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


**Authors:**

Zaida Chinchilla-Rodríguez – zaida.chinchilla@cchs.csic.es

Félix de Moya-Anegón – felix.demoya@cchs.csic.es

Benjamín Vargas-Quesada – benjamín@ugr.es

Elena Corera-Álvarez – ecorera@ugr.es

Yusef Hassan-Montero – yusef_hassan@yahoo.es

**Institution:**

Unidad Asociada Grupo SCImago, IPP- CSIC. Universidad de Granada. Departament of Library and Information Science. Campus of Cartuja. Colegio Máximo of Cartuja s/n 18071 Granada (Spain)


## 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[1]. 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] 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[2] that scientific collaboration became a state affair, a stepping stone toward the goals set forth in the Lisbon Strategy[3]. 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 based on co-authorship of scientific documents do indeed

[2] 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.

3 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.

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 (Shrum, and Mullins, 1988; Glänzel and Schubert, 2004), 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; Perianes, 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) [4], 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[5] 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

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

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

loading, modelling and treatment of data. The database contained the following 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) [6] 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

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

entails 26 great areas. The present paper focuses on one of them, 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

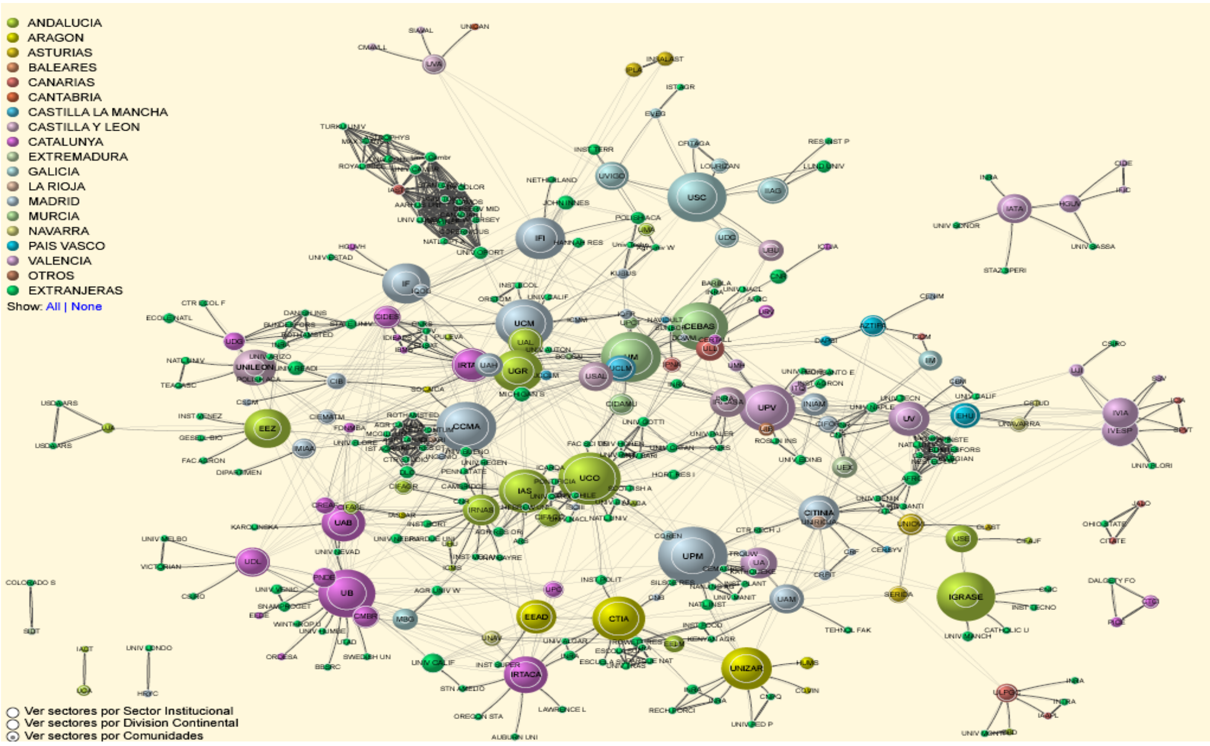
The techniques of graphical representation for the generation of maps are based on social networks. The standard matrix is processed using Pajek [21] and the algorithm of Kamada-Kawai is then applied (Kamada and Kawai, 1989) to locate the institutions in the graph based on its joint copublications. Once 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.

