Paper presented in the Prime-Latin America Conference at Mexico City, September 24-26 2008

Inter-institutional scientific collaboration: an approach from social network analysis.

Zaida Chinchilla-Rodríguez[1]
Félix de Moya-Anegón[2]
Benjamín Vargas-Quesada[3]
Elena Corera-Álvarez [4]
Yusef Hassan-Montero[5]

Introduction and motivation

The evaluation of scientific activity and technological innovation has become the norm in industrialized societies. Attesting to this fact is the regular publication of technical reports that have documented such developments since the end of the 70´s[6]. In addition to the traditional functions of evaluation (certification and detection of scientific excellence), we now see its importance as an added value in decision-making processes that involve science and technology, and as a tool used in strategically advancing systems of R+D and innovation. Evaluation plays a key role in building scientific and technological potential, making it essential for social well-being and economic competitiveness. For these reasons, Scientific Policy and Scientometrics are closely linked, and the assessment of science, technology and innovations at all levels calls for tools that will permit measurement in various dimensions (Rinia, 2000)

At present, politicians place great emphasis on innovation as a collective process of interaction and mutual instruction amid a group of actors that form part of the system of science and technology. From the standpoint of innovation, political intervention can be justified to overcome institutional paralysis and promote energetic incentives for cooperation, learning, and adaptative conduct among all the members of the system. Such actions have two objectives. On the one hand, they attempt to resolve the "systemic failures" that reflect deficiencies in interaction aimed toward technological development (Laranja, Uyarra and Flanagan, 2007), while on the other hand, they can enhance the efficiency of the system by lending it an architecture with the power of distribution of technological information and knowledge (David and Foray, 1995).

[1] zchinchi@ugr.es. Grupo SCImago – Universidad de Granada. Departament of Library and Information Science. Campus of Cartuja. Colegio Máximo of Cartuja s/n 18071 Granada (Spain)
[2] felix@ugr.es
[3] benjamín@ugr.es
[4] ecorera@ugr.es
[5] yusef_hassan@yahoo.es
[6] You can see the following reports: *Science and Engineering Indicators* of the National Foundation Science of United States from 1972; *Science and Technology Observatory* of France and *World Science Report* of UNESCO.

Generally speaking, administrations worldwide have made explicit the need to foment collaboration at all levels and in all productive sectors. Along these lines, together with the traditional instruments of support for research (the competitive financing of R+D projects), there are policies and programs to further the mobility of researchers and promote lasting associations among these actors. The goal, after all, is to favor scientific excellence, visibility, and the international reputation of one´s country, and enhance difussion and interchange of knowledge and innovation.

In the case of Spain in particular, ever since the "*IV Plan Nacional*" came out in the year 2000-2003, the emphasis has been on coordinating all the different agents under the Spanish System of Science and Technology (SECYT), especially when public programs and business initiatives are involved. Another foremost objective has been to make Spanish science more international, giving priority to projects within the framework of the European Research Space. Yet it was not until the initiative of Ingenio 2010[7] that scientific collaboration became a state affair, a stepping stone toward the goals set forth in the Lisbon Strategy[8]. At the European level overall, reinforcement of collaborative relations is considered a main via of cohesion and convergence on the road to constituting a transnational system (Maltrás, 2003). The Framework Programs insist, therefore, on the collaboration of plurinational research teams to bid for funding. Both initiatives obey global strategies to finance clearly common goals: integration on the front lines of research, forming versatile research teams, reducing redundant research efforts, and taking optimal advantage of infrastructures, among other beneficial aspects.

So then, what tools are available to those who evaluate in order to make decisions in the context of R+D collaboration and innovation? Traditional indicators have focused on the study of distributive effects of policy –that is, those affecting the actors as individuals, whether they be businesses, universities, or public research centers (Sanz, 2001). Policy makers and managers need methodological tools that can deal with the multidimensional and heterogeneous nature of the activities that generate knowledge and innovation (Buesa et al. 2007).

Evaluating and following the wake of scientific activity is no easy task, for one reason because its products may be tangible or intangible (Moravsick, 1989; Sancho, 2001). Collaborative efforts can be assessed by quantifying joint ventures, co-publications, informal contacts, interchange of research fellows or scholars, or attendance at international conferences (Fernandez, Gomez, and Sebastian, 1998) Whatever the "unit" of analysis, scientometric studies alone cannot do justice to the true dynamics behind the process of scientific collaboration (Wang, et. al., 2005), as not all activities end up in the form of joint publications of collaborating parties (Katz, and Martin, 1997). Notwithstanding, analyses

[7] The initiative Consolider-Ingenio 2010 promotes the creation/renovation of centres of excellence, and the configuration of major research teams. This action, geared toward high level research, requires a concentration of efforts on the part of groups and consortia in collaboration with Autonomous Communities, private firms, and international organizations. Most of the basic instruments (Cenit and Consolider) put special emphasis on collaboration.

8 During the European Council of Lisbon (March 2000), the heads of states and government propose that European Union must be the highest competitive economy in the world and reach the full employment before 2010.

based on co-authorship of scientific documents do indeed provide a good estimation of cooperative productivity (Okubo et. al., 1992; Bordons and Gómez, 2000).

