COLLABORATION NETWORKS IN *BIG SCIENCE*: THE *ATLAS* EXPERIMENT AT *CERN*

Redes de colaboración en *big science*: el experimento *ATLAS* en el *CERN*

Agustí Canals, Eva Ortoll and Markus Nordberg

Agustí Canals is an associate professor at the *Universitat Oberta de Catalunya* in Barcelona, where he also leads the *KIMO Research Group on Knowledge and Information Management in Organizations*. He also teaches Strategy at the *Esade Business School* (Barcelona) and serves as senior research fellow at the *I-Space Institute* (Philadelphia). He received his MBA and his PhD in Management Sciences from *Esade Business School* and a MSc in Physics from *Universitat Autònoma de Barcelona*. He has also held visiting research positions at the *Warwick Business School* (*University of Warwick*), at the *Sol C. Snider Entrepreneurial Research Center* (*Wharton School*, *University of Pennsylvania*) and at the *Saïd Business School* (*University of Oxford*). His research and professional interests are in the areas of information management, strategic knowledge management, innovation, social networks, and complexity.
http://kimo.uoc.edu
http://orcid.org/0000-0002-1700-3983

Universitat Oberta de Catalunya, *Information and Communication Sciences Department*, *KIMO Research Group*
Av. Tibidabo, 39-43. 08035 Barcelona, Spain
acanalsp@uoc.edu

Eva Ortoll is an associate professor of the *Information and Communication Sciences Department* at the *Universitat Oberta de Catalunya* (*UOC*). She is also a researcher of the *KIMO Research Group* (*Knowledge and Information Management in Organizations*). She holds a PhD in Information and Documentation Systems from the *Universidad de Zaragoza*. BSc in Documentation from *Universidad Carlos III de Madrid* and a graduate degree in Librarianship from *Universitat de Barcelona*. Her research interests are: competitive intelligence, collaboration networks and information transfer, user behavior and knowledge management. She has participated in several competitive research projects. In 2010 she was a visiting researcher at the *Loughborough University* (UK).
http://kimo.uoc.edu
http://orcid.org/0000-0002-1698-1347

Universitat Oberta de Catalunya, *Information and Communication Sciences Department*, *KIMO Research Group*
Av. Tibidabo, 39-43. 08035 Barcelona, Spain
eortoll@uoc.edu

Markus Nordberg is the head of *Resources Development* of the *Development and Innovation Unit* (*IPT-DI*) at *CERN*, Switzerland. He is currently involved in launching a sensor and imaging R&D initiative called *Attract* (*www.attract-eu.org*) aiming at both scientific and societal impact. He also manages the new *IdeaSquare* initiative at *CERN* (*cern.ch/Ideasquare*) that hosts detector R&D and society-driven MSc-student projects. Prior to this function, he served 12 years as the resources coordinator of the *ATLAS* project at *CERN* (*www.atlas.ch*). He has also served as visiting senior research fellow at the *Centrum voor Bedrijfseconomie*, *Faculty ESP-Solvay Business School*, *University of Brussels*, and as a member of the *Academy of Management, Strategic Management Society* and the *Association of Finnish Parliament Members and Scientists* (*Tutkas*). He has a degree both in Physics and in Business Administration.
https://nordberg.web.CERN.ch/nordberg
http://orcid.org/0000-0001-5187-9340

CERN
CH-1211 Geneva 23, Switzerland
markus.nordberg@CERN.ch

Manuscript received on 10-9-2016
Accepted on 3-7-2017

Abstract

Nowadays big scientific experiments require large organizations and hundreds of researchers who participate from several institutions. An interesting, yet rarely studied aspect of this new kind of scientific enterprise is the internal collaboration between the members of the participating institutions. Here we assess this matter in one of the most well-known examples of big science: the *ATLAS* experiment at *CERN*. Applying different network analysis techniques to data from internal *CERN* databases, we have identified several collaboration patterns in the experiment. We observe, on the one hand, the high level of collaboration between the institutions represented in *ATLAS*, higher than the average in the field of physics, and we identify the key institutions in the collaboration network. On the other hand, we notice that the collaboration network does not follow a scale-free or power-law model, contrary to what happens in other studied collaboration networks in physics and other areas. Finally, we observe that geographic distance between two institutions does not seem to affect the probability of establishing collaboration relationships, in contrast also to what happens in other kinds of collaboration networks.

Keywords

Scientific research; Scientific collaboration; Big science; High-energy physics; Co-authorship networks.

Resumen

Actualmente los grandes experimentos científicos necesitan una gran organización en la que participan cientos de investigadores de varias instituciones. Un aspecto interesante, aunque poco estudiado, de este nuevo tipo de iniciativas científicas es la forma que toma la colaboración interna entre los miembros de las instituciones participantes. En este trabajo se investiga este tema en uno de los ejemplos más conocidos de *big science*: el experimento *ATLAS* del *CERN*. Mediante la aplicación de varias técnicas de análisis de redes a información proveniente de bases de datos internas del *CERN*, se identifican varios patrones de colaboración dentro del experimento. Se comprueba, por una parte, el alto nivel de colaboración entre las instituciones representadas en *ATLAS*, superior al observado usualmente en el campo de la física, y se identifican los actores principales en la red de colaboración. Por otra parte, se observa que la red de colaboración no sigue un patrón "libre de escala" o de "ley de potencia", al contrario de lo que ocurre en otras redes de colaboración estudiadas en física y en otros campos. Finalmente, se advierte que la distancia geográfica entre dos instituciones no parece tener ningún efecto en la probabilidad de establecer relaciones de colaboración, contrastando también con lo que ocurre en otros tipos de redes de colaboración.

Palabras clave

Investigación científica; Colaboración científica; Grandes experimentos científicos; Física de altas energías; Redes de coautoría.

Canals, Agustí; **Ortoll, Eva**; **Nordberg, Markus** (2017). "Collaboration networks in *big science*: The *ATLAS* experiment at *CERN*". *El profesional de la información*, v. 26, n. 5, pp. 961-971.

https://doi.org/10.3145/epi.2017.sep.17

1. Introduction

One of the rapidly changing aspects of scientific activity today is the dynamic of collaboration, both between individual scientists and between institutions or countries. In recent decades, the frequency and importance of collaboration initiatives in most scientific disciplines has experienced a large increase (**Sonnenwald**, 2007).

The growth of scientific collaboration becomes more evident in big scientific experiments, or in so-called *big science* projects, where typically hundreds, or even thousands, of researchers coming from many institutions and countries join forces to tackle big scientific questions. One can find examples of *big science* in diverse disciplines such as genomics, astronomy, nuclear fusion, and high-energy physics. In these fields, a large part of the budget allocated to scientific research comes from governmental or international funding agencies.

Because of their increasing number and importance, it is highly relevant to understand the structure and dynamics of these collaborations. Current published research recognizes that they vary depending on the discipline (**Chompalov**; **Genuth**; **Shrum**, 2002), but one can hardly find in-depth case studies of modern *big science* experiments that examine internal organization and features.

