Benefits and obstacles of openness in science: an analytical framework

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 plenty of 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 elements 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 the specific aspect of the practice of open science that is being opened (or attempted to be opened). We propose an analytical framework building from RIN / NESTA's (2010) dimensions of openness (*what* is being opened, *under what conditions or how*, and *who* can enjoy the benefits of openness) and research stages. Our framework allows us to characterize different open science initiatives and to compare them with 'ideal' levels of openness in different elements so as to make the most of potential benefits. The framework also indicates what specific obstacles or barriers will need to be overcome when trying to opening-up specific elements of the practice of open science. We illustrate our framework by discussing and Argentinean initiative on open science in limnology studies.

* 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. This working paper was presented in the OCSDNet workshop, held in Bangkok Feb. 16th - 20th, 2016.

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 five sections. Next section 2 presents a discussion of the different dimensions of openness organised along research stages. Section 3 presents the discussion on benefits of open science practices and proposes an analytical framework to associate those benefits with levels of openness. Section 4 does the same but for obstacles and barriers. Section 5 concludes.

2. Dimensions and elements of openness

Building from the RIN / NESTA (2010) report we may identify three main dimensions of openness related to:

1. *What* it is being opened: while traditional open access movements focused on promoting access to the final outcome of scientific production (i.e. publications), more recently the literature and advocacy moved towards sharing intermediate products, such as data, research

protocols, laboratory notes, design proposals, etc., and also processes, such as interaction and participation in the research design and analysis with actors outside the laboratory environment (Grand, 2012; Kelty *et al.*, 2009).

2. *How* it is opened (under what conditions): the degree of openness on final and intermediate outputs and processes could vary depending on explicit restrictions. These restrictions could be more formal; as in the imposition of paid subscriptions or formal licenses (to read/use, to distribute, to reproduce, etc.) (Molloy, 2011) or informal; as in the necessary skills or other resources needed to fully enjoy the functions of what is being shared. Moreover, in the case of research processes, outsiders could participate i) on individual basis or collectively and interactively and ii) they could be more or less involved in the analysis or design of the research project. Regarding the former, 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. As for the participation, 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).

3. *To whom* the openness process is oriented. Researchers tend to share their research outputs with other scientists rather than with a diverse audience (while intermediate outputs would be only shared with very close colleagues). 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 extent of openness in these three dimensions contributes to creation of different type of benefits that the literature has associated to open science practices. For example, in dimension 1 opening up intermediary and not just final research outputs, multiplies the pool of evidence available to all scientists and therefore the generation of new discoveries accelerates (Tacke, 2010; Vision, 2010; Woelfle *et al.*, 2011). In dimension 2, when research outputs become not just fully available but also translated and communicated so that everyone could appropriate those results, democratization of science is definitely enhanced. In dimension 3, when research processes enable public participation of a wide community of actors (and not just scientists), it amplifies the power of inquiry and application because evidence and research questions are collected from a variety of contexts (Catlin-Groves, 2012).

Similarly, obstacles and barriers to open science also differ in relation to different dimensions of openness. For example, researchers might be more resistant to opening up data than publications (dimension 1); there may be different legal restrictions (dimension 2) raised to opening up data, publications or other research outcomes; as a community of experts researchers may be reluctant to share the research process with non-scientists while they look for opportunities to raise money doing collaborative research with other colleagues (dimension 3); etc.

In order to systematize the relation between dimensions of openness and benefits and obstacles we find it helpful to intertwine these dimensions across four different research stages of the scientific production processes: 1. Research design and analysis; 2. Collection of data; 3. Communication and use of research outcomes; and 4. Infrastructure (to facilitate openness in all previous stages)[2].

This crossing of dimensions and stages allows us to identify fourteen different elements of openness in Table 1 which are relevant to assess open science experiences in terms of potential benefits and obstacles. Section 3 describes benefits, and Section 4 describes obstacles

Table 1: Dimensions, research stages and elements of openness in science

		What		How	To whom
				Elements of openness	
Design		Research questions		i. Degree of participation ii. Degree of interaction	iii. Quantity of actors iv. Diversity of actors
		Methodology			
Analysis		Calculus			
		Concepts			
Collection		Observation		v. Lack of mediation / conditions to produce data	vi. Quantity of actors vii. Diversity of actors
		Experiments			
		Samples			
		Archives			
		Interviews			
		Surveys			
Communication		Intermediate outputs	Laboratory notes	viii. Lack of restrictions to access to data ix. Easiness to enjoy functions of shared outputs x. Lack of restriction to access to outputs	
			Data		
			Protocols		
		Final outputs	Publications		
			Prototypes		
Infrastructure		Software Hardware		xi. Degree of participation xii. Lack of restrictions to re-use and re-distribute	xii. Quantity of actors xiv. Diversity of actors

2 We took the idea of research cycle from RIN/NESTA (2010) with some modifications. In particular we have chosen to join the phases of design with analysis, since the elements included in these dimension are quite similar and they also interact similarly with benefits and obstacles.

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 what the specific elements of openness are interplaying in each case. These elements will be identified between brackets by their number in Table 1.

1. *Increased research productivity and dynamic efficiency*

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 openness and collaboration (*elements ii, viii and x*). 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 (*elements viii and x*), 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 (*element ii*). All researchers get to know faster and more accurately by interacting with other researchers in their fields.

d) Open access (*elements viii and x*) 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 (*element ii*). 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 (*element xii)*. 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).[3] 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 (*element xii*).

g) By increasing the quantity of actors participating in data collection (or analysis) (*elements iii and vi)* 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 (*elements iv and vii*) (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 one well inside the problem field. A fresh look into a problem could be especially relevant when scientific groups

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

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.

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 (both through improving access and accessibility to knowledge) (see Fecher and Friesike (2014))[4]. 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 more science during the process.

Open access (*elements viii and x*) 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 (*element ix*). These costs are inversely related to the investment in knowledge translation and communication efforts.

The same kind of electronic tools that encourage increases in productivity and dynamic efficiency might also promote the democratization of scientific knowledge (*element xii*). There are two key aspects of democratization here: first, electronic tools are based in a model of open sharing of knowledge and ideas. 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. The second key element in open science infrastructure is participation. 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 (*elements iii, iv, vi and vii*) 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 and democratise the process of scientific production, there should be efforts to free restrictions so as to increase the likelihood that everyone could participate in

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

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 (*element v*).

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 (*element ix*). 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) (*elements viii and x*). 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 (*element ii*). 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 software or open hardware. 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) (*element xii*).

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) (*element i and xi*). 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.

4. *Innovation or solution to problems*

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. 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.

In fact there are two mechanisms underlying the potential of opening science for innovation. The first mechanism is related to the effects of open science as an empowering tool, which increase the potential for learning and therefore for innovation. Relevant elements of openness here are those associated to participation, open access and communication (*elements i, viii, ix, and x*). More information available and more brains putting forward problems and solutions will render faster innovation rate. The second mechanism relates to the practice of open science as a way to improve scientific productivity that works as an input for innovation, including the elements of interaction, diversity and quantity of actors participating in the scientific endeavour, and open access (*elements ii, iii, iv, vi, vii, viii, x and xii*).

There are few examples of successful innovation based on open source characteristics -mainly in open software, open hardware and a few examples of alternative sharing mechanism such as open seeds- but universities and mainstream R&D institutions have been slow in taking notice of these possibilities.

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.) and open access (costs reduction and creation of commons). 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.

In sum, although open science could contribute to innovation and, in theory, innovative outcomes could be widely appropriated in the society, these kinds of open initiatives (in production and in appropriation) tend to face a lot of resistances and challenges for incumbent actors and established practices in the R&D networks of academia and companies. We will expand a bit on this in Section 4.

3.2. Analytical framework to associate levels of openness and associated benefits

We propose a methodological tool to assess level of openness and associated benefits of different experiences of open science around the world. To display the multidimensional concept of open science in a two-dimensional space, we use a radar chart. Each of the fourteen elements mentioned in Table 1 will be represented by an axis starting from the center of the graph. The relative position of the elements (spokes in the chart) follows the order as they were mentioned in Table 1. We will arbitrarily set the length of the spoke as 5 to be able to use a 5-point Likert scale when assessing empirically each of the fourteen elements.