We should bear in mind, meanwhile, that the generation of knowledge takes place over complex multidimensional networks that serve as evidence of the reticular and dynamic nature of the system of scientific communication. These activities appear as the result of a continuous process that gives rise to growing networks and increasingly complex mechanisms deriving from the interaction of the system and its setting (Fry, 2006). Collaboration is a reflection of interaction on the part of networks of individuals who, in turn, configure institutional and global networks (Kretschmer, 1993; Kyvik and Larsen, 1994). These networks are therefore condidtioned not only by scientific factors, but also by well-documented social and cultural factors related with the given field (Subramanyam, 1983; Beaver, 2001; Wagner and Leydesdorff, 2005). Depending on the level of aggregation to be analyzed, and the techniques applied to that end, such factors can be uncovered and identified. The context of scientific collaboration provides an opportunity to develop indicators that will reveal the essential organization of patterns of communication, as well as the structural effects of scientific policy and how it is related with its actors and their capacity to produce new knowledge.

The analysis of scientific systems looks beyond fragmented individual results to take in the panorama of production; and its characterization should therefore reflect the behaviour of aggregate components (institutions, sectors, autonomous communities, countries) as a by-product of participation in structured social relations. It is clearly beneficial to view and analyze the different levels (Laranja, Uyarra and Flanagan, 2007) that participate in the generation of knowledge and innovation, sometimes overlapping, sometimes active on more than one stage.

This paper presents a tool that can be used to characterize, analyze and interpret the patterns of collaboration among institutions by means of the visual display of scientific information. These graphic representations allow for a combined analysis of a given institution in the system of relations (network), and of the particular attributes of that institution (indicators). The tool affords the possibility of regenerating the network to make any number of aggregates appear or disappear, thus allowing one to focus on institutional sectors, geographic regions, etc. It also allows for analysis of sectorial interaction, institutional backing of research, and the influence of geographic proximity, linguistic affinity, or regional politics. This is indeed a versatile analytical tool, and it is bound to prove its potential for evaluating patterns of collaborative research, development and innovation.

2. Related Work

Over the last years the concepts and methods associated with Social Network Analysis (SNA) have undergone considerable development, giving rise to a new means of studying social structures. Analysis of the systems of science and technology based on the structural analysis of networks plays an important role as a complement to conventional analyses regarding scientific output and, in particular, scientific collaboration.

2.1. Social Network Analysis

The main difference between the explanations contributed by social network analysis (SNA) and more conventional bibliometric analyses lies in the inclusion of concepts and information about the relations between units. SNA is based on the premise that the relations between social actors can be described in a graph. SNA render the actors as nodes. Each pair of nodes is connected by lines, evidencing their social interaction (Wasserman & Faust, 1994). This methodology based on graph theory allows us to analyze the framework of relations within scientific activity.

Although SNA has been on the scene for over fifty years, and is applied in many fields of investigation, in the field of the Information Science its use has increased very significantly in the last 20 years (Otte and Rousseau, 2002, Scott 2000). The availability of vast amounts of data and the automatic processing thereof can facilitate analysis at meso and macro levels. Although the notion of building collaboration networks is not new (Erdos number) up to now the study of social networks appears most frequently in the context of cocitation networks (White, 2003; Börner, Chen and Boyack, 2003, Moya et.al, 2004). Accordingly, documents constitute the nodes, and the connections among them represent the network of references (citations). The authors of documents need not have any personal relation with the authors that cite them, only a disciplinary connection. Despite the interest sparked by such representation and analysis of scientometric studies, we consider that networks based on co-authored documents afford a better reflection of the social character of science. That is, they give reliable account of the voluntary nature (independent of underlying reasons) behind collaboration between individuals and, by extrapolation, between the institutions in which they are working.

In the present study, the network is constructed from bibliographical data about institutional affiliation. The nodes of the network will be the signatory institutions of documents, while the connections between the nodes will reflect the intensity of collaboration among them.

2.2. Scientific Collaboration

Whereas the sociologists of science carried out their initial studies of scientific collaboration in the 60´s, the use of co-authorship data to examine international scientific collaborative activities is a more recent phenomenon. It was not until the 90´s that the use of these data and the proposed methodologies diversified (Yamashita, and Okubo, 2006).

As we said before, one of the earliest works about collaboration networks is that by Paul Erdös, father of the graph theory. Scientific literature harbours studies that have analyzed the phenomenon of collaboration through co-authorship, and there are also works that describe mapping techniques used to visualize collaborations between or among countries in a specific sector (Dumont and Meussen, 1997). As far as sectorial collaboration and, in particular, the Triple Helix model are concerned, we would underline the work by Leydesdorff and that of Heimeriks, Hörlesberger and Van Der Besselaar, 2003). Others have defined indicators to describe the participation patterns of some countries within the Framework Programs of the European Union (Lukkonen et al., 1999; Gusmao, 2000). The work of Glänzel is concentrated on the networks of international collaboration. At the institutional level, we might highlight the study by Mählck and Persson (2000) based on co-authorship networks in two departments of different Swedish universities. Regarding disciplines, a solid point of reference is the study done in the field of mathematics and the neurosciences by Barabasi and Hungarian bibliometric investigators (Barabasi, Jeong et al., 2001), the analysis of Newman (2001). On the national level, meanwhile, we have the works of Molina, Muñoz and Domenech (2000); those of the SCImago Group (Moya et. al, 2004, 2005, 2006, 2007, 2008); those referring to the end purposes of science and technology policies (Sanz Menéndez, 2000); and finally, those signed by a team of Valencian authors in the field of drug addictions and neurology (González et al. 2006, 2008).