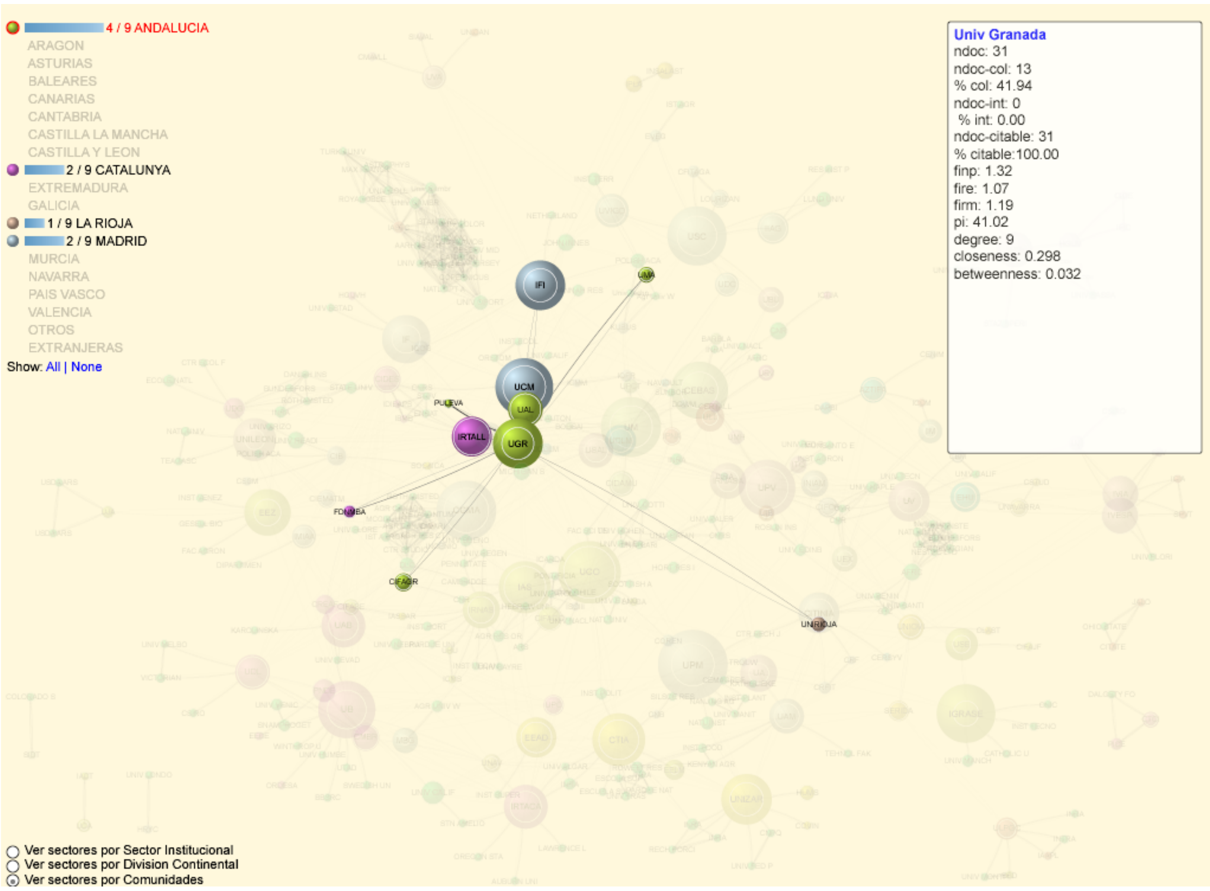
Map 1. Inter-institutional collaboration network of Spanish Agriculture (1995-1996)



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.

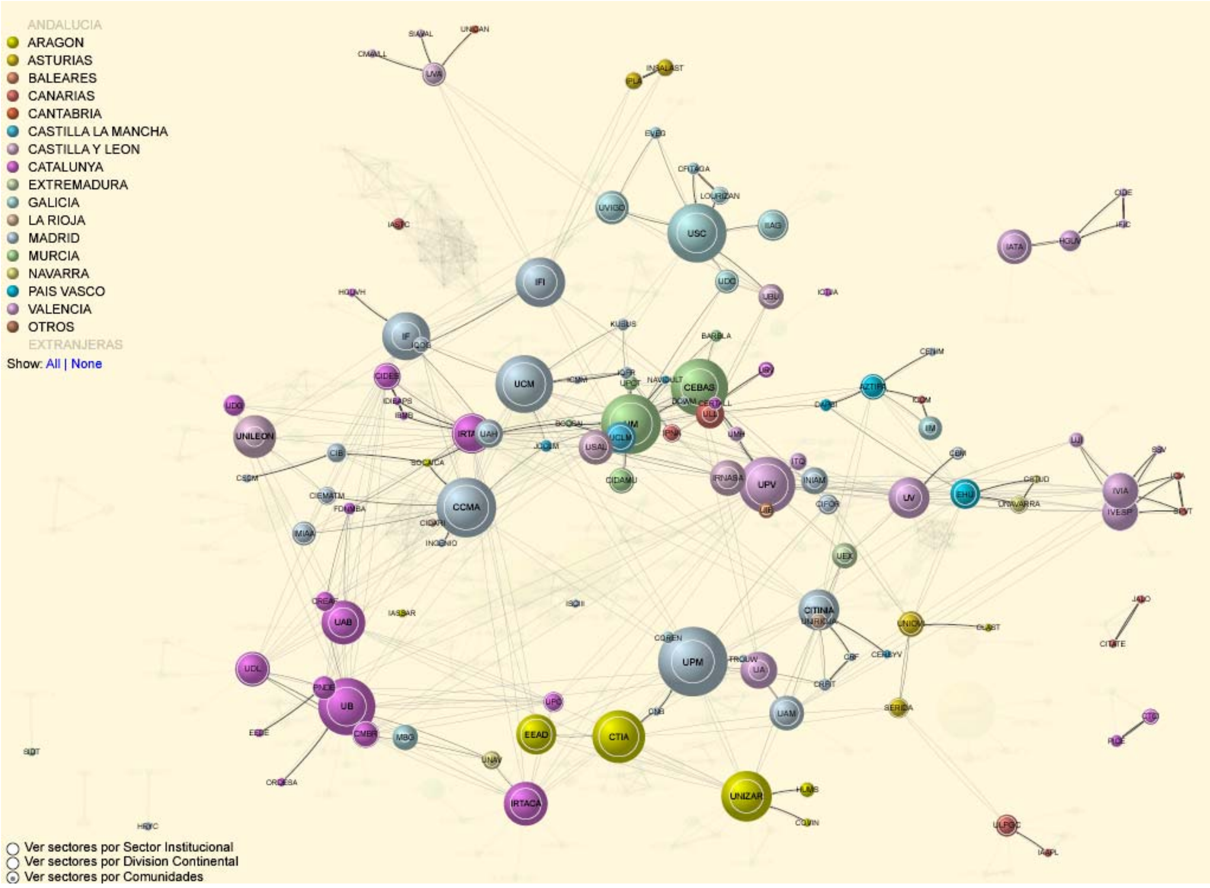
Meanwhile, placing the cursor on a node, we can see its relations with the other through a visual discrimination effect. Moreover, we can see in map number two: the nodal grade (number of links) for every level of aggregation (legend in the left higher part of the map) and institution as well as the 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.