The angles between spokes are all equal size. We will then be able to visualise each open science experience as a star-like plot, which makes it easier to characterise each experience by its relative degree of openness in the fourteen elements of openness.

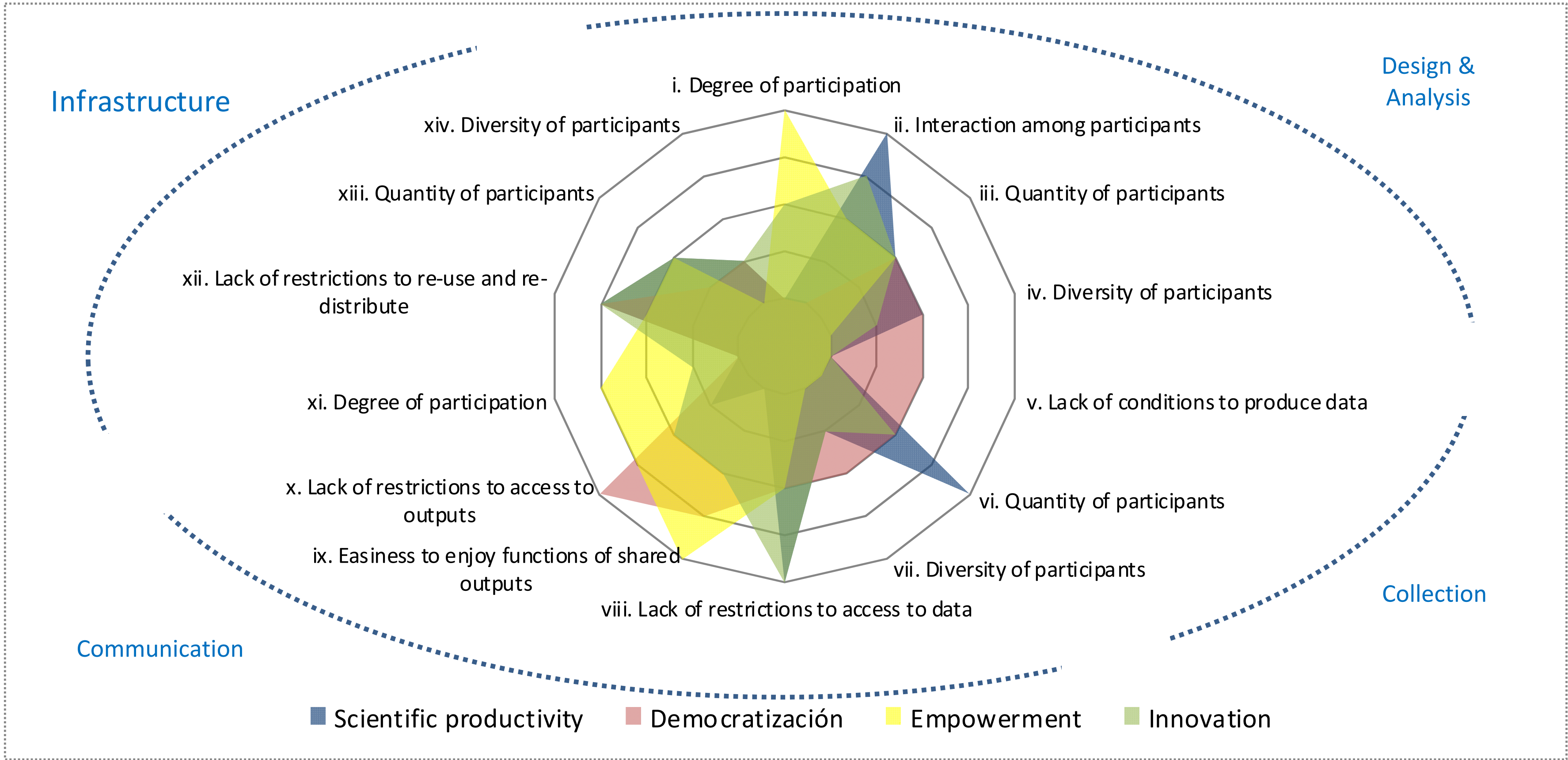
We can also use this tool to compare each open science experience with an ideal type. Based on the above discussion we could actually construct four ideal types, one for each of the benefits that have been associated to open science practices in the literature. To that end we present Table 2 that assigns weights (1-5 Likert scale) to the importance that each of the 14 elements of openness has in driving each of the four identified benefits of open science. The justification for each of these weight could be found in the narratives above.

This matrix is plot in Graph 1. As can be seen, each benefit is represented by a star-like area. These four star-like plots could be used as a sort of performance metrics. They could become a *standard* to be used to assess open science initiatives (i.e. to what extent benefits could be expected in terms scientific productivity, democratization, empowerment and innovation given the observed levels of openness in different elements). We could even create a single ideal type encompassing the four benefits taking the maximum weights in the four benefit as *the standard* for each level of openness (column *All* in Table 2), which correspond to the painted (with any colour) area in Graph 1.

Table 2: Elements of openness positively affecting scientific productivity, democratization, empowerment and innovation (weights in 1-5 Likert scale):

Elements of openness		Benefits				
Research Stages		**Scientific productivity**	**Democratización**	**Empowerment**	**Innovation**	**All benefits**
Design & Analysis	i. Degree of participation	1	1	5	3	5
	ii. Interaction among participants	5	1	3	4	5
	iii. Quantity of participants	3	3	3	3	3
	iv. Diversity of participants	3	3	1	2	3
Collection	v. Lack of conditions to produce data	1	3	1	1	3
	vi. Quantity of participants	5	3	1	3	5
	vii. Diversity of participants	2	3	1	2	3
Communication	viii. Lack of restrictions to access to data	5	3	3	5	5
	ix. Easiness to enjoy functions of shared outputs	1	4	5	3	5
	x. Lack of restrictions to access to outputs	2	5	4	3	5
Infrastructure	xi. Degree of participation	1	1	4	2	4
	xii. Lack of restrictions to re-use and re-distribute	4	4	3	4	4
	xiii. Quantity of participants	3	2	3	3	3
	xiv. Diversity of participants	2	2	1	2	2

Graph 1: Elements of openness and their expected benefits



The areas in radar graphs could be calculated as a sum of *n* triangles, using the following formula:

$$A = \frac{1}{2} ab \sin C$$

In our case, there are fourteen (*n*=14) triangles for each star-like area. Sides a and b are the spokes of two consecutive elements of openness (e.g. element i. degree of participation in design and analysis and element ii. degree of interaction in design and analysis) and are informed in Table 2. Angle C is the one formed between each of two consecutive elements of openness, which are all the same side equal to 360/14 radians.

Radar charts create some limitations when used as metrics to compare across series of outcomes. In our case, the relevant limitations are the following:

a) The data structure: On the one hand, we are creating some spurious connections (i.e. neighbour spokes are not necessarily related conceptually). Although, some of the connections are related (e.g. they are all elements occurring in the same research stage), some of them are not. On the other hand, we are imposing a cycle structure to the multidimensional characteristic of our data (i.e. first and last elements are drawn close together), when this is not really the case. Although some circularity in the research process does exist.
b) Using area as metrics: 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). This issue when combined with the above-mentioned issue of data structure implies that the order in which we organised the different elements matters for calculation of metrics. The calculated area will depend on whether elements ranked higher in each initiative happened to be plot together or not (which is quite an arbitrary decision).

However, we still decided to use it because, firstly, it is a simple tool to analyse performance of multidimensional outcomes; secondly, we do not plan to use this tool to decide over one experience against others but rather to understand the potential benefits of each experience given its degree of openness in different elements; thirdly, the connections and circularity of elements of openness are not totally arbitrary,[5] fourthly, we will not compare areas as such but percentage of coverage for each of the four ideal areas (one for each potential benefit) by each analysed open science initiative, and finally correlation analysis between percentage of covered area and a weight index calculated for each benefit using data for different open science initiatives is above 95%.