One of the interesting aspects of many big scientific experiments is that their findings tend to be attributed to the collaboration as a whole and not to individual researchers because of their publication policies. This offers only a partial view into the internal dynamics of collaboration and is precisely one of the main difficulties when it comes to understanding how big scientific collaborations work. The fact that the most important discoveries are published in articles that are signed by all members of the collaboration (Figure 1) makes it very difficult to ascertain who has taken part in each of the multiple sub-projects in which the global project is divided. It is impossible to know, by looking at the publications produced, what specific institutions or scientists have established a direct collaboration within the experiment. Thus, the use of established methodologies employed to study scientific collaborations, such as applying social network analysis techniques to identify co-authorship networks based on data originating from published journals' databases, is rather ineffective. Since articles are signed by everyone in the collaboration, co-authorship relations measured in the usual way do not represent real collaboration in actual research tasks carried out together. Applying the usual procedures, like the construction of co-authorship networks based on articles retrieved from scientific publication databases such as *Web of Science* or *Scopus,* do not tell us much about real collaboration. Indeed, it would lead to the wrong conclusion that every scientist and institution in *big science* experiments directly collaborate and work together, because they appear as joint authors in a large number of articles.

Figure 1. This article is signed by 5,154 authors from 344 institutions. The text occupies 9 pages and the signatures 24 pages.

2. Objectives

In this paper we aim to solve, to some extent, the aforementioned problem by using an alternative source of information about the inner scientific collaboration structures within one of the most important *big science* experiments: the *ATLAS* experiment at *CERN*. This source of information is the *ATLAS* internal publication database, which allows us to find out about the actual scientists involved in the research that led to each published paper. Thanks to this feature, we will investigate:

- the patterns of collaboration between the institutes and universities that participated in *ATLAS* and whether the main features of the collaboration differ from the patterns observed in smaller scale projects;
- the structure of their internal collaboration network;
- the most prominent actors in that collaboration network;
- the influence of geographical proximity of the institutes' headquarters on the probability of collaboration between them.

3. Scientific collaboration

3.1. Scientific collaboration in large experiments

Several studies conducted in the last several years suggest that research publications made by scientific collaborations have an increasing impact (**Benavent-Pérez** *et al.*, 2012). This effect appears to be even stronger in the case of international collaborations that involve institutions coming from various countries (**Kronegger**; **Ferligoj**; **Doreian**, 2011). It is, therefore, to be expected that an important part of public research funds goes to projects involving collaborations between different groups of researchers. This fact becomes even more accentuated in the calls for EU programs, where projects are seldom admitted unless the collaboration of scientists and institutions are from several member states.

This growth of scientific collaborations has gone hand in hand with the proliferation of the so-called big scientific experiments or *big science*. In different areas (genomics, high-energy physics, climate sciences, ecology, astronomy, nuclear fusion, etc.), scientific research has moved its center of gravity in the last decades from small or mid-size experiments to large and complex collaborations (**Galison**; **Hevly**, 1992).

The concept of *big science* is not new. It was proposed in the 1960s by **Weinberg** (1961) and **Price** (1963) after the big research projects that took place in the US during the Second World War, like the famous Manhattan project which resulted in the development of the first atomic bomb. However, in recent decades the term *big science* has increased its popularity with the emergence of more and more projects that fit with this idea (**Etzkowitz**; **Kemelgor**, 1998; **Hicks**; **Katz**, 1996; **Knorr-Cetina**, 1999). Thus, both academics and research funding agencies have an increased interest in better understanding the structures and characteristics of *big science* collaborations.

Scientific collaborations can be studied from different points of view, depending on the unit of analysis. For instance, we can look at collaboration between individuals, between departments or laboratories, between institutions (universities, research institutes, hospitals, etc.) or between countries (**Sonnenwald**, 2007). In the case of *big science* projects, the most relevant level is the institutional. Duration of this type of experiment ranges from years to decades. The complexity and variety of the required tasks suggest, as a consequence, that the individual researcher turnover is quite high. Institutions, however, usually maintain their participation in the experiments throughout their life cycle because they have made large capital investments. Therefore, we will base our analysis in the study of inter-institutional collaboration within big scientific experiments. Of course, it is always possible to aggregate institutional data to obtain meaningful results at the regional or country level.

The study of scientific collaboration also requires some level of contextualization. In order to understand the established collaborative dynamics, one has to take into account the discipline (or disciplines) involved, the geographic area where the project takes place, the kind of research performed, and other factors (**Gazni**; **Sugimoto**; **Didegah**, 2012; **Ortoll** *et al.*, 2014). The contexts where *big science* experiments are performed are very diverse, but there are some common characteristics between them. On the one hand, the size of the collaboration measured, in terms of number and variety of participant researchers, may influence the way in which people work in the collaboration, both at organizational level (**Santalainen** *et al.*, 2011) and at the level of the interaction of the individual researcher with the group (**Bressan**; **Boisot**, 2011; **Creus**; **Canals**, 2014). On the other hand, it is necessary to consider the geographic location of the experiment. In many cases, due to its special characteristics, the experiment must be situated in a specific place. This is the case, for instance, of big telescopes or high-energy physics detectors, which rely on supporting infrastructure. In other instances, like in many genomics-related projects, the experiments are distributed among several institutions taking part. What seems to be a common trend is that, once the data have been obtained, they can be analyzed in several locations in a distributed way thanks to the current possibilities of information and communication technologies. Thus, it is not necessary for all participants to be at the experiment site at all times. Instead, they can often work from their institutions of origin.

3.2. Collaboration in physics

In order to correctly interpret the results, it is necessary to take into account the characteristics of collaboration in current physics experiments. Thus, we will first review the relevant literature in the domain of physics.

Recently published studies in the field of physics suggest a high degree of collaboration: around 50% of the physics research findings come from inter-institutional collaboration and, from them, approximately 30% come from international collaborations (**Benavent-Pérez** *et al.*, 2012). In particular, physics shows a higher level of international collaboration than in other fields. Also, geographical distance turns out to be quite relevant when it comes to establishing collaborations: unsurprisingly, the frequency of collaborations between close institutions is clearly higher that between those that are farther away from each other.

In one of the few studies based on longitudinal data, i.e., corresponding to different points in time, **Lorigo** and **Pellacini** (2007) observe:

- an increase in the number of inter-institutional collaborations,
- an increase in the intensity of inter-institutional collaborations, measured in terms of the number of publications produced,
- an increase in the percentage of nodes of the co-authorship network of the field belonging to the principal component, which suggests that the density of collaborations between the institutions in the field is growing, and
- a slight loss of the centrality of *CERN* as an institutional node, implying a reduction in need for institutions to act as hubs for international collaboration initiatives.

The most recent successes of the *CERN*-based experiments may cast some doubt on the last point. In any case, as **Huang**, **Tang,** and **Chen** (2011) suggest, there is a need for an in-depth study about collaboration networks formed around institutions like *CERN* in order to better understand the collaboration patterns in the field of physics. Our aim is to contribute to this idea by studying the *ATLAS* collaboration network.