Map 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

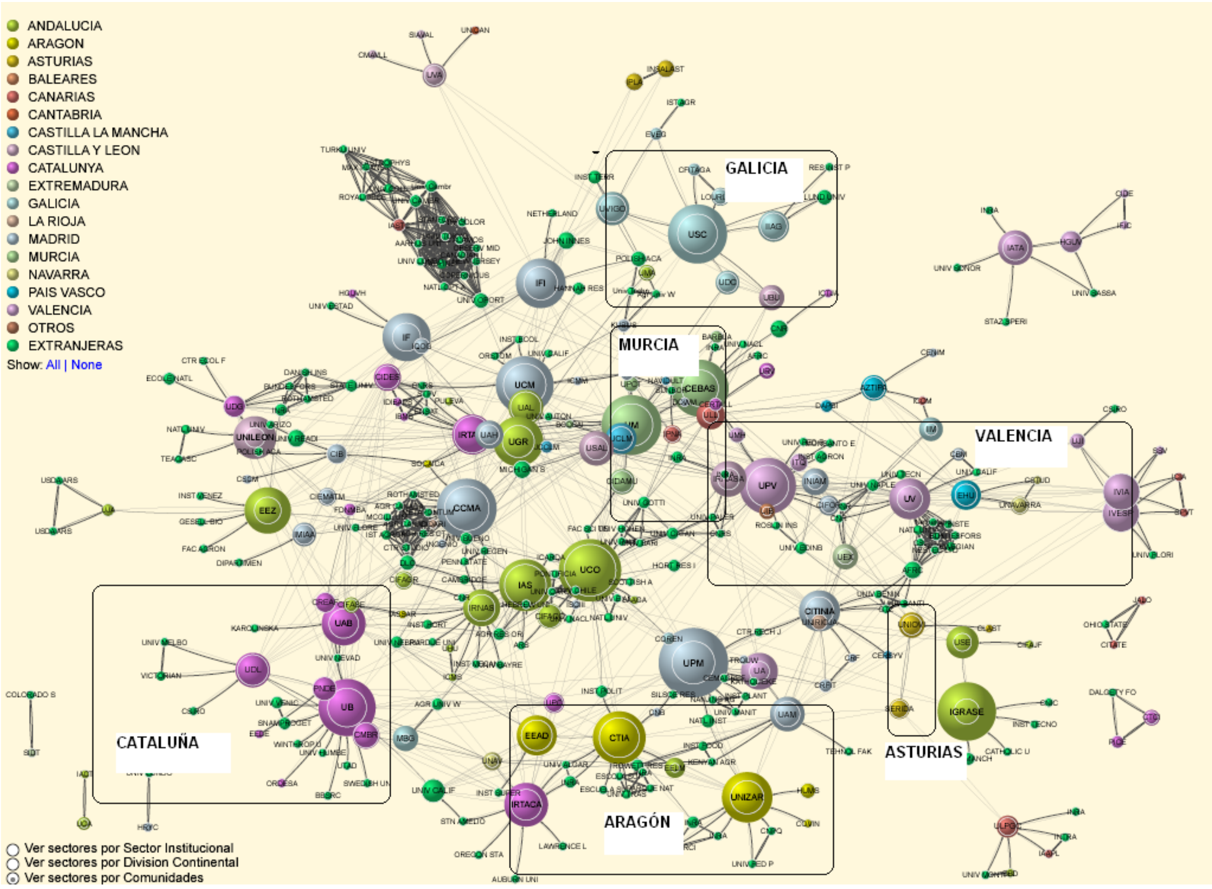
In the case of Spain, a study of the Autonomous Communities is perfectly justified because the scientific policies set forth a set of common conditions for the institutions in the domain of the same Administration. If we want to comparatively assess the effects of these policies, it is necessary to separate 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

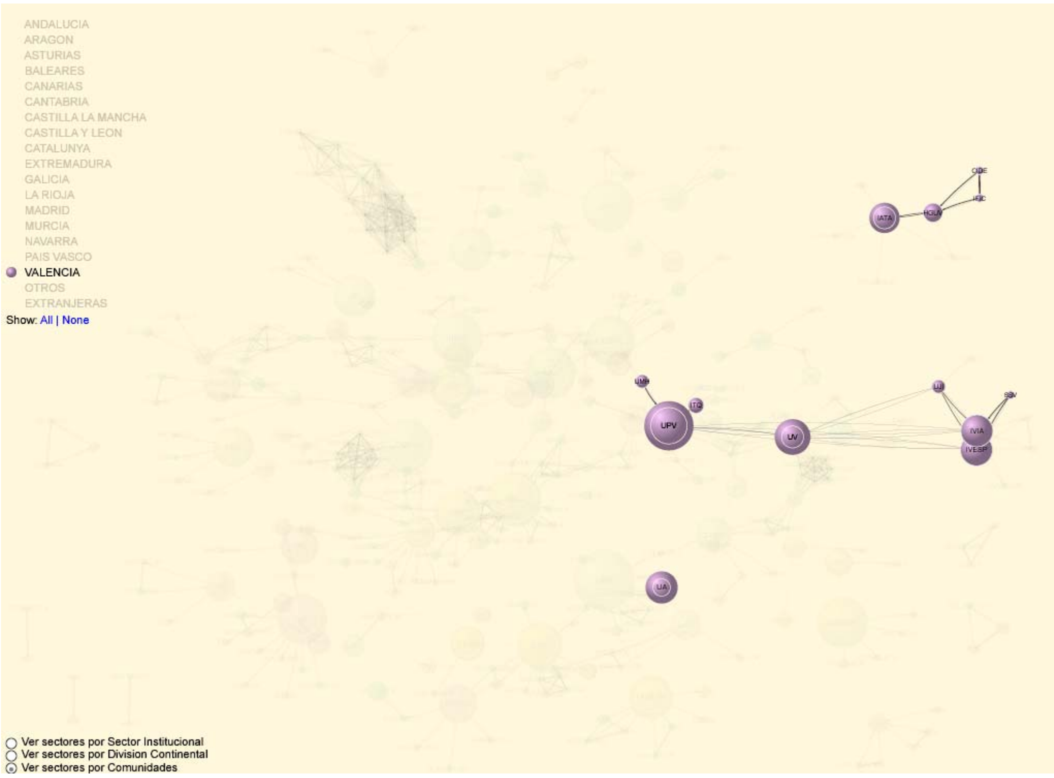


On the other hand, we can ask if the map has a centre-periphery structure. In this kind of network, we can identify a core of actors closely related and a periphery with disperse actors and weakly linked. Spanish Agriculture has not only one core, but several more whose actors are very related and occupy strategic positions (UCO and UPM with the rest).