3.3. Illustrative example using the analytical framework: The Pampa2 project: Pampa lagoons as sentinels of climate and anthropogenic change.

As an illustrative example of how this framework could be used we characterise one open science Argentinean initiative.

[5] Others in the past have also presented similar conceptual connections and circularity. See for example RIN/NESTA (2010)

Pampa2 is an interdisciplinary network of scientists from seven different research laboratories in Argentina aimed at monitoring how lagoons' ecosystems deal with climatic variation and other anthropogenic changes (e.g. agriculture). 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 5 years. To collect data the team uses buoys that measure temperature, atmospheric pressure, humidity, precipitations, oxygen and chlorophyte levels, among others and they complement that data using laboratory information from samples collected monthly or every six months from the lagoons. The buoys 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 buoys as well as other information generated by the project can be requested to the teams.

We will use our analytical framework to characterise Pampa2 project. All Graphs 2 produce the same star-like plot of the project, which is consecutively contrasted against target areas representing the ideal level of openness in different elements to achieve benefits associated to scientific productivity (2a), democratization (2b), empowerment (2c), and innovation (2d).

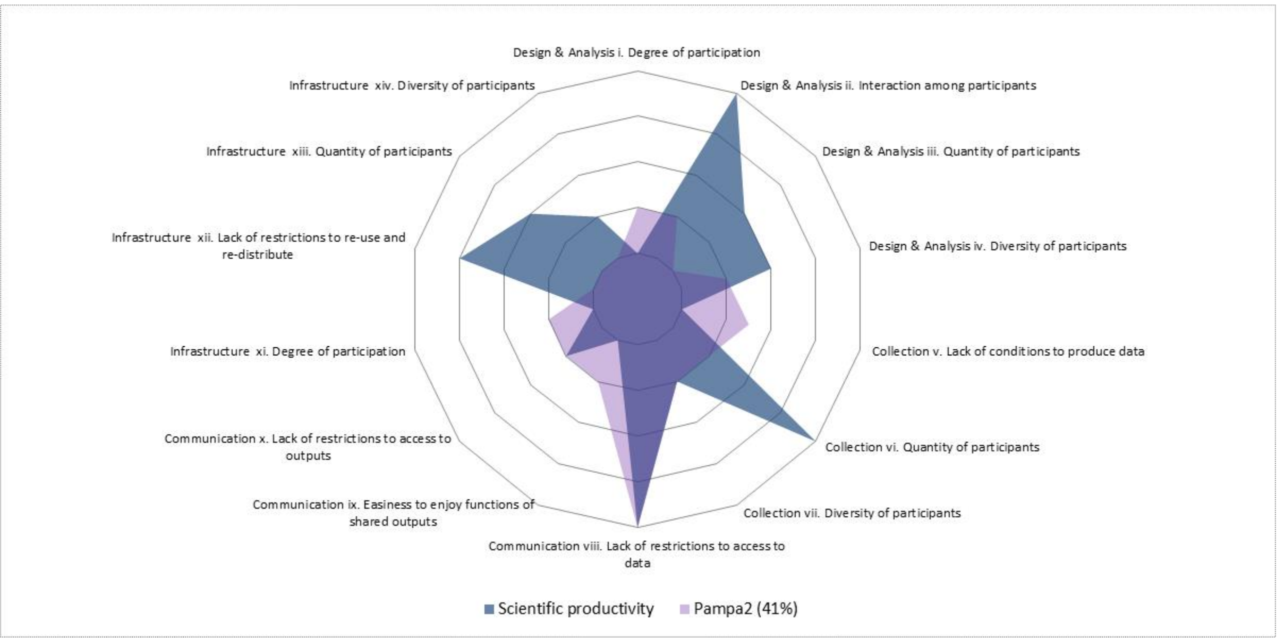
We now organize the description of openness and collaboration of this experience by research stage in order to justify the specific mark (in the 1-5 Likert scale) we assigned to Pampa2 in each of the fourteen elements of openness identified in our framework

In the Pampa2 project, only those teams that originally formed the network participated in the design, collection and analytical phases. Actually, the design of the project was carried out predominantly by one of the networked organizations. This explains the relatively low rank to quantity of participants in design and analysis (*element iii;* Likert point 1) and collection (element vi; Likert point 2). In terms of diversity, although there is some diversity in terms of disciplines, all participants are professional scientists from close scientific fields (*element iv and vii;* Likert point 2). In terms of interaction, the position is also fairly weak (*element ii*, Likert point 2) because there are no formal instances for interaction for all members of the network besides one workshop held every year. Otherwise, they would 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 (element i, Likert 2). 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 communication, there is open access to most of the data (element viii, Likert 5), but the team is not particularly prone to publishing open access (element x, Likert 2). Finally, the teams do very little efforts to communicate their results to a wider audience (element ix, Likert 1) in part because neither they have the required expertise nor the necessary resources to hire these services. They have not even written a protocol to allow users to work properly with the data they produce.

Lastly, in terms of infrastructure, it is worth highlighting that the buoys used in the project were designed and built by one of the research teams and in 2011 it won the second prize in the National Innovation contest organized by the Ministry in Science and Technology. The researcher that was the main responsible of this innovation, had not yet finished his PhD, so in this case, participation seemed to have been a little bit less hierarchical or more participative (element xi, Likert 3) although this is actually tricky to assess because there were too few participants in the design of the buoy (probably nor more than one or two people) (element xiii and xiv, Likert 1). So far the buoy was not designed following and open source approach (element xii, Likert 1) but the same teams was currently working in a new design of open source software for more ambitious projects, such as the design of buoys that can support more extreme environments, such as those in open seas. Therefore, open source infrastructure has been considered as something to achieve in the future but not as a priority for the project.

Graph 2a: Ideal level of openness to target **scientific productivity** and actual openness of Pampa2 project which covers 41% of targeted area.

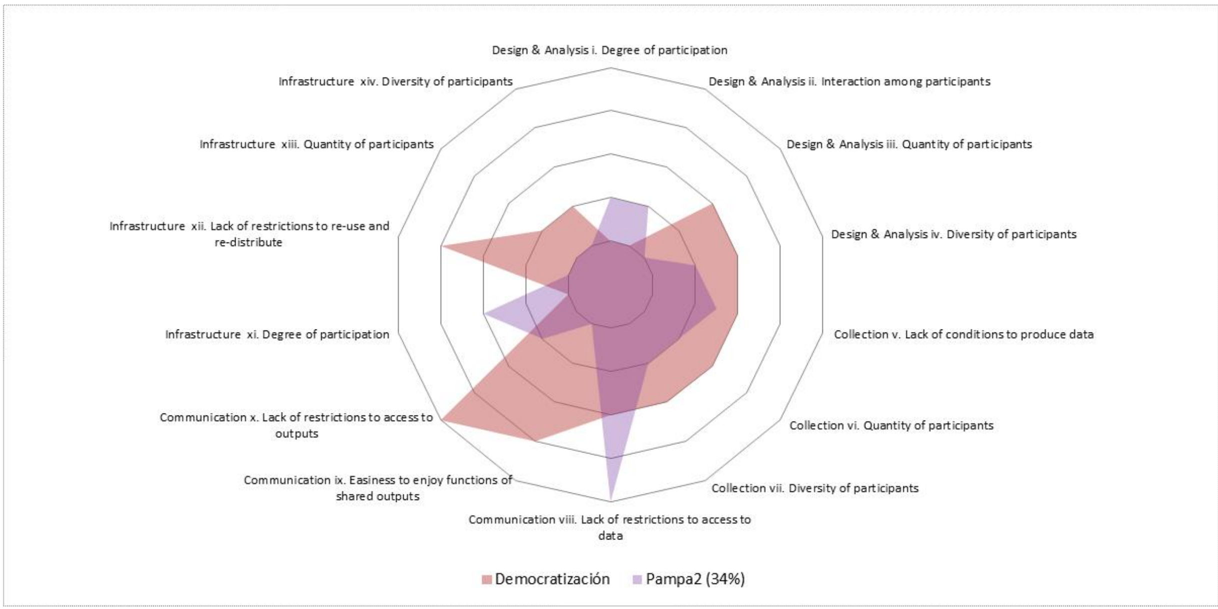


As can be seen in Graph 2a, 41% of the target area for achieving scientific productivity was covered by the Pampa2 project. The good performance in terms of data sharing explains a great deal of this metric. The role of data sharing for scientific productivity was confirmed in our interviews. On the one hand, some of the researchers increase the capacity to produce papers, because they could collaborate and therefore access to data produced by other teams in the network. Moreover, the open access nature of Pampa2 project allows it to be member of the Global Lake Ecological Observatory Network (GLEON), which aims to promote the exchange of data produced by high-resolution sensors to as to better predict the role and response of lakes to climate and environmental change. This network is creating new research opportunities for Pampa2 team members. For example, data generated by buoys of Pampa2 are used in an international project called SAFER (Sensing the Americas' Freshwater Ecosystem Risk from Climate Change) in which research groups from

