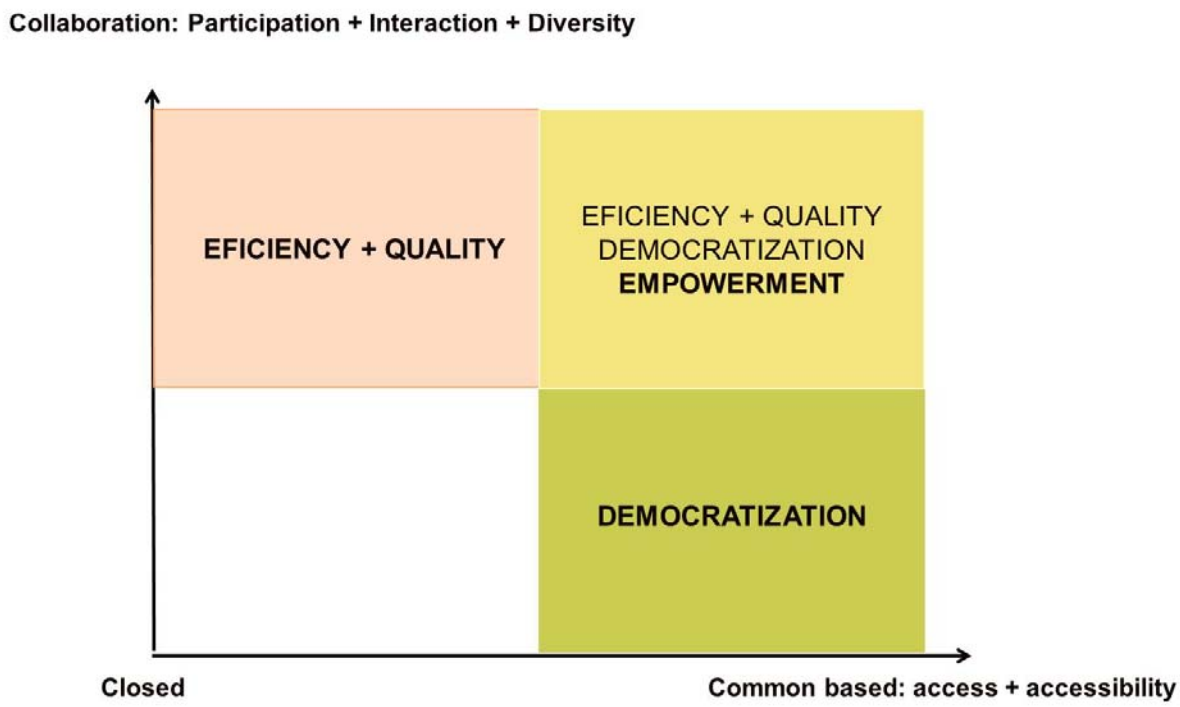
However, it is important to notice that not every open science projects are concerned with building capabilities -authors like Irwin (1995) suggest that some citizen science projects can even be disempowering. Therefore, while capacity building and increasing scientific literacy become key components in a strategy for the development of open science in the long term, the degrees of participation (and therefore their opportunities for learning) of different actors is also important. In this sense, the degree of learning can be used as a measure of how much an open science project allow different actors to gain participation in the different aspects of the research cycle (Wagner, 2009). This depends on both formal and informal restrictions (*dimension 2*). It also worth to keep in mind Arnstein (1969) ladder of participation to try to think critically how much participation really empowers the participants of an initiative. As Arnstein said it would depend on the actual commitment of participants and the distribution of power. Thus, not only learning but also commitment and recognition while participating in different stages of the scientific process are of key importance to promote empowerment (*dimension 1*).

Open science practices can foster innovation either because local problems became more visible and better communicated when using digital tools for sharing data and scientific outcomes and/or because 'the wisdom of the crowds' (Nielsen, 2012; Surowiecki, 2004) in all stages of knowledge production help solutions to come up more quickly and effectively. Open source tools that allow a wider participation and interaction among scientist -and in some cases including the public - speed interaction up and allow a more diverse set of eyes on the problem. Interactive and collaborative processes associated to open science contribute to a better understanding of local problems and the creation of innovative solutions to these problems (*dimension 1*). When this is combined with open access and open licenses such as creative commons, scientist and entrepreneurs can avoid barriers that hamper the process of turning scientific knowledge into concrete solutions to local problems (*dimension 2*).

The socio-economic impact of innovation depends, in turn, on the potential for promoting a wide appropriation of innovative outcomes. For instance, Masum and Harris (2011) discusses a series of initiatives using open source approaches to improve R&D for neglected diseases. The main mechanisms are those related to open collaboration (rapid verification, avoidance of duplication; amplified collective intelligence; burst of creativity, etc.) (*dimension 1*) and open access (costs reduction and creation of commons) (*dimension 2*). However, they found these approaches are heavily "weighted toward the discovery (or pre-competitive) stage of R&D, with little in development and none in delivery (e.g. clinical trials and filing)." (p. 11) and this is explained by the greater amount of investment required in those stages which push innovation towards more closed approaches such as incentives to obtain exclusive rights on patents. Thus, although open science could contribute to innovation in very relevant areas for developing countries (such as neglected diseases), if at some point on the innovation processes –such as the final stages needed for delivering a new drug- are consecutively limited by restrictive regulations their socio-economic impact could be quite limited. As a matter of fact, innovation processes by themselves do not prevent the increase in inequalities. It might be expected that since open science research shares a lot of values with the open software and commons ideology, their innovations will have similar characteristics. However, this is far from certain, as we will discuss in next section.

In sum, empowerment actually needs openness in both dimensions, because it needs collaboration, interaction and participation to guarantee capability enhancing activities and re-orientation of the research agenda towards societal demands and it also needs the common based aspects of openness to guarantee autonomy and common appropriation of research resources. This is represented in Figure 4.

Figure 4: Main benefits of open science in the two dimension space of openness



4. Obstacles faced by open science practices

A first set of obstacles is related to problems of rigidity within the scientific tradition, where scientists are reluctant or are not yet prepared to exploit open and collaborative research to the most. The scientific culture is reflected in norms and rules within scientific organization. It sometimes makes it complicated to do multidisciplinary work or to articulate a multiplicity of knowledge when interacting with other social actors (Wagner, 2009).