3. Objectives

First of all, we shall try to characterize the production of knowledge of the different productive sectors in the field of Agricultural studies. A second objective is to project inter-institutional networks of collaboration through visualization techniques. Thirdly, we intend to analyze this network, looking more carefully at aspects such as the degree of interaction between/among agents, the appearance or disappearance of actors, or potential effects on the increase of the internationalization of scientific output. We also try to elaborate an interpretation protocol for collaboration networks in R&D that could be applied to any field of knowledge or upper level of aggregation (i.e. countries or regions). Finally, in order to equip this graphic representation with other functionalities, we lend it an added value by allowing diverse levels of analysis and its service as an interface for domain analysis and information retrieval.

4. Source Data and Information Processing

4.1. Source data

The data was extracted from the Web of Science, a product developed by Thomson Scientific. Its use and operation have been possible thanks to the no-charge access made readily available by the Spain´s Ministry of Education and Science, and entrusted to the Spanish Foundation of Science and Technology (FECYT) [9], as a public service provided to academic and research institutions. For the specific case of Spanish science, the selection of this source for the purpose of analysis and evaluation would agree with the current norms[10] that establish the criteria regarding the system of incentives for investigators in all the scientific fields, except in Law and Jurisprudence, History, Art, Philosophy, Linguistic Philology, and Linguistics in general.

4.2. Selection data

Within the project “Atlas of Science” (http://www.atlasofscience.net) and with strictly academic pouposes, we downloaded the data of the Spanish scientific output from the Science Citation Expanded Index (SCI-Expanded), Social Science Citation Index (SSCI) and Arts and Humanities Citation Index (AHCI). A total of 375,256 registries corresponding to the period 1990-2005 were recovered, in which Spain appears as the country under the Address field. After retrieval, a relational database was built to allow us to operate in a simple, flexible and fast way for the different analyses needed. For the construction of the database, we developed an ad-hoc software specifically for the loading, modelling and treatment of data. The database contained the following

[9] Fundación Española de Ciencia y Tecnología. Access to the Web of Knowledge available in: http://www.accesowok.fecyt.es/

[10] Decisión from 28 August 1989, modified and completed through Real Decreto 1325/2002

information: authors; institutional affiliation, publication title, information about source data (title of magazine, year of publication, volume, number, first and last pages, and publication type), as well as the bibliographical references contained in each publication. We added to this database all the bibliometric information from the Journal Citation Report (SCI and SSCI versions) corresponding to the journals processed during the period 1995-2005. The information added for each one of the journals was: bibliographical identity data, number of works published per year, thematic categories of pertainance, and impact factor for that year. This information conformed the international comparative reference, giving the total number of publications aggregated chronologically and thematically at the world-wide level.

Institutional and Regional Normalization

A well-known limitation of the source of data is the lack of normalization of the institutional field (among others); this is even more serious in the case of countries where the English is not the *lingua franca* (Russell, 2000). In recent years various descriptions of projects or pilot projects have attempted to standardize addresses to make it possible, on a broad scale, to analyze citations and collaboration in the output of institutional papers(Bruin and Moed, 1990, 1993; Fernández, et.al, 1993; Katz and Hicks, 1997; Gálvez and Moya, 2006, 2007). In our study, normalization of institutions as units of analysis was effected along with their correspondence with the city in which each institution is situated, and by extension its autonomous region. As a general rule, the structure of the institutional field contains four parts. The country is usually well standardized, and the information on the city can be standardized from postal codes. At all these levels we can find a great number of variants that, once located, are unified under a single entry, and are then assigned to the corresponding autonomous region. Finally, we added several tables with information on the autonomous regions (CCAA) to the relational structure that was derived from the downloaded information.

4.3. Aggregation levels / Unit of analysis

4.3.1. Temporal distribution

Although the extracted data set goes from year 1990 to 2005, in order to include each work, the year of publication of the number of the magazine in which it appears was taken as the reference rather than the year of its entry in the database, which helped avoid delays in its inclusion (Moed et.al., 1989). This work only renders the information for the biennium 1995-1996.

4.3.2. Topic Distribution

The information available for the assignment of documents to a certain area is given by the thematic categories (Subject Category, SC) [11] that the JCR offers. Once the category or categories of a journal are determined, all the documents published in it are considered to belong to that thematic discipline. In addition, this categorisation is used to classify each one of the documents in more general areas, in following the classification elaborated by the National Agency of Evaluation and Prospective (ANEP). This new classification entails 26 great areas. The present paper focuses on one of them,

[11] From here onward, references will be to ISI subject categories.