Argentina, Canada, Chile, Colombia, Uruguay and the USA participate. This project shares similar goals as those of Pampa2's but with a highly important social component (i.e. how the population is affected and how it could be better involved research and monitoring) not included in Pampa2.

The shortfalls in terms of scientific productivity are mostly related to the low performance in openness and collaboration in Design and Analysis and in Infrastructure. Interaction and participation in these phases can yet be much improved, as it has been recognized in our interviews: there is insufficient communication intra-network.

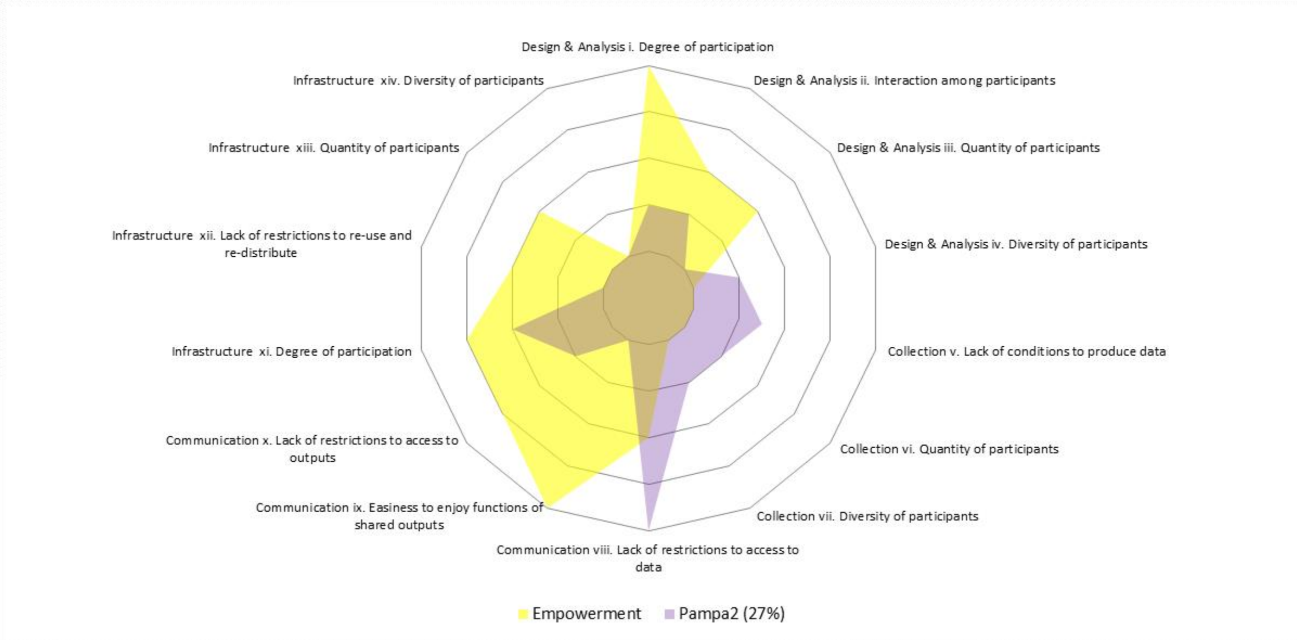
Graph 2b: Ideal level of openness to target **democratization** and actual openness of Pampa2 project which covers 34% of targeted area.



Graphs 2b shows that Pampa2 covers 34% of the democratization target. In this case the short falls are related to the lack of efforts to communicate data and outputs to the wide audience.

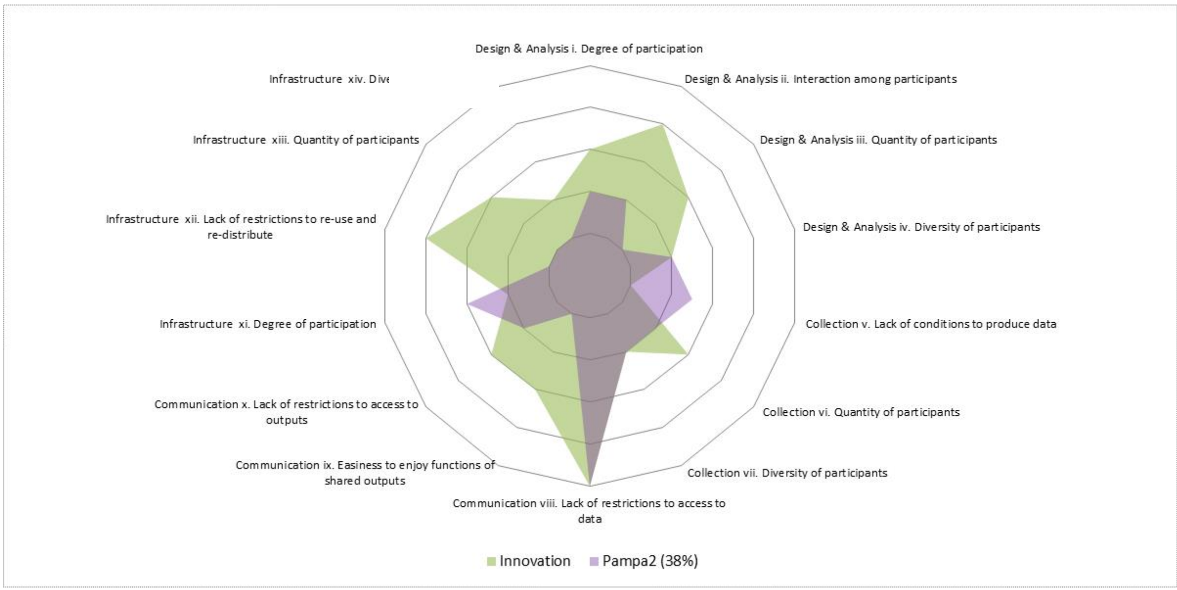
Although data collected by the automatic buoys of Pampa2 is available to the public through a website, there is no clear protocol on how to use this data. We understand that researchers intend to share project's information and results with the local population but they did not manage to do so due to lack of resources. Although the project funder seems to be very supportive of open access approaches and open repositories, in concrete terms, intermediate and final outputs of the project hardly reach the public.

Graph 2c: Ideal level of openness to target **empowerment** and actual openness of Pampa2 project which covers 27% of targeted area.



Graphs 2c shows that only 27% of the empowerment target area is covered by Pampa2 project. Actually, the intuition we got is that the only aspect of empowerment that is present in the project is scientific capacity building. The project has an important component of formation of junior scientists and students, that learn to use the tools and work in interdisciplinary groups. It runs short in terms quantity of participants and degree of participation, both in the Design and Analysis and the Infrastructure stages. We know from the interviews that they aim at improving openness in these elements, by means of adopting open source approaches in the design of buoys and by incorporating other participants from the local population in some of the research stages as it is expected by SAFER project.

Graph 2d: Ideal level of openness to target **innovation** and actual openness of Pampa2 project which covers 38% of targeted area.