A second set of obstacles is related to power asymmetries in the capacity to benefit from openness. We identified two versions of this same problem. One of them is related to asymmetries in the distribution of resources needed to draw from knowledge pools that become freely available. There is an obscure side of the new mantra "data is the new oil"[4]: this is how easily some actors with unequal resources and capabilities could obtain complete access and they could disproportionally benefit from that oil; or in other words some actors have the necessary resources to make the most of available knowledge while others have not. In that sense, open knowledge practices could well exacerbate current inequalities. Another version of power asymmetry is related to research performed in politically disputed context, when certain minorities may be affected if specific knowledge resources become appropriated by the wrong hands.

The following narrative relates both types of obstacles with the two-dimensional framework of openness. This should be read as the obstacles open science initiatives may face when trying to increasing openness in each of the two dimensions.

[4] This quote was originally said by Clive Humby, an English mathematician in 2006, although it has become now common parlance.See Arthur (2013)

1. Cultural or institutional rigidity

Incumbent actors and institutions struggle to understand the potential and the benefits of a broad range of open science practices. Cultural change is always a difficult process and people used to some practices will probably resist new forms of doing things.

Traditional models of producing science do not favour open interaction with broader networks (i.e. *dimension 1*). Policies of science and technology are sometimes attached to a model of endogenous capability building and international competition that leaves little space for collaboration. One clear example is the incentive and evaluation system in research that privileges publications in peer-reviewed journals. Collaboration among scientists without formal connections (i.e. from different laboratories that do not participate in the same project) is not a common practice in the scientific community, because scientists compete for grants, resources, and publications in journals that require original results.[5] Therefore, they might look with disdain or fear to those tools that allow open sharing of ideas, protocols and data (Nielsen, 2012).

Open science practices often challenges these traditions, which are widely used for scientific assessment but also deeply embedded in the scientific culture. This culture discourages collaboration (*dimension 1).* In the case of open data and open access (*dimension 2*), scientist might argue that data might be "scooped" by other scientists before the analysis is completed (Bishop, 2015). Collaboration is particularly difficult when it involves a wider participation of diverse actors (*dimension 1*). This requires significant amounts of time and leadership which are not always recognized by the evaluators (RIN/NESTA, 2010). Another common concern of opening up participation in *dimension 1* is the scientific reliability of data collection or other scientific process where citizens may become involved (RIN/NESTA, 2010) (Wiggins and Crowston, 2011).

Finally, typical efforts from open science practices to widen the audiences that may use scientific outcomes (such as writing for wide public, or giving talks to school children, community organisations, and the sort) do not count in scientific assessment schemes. This affects attempts to openness in *dimension 2.*

In sum, cultural rigidity puts obstacles to openness in both dimensions but it may weight more heavily on opening up collaboration, participation and interaction with diverse actors (*dimension 1*), as represented in Figure 5.

2. Power asymmetries in the capacity to benefit from openness.

The flipside of opening and sharing indiscriminately is that some actors are better prepared or have better access to complementary resources that enable them to make the most of freely available data. If this appropriation increased power asymmetries in detriment of powerless actors, inequality might increase rather than decrease as a consequence of open access.

In the context of scientific production in developing countries, the risk is that the opening-up scientific outcomes (*dimension 2*) produces a paradoxical effect in which increasing scientific

[5] It takes a lot of time and effort to publish a scientific finding. The time between obtaining the research result and its publication is usually very long. The process is tedious. In order to surpass it, the scientist invests resources motivated more by the need to make progress in their career rather than by the desire to share information with the interested public (which surely is broader than the scientific group researching in any specific knowledge field).

production end-up increasing technological dependence due to asymmetric appropriation of open data by powerful and dominant actors.

Asymmetric appropriation of scientific knowledge is not the only risk associated to power asymmetries in open science practices. Other tangible risk is the dishonest use of scientific information and knowledge regarding contentious social and environmental issues. Such risks usually turned up linked to openness and collaboration process taking place in the context of "undone science" (Hess, 2007). We refer here to those research projects that address problems that have not been addressed by mainstream scientific agenda.

In these contexts often there are strong power asymmetries between actors involved in those alternative research agenda[6] and other groups whose interests are attached to keep that agenda off or hidden. The former fear that their results were discredited (considered invalid) or that obstacles were being placed into the research process. Thus, they often prefer to restrict openness or research outputs and collaboration to a close group of trustable allies. This affects openness in both *dimensions 1 and 2*.

The politicization of scientific practices is not new for socially committed science. Since the publication of Silent Spring by Carson (2002), research on environmental and/or social factors that affect power interests (such as companies or industries or government regulation and practices) became an inherently political activity. Before publishing Silent Spring, Carson herself took a lot of care in gathering relevant data to support her research and send the manuscript to other experts to revised it to make sure the argument will stand up contestation from incumbent powers (Millstone, 2015). Equally, in some cases, committed scientists have to defend their research problems against attacks by their scientific institutions that do not consider them as scientific problems (Hess, 2007).
Moreover, in those contexts, scientific validity issues come up very often as tools to question scientific production. This is especially the case when researchers open up the research process and work side by side the affected community (Allen, 2004). Another important issue is to ensure that the people who are collaborating with the project do not become exposed to any reprisal by the authorities or other incumbent powers.