Moreover, we can ask if there is some relation between production volume and collaboration (Melin and Persson, 1996; Melin, 1999). In the case of Spanish output this relation appears (Chinchilla and Moya, 2007) and, concretely in Agriculture, we can observe that the smaller communities (La Rioja, Castilla La Mancha, Canarias) collaborate more than the bigger ones, in which more human resources and infrastructure potential exist, and in which collaboration is hopefully the pattern. Furthermore, we can appreciate this patter through the sphere and concentric circles volume in the graph.

Finally, it is logic to think that network position is related with the different research fronts. For example, Valencia region position indicates that there are three possible thematic areas in Agriculture (map 5). Moreover, this pattern can explain the distribution of institutions of the same sector or Autonomous Communities in the graph, as we said before. Future works will follow this research line in order to identify the research fronts in a discipline and to complement the analysis.

Map 5. Institutions of an autonomous community (Valencia)



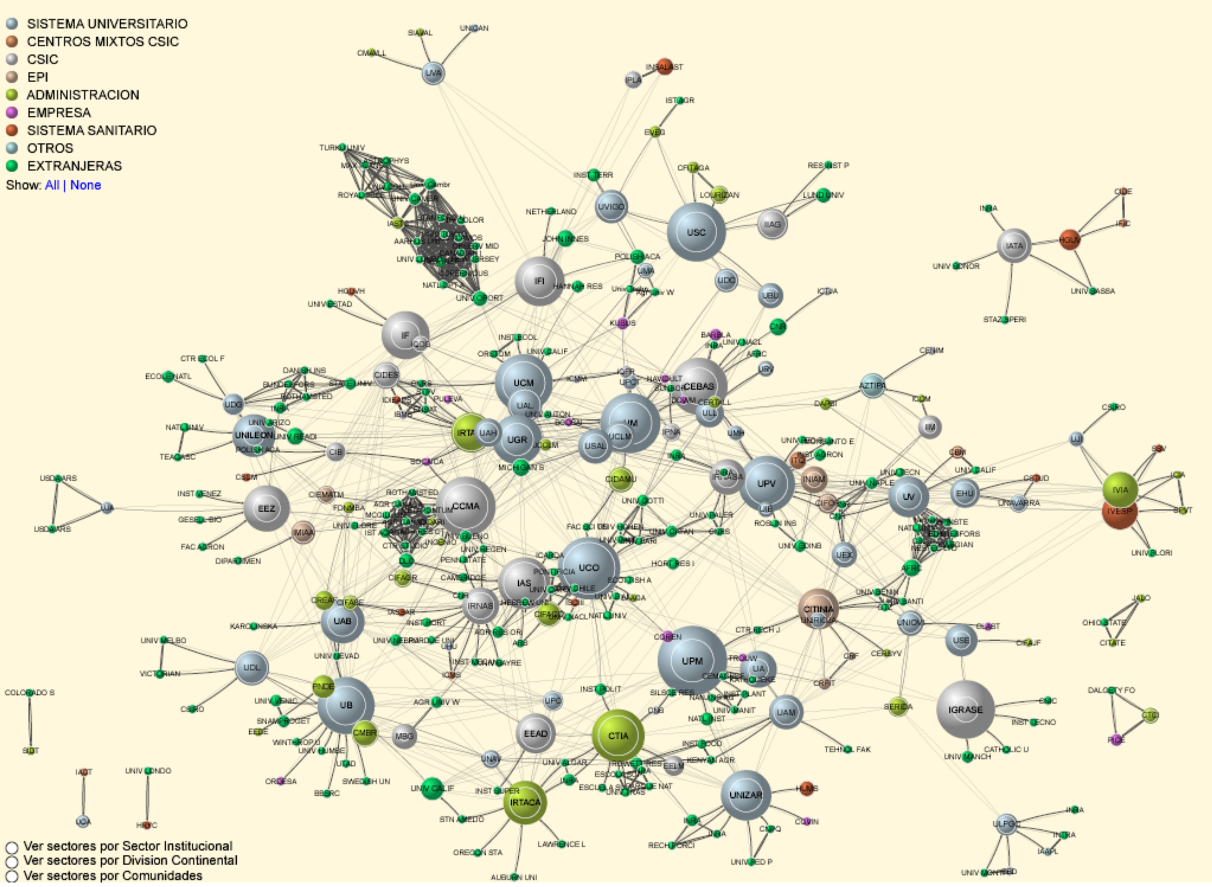

### 7.4. Network analysis by institutional sector

The transfer of knowledge between all the productive sectors is becoming an essential question. This transfer is interesting at the intra-organizational and inter-organizational level because it creates more social capital. In this scenario, the associations between the institutions in different sectors play a more important role in the production of relational social capital. That is because the transfer of knowledge between sectors normally ends in innovation processes (and this is not so frequent in the transfers inside the same sector). The theoretical models than support these arguments are known as Triple Helix.

With this idea, we can define sector degree as the indicator of the inbreeding character of an institution or, on the contrary, of its capacity to export knowledge. Moreover, we can investigate the effect of this indicator in the visibility of the institutional research. In the case of Agriculture in Spain, it seems that the Administration is the sector with the most heterogeneous collaboration, that is, with higher sector degree. Moreover, they have a strategic position in the net through their relations with the more productive actors, and they are some of the institutions with higher values of impact. In this situation, we can show the CIFA institutions (Centros de Investigación y Formación Agraria) and the IRTA institute (Institut de Recerca i Tecnologia Agroalimentàries) that maintain a relation with the big producers as the Universidad de Cordoba or the Universidad Politécnica de Madrid.

In the case of the universities, we can see that this sector is the more important due to the number of institutions and the volume of documents. The Universities with more production (UCO and UPM) are the institutions with more sector degree with foreign institutions (their production oscillates between 15 and 20%, respectively) or, in the case of the Universidad de Murcia, with higher collaborations with firms.