4. Research setting: The *ATLAS* experiment

ATLAS is one of the four experiments associated with the LHC (Large Hadron Collider) at *CERN*, which became operational in 2008. The LHC is a circular accelerator housed inside a 27 km long underground tunnel that straddles the Franco-Swiss border near Geneva, Switzerland. Inside the LHC, beams of protons (or alternatively heavy nuclei) travel in both directions at speeds very close to the speed of light. At four specific points along the ring these beams collide at very high energies, designed to reach 14 TeV[1]. This makes it possible to reproduce conditions of our universe just after the Big Bang. The aim of the LHC experiments has been to test the present Standard Model of elementary particles as well as other theories that go beyond it. The first important discovery of the LHC experiments was the Higgs boson, announced in 2012. Currently, there is an aim to solve even more difficult challenges, like trying to find answers to puzzling questions such as, "What is dark matter and dark energy?" Dark matter and dark energy are believed to exist in the universe, but somehow have remained undetected.

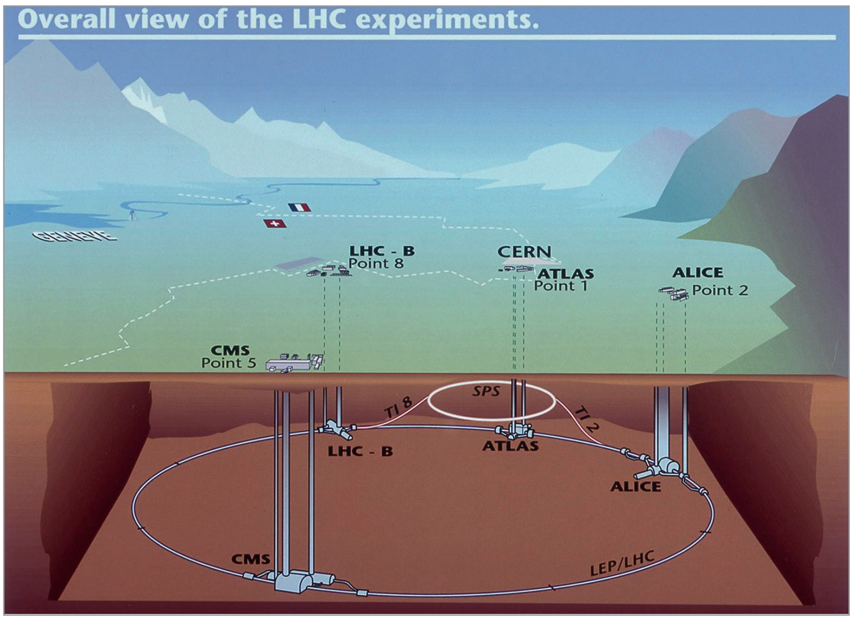


Figure 2. The Large Hadron Collider (LHC) at *CERN*. (Source: *CERN*)

For that it is necessary to register what happens at the collision points. Four detectors built for that purpose take care of it: *ATLAS* and its "siblings" CMS, *Alice* and LHCb (Figure 2). The *ATLAS* (A Toroidal LHC ApparatuS) detector is situated in a point of LHC in French territory, some 80 meters underground, and consists of a complex instrument capable of measuring the trajectories and energies of the particles produced from the proton collisions by the LHC. *ATLAS* has the approximate shape of a cylinder having the LHC beam pipe as its axis, measuring 25 meters in diameter and 45 meters in length and weighing about 7000 tons (Figure 3). It is arguably one of the most complex instruments ever built by the human being, capable of gathering data corresponding to the 40 million collisions produced every second. Even after discarding those events without scientific interest, 100 Mbytes of data are stored per second. The later analysis of those data permits identifying the production of new particles, and, therefore, verifying the proposals put forward by theoretical physicists.

a)

b)

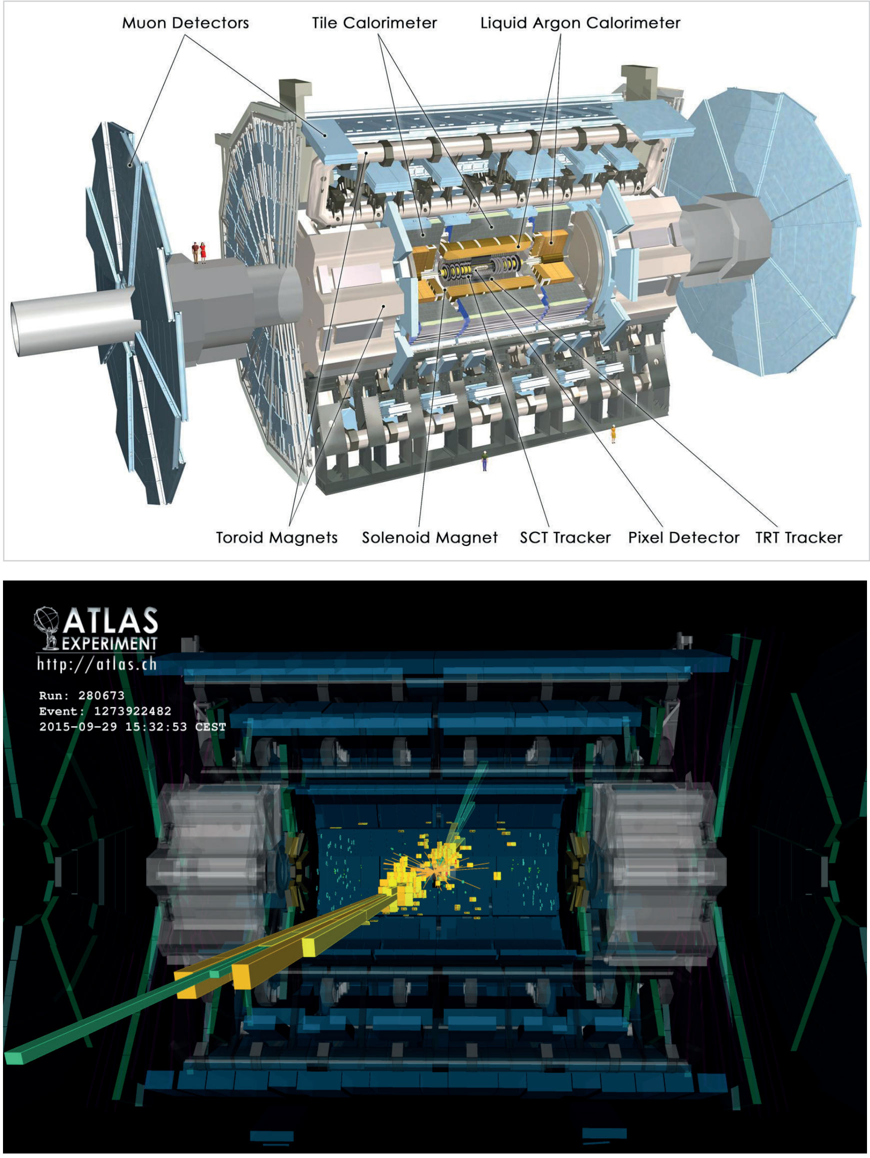


Figure 3. The *ATLAS* detector (a) and the reproduction of one of the events produced in it (b). (Source: *ATLAS Collaboration*)

With the objective of building and operating the *ATLAS* detector, in 1995 the *ATLAS Collaboration* was created. Today, it is comprised of more than 5,000 researchers, most of them physicists, but with strong participation also from engineers and computer scientists from about 180 institutes and universities in 38 countries all over the world, representing four continents (Figure 3). The collaboration adopts a peculiar and unusual organization structure, with a very low degree of hierarchy, where meritocracy rules and decisions are usually made by consensus according to scientific and rational criteria (**Santalainen** *et al.*, 2011).