Graph 2d shows the project's potential for innovation and it reached 38% of the area. As in research productivity, the main contribution to the target area is given by open access. As said above, the open access characteristic of the project allows its participation in big international networks doing research in the subject. The laboratory that was responsible for the design of the buoy developed capabilities in sensor production and its software, which later led to the production of some commercial buoys that were sold to the Uruguayan component of the SAFER project. Moreover, the researcher that did the lion share of the design of the buoy is currently involved in other projects with collegues in the US where he is finishing his PHD. One of the interesting challenges put forward in this further collaboration was the need to design another buoy that could support harder weather conditions. New innovations are ready to come, but as it was recognized by our interviewees there is still room for improving, especially by adopting an open source approaches in buoys design and development.

In sum, the main benefits of Pampa2 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.

4. Barriers, obstacles and challenges

4.1. Barriers to openness in the literature

This section briefly analyses the main obstacles and challenges faced by open science initiatives.

A first set of barriers 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 barriers 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"[6]: 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.

There is a third set of obstacles and barriers that are pragmatic in nature and had to do with problems of coordination that become the more apparent when open science initiatives search for expansion or replication.

As we have done with benefits above, in the following narrative we relate obstacles and barriers with specific elements of openness (between brackets). This should be read as the specific obstacles and barriers that come up when trying to increasing openness for different elements of the research process.

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.

One clear example is the incentive and evaluation system in research that privileges publications in peer-reviewed journals. This scheme does not promote openness and collaboration in science, and it does not promote scientific communication and dissemination either (*elements viii, ix and x*).

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.[7] In order to surpass it,

---

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

[7] A recent study shows that 21% of papers are rejected without peer review, either because of lack of merit or lack of editorial interest. An additional 39% are rejected during the peer review process and only 40% are normally accepted, prior to suggested changes, by two or three reviewers. If the manuscript is within this last group, the researcher will have

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). Moreover, as the journals require original results, to come first is essential. This discourages collaboration (*elements ii and iii)*.

Open science practices often challenges these traditions, which are widely used for scientific assessment but also deeply embedded in the scientific culture.

For example, to organize activities and to promote wide participation of diverse actors require significant amounts of time and leadership which are not always recognized by the evaluators (RIN/NESTA, 2010) (*elements iv and vii*). Similarly, efforts 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 (*element ix*).

Moreover, open collaboration among scientists without formal connections (i.e. from different laboratories that do not participate in the same project) faces a lot of resistance within the scientific community. Scientists are still very much attached to a culture of scientific competition for grants, publications and resources. Therefore, they might look with disdain or fear to those tools that allow open sharing of ideas, protocols and data (Nielsen, 2012) (*element i*). In the case of open data and open access, scientist might argue that data might be "scooped" by other scientists before the analysis is completed (Bishop, 2015) (*element viii*). Another common concern is the reliability of open data, in particular when its collection involved process of citizen or crowdscience (RIN/NESTA, 2010) (Wiggins and Crowston, 2011) (*element v*).

Finally, traditional models of producing science do not favour open interaction with broader networks (*elements iv and vii*). This could be due to a couple of reasons. On the one hand, policies of science and technology are sometimes attached to a model of endogenous capability building and international competition that leaves little space for collaboration. Since scientific production is changing towards an invisible international college (Wagner, 2009) this obstacle emanating from cultural rigidity implies in fact losing opportunities to tap into these international networks making more inefficient national investment in science and technology. On the other hand, policies aiming fostering innovation and entrepreneurship through commercialization of science (see for instance Etzkowitz (1998)) are well embedded in scientific organisations and also can present obstacles: they discourage policies of open access and that are the backbone of open participation (*elements viii, ii, iii, and iv)*.

Nevertheless, something may be changing. Although there is still resentment to open science due to cultural rigidity, we currently assist to wider spread of open science practices within some scientific disciplines (e.g. environmental science, astronomy, ecology). Moreover, scientific funding agencies and international organizations have begun to promote openness in their call for proposals.

to wait an average of 12 months to see their article published. Nevertheless, it is more likely that a manuscript will not be accepted by the first journal where it is submitted.

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

2.1. Unequal opportunities in data appropriation

Open access policies intend to restore the public nature of scientific production. In Argentina, these policies are shaped by the National Law for Digital Open Repositories forcing researchers to file all research outputs (publication and data) associated to projects funded through public sources. Some disciplines -particularly as physics, mathematics and astronomy- had already incorporated the use of open repositories as part of the practice of online collaboration using sites such as arXiv.org.

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 (*element x*). If this appropriation increased power asymmetries in detriment of powerless actors, inequality may increase rather than decrease as a consequence of open access. In that case, it may be more convenient that policy-makers were able to discriminate between potential users of knowledge.

In the context of scientific production in developing countries, the risk is that the opening-up produces a paradoxical effect in which increasing scientific production end-up increasing technological dependence due to asymmetric appropriation of open data by powerful and dominant actors. In our case studies we have found some resistance and doubts about open access policy in this regard.

2.2. Political disputed contexts

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 "undone science" (Hess, 2007) or "alternative science". 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 the “alternative science” agenda[8] 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 and collaboration to a close group of trustable allies (*elements iii, iv, vi, vii and x*), at least until they manage to gather stronger political support for their scientific endeavour which would eventually support them should conflicts turned up.

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

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, validity issues (*element v*) 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 (*element i*) (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 (*element v*) and questions about when is the right time to opening and disseminate data and analysis and how it should be done (*elements viii and x)* (McAllister, 2012). As the case known as the Climategate shows, information that is made public before a proper analysis and curation (*element v*) could be detrimental of the causes advocated by the proponents of the initiatives (McAllister, 2012). [9]

3. *Coordination problems and other obstacles that turn up in the context of expansion and replication*

One of the strong assumptions of open science is that increasing participation and interaction will improve the process of knowledge creation. However, to foster interaction among scientists it is not always easy and these intentions can become increasingly complex if openness includes lay public participation.

Some of the obstacles that are mention in the literature are: a) difficulties to foster interaction among researchers; b) difficulties to harness and take full advantage of wider participation from the public, and c) difficulties associated to global open sources approaches needed to develop open science infrastructure.

a) Although there are plenty of arguments in favour of interaction and collaboration for knowledge creation, in practice this is not always easy. In interdisciplinary projects for example, a lot of time is spent in the creation of a common language across disciplines (*elements ii, iii, iv, vi, and vii*) (Wagner, 2009). Actually, Nielsen (2012) argues that open

[9] 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).

collaboration is easier in cases where the different participants share a common set of knowledge and techniques. Without the existence of shared practices, collaboration is still possible but it might take longer to create a common praxis. Moreover, coordination is also needed to fairly acknowledge contributions into scientific production, motivating scientists to participate.

b) Allowing open participation of non-scientist can improve several phases of the research cycle while at the same time improving the overall scientific literacy. There are plenty of cases of open science that have successfully implemented massive schemes of online open participation such as Galaxy zoo, e-Bird, Foldit, etc. However, the design and implementation of cases like these it is not an easy task for scientists (*element iii and iv)*. To operate at large scale, open science initiatives require a broad range of issues, ranging from organizational to technological skills that are not commonly part of scientific laboratory[10]. Elements that guarantee the validity of data producers by the public are also a key part of the design (RIN-NESTA) (*element v)*. Initiatives like Galaxy Zoo or e-Bird are similar to what Ismail *et al.* (2014) regards as *exponential organization*. As such, in order to scale up these initiatives required elements of social networks like a sense of community, but also good interfaces, gaming aspects, etc. . Without these elements, participation can easily slip out of the interest of citizens and initiatives become abandoned.

c) One common problem in open source technologies is the coordination of small tasks (elements xi and xii). Open source works through modularity, that is, reducing projects into small task that can be easily completed by volunteers. This is already difficult to solve in open source technologies like software, requiring a strong and clear leadership, some form of hierarchical organization and a lots of motivation (Bonaccorsi and Rossi, 2003) This would as well happen when developing of open source infrastructure for scientific purposes; as it is mentioned in a) above it could work better in certain scientific areas than in others. Another risk in open source projects is forking (or speciation), that is dividing the project into smaller teams due to differences about the direction of development (*element xiii and xiv*). The risk of forking is that open tools could become incompatible with each other and synergies become lost (Weber, 2004) Therefore, when forking occurs, the effort of coordination increases to so much extend that in some cases it becomes unattainable. Finally, open source technologies are helping to create a dedicated infrastructure for open science where wider interaction and cooperation among non-related scientists and the public becomes easier. Some of the aforementioned examples like e-Bird or Galaxy Zoo are indeed global endeavors with hundreds of thousands of collaborations, allowing scientists and citizen scientists around the word to participate in data production and openly accessing the data. However, the global character of these initiatives also creates some risks. One of them is the fact that the server and the database where the information is stored is not always decentralized and scientist around the world have to ask for this data to be sent, especially when asking for sizeable series of data.