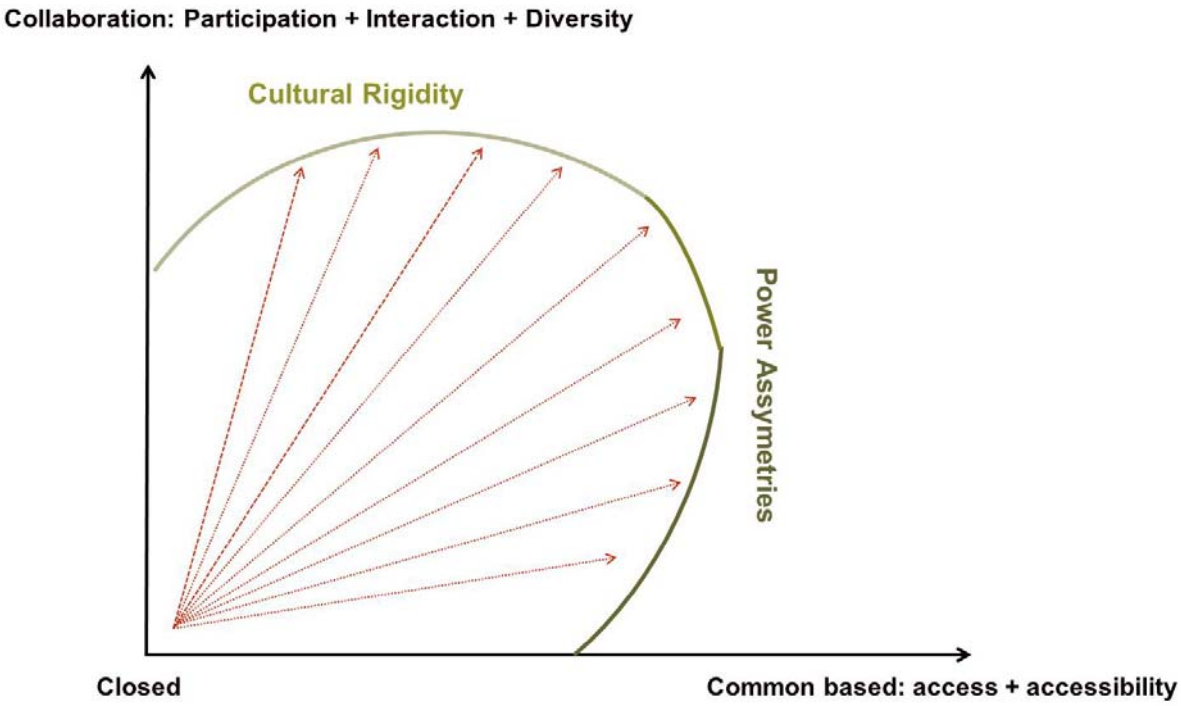
Therefore, socially committed scientists working in disputable arenas normally come into tension with open science practices. This tension involves both questions about the validity of collected data and questions about when is the right time to opening and disseminate data and analysis and how it should be done (McAllister, 2012). As the case known as the Climategate shows, information that is made public before a proper analysis and curation could be detrimental of the causes advocated by the proponents of the initiatives (McAllister, 2012). [7]

In sum, power asymmetries affects openness in both dimensions weighting possibly more on *dimension 2*, as showed in Figure 5.

[6] Both, as researchers or as communities or individuals directly affected by the issues being studied.

[7] In 2009 the servers where a series of emails and information from the Climate Research Unit from University of East Anglia were hacked. According to climate change skeptics these emails showed that scientists were hiding information that could be interpreted as opposing theories of global warming. The main criticism was that climate change scientists decided not to openly share some data while they did decide to disseminate other. Scientists of East Anglia instead argued that they were entitled to take care of the information they generate, which among other things also meant i) to decide about the right time to present the data they collected; ii) to identify how data should be cured prior to any analysis; and iii) to suggest interpretations the scientific team considered appropriate for each case. An independent investigation absolved scientists in 2010 (McAllister, 2012).

Figure 5: Main barrier to opening up in the two dimension space of open science



5. Analytical methods in case study research

We propose a methodological tool to assess levels of openness, associated benefits and expected obstacles of different experiences of open science and we test our methodology using case study information from four Argentinean open science projects.

5.1. Data sources and data collection methods

With chose case study methodology to analyze processes, benefits and obstacles of open science. We selected the cases from a survey we conducted to all researchers working in the national scientific system. We used an online short questionnaire (just four closed questions plus an open one). In the open field respondents had to describe the most relevant experience of open science in which they had participated. A definition of open science was offered as part of the questionnaire. 1453 researchers (response rate of 8%) responded the survey and 70% of them completed the open field. We read those entries and selected about 40 as potentially interesting case studies. This group was then enriched from online inquiries and discussions and interviews with four key informants.

Finally four case studies were selected for this research taking into account the need to cover the widest possible diversity of situations and openness processes. This selection strategy helped us build a more robust body of evidence (Yin, 2014) while allowing us to explore the heterogeneous spaces in which open science is being implemented in the country. Among factors of heterogeneity we considered: research disciplines; socio-political contexts in which research was carried out (i.e. more or less subject to political disputes); processes of knowledge production (i.e. uni-disciplinary or transdisciplinary); techniques of participation (i.e. citizen science techniques, participatory action research, workshops, etc); type of infrastructure (e.g. open databases; use of remote sensors, mobile applications, etc.).

The selected projects were: New Argentinean Virtual Observatory - NOVA (astronomy); Argentinean Project of Monitoring and Prospecting the Aquatic Environment - PAMPA2 (limnology); e-Bird Argentina (ornithology); and Integral Management of the Territory – IT (geography-chemistry)-.

To build information on these case studies, we use secondary sources such as scientific papers, reports, media stories and material available on Web and semi-structured interviews. Interviews were performed during 2015: three interviews each for NOVA, PAMPA2, e-Bird and IT. Except for e-Bird, all interviewees were scientists. In e-Bird the interviewees were project coordinators. In addition, four in-depth interviews just mentioned were conducted in 2015 with key informants as benchmarks in the areas of digital repositories, programming and open source and technology transfer, which helped us to put our cases selected in context and have a more global perspective on the processes of opening up.

The information collected in each interviews was triangulated with information obtained in other interviews and from secondary sources. The semi-structured questionnaire used covered issues such as the origins and motivations of the project, research and collaboration activities, training, infrastructure, financing and evaluation, dissemination of results, benefits and obstacles.

Besides, this case study information was complemented with structured interviewed performed during 2016 to one leader of each of the above-mentioned initiative. Closed questions were designed to assess levels of openness in a 4-points Likert scale in terms of 1. participation, 2. interaction, 3. diversity, 4. access and 5. accessibility (i.e. representing our conceptual framework) along six different research stages: 1. Research design; 2. Collection of data; 3. Analysis; 4. Documentation and Publication; 5. Public/Social communication & engagement; and 6 Infrastructure. [8]

5.2. Analytical methods

Following our conceptual framework we characterize open science process in a two-dimensional space. One dimension covers aspects of peer-production (i.e. participation, interaction and diversity) and the other on common-based appropriation (i.e. open access and accessibility).