5. Methodology

As stated before, we were granted access to an internal database used by the collaboration where all the generated physics preprints are registered. Once discussed and completed, those preprints are submitted for publication and, in due time, become published articles in refereed journals. In the above-mentioned internal database one can find useful information about each paper such as the date of collected data for analysis, the type of physics process analyzed or the names of the 'editors'. These 'editors' are the scientists in charge of writing the initial draft of the preprint or reviewing it thoroughly, because they have either led the analysis teams or have been directly involved in that specific analysis process. Of course, there are other scientists involved in the research leading to the publication apart from the editors, but at least one editor is always representing each one of the main institutions involved. Thus, by knowing the names of the editors that collaborate on the papers and which institution each *ATLAS* member belongs to, it is possible to obtain an accurate map of the structure of the internal collaboration within *ATLAS* and between the institutions that form it. Table 1 describes the fields in the database that we used for this study. Notice that there is one register in the database for each combination of Publication.ID and Editor, since each publication may have (and usually has) more than one editor.

The analyzed preprint data corresponds to documents initiated during the period between 2008 (including results coming from detector tests using cosmic rays data before the

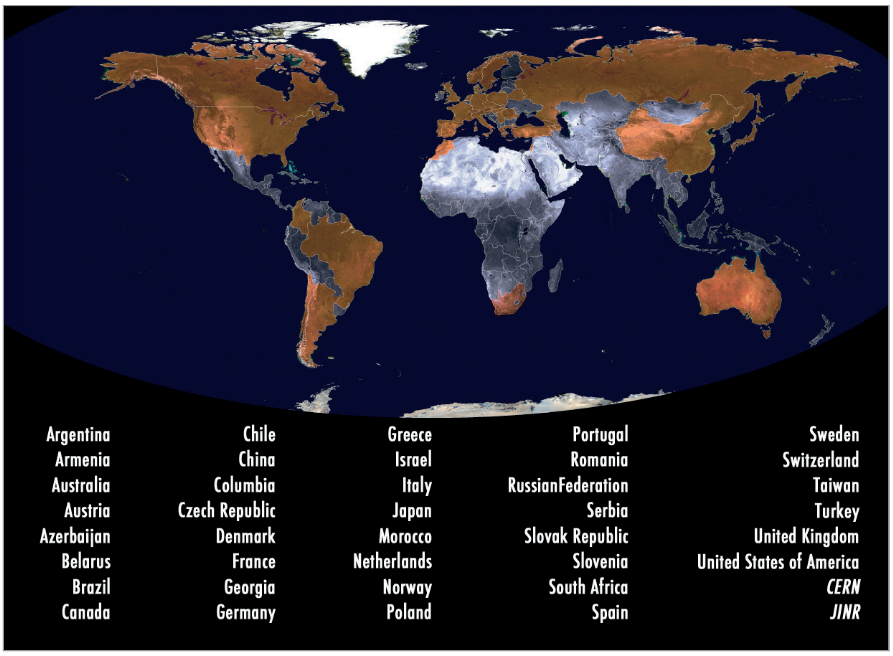


Figure 4. Member countries of the *ATLAS Collaboration.* (Source: *ATLAS Collaboration*)

LHC was commissioned) and the 31st of December of 2013. These data were gathered between 2013 and 2015 and contain most of the analysis performed on *ATLAS* data before the two-year shutdown of the collider in January of 2013. We analyzed a total amount of 508 preprints produced by 3,093 authors from 167 of the 217 institutes or sub-institutes involved in the *ATLAS* Collaboration. After we had performed our analysis, we contrasted the results obtained and their interpretation with several *ATLAS* researchers.

We should stress the fact that when the studied preprints are published as journal articles, they are signed by the whole *ATLAS* Collaboration. There is no way, thus, to ascertain from the published articles the internal collaboration patterns within the experiment. Although this practice may seem unethical in some research fields, in experimental high-energy physics it is the standard way to proceed. Physicists believe that all members in a collaboration deserve the credit for any discovery.

Table 1. Database structure

Database field	Description
Publication.ID	Internal identification number of the paper
Editor	Name of editor
Title	Title of the paper
Creation.Date	Time of creation of the paper
Lead.Group	Work group where the publication has been created
Num.Editors	Number of editors of the paper
Institution	Research institute or university to which the editor belongs
Institution. Country	Country where the institution is based

For the manipulation of the database we used the *R* statistics software. The co-authorship network was generated with the *Igraph* package for *R*. The networks measures and the algorithm applied for the detection of communities were calculated with *Igraph* and *Mathematica*™, and these tools were also used for the generation of the visualization images.