Map 6. Inter-institutional network of collaboration – Analysis by institutional sector

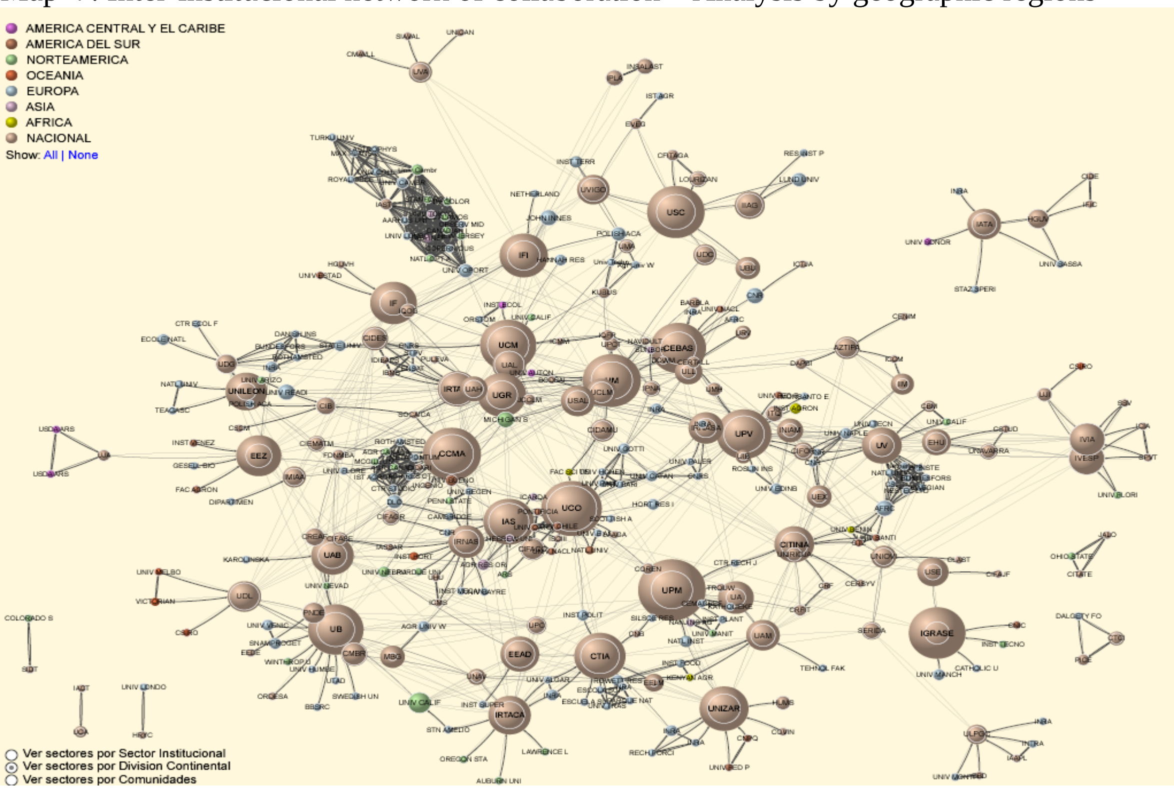


The business firms sector is scarcely present in Spanish Agriculture. They represent less than 4% of the total institutions with production. At firs, we can think that Research Public Centres, as the INIA or Consejo Superior de Investigaciones Científicas institutes, collaborate more with the firms due to their specialised character in the field. However, this is not the case for the agriculture in this period. Only the CEBAS (Centro de Edafología y Biología Aplicada del Segura) has a relation with a firm. The relations are basically produced with universities. That is proof of the scarce transfer of knowledge from the university to industry. Is it possible an evolution toward higher university-firm collaboration? Future works will analyze the effects across a time period, comparing the behaviours in different years.

Another kind of analysis allows us to study the effect of the closeness between institutions on the visibility of the scientific production. In the case of Agriculture, of the 10 institutions with a higher closeness degree, nine of them surpass the average world impact.

### 7.5. Network Analysis by continental division

Map 7. Inter-institucional network of collaboration – Analysis by geographic regions




The increase of the international collaboration in research and development is explained by direct profit (mainly the scientific excellence) and the indirect ones (associated with the political, social and economic benefits of the involved countries). (Georghiou, 1998). In the case of the Agriculture, Spain has a higher collaboration level with Europe, North America and South America, in terms of number of institutions and volume of productions of these. However, Africa together with Central America and Oceania are less frequent partners in the Spanish network of collaboration.

We may hold that the effort of the European Commission by means of the Framework Programmes to integrate more number of countries is one the elements that would explain the main collaborators with Spain in Agriculture. If we could introduce evolutionary information, we could monitor the efforts of the scientific policies to increase the number of countries of different regions in the network. Moreover, we want to know why several regions are outside of the circles of relations of Spanish Agriculture. It can be explained by the thematic specialization of the country or by "institutional imbalances". The latter are mainly related with the different research priorization???? in the scientific policy agendas of the governments.

On the other hand, the international collaboration of Spain is produced basically with academic institutions and it could be associated to the scarce economic competitiveness of this productive sector. There is almost no relation with the firms at national and international level.

Regarding the relation between collaboration and scientific visibility: In the Spanish case we can say that although the higher values of internationalization of the impact are produced with national collaborations, this is not a normal phenomenon in the patterns of publication. The international relations are normally an additional value that is favourable to the visibility and relevance on the international scientific community.

There are works in the literature where the international collaboration to the individual level is analysed. Concretely, the work of SCImago group as a complementary tool for the lecture of the national collaboration networks will allow us to identify the main geographical axis, to show the relations of the analysed domain with other countries (and also the countries with whom there is more collaboration, at which level and what the effects of these relations on the visibility are depending on the types of collaboration).

## 8. Conclusions

Scientific collaboration is a major aspect of the process of generating knowledge, and in divulging and making use of its results. With the understanding that the worldwide aim is to maximize resources, establish contacts and networks, increase productivity and visibility, we must not underestimate the importance of collaborative efforts. Whether crossing academic frontiers or political, economic or social borders, collaboration in scientific endeavours is a sign of maturity and efficiency, evidencing ripe branches of knowledge, of infrastructure, and of coordinated participants.