Using our structured interview we could place each open science initiative on one particular location of the Cartesian diagram and we could calculate how far it went into covering the different areas identified in Figures 2, 3, and 4. Of course, not every initiative is motivated to achieve all of these benefits, so the assessment of any actual initiative has to consider on its own ground. We could then relate that information with the qualitative data on processes, barriers and obstacles we collected through semi-structure interviews, in an attempt to capture the heterogeneous space of open science in terms of motivation and processes and therefore also benefits and obstacles.

We should acknowledge that the use of area as metrics is problematic because it disproportionally weights higher ranks (a 3x3 square is 9 times higher than a 1x1 square). However, we still decided to use it because we do not plan to use this tool to decide over one experience against others but rather to understand the diversity of potential benefits and how they relate to openness in different dimensions.

[8] The identification of research stages was inspired in RIN / NESTA (2010), which identified a research cycle including seven different stages: Conceptualising and networking, Proposal writing and design, Conducting and presenting, Documenting and sharing, Publishing and reporting, Engaging and translating, and Infrastructuring.

6. The four case studies

6.1. NOVA - New Argentinean Virtual Observatory

NOVA aims at centralizing astronomical data and making them available to all users. It was created by researchers from various institutions in the country as a digital platform that aims to store and share already processed astronomical data. It facilitates collaboration of local and international astronomical community, through documentation, digitization and open access to data.

It is a country-wide project that encourages the participation of all astronomical institutions in the country promoting the generation and integration of information technology with special emphasis on statistical analysis and data management of astronomical images. It also aims at integrating of local data to international standards agreed by the International Virtual Observatory, the coordination of astronomical resources and the dissemination of Astronomy as an educational tool.

The astronomical information stored in the database is open access and can be used by astronomers, researchers from other fields, students and the general public. However, it requires certain level of expertise to use specific software for image visualization.

The experience of NOVA and the aim of its founders to use it an educational tool, triggered the conception of a related Project called Galaxy Conqueror. This is a game that motivates citizen to mark possible galaxies surfing on sky image as if it were Google map. It offers a brief tutorial that teaches basic characteristics of galaxies. It was inspired by Galaxy Zoo. Galaxies identified by users are then checked by volunteers from NOVA. Since the creation of the game in 2015, 50 new galaxies were identified. The game is part of a Citizen Science platform called Cientópolis, managed by some of the institutions that participate in NOVA.

Box1: NOVA data sheet

Research Area: Astronomy / Astro-informatics

Founding institutions: El Leoncito astronomical complex (CASLEO); Faculty of Astronomy and Geophysics of La Plata / National University of La Plata (FCAGLP / UNLP); Argentine Institute for Radio Astronomy (IAR); Astrophysics Institute of La Plata (IALP);); Astronomy Institute of Theoretical and Experimental (IATE); Institute of Astronomy and Space Physics (IAFE); Institute of Astronomical Sciences, Earth and Space (ICATE) and Córdoba Astronomical Observatory (OAC).

Start: 2009

Funding: CONICET (National Science and Technical Research Council)

Features: digital platform that facilitates collaboration of local and international astronomical community, through documentation, digitization and open access to data

Benefits: To promote efficient use of information; to generate an overall view of it; to prevent data loss and duplication of efforts by different scientific groups and to encourage public participation in general and amateur astronomers

Difficulties: Lack of technical staff and funding. The project only has one person in charge of computer development and it receives financial resources just from CONICET

6.2. PAMPA2 - Argentinean Project of Monitoring and Prospecting the Aquatic Environment

The Argentine Monitoring Project and Exploration of aquatic environments, better known as PAMPA2, is an initiative that seeks to understand the reaction and behavior of water from lakes and ponds to certain natural and human events, to improve the design of management plans that may prevent deterioration and, at the same time they would preserve the population health

PAMPA2 is an interdisciplinary network of scientists from seven different research laboratories. Lagoons are regarded by these scientists as early warning systems; thus, by analyzing them the project could contribute to detect changes that would eventually affect the whole region. This, in turn, could help to design technical and financially more viable resource management, mitigation or adaptation plans that take better care of the environment and the health of the population located in the nearby. To monitor the lagoons properly, diverse type of data are needed. So an interdisciplinary team of oceanographers, meteorologists, biologists, zoologists and engineers was formed to monitor thirteen lagoons distributed in the Pampa region during five years. In five of these lagoons buoys equipped with automatic sensors capable of measuring temperature, pressure, wind, rainfall, humidity, oxygen, chlorophyll and depth they have been installed. These devices are connected to a processor that stores information and then transmits it in real time to the laboratories responsible for its operation. Researchers complement these data using laboratory information from samples collected monthly or every six months from the lagoons.

These buoys were distinguished in Innovate 2011, the National Innovation Competition organized annually by the Ministry of Science, Technology and Innovation (MINCyT). Specifically, the development reached the second place in the Applied Research category because it was considered a useful tool developed at a lower cost of imported commercial ones. Originally, the buoy was not designed following an open source approach; but the team is currently working in a new design based on open source software for more ambitious monitoring projects (i.e. buoys that can support more extreme environments, such as those in open seas).

Sensors transmit data in real time to a server that shares the information with the whole team. Information can be openly accessed for free in a website but only for the present month, given restrictions in their infrastructure. Historical data generated by the sensors as well as other information generated by the project can be requested to the teams.

In PAMPA2 only those teams that originally formed the network participate in the design, collection and analytical phases. Actually, the design of the project was carried out predominantly by one of the networked organizations. In terms of diversity, although there is some diversity in terms of disciplines, all participants are professional scientists from close scientific fields. In terms of interaction, there are no formal instances for interaction for all members: just one workshop held every year. They contact each other spontaneously to achieve specific objectives (e.g. few of them have co-authored papers). Similarly, the degree of participation by the different members is low in Arnstein terms. There are clear hierarchies that resemble the traditional hierarchies present in research teams (principal researchers, senior researchers, junior researchers, research assistant, students) and the command of power is related to that.