6. Results

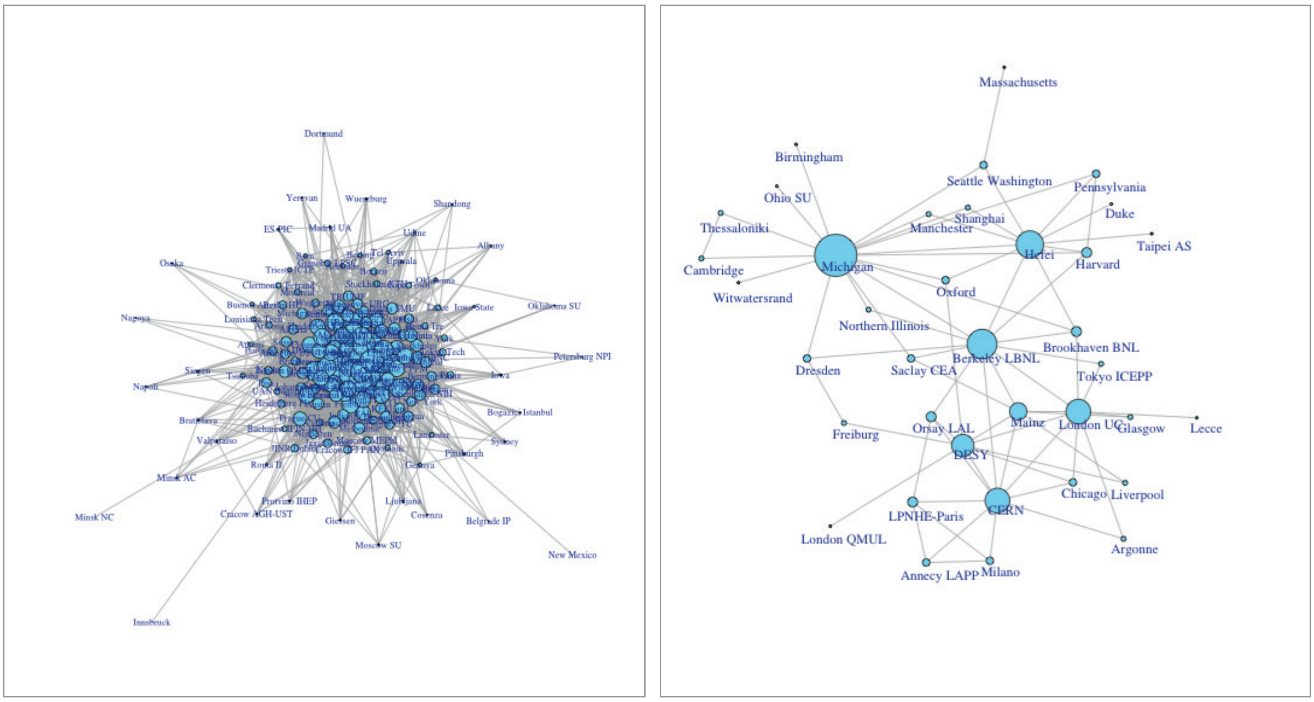
6.1. Co-authorship network

The co-authorship network resulting from the analyzed set of preprints is shown in Figure 5. In a) we represent the whole co-authorship network. In it, nodes represent institutions and links represent collaboration between them. Those links have a weight corresponding to the number of papers co-authored by the two institutions linked. The size of the node is proportional to the number of papers in which the institution has taken part. In order to facilitate the interpretation of the visualization, in b) we represent a subgraph of the whole network where we have pruned the institutions with less than 20 collaborating peer institutions.

The whole network contains 167 vertices in total, corresponding to the different institutes of researchers' origins, and 3470 links that represent collaborations between members of the institutes linked. The network is completely connected: there are not separate components. Network density is quite high (ρ = 0.25) for a network of this size. The mean degree $\langle k \rangle$, which represents the average number of collaborations per institute, *is* 41.56. This is also very large. We observe, thus, that there exists a real, high level of internal collaboration in *ATLAS*.

The observed clustering coefficient is also quite high (C = 0.64) compared to a random network of the same size. This indicates a high level of transitivity in the collaboration. That is, if two institutions are collaborating at the same time with a third one, it is likely that they are also collaborating with each other. The mean distance (average of the shortest path between all pairs of nodes in the network) is quite short (ℓ = 1.82), which is obviously much lower than the famous "six degrees of separation" in the well-known study by **Milgram** (1967). Something similar happens with the maximum diameter, which is relatively low (d = 4). This combination of short average distance and large transitivity allows us to qualify the *ATLAS* collaboration network as a *small world network* (**Watts**; **Strogatz**, 1998).

However, in contrast to most networks (**Dorogovtsev**; **Mendes**, 2003), the *ATLAS* collaboration network is not *scale-free*, i.e. its degree distribution does not follow a power-law. This differentiates it from the majority of collaboration networks studied through the analysis of co-authorship (**Barabási**;



a) Whole network

b) Simplified network

Figure 5. Collaboration network between the *ATLAS* experiment institutes.

Albert, 1999; **Newman**, 2001; **Redner**, 1998). Indeed, after analyzing the accumulated degree distribution

$$P_k = \sum_{k'=k}^{\infty} p_{k'}$$

following the method proposed by **Newman** (2010, p. 250), we observe that it clearly does not adjust to a power law. In Figure 6 we observe that the graphical representation of the distribution in a bilogarithmic plot adopts a shape that clearly deviates from a straight line. However, these findings should be taken with care, since they come from the analysis of a very specific and singular case: a high-energy physics *big science* experiment. Therefore, they may serve as one instance of what can happen in big science collaborations, but they may not be generalized.

6.2. Community analysis

In order to probe more deeply into the collaboration network structure, it is interesting to identify the clustered *communities* that emerge, i.e., the groups of vertices that show a higher degree of interaction between each other. In the case of *ATLAS*, and after testing various methodologies, we adopt the algorithm that maximizes the *Q* measure of modularity proposed by Newman and Girvan (**Clauset**; **Newman**; **Moore**, 2004; **Newman**; **Girvan**, 2004)

$$Q = \frac{1}{2m}\sum_{ij}\left(A_{ij} - \frac{k_i k_j}{2m}\right)\delta(c_i, c_j)$$

where A_{ij} is the adjacency matrix, k_i is the degree of vertex i, c_i is the class to which vertex has been assigned and m is the total number of links in the network.

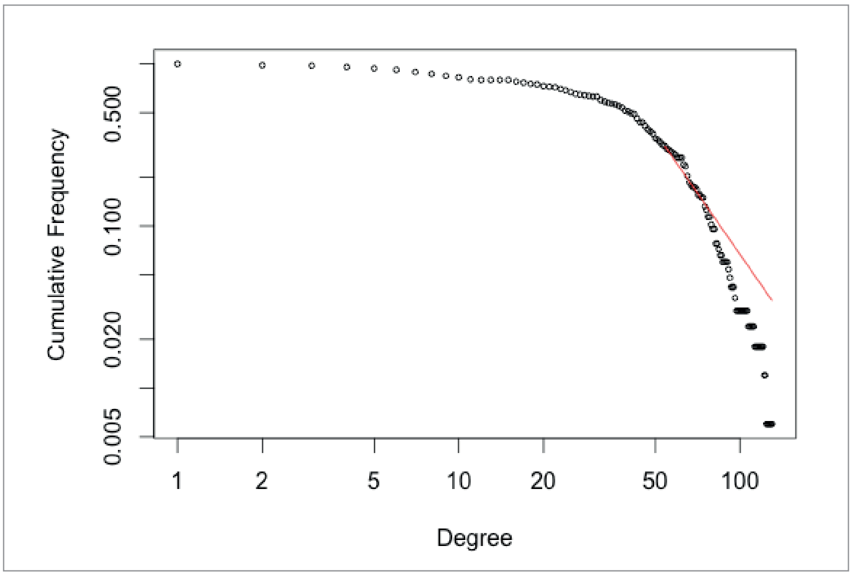


Figure 6. Cumulative degree distribution of the collaboration network between the institutes participating in the *ATLAS* experiment. The red line in the figure represents the best possible attempt to adjust the distribution to a power law.