Agriculture. Its election is justified by the role it plays in the subject distribution of the domain in particular, in terms of collaboration patterns (Moya et. al., 2005) as well as thematic specialization, impact factor and output Moya et. al., 2005).

4.3.3. Geographic Distribution

This division corresponds to the 17 autonomous regions. Ceuta and Melilla, with a total of twenty documented works, were included under the autonomous region of Andalusia. In addition, seven major geographic regions for analysis of continental analysis output were defined.

4.3.4. Sector Distribution

The sectorial classification used is based on the Manual of Frascati of the OECD and on the definitions established by the Ministry of Education and Science. In the normalization of the institutions, as with the independent communities, each institution is assigned to a sector, with the difference residing in its unique ascription. That is, an institution belongs to only one sector, which encompasses groups from institutions that share common profiles (SCImago Group, 2006). The sectors dealt with here are: Administration, Mixed Centers (CM), CSIC, Companies, Public Research Organizations (EPI); Medical System (SS), University System (Univ) and Others.

4.3.5. Denomination of the Documents

Finally, the assignment of documents at these levels of aggregation called for use of the system of complete account. This system ascribes a given document to each and every one of the signatory institutions and autonomous regions. This system of total count was chosen because it allows for quantifation of participation by the different institutions in research efforts, offers a more complete vision than the count by first author, and its reliability has been verified on more than one occasion (Moed, wt.al., 1989). The disadvantage stem from document duplication, which causes sum totals to be higher than the real total number of documents. In order to avoid this bias, the percentage was calculated over the real total number of documents.

5. Methodology

Having standardized the data from the relational database constructed with all the bibliographical information, some conventional bibliometric indicators were extracted are placed into relation with others based on social network analysis. For each one of the institutions a battery of indicators (Chinchilla and Moya, 2007) with the following information appears:

- ndoc: number of documents number;
- ndoc-col: documents in collaboration;
- % col: percentage of documents in collaboration in the area;
- ndoc-int: documents in international collaboration;
- % int: percentage of documents in international collaboration in the area;
- ndoc-citable: articles with impact factor;
- % citable: percentage of articles with impact in the area;
- finp: standardized impact factor (weighted) of the journals in which it is published;
- fire: relative impact factor of Spain;

• firm: relative impact factor in the world;
• pi: investigating potential;
• degree: nodal degree;
• closeness: proximity degree;
• betweenness: intermediation degree.

In order to analyze the institutional collaboration, we distinguished several types of collaboration. We defined documents without collaboration as those in which only one institutional address appears, regardless of the fact that they were signed by one or more authors of a single institution; therefore, this does not constitute intra-institutional collaboration. For national collaboration, we considered only the documents produced in collaboration, within the same country, between authors who work in different institutions. International collaboration is that involving groups and output in which the authors are of at least two different countries.

From the information on copublications, a matrix of double entery of inter-institutional collaboration was created, representing on a national level the collaboration between Spanish institutions in the field of Agriculture. The result is a symmetrical matrix of 333 by 333 to be used for the representation of the graph and its later analysis.

A key step previous to representation consists of normalization of the values of the matrix. Scientific Literature gathers diverse indexes for the creation of collaboration maps that reflect the natural topology of the variables of study, such as that presented by Salton or Jaccard. These indexes can reflect similarity in the collaboration of different agents, locating them in the spatial representation, so that the position occupied is exemplary of the “natural geographic order”, (for that reason, also denominated proximity index), while at the same time it offers information on the structure defined by the copublication connections (Arunachalam and Doss, 2000; Schubert and Braun, 1990).

Yet despite the fact that this normalization proves useful, it does not reflect the asymmetry that can exist between the connections. In other words, there exists the possibility that an institution may be a very important partner for another one, but that reciprocity does not necessarily characterize that association (Glänzel and Schubert, 2001; Zitt, Bassecoulard and Okubo, 2000). This is one of the limitations of the indicator of symmetrical collaboration, in addition to the fact that it is strongly affected by the size of the agents. In order to correct these deficiencies --the bidirectional intensity and the lack of normalization with respect to the size of the agents (Boyack and Börner, 2003) -- the literature provides indexes of asymmetric collaboration. Zitt, Bassecoulard and Okubo present one possible way to characterize the relative importance of the connections of a country with respect to another one:

$$asi = \frac{cop}{co(m-p)} * 100$$

cop = total number of copublications of a Country

Co (m - p) = total number of copublications of the rest of countries

Ideally, these two values would have to be identical, but this is not true. This index shows the attraction, or the absence thereof, at the time of collaboration among countries, regions or institutions (Glänzel, 2001; Glänzel, 2001). With a view to making comparisons in coherent manner, the authors propose the use of the ratio of the percentage of both agents. Here it is applied to publications involving collaboration between/among institutions.