At the same time, analysis of scientific collaboration provides relational information that complements the conventional bibliometric register, resorting to social network theory within the Domain Analysis paradigm (Horjland and Albretchsen, 1995). Such a focus combines bibliometric data with the schematic representation of the domain, so that one may view the flow of information and interpret it in intuitive fashion, or even re-arrange it on an alternative level of aggregation (regional, national, thematic, institutional, individual, etc). Furthermore, its graphic depiction allows it to be used as an interface for information retrieval.

The use of institutions as units of measure can be justified on the basis that, although research is carried out by scientists per se, they are necessarily associated through administrative structures –that is, institutions. These bodies occupy a central position in the system of science and technology, and play a key role in the emergence, consolidation, specialization, visibility and expansion of research findings and documented output. Identifying the topological structure of active institutions and the position they occupy within that structure will provide crucial new evidence of the potential for receiving knowledge or transmitting knowledge to other institutions.

This is a truly useful prospect for scientific policy-makers, lending them the opportunity to harvest the potential of a collective force in the distribution and absorption of recent knowledge, thereby compensating for limited resources in other terms (as would be the case of the SECYT)[7].

This type of institutional analysis makes it possible to discern patterns of collaboration among different subject areas. We can analyze, for instance, how the public or private nature of an institution may affect output and visibility, or look at the diversity of these institutions in productive sectors; and the type of inter-institutional networks that are created (dense or fragmented) give localized information regarding the rules that guide the generation of scientific knowledge, and how it can best be managed by political means.

Meanwhile, an analysis of systems helps identify the axes and key structures within networks, leading us to raise questions as to the major areas of specialization and their distinctive role in collaboration, or which particular institutions are behind the countries with outstanding production, and what sort of policies they use to forment research collaboration. The graphic representations clearly show which particular features of the network (density, cohesion or centrality) are involved in innovative processes (social capital)[8], revealing clusters of actors with strategic positioning in the institutional context of knowledge generation.

It is our belief that this tool will prove useful for researchers and policy-makers alike. More specifically, it may be used to decide upon strategic scientific alliances on the basis of domain receptiveness, or the effects of geographical proximity in terms of production and visibility. It yields social as well as scientific benefits. Socially, it aids in the evaluation of systems of science, technology and innovation, and of related policies and decisions. Scientifically, it facilitates comparative studies that are greatly needed to comprehend and emulate good practice.

Finally, its application as an interface and tool of analysis at diverse levels of aggregation makes it appropriate for use as a prototype of the Informational System at the core of the project “The Atlas of Science” (Moya, et.al., 2004) developed by the SCImago Group (http://www.scimago.es).

Future lines of work might delve further into the dynamic nature of these networks so as to analyze whether there is a transformation of relational patterns over time, or a reinforcement of the central-peripheral structure in the network of collaboration takes

[7] We can refer to national R+D plans or *Planes Nacionales* from number IV (2000-2003) to the present, within European Program. For instance, since the *IV Plan Nacional de Investigación y Desarrollo* (2000-2003), coordination of public programs and business initiatives has been promoted, to establish better ties among research groups and the business sector, while favouring internationalization of R+D activities.

[8] Cook and his collaborators described, in 1997, the creation of relational social capital, tied to the generation of innovation, using resources available on the Web.

place; or, contrariwise, there is a tendency for multiple nuclei to appear, with a less hierarchic distribution of the flux of scientific communication.

Such an analysis would ideally be complemented with further data concerning the context and reliability of new knowledge. Aside from integrating different methodologies, we need to broaden the array of network indicators and the levels of analysis. For example, data on competitive funding of projects, human resources, and the evaluation of results would be immensely useful.

Additionally, we should be able to analyze the scientific panorama while maintaining a fixed position of the main actors, as the dynamics of collaboration may have an effect on the size of the node. Such a development would allow us to bring many other variables into view, such as citations per document, impact factor relative to the domain, H-index, citable documents, number of documents produced under different types of collaboration, and other aspects perhaps unforeseen at present, all thoroughly contextualized within the universe of scientific collaboration and production.

## 9. Bibliography

Barabasi, A; Jeong, H; Néda, Z; Ravasz, E; Schubert, A; Viscseck, T. Evolution of the social network of scientific collaborations, 2001, Arxiv 1 (Web Page).

Beaver, D. Reflections on Scientific Collaboration (and its study): Past, Present and Future. Scientometrics. 2001; 52(3):179-199.

Bordons, M. and Gómez, I. Collaboration Networks in Science. Cronin, B. and Atkins, H. B. Eds. The Web of Knowledge: A Festschrift in Honor of Eugene Garfield. Medford: Information Today; 2000; pp. 197-213.

Börner, K., Chen, C., and Boyack, KW. Visualizing knowledge domains. Annual Review of Information Science and Technology, 2003; 37: 179-255

Boschma, R. A. Proximity and Innovation: A Critical Assessment. Regional Studies, 2005; 39(1):61-74

Buesa, M; Heijs, J; Asheim, B; Baumert, T; Navarro, M; Martínez, M. Sistemas regionales de innovación: nuevas formas de análisis. FUNCAS, Madrid, 2007.

Chinchilla, Z y Moya Anegón, F La investigación científica española (1995-2002): Una aproximación métrica. Universidad de Granada, Granada (2007).

Chinchilla-Rodríguez, Z. y Moya-Anegón, F. (2007) La Investigación Científica Española: Una Aproximación Métrica. Granada: Universidad de Granada

Cook, P; Gomez Uranga, M; Etxebarria, G. Regional innovation systems: Institucional and organisational dimensions. Resaerch Policy, 1997; 26: 475-491.

David, P y Foray, D. Accessing and expanding the science and technology knowledge base”. Science, Technology and Industry, 1995; 16: 13-68.

De Bruin, R. E. and Moed, H. F. (1990), The unification of addresses in scientific publications. In: L. Egghe, R. Rousseau (Eds), *Informetrics 1989/90*. Elsevier Science Publishers, Amsterdam, pp. 65–78.

De Bruin, R. E. and Moed, H. F. Delimitation of scientific subfields using cognitive words from corporate addresses in scientific publications. Scientometrics. 1993; 26(1):65-80.