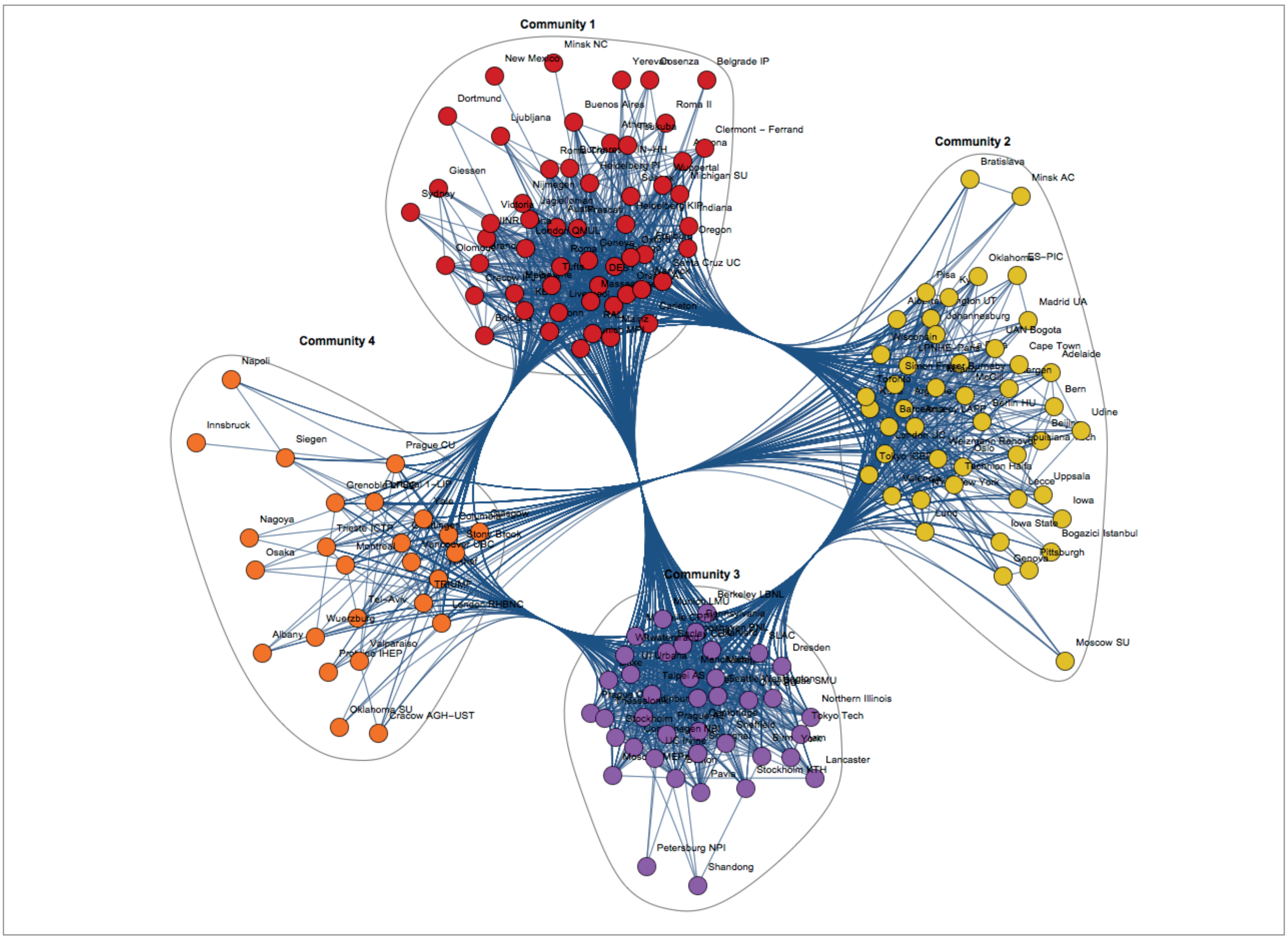


Figure 7. Emerging communities in the *ATLAS* collaboration network

The application of this method detects four separate communities (Figure 7). They are relatively equilibrated in terms of their number of members (55, 46, 40 and 26, respectively). After inspecting their composition, one cannot signal in any case geographical proximity as a factor to explain the aggregation since each community groups research institutes of many parts of the world. Table 2 shows the composition of each community.

7. Discussion

We extract results from the analysis of the *ATLAS* collaboration network, which questions some of the trends observed in collaborations related to smaller experiments or projects. Due to the collective signature policy in large scientific experiments such as *ATLAS*, the details of collaboration between scientists and institutions are not registered in published articles' databases. Since all participants sign all major articles regardless of whether they have directly contributed to the findings described in each paper or not, published journals do not contain information about that aspect. As a consequence, this type of project may be overlooked when studying scientific collaborations, despite the fact that the number of researchers involved and amount of funds deployed are extraordinarily relevant. Our work intends to contribute knowledge and avoid this bias, examining some of the particularities of collaboration in these kinds of scientific endeavors.

First, it is interesting to observe that the level of collaboration within the experiment was extremely high. Of course, this is not a surprising result since collaboration is the raison d'être of *big science* projects. However, both the average degree of the nodes in the co-authorship network and the relative size of its principal component indicate that the degree of collaboration within the experiment was much higher than outside of it in the same discipline, while a comparable clustering coefficient reflects a similar level of transitivity (**Newman**, 2001). This confirms the expectation that these kinds of collaborations constitute a mean to pull together the diverse knowledge and experience from scientists and institutions from all over the world to fulfil the set of objectives. Relatedly, they also offer a way to join forces in order to acquire influence and get funding from public and private sources.

Second, the *ATLAS* co-authorship network we have studied does not show a power-law shaped degree distribution, in contrast to what has been found in scientific collaborations in general (**Barabási**; **Albert**, 1999; **Redner**, 1998). Neither does it follow a power law with exponential cutoff found in other studies (**Newman**, 2001). The absence of a long tail signature of power law distributions indicates that the network did not adopt the characteristic topology of a 'hub and spoke'. This may be related to time constraints in the development of the project, the diversity of the level of specialization required, and limitations concerning the size of

Table 2. Composition of each of the communities in the *ATLAS* collaboration network