[10] Organizational task might include attracting and managing a large amount of volunteers while at the same time getting enough funds to support the initiative. Technological task such as designing a website, selecting collection tools, allowing user participation and generating tools to supervise user’s participation, etc. (see Wiggins and Crowston 2011)

4.2. Analytical framework to anticipate obstacles and barriers when pushing openness above certain thresholds

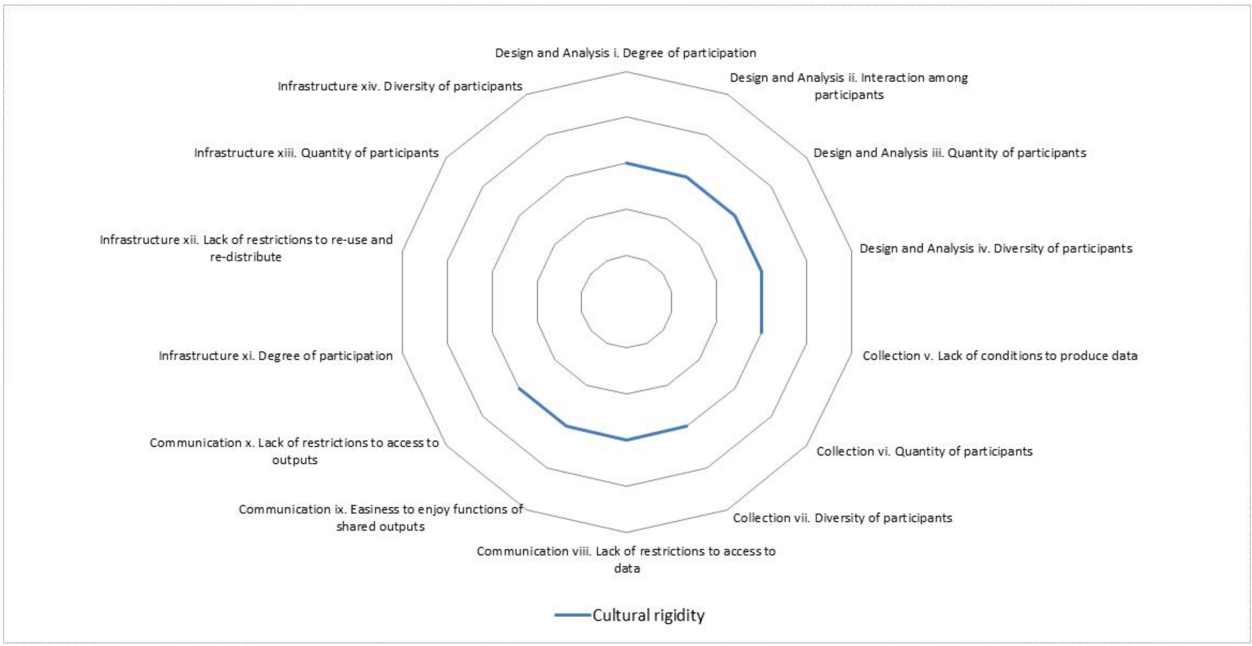
The discussion above suggests that there are three types of barriers operating against *attempts to opening up* different elements of openness. The association between barriers and elements are informed in Table 3.

Table 3: Barriers and obstacles that need to be overcome when attempting to opening up different elements

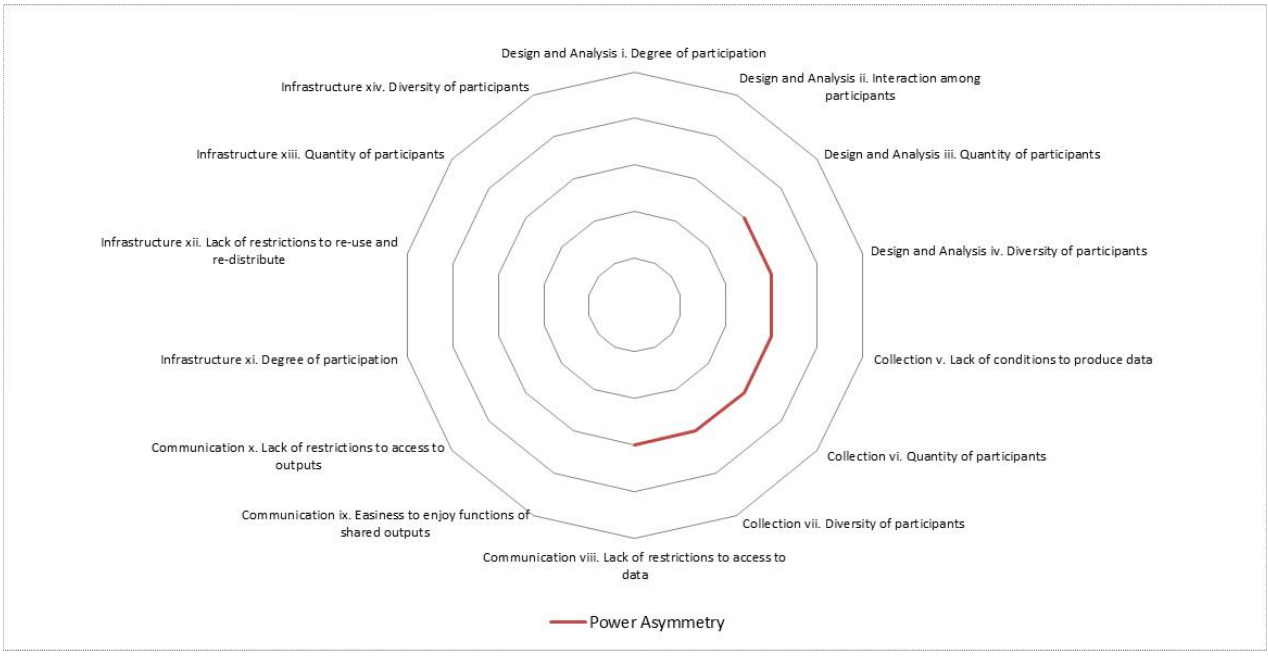
Elements of openness		Obstacles		
Research Stages		Cultural rigidity	Power Asymmetry	Coordination
Design and Analysis (D&A)	i. Degree of participation	x	x	x
	ii. Interaction among participants	x		x
	iii. Quantity of participants	x	x	x
	iv. Diversity of participants	x	x	x
Collection (C)	v. Lack of conditions to produce data	x	x	x
	vi. Quantity of participants		x	x
	vii. Diversity of participants	x	x	x
Communication of science (Comm)	viii. Lack of restrictions to access to data	x	x	
	ix. Easiness to enjoy functions of shared outputs	x		
	x. Lack of restrictions to access to outputs	x		
Infrastructure (I)	xi. Degree of participation			x
	xii. Lack of restrictions to re-use and re-distribute			x
	xiii. Quantity of participants			x
	xiv. Diversity of participants			x

The presented discussion, however, is not very informative on when these barriers start operating. Do they operate in the same way when pushing openness form low (1) or high (5) departing levels? We may assume that the barriers are only raised at medium theoretical (3) level; we may assume that they operate always; or we may assume they operate when pushing further a relevant empirical threshold. Graphs 3 draw the first of these options. These graphical representations allow us to identify the obstacles that have to be overcome by different experiences when opening up different elements of the research process.