In terms of accessibility, one of the goals of the project was to disseminate results to a wider audience, especially the population living close to the lagoons. However, these activities were not performed so far because they team does not have the required expertise for doing public communication nor can they get the necessary resources to hire these services. Another shortcoming in terms of diffusion is that the website has not been designed to as to be easily used by outsiders.

Moreover, there is no written a protocol to allow users to work properly with the data the project produces. However, researchers do receive frequent requests from people that look at available data, for example for recreational or productive purposes.

PAMPA2 enabled increased interaction with other similar research projects around the world. It became integrated to the GLEON Network (Global Lake Ecological Observatory Network), an umbrella organization of institutions around the world that monitor lakes continuously through instrumented buoys. Similarly, some of the participants of PAMPA2 are also involved in the SAFER Project (*Sensing the Americas´ Freshwater Ecosystem Risk from Climate Change*), an initiative that integrates scientists from different disciplines from Argentina, USA, Canada, Chile, Uruguay and Colombia.[9]

Box 2: PAMPA2 data sheet

Research Area: limnology - multidisciplinary

Founding institutions: Technological Institute of Chascomús (INTECH), Argentine Oceanographic Institute (IADO); National University of Central Buenos Aires (UNICEN); Limnology Laboratory of the Faculty of Natural Sciences of the University of Buenos Aires (FCEN / UBA); Ecology and Remote Sensing Laboratory Ecoinformatics University of San Martin (LETYE / UNSAM); Group of Environmental Studies at the National University of San Luis (GEA / UNSL) and the Ministry of Land Affairs of the Province of Buenos Aires.

Start: 2011

Funding: CONICET

Features: Project for monitoring inland aquatic environments to make local and regional climate prospects. Data is collected through remote sensors located in inland water and with laboratory information from samples collected in monthly or every six months from the lagoons. Sensor data is open access.

Benefits: To learn more about the response of inland water ecosystems to changes int the environment; to help in specialized human resources formation; and to increase scientific productivity and efficiency since it avoids duplication of resources.

Difficulties: Lack of budget to keep equipment in good conditions and to ensure wider dissemination of detected problems and other achievements

The SAFER project includes an educational component, and it includes some citizen science practices. For example, students from secondary school were distributed a tool kit to be able to monitory lagoons by themselves. Students collect some information on Ph, temperature, colour and they took some pictures under the supervision of the Argentinean scientific organization involved in SAFER (The Argentinean Oceanography Institute - IADO).

[9] The objectives SAFER are to: 1) employ freshwater ecosystems as "sentinels" or "sensors" of climate variability and watershed processes and investigate their interaction with multiple stressors to assess risks to ecosystem services in the Americas, and 2) determine management and mitigation strategies which are both technically and economically feasible as well as culturally acceptable. To accomplish these, the project: i) investigates the regional distribution of ecosystem services associated to continental waters; ii) estimates how climate variability will affect lakes and their watersheds in different climate regions with varying geomorphologic characteristics using both neo- and paleo-limnological techniques; iii) defines the proxies that can be employed to measure the ecosystem services of the lakes and their watersheds; iv) fills gaps in knowledge and methods in order to enable implementation of a harmonized ecosystem-based methodology and database system approach for American lakes and watersheds; v) enables assessments of climate change effects on water quality and quantity as well as sanitation conditions for underprivileged communities as a means to reduce poverty.

6.3. Integral Management of the Territory - IT

After the tragic floods in 2013 left the city of La Plata under water and caused nearly a hundred deaths, an interdisciplinary group of researchers designed a project for Integrated Land Management seeking to relieve, along with neighbors, the needs of two particularly affected areas. Thus, they expect to identify environmental consequences of this phenomenon to start thinking and developing appropriate technologies to help to reverse them.

The researchers group is formed by geographers and historians from the Faculty of Humanities of the National University of La Plata (UNLP) and environmental chemists form the Network of Environmental Studies. They started working in the Maldonado basin and in the area just next to the YPF refinery that went on fire during the floods. Thus, the project worked on two vulnerable areas that have been particularly affected by the flood events.

The project proposal included the development of what is called an Integrated Management of the Territory (GIT), which means to achieve an orderly, planned and sustainable land management, which in this case was raised in two stages: diagnosis and implementation of proposed solutions. At the time we did the case study they were half way through the first stage.

The neighbors were involved in two ways during the first stage: in the so-called Catalyse method, by collectively designing the survey so that their views and needs were included from the beginning in the questionnaire, and in the sampling of rainwater, which measure their level Ph (to detect the acidity or alkalinity of water). These samples were then delivered to investigators.

The analysis of all collected data is performed by researchers (without the participation of the neighbors).

Box 3: IT data sheet

Research Area: Interdisciplinary: mainly geography and chemistry

Promotion agencies: Faculty of Humanities and Faculty of Exact Sciences of the National University of La Plata (UNLP), along with various faculties and research centers CONICET and CIC.

Start: 2014

Funding: UNLP and CONICET

Features: Methods and techniques of social and natural sciences to collect, process and interpret data from the territory surveyed with citizens, to think and implement solutions for environmental and social problems.

Benefits: To communicate scientific outcomes in a friendly way. To better link researchers, citizens, policy makers and businessmen. To promote the participation of neighbors, who through a collective process of planning and management of the territory gain scientific evidence to back their claims.

Difficulties: several territorial shortcomings. Insufficient consideration of the environment by main stakeholders. Insufficient economic incentives for deepening this research with scientists and scholars in applied subjects. Insufficient links and interaction with policy makers and businessmen.

6.4. e-Bird Argentina

eBird is a citizen science project that receives bird sightings from anybody in any part of the world. The online platform was developed in the United States in 2002 by the Ornithology Laboratory at Cornell University and the National Audubon Society. It is an open access tool to manage and share online data of bird sightings made by amateur and professional watchers, built on the simple concept that each time a watcher grabs their binoculars they have the chance to gather useful information about bird sightings.

eBird makes use of free software tools and online collaboration to efficiently gather, archive, and distribute information about birds to a much wider audience. eBird’s regional portals are customizable, in response to the need to meet the demands of local users. Each portal is integrated into the application infrastructure, and its database is in the United States. eBird is an open platform, whereby data can be shared and analysed freely across political and geographical borders.