Dumont, M., and Meussen, W. Some results on the graph-theoretical identification of micro-clusters in the Belgian innovation system. OECD/DSTI Workshop on Cluster Analysis and Cluster-Based Policies, Amsterdam.

Etkowitz, H y Leydesdorff, L. Universities and theGlobal Knowledge Economy- A triple Helix of University-Industry-Government Relations. Pinter, London (1997).

Fernández, M. T.; Cabrero, A.; Zulueta, M. A., and Gomez, I. Constructing a Relational Database for Bibliometric Analysis. Research Evaluation. 1993; 3(1):55-62.

Fernandez, M. T.; Gomez, I., and Sebastian, J. Scientific Cooperation of Latin-American Countries Through Bibliometric Indicators. Interciencia. 1998; 23(6):328-337.

Fernández, M.T., Cabrero, A., Zulueta, M.A. y Gomez, I. (1993) Constructing a Relational Database for Bibliometric Analysis. Research Evaluation 3, 55-62.

Fry, J. (2006) Scholarly Research and Information Practices: A Domain Analytic. Information Processing & Management, 42 (1), 299-316

Fundación Española de Ciencia y Tecnología. Portal de acceso a la Web of Knowledge. Avaible in:http://www.accesowok.fecyt.es/

Gálvez Martínez, C. and Moya Anegón, F. The unification of institutional addresses applying parametrized finite-state graphs (P-FSG), *Scientometrics*, 2006; 69 (2): 323–345.

Gálvez Martínez, C. and Moya Anegón, F. Standardizing formats of corporate source data. Scientometrics. 2007; 70(1):3-26.

Georghiou, L. Global Cooperation in research. Resarch Policy, 1998; 27: 611-626

Glänzel, W. International Scientific Collaboration in a Changing Europe. A Bibliometric Analysis of Co-Authorship Patterns of Eleven East Central European Countries in the 90s. Second Berlin Workshop on Scientometrics and Informetrics: Collaboration in Science and Technology; 2000.

Glänzel, W. National Characteristics in International Scientific Co-Autorship Relations. Scientometrics, 2001; 51(1): 69-115.

Glänzel, W., and Schubert, A. Analyzing scientific networks through co-authorship. In H.F. Moed et. al. (Eds.), Handbook of QuantitativeScience and Technology Research, (pp.257-276), The Netherlands: Kluwer Academic Publishers, 2004

González Alcaide, G., Alonso Arroyo, J., González de Dios, A. P., Sempere, J. C., Valderrama Zurián, J. C., Aleixandre Benavent, R. Redes de coautoría y colaboración institucional en Revista de Neurología. Revista de Neurología, 2008, 46: 642-651.

González Alcaide, G., Alonso Arroyo, J., González de Dios, A. P., Sempere, J. C., Valderrama Zurián, J. C., Aleixandre Benavent, R. Redes de coautoría y colaboración institucional en Revista de Neurología. Revista de Neurología, 2008, 46: 642-651

Hamers, L., Hemeryck, Y., Herweyers, G. Janssen, M., Keters, H., Rousseau, R., and Vanhoutte, A. (1989) Similarity Measures in Scientometric Research – The Jaccard Index versus Salton Cosine Formula. Information Processing & Management, Vol. 25, No. 3, 315-318

Heimeriks, Hörlesberger y Van Der Besselaar. Mapping communication and collaboration in heterogeneus research network. Scientometric, 2003; 58(2): 391-413.

Katz, J. S. and Hicks, D. Desktop Scientometrics. Scientometrics. 1997; 38(1):141-153.

Katz, J. S. and Martin, B. R. What Is Research Collaboration. Research Policy. 1997; 26(1):1-18.

Katz, J.S. y Hicks, D. How Much Is a Collaboration Worth? A Calibrated Bibliometric Model. Scientometrics, 1997; 40, (3): 541-554.

Kretschmer, H. (1993) Measurement of Social-Stratification - A Contribution to the Dispute on the Ortega Hypothesis. Scientometrics. 26(1):97-113.

Kyvik, S. y Larsen, I. M. (1994) International Contact and Research Performance. Scientometrics 29(1):161-172.

Laranja, M; Elvira, U; Flanagan, K. Policies for Science, Technology and Innovation: Traslating Rationales into Regional Policies in a Multi-level Setting. Manchester Business School Working Paper, Number 527, avaible: http://www.mbs.ac.uk/research/workingpapers/

Lewison, G. y Cunningham, P. (1991) Bibliometric Studies for the Evaluation of Trans-National Research. Scientometrics, 2004; 21, (2):223-244

Lukkonen, T; Tijseen, Persson, O; Silverstsen, G. The measurement of International Scientific Collaboration, Science, Technology and Human Values, 2003; 28(1): 15-36.

Mählck, P y Persson, O. Socio-bibliometric mapping of intra-departamental networks. Scientometrics, 2000; 49(1): 81-91.

Maltras, Bruno. Los indicadores bibliométricos: fundamentos y aplicación al análisis de la ciencia. Trea, 2003.

Melin, G. and Persson, O. Studying Research Collaboration Using Co-Authorships. Scientometrics. 1996; 36(3):363-377.

Melin, G. Impact of National Size on Research Collaboration - A Comparison Between Northern European and American Universities. Scientometrics. 1999; 46(1):161-170.

Moed, H. F.; Burger, W. J. M.; Frankfort, J. G., and Van Raan, A. F. J. Moed, H. F. The Use of Bibliometric Indicator for the Assessment of Research Performance in the Natural and Life Sciences. Leiden: DSWO Press; 1989.

Molina, J; Muñoz, M; Doménech, M. Networking Scientific Publications; An analysis of the co-autorship Structure. International Sunbelt Social Network Conference, 2000.

Moravsick, M. J. ¿Cómo evaluar a la Ciencia y a los Científicos? Revista Española De Documentación Científica.1989; 12:313-325.