Community 1		**Community 2**	**Community 3**	**Community 4**
Arizona (USA)	Santa Cruz UC (USA)	Adelaide (Australia)	Berkeley LBNL (USA)	Albany (USA)
Athens (Greece)	Sussex (UK)	Alberta (Canada)	Birmingham (UK)	Columbia (USA)
Austin (USA)	Sydney (Australia)	Annecy LAPP (France)	Boston (USA)	Cracow AGH-U (Poland)
Belgrade IP (Serbia)	Tsukuba (Japan)	Argonne (USA)	Brookhaven BNL (USA)	Glasgow (UK)
Bologna (Italy)	Tufts (USA)	Arlington UT (USA)	Cambridge (UK)	Goettingen (Germany)
Bonn (Germany)	Victoria (Canada)	Barcelona (Spain)	Copenhagen (Denmark)	Grenoble LPSC (France)
Brandeis (USA)	Warwick (UK)	Beijing (China)	Dallas SMU (USA)	Innsbruck (Austria)
Bucharest IFIN (Romania)	Wuppertal (Germany)	Bergen (Norway)	Dresden (Germany)	London Rhbnc (UK)
Buenos Aires (Argentina)	Yerevan (Armenia)	Berlin HU (Germany)	Duke (USA)	Montreal (Canada)
CERN (Switzerland)		Bern (Switzerland)	Edinburgh (UK)	Nagoya (Japan)
Carleton (Canada)		Bog. Istanbul (Turkey)	Harvard (USA)	Napoli (Italy)
Chicago (USA)		Bratislava (Slovakia)	Hefei (China)	Nikhef (Netherlands)
Clermont – Ferrand (France)		Cape Town (S. Africa)	Lancaster (UK)	Oklahoma SU (USA)
Cosenza (Italy)		ES-PIC (Spain)	Manchester (UK)	Osaka (Japan)
Cracow IFJ PAN (Poland)		Genova (Italy)	Marseille CPPM (France)	Portugal 1-LIP (Portugal)
DESY (Germany)		Iowa (USA)	Michigan (USA)	Prague CU (Czech Rep.)
Dortmund (Germany)		Iowa State (USA)	Moscow MEPhI (Russia)	Protvino IHEP (Russia)
Frascati (Italy)		Johannesburg (S. Africa)	Munich LMU (Germany)	Siegen (Germany)
Freiburg (Germany)		Kobe (Japan)	Northern Illinois (USA)	Stony Brook (USA)
Geneva (Switzerland)		Kyoto (Japan)	Ohio SU (USA)	Triumf (Canada)
Giessen (Germany)		Lpnhe-Paris (France)	Pavia (Italy)	Tel-Aviv (Israel)
Heidelberg KIP (Germany)		La Plata (Argentina)	Pennsylvania (USA)	Trieste ICTP (Italy)
Heidelberg PI (Germany)		Lecce (Italy)	Petersburg NPI (Russia)	Valparaiso (Chile)
Indiana (USA)		London UC (UK)	Prague AS (Czech Rep.)	Vancouver UBC (Canada)
JINR Dubna (Russia)		Lousiana Tech (USA)	Prague CTU (Czech Rep.)	Wuerzburg (Germany)
Jagiellonian (Poland)		Lund (Sweden)	SLAC (USA)	Yale (USA)
KEK (Japan)		Madrid UA (Spain)	Saclay CEA (France)	
Liverpool (UK)		McGill (Canada)	Seattle Washington (USA)	
Ljubljana (Slovenia)		Milano (Italy)	Shandong (China)	
London QMUL (UK)		Minsk AC (Belarus)	Shanghai (China)	
Mainz (Germany)		Moscow SU (Russia)	Sheffield (UK)	
Massachusetts (USA)		NYU New York (USA)	Stockholm (Sweden)	
Melbourne (Australia)		Oklahoma (USA)	StockholmKTH (Sweden)	
Michigan SU (USA)		Oslo (Norway)	Taipei AS (Taiwan)	
Minsk NC (Belarus)		Pisa (Italy)	Thessaloniki (Greece)	
Munich MPI (Germany)		Pittsburgh (USA)	Tokyo Tech (Japan)	
New Mexico (USA)		Simon Fraser B. (Canada)	UC Irvine (USA)	
Nijmegen (Netherlands)		Technion Haifa (Israel)	UI Urbana (USA)	
Olomuc (Czech Republic)		Tokyo ICEPP (Japan)	Witwatersrand (S. Africa)	
Oregon (USA)		Toronto (Canada)	York (UK)	
Orsay LAL (France)		UAN Bogota (Colombia)		
Oxford (UK)		Udine (Italy)		
RAL (UK)		Uppsala (Sweden)		
Roma I (Italy)		Valencia (Spain)		
Roma II (Italy)		Weizmann R. (Israel)		
Roma Tre (Italy)		Wisconsin (USA)		

the teams that the participating institutions employed. In these conditions, it is difficult for any institution to accumulate a significantly higher number of collaborative relationships to become a *hub* in the co-authorship network.

Third, there were a small kernel of institutions that had a prominent situation within the *ATLAS* collaboration network (see Figure 5 b). This kernel is formed by *Michigan*, *Berkeley LBNL*, *London*, *DESY*, and *CERN*. Among them, *Berkeley LBNL* seemed to play a central role. Other institutions around them appeared in this simplified version of the network including *Harvard*, *Mainz*, *Oxford*, *Orsay LAL*, *Brookhaven BNL,* or *Lpnhe- Paris* also played an important role.

Fourth, geographical distance between *ATLAS* institutions did not seem to be a determining factor for the establishment of collaborations. After inspecting the composition of the communities we found in the *ATLAS* collaboration network, nothing seems to indicate that institutions that are closer tend to group in the same communities. This is not what happens in collaboration networks outside big experiments. In that case it has been shown that the importance of collaborative relationships decreases with the distance between the geographical sites of institutions (**Pan**; **Kaski**; **Fortunato**, 2012). The explanation for this discrepancy is probably related to the existence of one physical location only for the *ATLAS* experiment. At *CERN* researchers from different institutions meet periodically and may know each other, solve misunderstandings, or transmit tacit knowledge, all of them aspects that act as lubricants in any collaboration. Thus, although some of them usually work from their home institutions —what happens in the current phase of data analysis, distance, is not a determining factor for the successful development of the project.

8. Conclusions

ATLAS is one of the most paradigmatic examples of the large scientific experiments that take place nowadays and are known as *big science*. Although it is obvious that not all big experiments work in the same way, we believe that the conclusions we extract from this study help to shed light on the structure and functioning patterns of collaboration of this kind of experiment.

Based on the findings of this research, it appears that the collaboration patterns in large scientific experiments slightly differ from other, smaller collaborations. Taking into account the increasing number of these kinds of projects in several areas of science and the growing amount of public funds allocated to them, collaboration in *big science* deserves a deeper study. That often requires the adoption of new methodological approaches, since the traditional co-authorship analysis is not worth much when articles are signed by all participants in the project.

In this paper we have studied the case of a representative big experiment in the field of high-energy physics, the *ATLAS* Collaboration at *CERN*. *ATLAS* is, of course, a particular case and big experiments in this field have features that differentiate them from those in other fields. However, we believe that our findings constitute a modest contribution to knowledge about the nature of modern scientific collaboration. Our results hint at interesting features of *big science* that need to be studied in more depth in future research, taking advantage of new methodological tools, both quantitative and qualitative, that will provide us with a deeper understanding of the nature of collaboration in large scientific experiments.

Aknowledgements

The authors would like to thank the participation and help in the present research of the *ATLAS* Collaboration, of *CERN*, and of *IFAE* (*Institut de Física d'Altes Energies*) in Barcelona. AC and EO also thank the financial support from the *Generalitat de Catalunya* to the *KIMO* research group (2014-SGR-1486) and from *Mineco* through the funding of the projects *Kesir* (CSO2009-09194), *Kibis* (CSO2012-33959) and the *Maredata Network* (CSO2015-71867-REDT). We would like also to thank an anonymous reviewer for her/his suggestions.

Note

1. The energy unit of one TeV (Tera-electron volts) corresponds to the mass-energy of more than a thousand protons equivalent.