The large amount of data collected by eBird, which contribute information about the spatial distribution of species and allow population trends to be followed, can help in the identification of important areas and sites for the conservation of birds and contribute, in this way, to the design of better plans for managing or recovering threatened species or those in danger of extinction.

eBird collects data about the appearance and relative abundance of birds in specific locations through websites available in various languages. Bird watchers who use eBird to report their sightings must follow a standardised protocol to load the information to guarantee the uniformity and quality of the registers. This protocol is quite dynamic and has improved with time, successively adding different characteristics that allow the watchers’ data to be classified in a more precise way.

When uploading data currently users must indicate with the greatest possible precision the location and protocol that was followed to count birds: if the birds were sighted a) travelling, that’s to say in movement, b) at a point in space (motionless), c) touring around an area (in which case the different ecosystems must be specified), and d) if it was a random sighting. With each protocol and the additional information required (distance travelled and time dedicated to the sighting in the case of point c), for example) an indirect measurement of the effort made by the watcher is sought.

Once the location and protocol have been selected, the site displays a verification list including the species that are most likely to be spotted in the reported location at the given time of year. In this list the users must indicate the number of each species sighted, and the information is sent. Once it has been sent, the list is subjected to some automatic control filters that seek to detect “unusual” registers. These are resent, also automatically, to the user who created them to check the data that has been flagged. If the data is confirmed to be correct, the list will then be passed to a regional expert, called an “inspector”, for evaluation, who can get in touch with the watcher to ask for additional information to help to determine the validity of the register. If the register is rejected it will not form part of the eBird database, although it will be saved in the user’s personal register.

The automatic filters are built, maintained, and updated by the regional experts. Interaction with the watchers is crucial for improving the quality of the controls, especially in regions where there is only one inspector for a very extensive area. In Argentina there are currently 20 experts who do the work of inspectors on a voluntary basis. The short-term objective is to reach one expert for each province. Beyond the voluntary work of the experts, the Aves Argentinas personnel dedicated to the project is minimal (four people), and as such it is entirely a citizen science project, depending on the voluntary participation of an amateur public.

The site appeals to amateur bird watchers who traditionally made their own lists of birds. One of the attractions of eBird for them is the ability to track their personal bird listings, share their data with other users, receive alerts about rare birds, upload their old sightings lists, explore information about when and where to find birds (which could be useful, for example, in planning a field trip), and play games that appeal to the competitive spirit [5]. More recently, eBird Argentina has included the option to upload images and sounds (as requested by the users) and use mobile phone applications, which simplifies the task of registering the birds. The site also gives users recognition for their sightings. For example, it publishes the top 10 or 100 eBirders, as much for the species spotted as for the list that was uploaded.

Box 4: eBird data sheet

Research Area: Ornithology

Partners: The Plata Ornithology Association, known as Aves Argentinas.

Start: 2013

Funding: Ministry of Science and Technology (MINCyT) and Aves Argentinas.

Features: It is an open access tool to manage and share online data of bird sightings made by amateur and professional watchers.

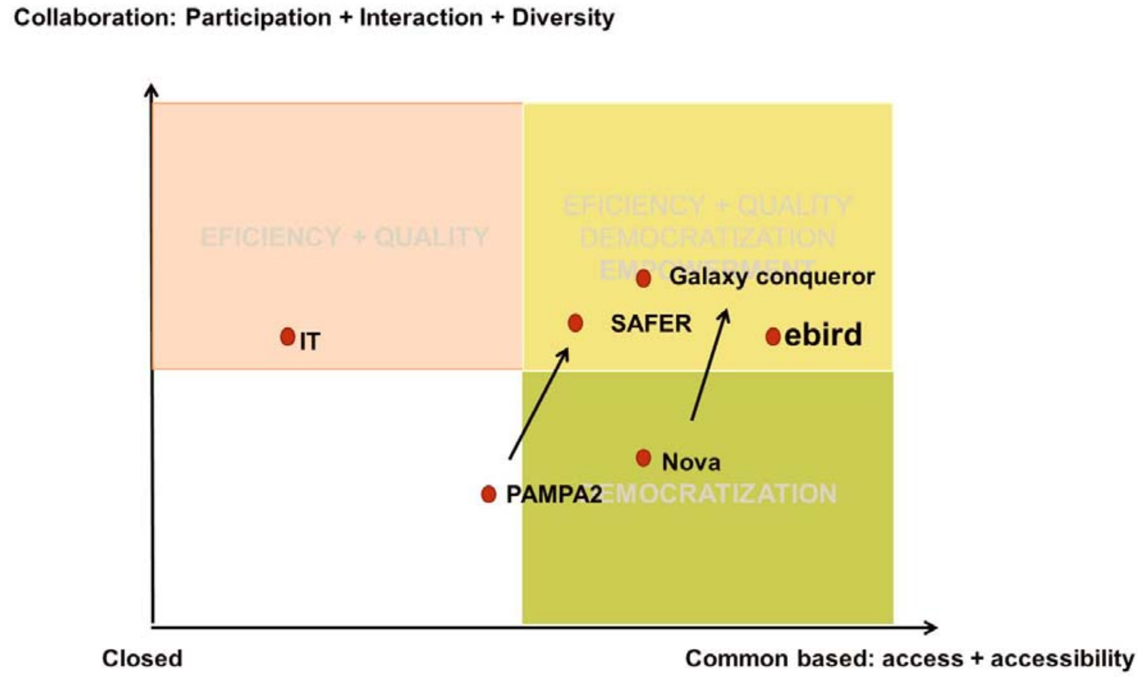
Benefits: To familiarise users with the use of standardised techniques of data collection; to increase their knowledge about birds, habitat, ecology, etc. through the interactive visualisation tools, and to improve their ability to watch through interaction with regional experts. To efficiently generate large database that is updated on a daily basis. This dataset can be used for the identification of areas that are critical for conservation of birds, and improve knowledge about different bird species in the country. In the brief period that eBird Argentina has been running, 967 of the country's approximately thousand species of birds have been detected

Difficulties: eBird depends on the work of volunteers to make revisions in the lists that do not make it through the automatic filter. Currently the number of inspectors is limited compared to the number of lists that have been uploaded and the geographical area to be covered. Another challenge for the project is the data archiving. Currently, the data is only harvested in Cornell University. And whilst the possibility of mirroring the data of eBird Argentina with the National System of Biological Data (MINCyT) is pending, it has not yet happened.