6. Visualization

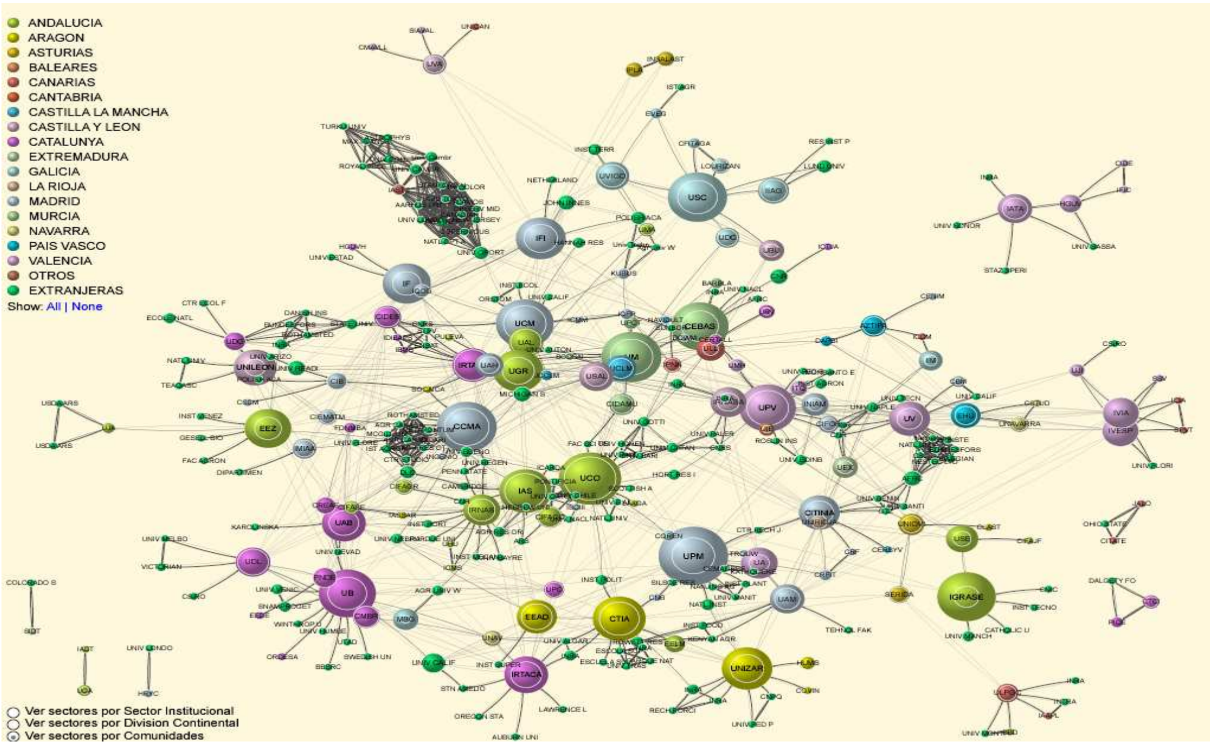
The techniques of graphical representation for the generation of maps are based on social networks. The standard matrix is processed using Pajek_[21] and the algorithm of Kamada-Kawai is then applied (Kamada and Kawai, 1989) to locate the institutions in the graph based on its joint copublications. Once the spatial coordinates are defined for each institution, they are exported to ad-hoc software in which, next to the bibliometric information and that of the different levels of aggregation, the information is processed adding scripts and an interactive navigation option in each one of the nodes is obtained. This tool gives back the final network in a SVG (Scalable Vector Graphics) format that allows one to zoom in on an element, displace elements in any direction upon the screen, etc. This procedure has been meticulously detailed in a recent work published in *Scientometrics* (Moya, et. al., 2004b).

7. Results

7.1 General Analysis of the network

In every graphic representation, the relations between institutions show the bidirectional intensity of asymmetric collaboration. The node size is proportional to the production volume and the concentric circle reflects the collaboration rate. The node colour means that it belongs to an aggregate (institutional sector, autonomous community and geographical region). Every node is placed in the space according to its dependence. The closeness or distance is related to the total number of links of an actor with the others.