9. References

Barabási, Albert-László; **Albert, Réka** (1999). "Emergence of scaling in random networks". *Science*, v. 286, pp. 509-512.
http://barabasi.com/f/67.pdf
https://doi.org/10.1126/science.286.5439.509

Benavent-Pérez, María; **Gorraiz, Juan**; **Gumpenberger, Christian**; **De-Moya-Anegón, Félix** (2012). "The different flavors of research collaboration: A case study of their influence on university excellence in four world regions". *Scientometrics*, v. 93, n. 1, pp. 41-58.
https://doi.org/10.1007/s11192-012-0638-4

Bressan, Beatrice; **Boisot, Max** (2011). "The individual in the *ATLAS* Collaboration: A learning perspective". In: Boisot, Max; Nordberg, Markus; Yami, Saïd; Nicquevert, Bertrand (eds.). *Collisions and collaboration: The organization of learning in the ATLAS experiment at the LHC*. Oxford, UK: Oxford University Press, pp. 201-225. ISBN: 978 0 199567928

Chompalov, Ivan; **Genuth, Joel**; **Shrum, Wesley**. (2002). "The organization of scientific collaborations". *Research policy*, v. 31, n. 5, pp. 749-767.
https://goo.gl/WdeCcx
https://doi.org/10.1016/S0048-7333(01)00145-7

Clauset, Aaron; **Newman, Mark E. J.**; **Moore, Cristopher** (2004). "Finding community structure in very large networks". *Physical review E*, v. 70 (066111).
http://ece-research.unm.edu/ifis/papers/community-moore.pdf

Creus, Amalia; **Canals, Agustí** (2014). "Desarrollo profesional e intercambio de conocimiento en los grandes experimentos científicos". *Revista española de documentación científica*, v. 37, n. 4, e070.
https://doi.org/10.3989/redc.2014.4.1140

Dorogovtsev, Sergei; **Mendes, Jose** (2003). *Evolution of networks: From biological nets to the internet and WWW*.

Oxford, UK: Oxford University Press.
http://www.cs.unibo.it/~babaoglu/courses/cas03-04/tutorials/evolution_of_networks.pdf

Etzkowitz, Henry; **Kemelgor, Carol** (1998). "The role of research centres in the collectivisation of academic science". *Minerva,* v. 36, n. 3, pp. 271-288.
https://doi.org/10.1023/A:1004348123030

Galison, Peter; **Hevly, Bruce** (eds.) (1992). *Big science: The growth of large scale research*. Stanford, CA: Stanford University Press. ISBN: 978 0 804718790

Gazni, Ali; **Sugimoto, Cassidy R.**; **Didegah, Fereshteh** (2012). "Mapping world scientific collaboration: Authors, institutions, and countries". *Journal of the American Society for Information Science and Technology*, v. 63, n. 2, pp. 323-335.
https://goo.gl/6jmyS4
https://doi.org/10.1002/asi.21688

Hicks, Diana M.; **Katz, J. Sylvan** (1996). "Where is science going?". *Science, technology & human values*, v. 21, n. 4, pp. 379-406.
https://doi.org/10.1177/016224399602100401

Huang, Mu-Hsuan; **Tang, Muh-Chyun**; **Chen, Dar-Zen** (2011). "Inequality of publishing performance and international collaboration in physics". *Journal of the American Society for Information Science and Technology*, v. 62, n. 6, pp. 1156-1165.
https://doi.org/10.1002/asi.21516

Knorr-Cetina, Karin (1999). *Epistemic cultures: How the sciences make knowledge*. Cambridge, MA: Harvard University Press. ISBN: 978 0 674258945

Kronegger, Luka; **Ferligoj, Anuska**; **Doreian, Patrick** (2011). "On the dynamics of national scientific systems". *Qual quant*, v. 45, n. 5, pp. 989-1015.
http://patrickdoreian.com/NEW/wp-content/papers_resources/new_papers_4-13/on_the_dynamics_of_scientific_systems.pdf
https://doi.org/10.1007/s11135-011-9484-3

Lorigo, Lori; **Pellacini, Fabio** (2007). "Frequency and structure of long distance scholarly collaborations in a physics community". *Journal of the American Society for Information Science and Technology*, v. 58, n. 10, pp. 1497-1502.
https://goo.gl/ewp4og
https://doi.org/10.1002/asi.20638

Milgram, Stanley (1967). "The small-world problem". *Psychology today*, v. 1, n. 1, pp. 61-67.
https://goo.gl/aGgERU

Newman, Mark E. J. (2001). "The structure of scientific collaborations". *Proceedings of the National Academy of Sciences*, v. 98, pp. 404-409.
https://doi.org/10.1073/pnas.98.2.404

Newman, Mark E. J. (2010). *Networks: An introduction*. Oxford, UK: Oxford University Press. ISBN: 978 0 199206650

Newman, Mark E. J.; **Girvan, Michelle** (2004). "Finding and evaluating community structure in networks". *Physical review E*, v. 69 (026113).
https://arxiv.org/abs/cond-mat/0308217
https://doi.org/10.1103/PhysRevE.69.026113

Ortoll, Eva; **Canals, Agustí**; **García-Alsina, Montserrat**; **Cobarsí-Morales, Josep** (2014). "Principales parámetros para el estudio de la colaboración científica en 'big science'". *Revista española de documentación científica*, v. 37, n. 4, e069.
https://doi.org/10.3989/redc.2014.4.1142

Pan, Raj-Kumar; **Kaski, Kimmo**; **Fortunato, Santo** (2012). "World citation and collaboration networks: uncovering the role of geography in science". *Scientific reports*, n. 2, 902.
https://doi.org/10.1038/srep00902

Price, Derek J. De-Solla (1963). *Little science, big science*. New York: Columbia University Press. ISBN: 0231049560

Redner, Sidney (1998). "How popular is your paper? An empirical study of the citation distribution". *European physical journal B*, v. 4, n. 2, pp. 131-134.
http://physics.bu.edu/~redner/pubs/pdf/citation.pdf
https://doi.org/10.1007/s100510050359

Santalainen, Timo J.; **Nordberg, Markus**; **Baliga, Ram B.**; **Boisot, Max H.** (2011). "New management research models: Lessons from the *ATLAS* adhocracy". In: Boisot, Max; Nordberg, Markus; Yami, Saïd; Nicquevert, Bertrand (eds.). *Collisions and collaboration: The organization of learning in the ATLAS experiment at the LHC*. Oxford, UK: Oxford University Press, pp. 55-76. ISBN: 978 0 199567928

Sonnenwald, Diane H. (2007). "Scientific collaboration". *Annual review of information, science and technology*, v. 41, pp. 643-680.
https://doi.org/10.1002/aris.2007.1440410121

Watts, Duncan J.; **Strogatz, Steven** (1998). "Collective dynamics of 'small-world' networks". *Nature*, v. 393, pp. 440-442.
https://doi.org/10.1038/30918

Weinberg, Alvin M. (1961). "Impact of the large-scale science on the United States". *Science*, v. 134, n. 3473, pp. 161-164.
https://doi.org/10.1126/science.134.3473.161