7. Openness, benefits and obstacles in a two dimensional space

Based on information from the case studies, combined with answers to the structured interviews to capture openness of these four initiatives, we plot them in the Cartesian diagram as showed in Figure 6

Figure 6: Four open science initiatives located in the two dimension space of open science



As can be seen in Figure 6 the four initiatives are heterogeneous in terms of openness. Some targeted to open up collaboration while others were mostly based on creating the conditions for a common appropriation of research resources. Thus, benefits and obstacles faced were also different.

eBird is the case that ranks higher in terms of efficiency and democratization. In fact, eBird familiarises users with the use of standardised techniques of data collection, increases their knowledge about birds, habitat, ecology, etc. through the interactive visualisation tools, and improves their ability to watch through interaction with regional experts. In sum, it leads to expertise building in amateur bird watchers. The platform promotes collaboration between professionals and the community of amateur bird watchers, and among professionals. It puts professionals from diverse regions of the country in touch, which increase efficiency since it avoids duplication and common use of shared resources. It allows the generation of a large database that is updated on a daily basis, which can be used for the identification of areas that are critical for conservation of birds, and improve knowledge about different bird species in the country. In the brief period that eBird Argentina has been running, 967 of the country's approximately thousand species of birds have been detected. Finally, it promotes the generation of processes of data driven intelligence, which allow the reuse of data and infer new hypotheses or uses by scientists and / or users who are not necessarily bird watchers (e.g. studies on ecology, human computing, and recreational users of the data).

The main obstacle based by eBird is that it depends on the work of volunteers to make revisions in the lists that do not make it through the automatic filter. Currently the number of inspectors is limited compared to the number of lists that have been uploaded and the geographical area to be covered. This situation impacts on the quality of the automatic filters because the inspectors do not have enough available time or because they do not have sufficient knowledge on the birds that inhabit the determined places. A latent challenge for the project is the data archiving. Currently, the data is only harvested in Cornell University. And whilst the possibility of mirroring the data of eBird Argentina with the National System of Biological Data (MINCyT) is pending, it has not yet happened. The risk of this set up is that Cornell stops making the data available or the project is simply suspended and the data collected in Argentina could be lost.

NOVA has been very beneficial in terms of data sharing and data re-use among astronomers, which has improved the efficiency of information produced worldwide, avoiding duplication of efforts. New data have been accessed and participants became better trained in data management and analyses. So far, the main challenge that NOVA has to overcome is the generation of a platform that could be accessible by a wide public. NOVA runs short in terms of technical personal, with one single person working on computing development, which makes it hard to increase the number of resources available, or to diversify activities to develop tools that make the process of data sharing more accessible. In part this is related to the fact that the only source of fund they got comes from CONICET, which does not include public communication of sciences as one of its priorities. However, although NOVA originally ran a bit short in terms of amplifying this impact beyond the scientific community, the creation of a new related project using citizen science practices (Galaxy conqueror) has definitely improved the diffusion of astronomy among the wide public, democratizing science.

The main benefits of PAMPA2, in turn, are related to the scientific aspects of the project: that is the creation of interdisciplinary collaboration among local and international scientist in part as a consequence of producing open access data, fostering capacity building between local scientists and improving efficiency in the use of locally generated data. Another interesting aspect is the creation of local tools. However, in terms of democratization and empowerment of the local communities who are the most affected by changes in the lakes, they have only partially been materialized. To open-up these elements further, the binding constraints that seem to be operating are cultural rigidity. The obstacles and barriers we identified had to do with lack of vision of the potential benefits associated to opening-up the project to a wider audience. This is in part explained by the scarcity of resources available in the scientific system to finance those efforts (e.g. normally there are neither funds nor incentives to devote efforts to communication of science). However, it was not just a question of resources. For example, the buoy could have been produced using an open source approach from scratch, but it was not. The team could have used online platforms to increase the frequency of their interactions, but they did not. They could have attempted to publish open access, but they did not. As far as we can see, these open science practices were originally out of reach/vision of the involved researchers, but they may start looking beyond, as these practices become better known, diffused and accepted, and as long as they start experiencing with them. The fact that one of the teams involved in PAMPA2 started to be part of an international project that seems to be more committed to the common-based aspects of open science, could be start point towards that direction.

The IT project was an ongoing project at the time of our case study, so we cannot really assess its benefits. The local community that participated in the project (which is a small proportion of the community affected by the floods) has increased their knowledge about territorial planning and they have also collected some data that could back their claims in the future. Researchers claimed that multidisciplinary work was useful to look at different relevant aspects of the socioenvironmental problems, but in order to interact, they had to overcome several aspects related to scientific cultural rigidities. They claimed that multidisciplinariety somehow risked the likelihood of obtaining publishable outcomes, partly because specialized journals normally belonged to certain disciplines and also because the final outcomes depended on the commitment of other researchers in a context where quality could not be cross-checked due to lack of specific skills. Power asymmetry was also an outstanding obstacle, because local authorities were the ones who decided when it was the right political time to show certain results and to reach consensus on what and when solutions would be carried out. Moreover, researchers also said that it was puzzling to work with local communities in the context of political disputes, because they (the researchers) did not want to create false expectations on the outcomes of the project, while at the same time they needed to motivate the community to be part and committed to it.

8. Conclusions

This paper is an attempt to organise different elements of openness in order to relate them to specific benefits and barriers attached in the literature to open science. Our claim is that both benefits and barriers are somehow related to the specific characteristics of the opening up process. We built an analytical framework that characterised the level of openness of particular initiatives of open science in a two dimensional space, one related to the process of peer-production (i.e. interaction, participation and diversity of participants) and the other on the common-based characteristics of resource appropriation (i.e. access and accessibility) The framework also indicated what obstacles or barriers would be more likely to turn up when trying to open-up science in these two dimensions.