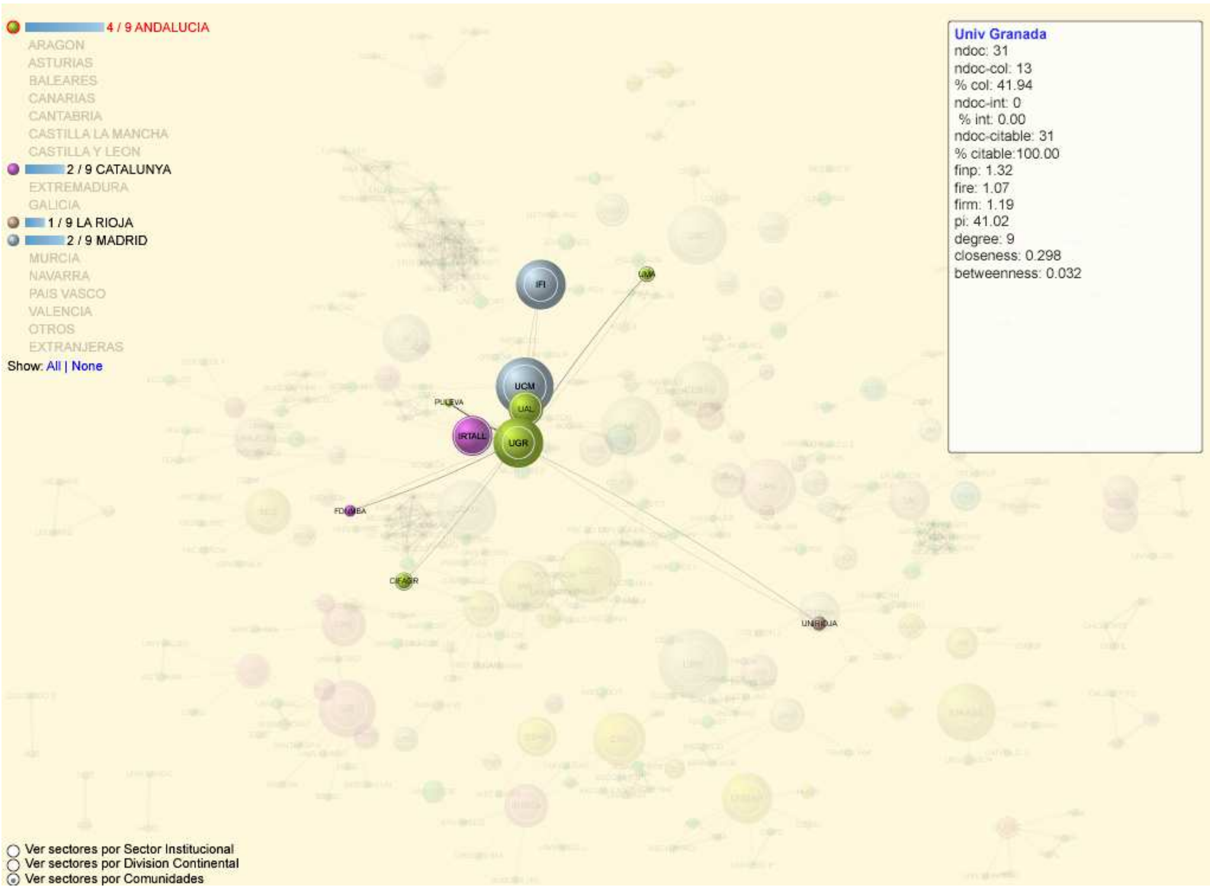
Map 1. Inter-institutional collaboration network of Spanish Agriculture (1995-1996)



The lower left part of the representation shows the different levels of aggregation that the user can choose in order to develop analysis by sector, autonomous communities or geographical regions. Once the level of analysis is selected, the legend (situated in the upper left part of the map) shows the list of aggregates depending on the selected level (institutional sectors, autonomous communities and geographical regions). The graphic representations allow us to choose one or several items. Moreover, we can show or hide the institutions of the selected aggregation level.

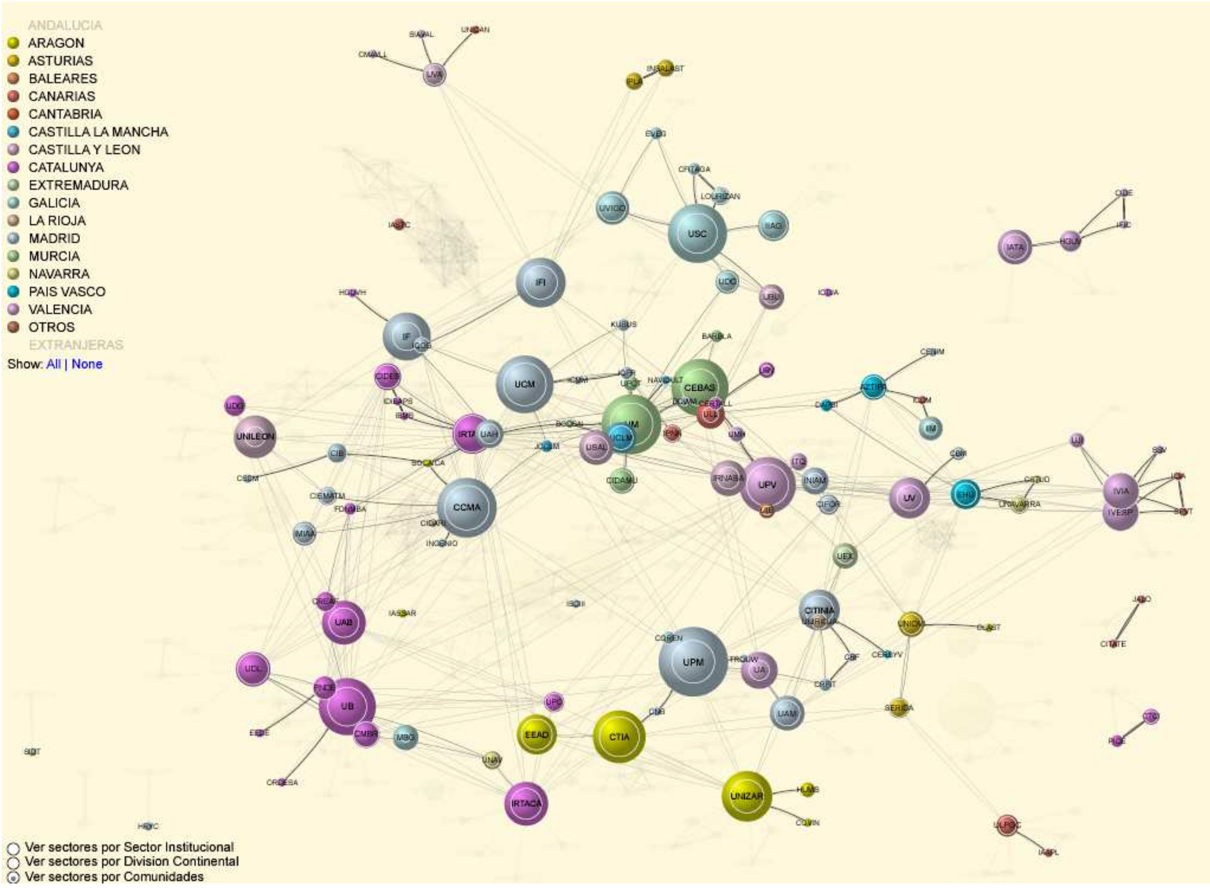
Meanwhile, placing the cursor on a node, we can see its relations with the other through a visual discrimination effect. Moreover, we can see in map number two: the nodal grade (number of links) for every level of aggregation (legend in the left higher part of the map) and institution as well as the array of indicators described in the methodological section. Thus we can make a combined analysis of the role played by every institution in the relations system and of its particular attributes.