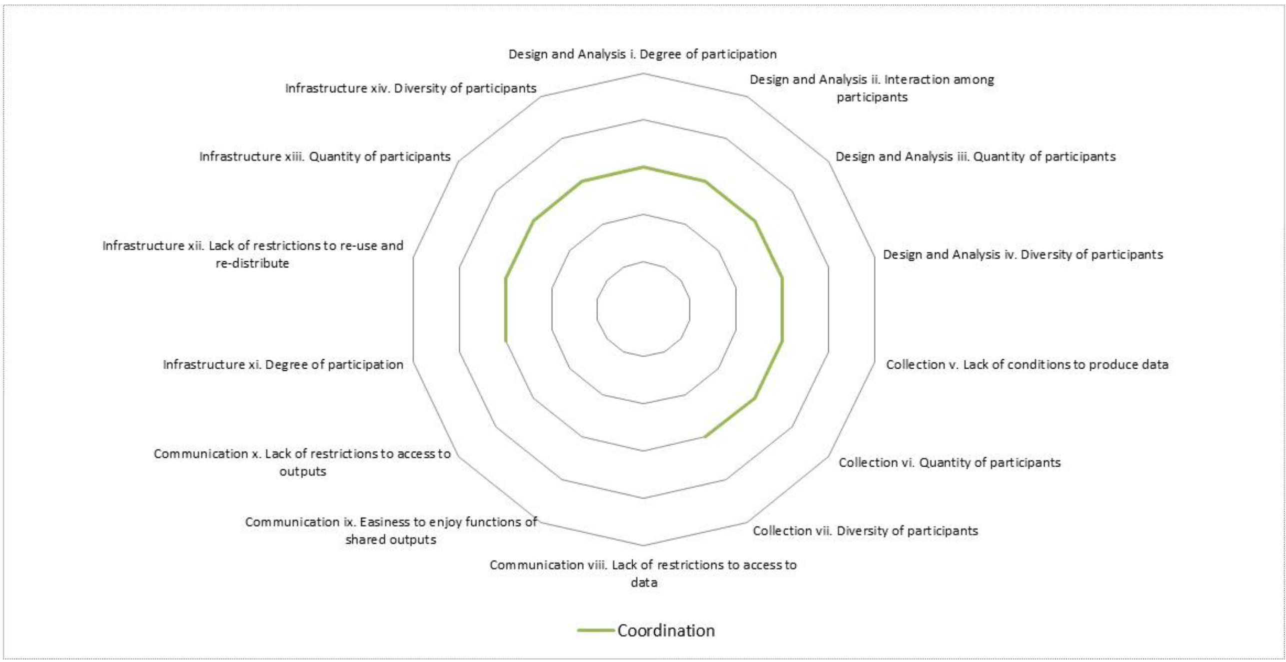
Graph 3a: Cultural rigidity as a barrier to opening up some elements of openness



Graph 3b: Power asymmetry as a barrier to opening up some elements of openness



Graph 3c: Coordination problems as barriers to opening up some elements of openness

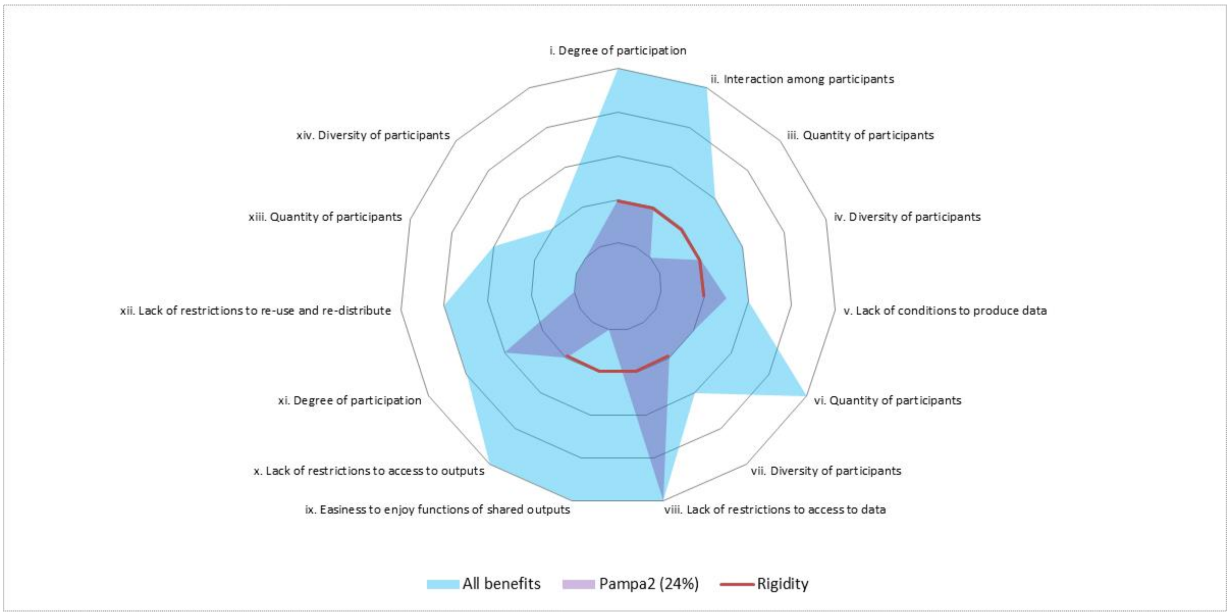


4.3. Illustrative example using the analytical framework of openness and obstacles for Pampa2

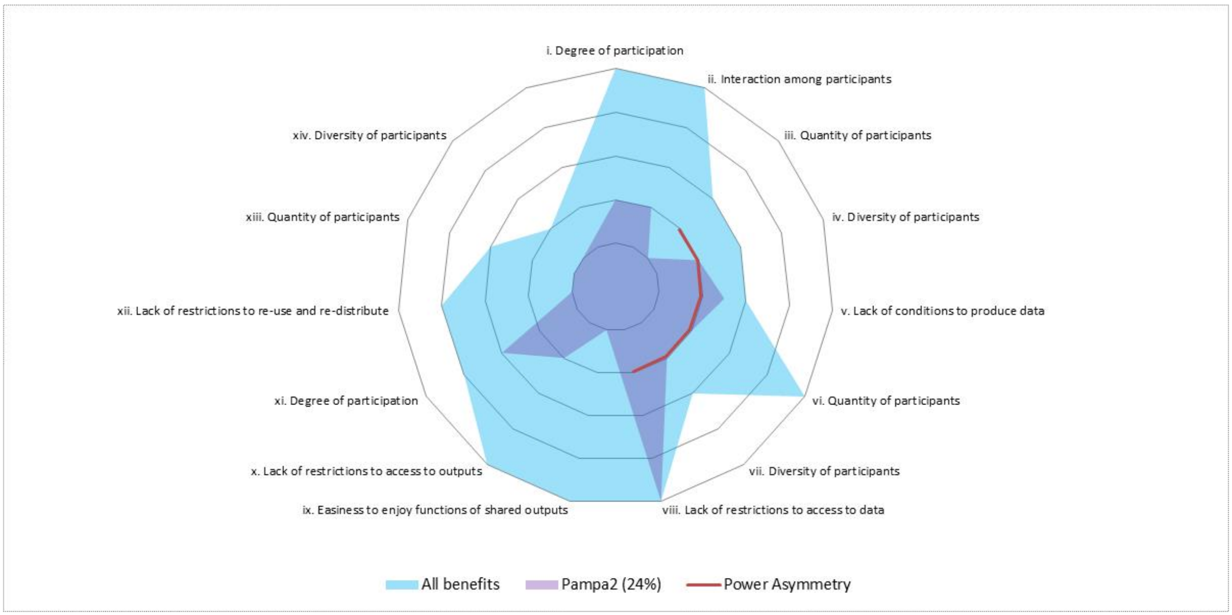
To show an example of how this graphical representation can be used, let's go back to the Pampa2 project. Each Graph 4 represents three series of data: i) the actual level of openness in the fourteen elements by the initiative Pampa2 (in violet); ii) the needed openness to achieve all potential benefits of open science coming from Graphs 2 taken all together (in light blue); and iii) in turn, the three barriers: cultural rigidity (Graph 4a), power asymmetry (Graph 4b) and coordination issues (Graph 4c), that need to be overcome when pushing openness forward (in red). The barriers were graphed using an empirical threshold calculated as the average level of openness of this initiative in all fourteen elements of openness.