The analytical framework was illustrated using data of four case studies of open science initiatives from Argentina. All the selected initiatives did a great job in opening up some aspects of their research project, which redounded in improving benefits in terms of scientific efficiency and quality, democratization of science and empowerment of the population related to the research project. Moreover, in some cases, the experience of openness was so good, that project leaders initiated new related project involving a deeper commitment towards openness. Yet, we found there was space for going further in the opening processes, which might be highly beneficial for the cases we analysed.

Nevertheless, our point is that there are several directions of openness, and choosing each one depends on different factors. This means that the expected benefits and obstacles of open science are also diverse. We believe our analytical framework could be informative for researchers, policy makers and practitioners. It works as a guide for characterising open science experiences and it helps to identify specific aspects of open science practices that could be opened-up further. It also anticipates the specific barriers that will be needed to overcome when trying to do so.

References

Arnstein, S. R., (1969). 'A Ladder of Citizen Participation', *Journal of the American Institute of planners*, 35(4), pp. 216-24.

Arthur, C., (2013). 'Tech Giants May Be Huge, but Nothing Matches Big Data', *The Guardian,* 23 August 2013, London:

Baden, T., Chagas, A. M., Gage, G. and Marzullo, T., (2015). 'Open Labware : 3-D Printing Your Own Lab Equipment', *PLoS Biology*, pp. 1-12.

Bartling, S. and Friesike, S., (2014). 'Towards Another Scientific Revolution', in S. Bartling and S. Friesike (ed.), *Opening Science,* Springer International Publishing.

Ben-David, J., (1960). 'Roles and Innovations in Medicine', *American Journal of Sociology*, pp. 557-68.

Benkler, Y., (2006). *The Wealth of Networks: How Social Production Transforms Markets and Freedom*, Yale University Press.

Bijker, W., (1997). *Of Bicycles, Bakelites, and Bulbs. Toward a Theory of Sociotechnical Change Cambrige*, MA: The MIT Press.

Bishop, F., (2015). 'Who's Afraid of Open Data: Scientists' Objections to Data Sharing Don't Stand up to Scrutiny, ' *The Impact Blog - LSE. ,* available at http://blogs.lse.ac.uk/impactofsocialsciences/2015/12/16/whos-afraid-of-open-data-dorothy-bishop/:

Carson, R., (2002). *Silent Spring*, Houghton Mifflin Harcourt.

David, P. A., (2003). 'The Economic Logic of Open Science and the Balance between Private Property Rights and the Public Domain in Scientific Data and Information: A Primer', in J. M. Esanu and P. F. Uhlir (ed.), *The Role of the Public Domain in Scientific and Technical Data and Information,* National Academies Press.

Fecher, B. and Friesike, S., (2014). 'Open Science: One Term, Five Schools of Thought', in (ed.), *Opening Science,* Springer.

Gregson, J., Brownlee, J. M., Playforth, R. and Bimbe, N., (2015). 'The Future of Knowledge Sharing in a Digital Age: Exploring Impacts and Policy Implications for Development', *Evidence Report Nº 125*.

Hess, D., (2007). *Alternative Pathways in Science and Industry. Activism, Innovation and the Environment in the Era of Globalization* Cambridge, MA: The MIT Press.

Irwin, A., (1995). *Citizen Science, a Study of People, Expertice and Sustainable Development*, London: Routledge.

Jeppensen, L. B. and Lakhani, K., (2010). 'Marginality and Problem-Solving Effectiveness in Broadcast Search', *Organization Science*, 21(5), pp. 1016-33.

Lakhani, K. R. and Wolf, R. G., (2005). 'Why Hackers Do What They Do: Understanding Motivation and Effort in Free/Open Source Software Projects', in J. Feller, B. Fitzgerald, S. Hissam and K. R. Lakhani (ed.), *Perspectives on Free and Open Source Software,* Cam-
bridge, MA: MIT Press.

Lundvall, B.-Å., (1992). *National Systems of Innovation : Toward a Theory of Innovation and Interactive Learning*, London ; New York: Pinter Publishers; Distributed exclusively in the USA and Canada by St. Martin's Press.

Masum, H. and Harris, R., (2011). *Open Source for Negleted Diseases: Challenges and Opportunities*, Center for global health R&D Policy Assessment.

McAllister, J. W., (2012). 'Climate Science Controversies and the Demand for Access to Empirical Data', *Philosophy of Science*, 79(5), pp. 871-80.

Millstone, E., (2015). 'Invoking Science in Green Transformations', in I. Scoones, M. Leach and P. Newell (ed.), *The Politics of Green Transformations,* New York, : Earthscan from Routledge.

Molloy, J., (2014). 'Open Training for Open Science', *OKF - Open Science Working Group,* available at: http://science.okfn.org/2014/12/21/open-training-for-open-science/:

Molloy, J. C., (2011). 'The Open Knowledge Foundation: Open Data Means Better Science', *PLoS Biology*, 9(12), pp. p.e1001195.

Nielsen, M., (2012). *Reinventing Discovery: The New Era of Networked Science*, New Jersey: Princeton University Press.
RIN/NESTA, (2010). 'Open to All? Case Studies of Openness in Research'.
Shah, S. K., (2006). 'Motivation, Governance, and the Viability of Hybrid Forms in Open Source Software Development', *Management Science*, 52(7), pp. 1000-10014.
Surowiecki, J., (2004). *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations Little*, New York: Anchor Books.
UN Independent Expert Advisory Group Secretary, (2014). 'A World That Counts: Mobilising the Data Revolution for Sustainable Development, United Nations Independent Expert Advisory Group on a Data Revolution for Sustainable Development'.
Wagner, C. S., (2009). *The New Invisible College: Science for Development*, Brookings Institution Press.
Wiggins, A. and Crowston, K., (2011). 'From Conservation to Crowdsourcing: A Typology of Citizen Science', *Paper presented at System Sciences (HICSS), ,* Hawai:
World Bank, (2015). 'Open Data for Sustainable Development', *Public Note ICT 001*.