Mapa 2. Indicators Battery and nodal degree for a network institution (Universidad de Granada)



Moreover, the advantage of showing or hiding the aggregates (institutional sector, autonomous communities or geographical regions) allows us to make combined analysis about sector interaction, institutional structuring of research, geographical proximity, linguistic affinity, regional policies, fragmentation/cohesion of the knowledge production in the field, etc.

Map 3. Disappearance of two actors (Andalucia and foreign institutions).



7.2. Conventional Scientometric Analysis

We must bear in mind that the output indicators in the institutional level reflect the activity of the researchers in these institutions. These indicators appear in the evaluation and accreditation processes, and so the finality is to improve these outputs (productivity, visibility, prestige, etc). Doing that will enhance the scientific reputation of the institution. We may ask: Is there any relation between the position of the institution in the network with its scientific relevance as measured in terms of output of their researchers? What are the effects of closeness, between-ness or degree with the visibility of scientific production?

Traditionally, it is thought that increases in inputs are related to an increase of the outputs. The capacities of an institution to collaborate are their social capital. Does the accumulation of social capital have a multiplying effect on the development of the output having the same inputs? What are their qualitative and quantitative effects on the outputs? This is the main utility of the tool: showing the relation between social capital and production of knowledge and innovation. Moreover, it is intended to relate the net indicators with the outputs in order to see what kind of effect the scientific collaboration (in all its forms: national, international, regional, etc) has on the increase of institutional prestige.

In the beta version of this tool, we are going to set three levels of analysis: regional, sector and international divisions. The aim is to emphasize the most interesting aspects in the network, always keeping in mind their relations with the output indicators of the institutions.

For example, we can ask about the effect of the production size on the collaboration. In the literature it is shown that there is a negative correlation between the size of any geographical domain and the publications in collaboration rate. This affirmation suggests that the small knowledge producers have a big incentive to collaborate. That is because if they want to participate in the research network dynamic of a particular scientific community, they must do so with national or international partners.

In relation with the more or less heterogeneous character of the relations (that link institutions of different sectors), it seems that there is a relation with the capacity of the institution to transfer knowledge and to create social capital. The diversification in the relations is tied to the institution size, the consolidation of the scientific field, etc.

Another aspect to take into account is the positive correlation between the impact factor of the journal in which the researchers publish and, by extension, of the number of cites received by work and the participation of more than one author (individual or institutional). This is particularly important in the case of foreign partners (international collaboration). This front has created a very considerable amount of academic works to date.