Moya-Anegón, F., Vargas-Quesada, B., Herrero-Solana, V., Chinchilla-Rodríguez, Z., Corera-Álvarez, E., Muñoz-Fernández, F. J. A new technique for building maps of large scientific domains based on the cocitation of classes and categories. Scientometrics, 61 (1):129-145, 2004

Moya-Anegón, F., Herrero-Solana, V., Vargas-Quesada, B.,Chinchilla-Rodríguez, Z., Corera-Álvarez, E., Muñoz-Fernández, F. J., Guerrero-Bote, V. , and Olmeda-Gómez, C. (2004) Atlas de la Ciencia Española: propuesta de un sistema de información científica. Revista Española de Documentación Científica, 27 (1): 11-29. Avalaible at: http://www.atlasofscience.net

Moya-Anegón, F., Chinchilla-Rodríguez, Z., Corera-Álvarez, E., Herrero-Solana, V., Muñoz-Fernández, F., Vargas-Quesada, B. (2005). Indicadores Bibliométricos de la Actividad Científica Española - 2004. Madrid: Fecyt

Moya-Anegón, F., Chinchilla-Rodríguez, Z., Vargas-Quesada, B. and González-Molina, A. Visualization of networks of international collaboration. I International Conference on Multidisciplinary Information Sciences & Technologies, Mérida, Spain, 25-28 October. 2006.

Moya-Anegón, F., Chinchilla-Rodríguez, Z., Corera-Álvarez, E., Herrero-Solana, V., Muñoz-Fernández, F., Vargas-Quesada, B. (2007). Indicadores Bibliométricos de la Actividad Científica Española – (1990-2004). Madrid: Fecyt

Moya-Anegón, F., Chinchilla-Rodríguez, Z., Corera-Álvarez, E., González-Molina, A. y Vargas-Quesada, B. (2007) Galicia: Patrones de Colaboración Científica (ISI-WOS, 2004). Santiago de Compostela: Xunta de Galicia available at: http://www.bugalicia.org/adjuntos/estudios/Borrador_PatronesColaboracionCientificaGalicia_ISI-WOS_2004.pdf

Moya-Anegón, F., Chinchilla-Rodríguez, Z., Corera-Álvarez, E., González-Molina, A., Hassan-Montero, Y., Vargas-Quesada, B. Indicadores Bibliométricos de la Actividad Científica Española: 2002-2006. Madrid: FECYT (2008).

Narin, F., Stevens, K. y Whitlow, E.S. Scientific Cooperation in Europe and the Citation of Multinationally Authored Papers. Scientometrics, 1991; 21, (3):313-323.

Newman MEJ. Scientific collaboration networks. I. Network construction and fundamental results. Physical Review E. 2001;64(1):016131.

Newman MEJ. Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. Physical Review E. 2001;64(1):016132.

OCDE. Manual de Frascati: Propuesta de Norma Práctica para Encuestas de Investigación y Desarrollo Experimental. Paris: OCDE; 1993.

Okubo, Y.; Miquel, J. F.; Frigoletto, L., and Dore, J. C. Structure of International Collaboration in Science - Typology of Countries Through Multivariate Techniques Using a Link Indicator. Scientometrics. 1992; 25(2):321-351.
Otte, E., and Rousseau, R. Social network analysis: a powerful strategy, also for Information sciences. Journal of information science, thousand oaks, 2002; 28 (6): 441-453

Perianes-Rodriguez, A. Analysis and Visualization of Networks Applied to the Study of Scientific Collaboration in Research Groups: Carlos III University of Madrid (ISI, Web of Science, 1990-2004). [PhD Dissertation]. Madrid: Carlos III University, 2007

Rinia, Ed J. Scientometrics Studies and their Role in Research Policy of Two Research Councils in the Netherlands. Scientometrics, 2000; 47(2):363-378.

Russell, J. Publication Indicators in Latin America Revisited. Cronin, B. y Atkins, H. B. Eds. The Web of Knowledge: A Festchrift in Honor of Eugene Garfield. Medford: Information Today; 2000.

Sancho Lozano, R. Medición de las Actividades de Ciencia y Tecnología. Estadísticas e Indicadores empleados. Revista Española De Documentación Científica 2001; 24(4):842-865.

Sanz Menéndez, L. and Cruz-Castro, L. Explaining the Science and Technology Policies of Regional Governments. Regional Studies, 2005; 39(7):939-954

Sanz Menéndez, L. Indicadores relacionales y redes sociales en el estudio de los efectos de las políticas de ciencia y tecnología. II Taller de Obtención de Indicadores Bibliométricos y de Actividad Científica. Cyndoc-Rycit (2000).

SCImago Group. Manual de criterios y procedimiento para la normalización, control de calidad y análisis sectorial de las instituciones españolas incluidas en las bases de datos de *Thomson Scientific*. Granada: Grupo SCImago – Universidad de Granada; 2006; Working Paper 2006-01.

Scott, 2000 J. Scott, Social network analysis: A handbook, Sage Publications, London (2000).

Shrum, W., and Mullins, N. (1988). Network analysis in the study of science and technology. In: Vann Raan, AFJ. Handbook of quantitative studies of science and technology, 1988 North-Holland, Elsevier.

Subramanyam, K. Bibliometric Studies of Research Collaboration: A Review. Journal of Information Science. 1983; 6(35).

Wagner, C. S. and Leydesdorff, L. Network Structure, Self-Organization and the Grow of International Collaboration in Science. Research Policy. 2005; 34(10):1608-1618.

Wang, Yan; Wu, Yishan; Pan, Yuntao; Ma, Zheng, and Rousseau, Ronald. Scientific Collaboration in China as refected in co-authorship. Scientometrics. 2005; 62(2):183-198.

Wasserman and Faust, 1994 S. Wasserman and K. Faust, Social network analysis: Methods and applications, Cambridge University Press (1994).

White, HD. (2003) Pathfinder networks and author cocitation analysis: a remapping of paradigmatic information scientists. Journal of the American Society for Information Science and Technology, 2003; 54(5):423-434

Yamashita, Y., and Okubo, Y. (2006) Patterns of Scientific Collaboration between Japan and France: Inter-sectorial Analysis using Probabilistic Partnership Index (PPI). Scientometrics, 2006; 68 (2): 303-324