As can be seen, the larger gap between actual openness and the one that renders all potential benefits together is explained by shortages in opening up elements i to iii, ix and xi to xiv. The largest gaps fall in the Infrastructure, Communication and Design and Analysis stages (in that order). To open-up these elements further, the binding constraints that seem to be operating are cultural rigidity and coordination issues. In fact, there are not very important elements of power asymmetries or power disputes that were revealed as important in our interviews. 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). Also, there were coordination issues that come up since the very beginning of the project, for example regarding how to deal with broken buoys, where to get money from maintenance, who pay for it, etc. 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 out of reach/vision of the involved researchers, but they may start looking beyond, as these practices become better known, diffused and accepted.

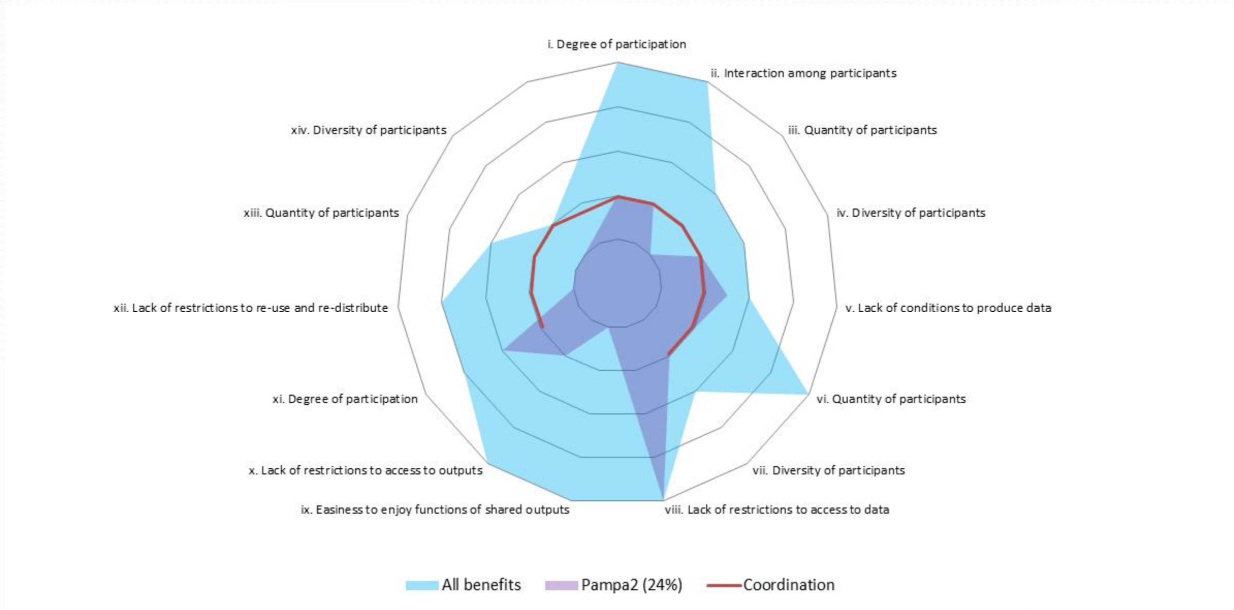
Graph 4a: Cultural rigidity as a barrier to opening up some elements of openness, actual openness of Pampa2 and ideal level of openness to target all benefits together



Graph 4b: Power asymmetry as a barrier to opening up some elements of openness, actual openness of Pampa2 and ideal level of openness to target all benefits together



Graph 4c: Coordination problems as barriers to opening up some elements of openness, actual openness of Pampa2 and ideal level of openness to target all benefits together



5. Conclusions

Doing 'open science' is producing scientific knowledge collaboratively while sharing openly the outcomes of that process. There many different examples of open science: mathematicians around the world solve complex problems using web platforms; amateur astronomers interact with scientist classifying galaxies; citizens collect data to be used by scientist or other actors; open access repositories gather data and other scientific outputs within a knowledge fields, etc. These examples (and others) involve different elements on *what* is being opened, *under what conditions or how*, and *who* can enjoy the benefits of openness.

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 aspect of open science practices that is being opened (or attempted to be opened). We build an analytical framework that characterises the level of openness of particular initiatives of open science in a way that can be compared with 'ideal' levels of openness in different dimensions so as to achieve the potential benefits to the most. The framework also indicates what specific obstacles or barriers will need to be overcome when trying to open-up specific elements of the practice of open science.

The analytical framework is illustrated using data of one open science project from Argentina that gathers and analyses limnology data of The Pampa's lagoons. Limnology data in that project are used to anticipate the effect of anthropogenic or climate change. According to our analytical framework, there is a wide space for improving openness that would be highly beneficial in terms of empowerment of affected population in the surroundings of the analysed lagoons and also in terms of democratization and potential innovations. However, to opening up the relevant elements (for example, allowing higher participation of the affected population in the research process or lifting restrictions to access to raw data) the project will need to sort out some coordination issues and overcome institutional and cultural rigidity.

We believe our analytical framework could be informative for researchers, policy makers and practitioners. It works as a guide for characterising open science experiences. It identifies the specific elements of open science practices that need to be opened-up further so as to close the gaps between current situation and potential benefits. It also anticipates the specific barriers that need to be 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:

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/:

Bonaccorsi, A. and Rossi, C., (2003). 'Why Open Source Software Can Succeed', *Research policy*, 32(7), pp. 1243-58.

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

Catlin-Groves, C. L., (2012). 'The Citizen Science Landscape: From Volunteers to Citizen Sensors and Beyond', *International Journal of Zoology*, 2012(pp. 1–14.

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.

Etzkowitz, H., (1998). 'The Norms of Entrepreneurial Science: Cognitive Effects of the New University-Industry Linkages', *Research policy*, 27(8), pp. 823-33.

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

Grand, A., et al., (2012). 'Open Science: A New "Trust Technology"? ', *Science Communication*, 34(5), pp. 679–89.

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.

Ismail, S., Malone, M., van Geest, Y. and Diamandis, P., (2014). 'Exponential Organizations: Why New Organizations Are Ten Times Better, Faster and Cheaper Than Yours (and What to Do About It)', Diversion Books New York, NY.

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

Kelty, C., Faubion, J. D. and Marcus, G. E., (2009). 'Collaboration, Coordination, and Composition: Fieldwork after the Internet', *Fieldwork Is Not What It Used to Be*, pp. 184-206.

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.
Tacke, O., (2010). 'Open Science 2.0: How Research and Education Can Benefit from Open Innovation and Web 2.0. ', in T. J. Bastiaens, U. Baumöl and B. J. Krämer (ed.), *On Collective Intelligence,* Berlin: Springer Berlin Heidelberg.
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'.
Vision, T. J., (2010). 'Open Data and the Social Contract of Scientific Publishing', *BioScience*, 60(5), pp. 330-31.
Wagner, C. S., (2009). *The New Invisible College: Science for Development*, Brookings Institution Press.
Weber, S., (2004). *The Success of Open Source*, Cambridge Univ Press.
Wiggins, A. and Crowston, K., (2011). 'From Conservation to Crowdsourcing: A Typology of Citizen Science', *Paper presented at System Sciences (HICSS), ,* Hawai:
Woelfle, M., Olliaro, P. and Todd, M. H., (2011). 'Open Science Is a Research Accelerator', *Nature chemistry*, 3(10), pp. 745-48.
World Bank, (2015). 'Open Data for Sustainable Development', *Public Note ICT 001*.