7.3. Network analysis by Autonomous Communities

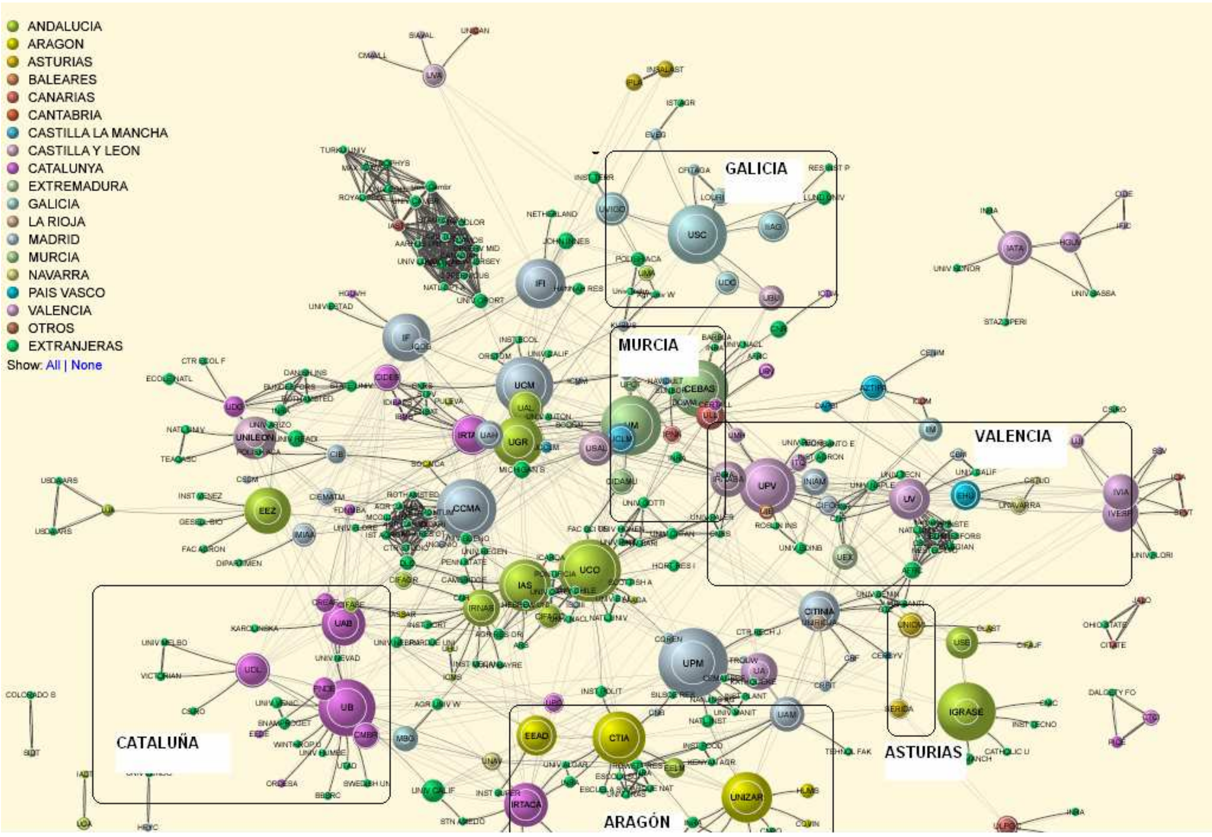
In the case of Spain, a study of the Autonomous Communities is perfectly justified because the scientific policies set forth a set of common conditions for the institutions in the domain of the same Administration. If we want to comparatively assess the effects of these policies, it is necessary to separate the relative part of every one. That is, it is possible to identify autonomous communities that follow an academic model focus on the funding and promotion of public research, especially the university research. It is also possible to identify an entrepreneurial model directed to the promotion of applied research, the transfer of results to the private sector and, finally, the relation between all the productive sectors.

Map number 1 depicts the network of inter-institutional collaboration of Agriculture in Spain by regions. From an analytical point of view, the places of publication show the more prolific regions in this field in terms of number of institutions (identified by the same colour) and their volume of production. For example, Extremadura has few centres with production in the subject area of Agriculture. However, its relations are produced with strategic regions (Valencia, Andalucia, Madrid). Moreover, this community has very high research internationalization, because 8 of the 10 institutions with which Extremadura collaborates are foreign. We could think that Extremadura strategic relations could play an important role in its internationalization degree. In the case of Cataluña, on the other hand, endogamic relations between institutions are seen (relations that are produced fully inside the region).

In the analysis of the inter-regional networks, we can see the effect of the geographical proximity between the Autonomous Communities. Until now, the proximity has been considered like an important factor in the creation of collaboration networks. Nowadays, the technologies break down these geographical barriers. The literature points out that the geographical proximity *per se* is not a necessary or sufficient condition for the innovation. Moreover, its positive or negative effects are mediated by the dynamics of the social relations inside the system.

Other kind of analysis is positional, associated with the place that the institutions occupy in the network, the type of relations that they maintain and, by extension, the position of the Autonomous Communities to which these institutions belong. Normally, institutions in central positions in the network have a higher degree of relations. We can see this in the UCO, CCMA, UPC, and UPM cases. However, the numbers of relations considered are for all the studied period. Then, it is possible than only one work could be co-authored with a large number of institutions, but it does not mean that the relations must be permanent and present in every publication. Moreover, the publications with more visibility do not necessarily coincide with those of higher relational degree. However, the combined analysis between closeness and scientific production visibility indicates that 9 of the 10 institutions with the highest closeness degree overcome the world impact average in Agriculture. At the same time, we can identify clusters positioned in the graph institutions of the same Autonomous Community. In map number 4, some Autonomous Communities with relatively high concentration and closeness are seen. However, Andalucia or Madrid present institutions very dispersed in the map. In map number 3, we can observe the gap that the absence of institutions of Andalusia produces in the graph.

Map 4. Identification of cluster by autonomous comunity



ERROR: ioerror
OFFENDING COMMAND: image

STACK: