RESEARCH FRONTS IN LIBRARY AND INFORMATION SCIENCE IN SPAIN (1985-1994)

F. DE MOYA ANEGÓN,[+] E. JIMÉNEZ CONTRERAS,
MERCEDES DE LA MONEDA CORROCHANO

Departamento de Biblioteconomía y Documentación,
Universidad de Granada, Campus de Cartuja, 18071 Granada (Spain)

Publications and author cocitations in library and information science in Spain during the period from 1985 to 1994 were analyzed as a measure of the structure, specificity and composition of research fronts in this country. A cocitation matrix developed from an ad hoc database was subjected to cluster analysis, multidimensional scaling and principal components analysis. The resulting cocitation maps identified specific areas of research and their knowledge bases. We inferred the degree of consolidation of the discipline of library and information science, and of the subdisciplines informetrics, librarianship and university affiliation, from the research activities revealed. In this respect, the conclusions from the study show the existence of several research fronts in Spanish literature the contents of which are in most cases difficult to compare with those in other countries. A lesser degree of maturity of research in this field is shown.

Introduction

Most researchers are familiar with the authors, specialties, and literature produced in their own field; however, the conceptual and methodological framework that makes it possible to objectively elucidate the structure of different fields of science has been available only since the 1960s. Analyses of these concepts and methods have been aimed at tracing the development of scientific activity and improving its efficiency. The first half of the 1960s saw the establishment of a nexus between the network of relations among scientific papers through references and citations, and the structure of specialities within a given field (*Price*[1]). This relationship led to the development of concepts such as research fronts, invisible colleges, and science maps; current data representation techniques based on multivariate analysis for reducing the number of

---

[+] To whom correspondence should be addressed. E-mail: felix@goliat.ugr.es

dimensions in a graphic display now make it possible to draw precise conclusions on the structure of and relationships between scientific specialties.

According to *Price* (1965), an area that is highly active during a certain period reflects this activity in the references cited in papers published during this period. This makes it possible to identify so-called research fronts within the discipline by analyzing the network of citations. Areas showing the greatest density (i.e., agglomerations of points on the map) are associated with research fronts (*Braam*[2]). This idea has two favorable consequences. Research fronts are identified with specific areas of research within a discipline; in other words, fields of activity within one branch of science are identified by the references researchers use in a given article. Moreover, research fronts are not defined by one person alone, but are identified by the sum of the contributions of the set of authors within the discipline. This considerably offsets the biases that can arise from partial analyses of individual authors. Many of the criticisms aimed at citation analysis (*MacRoberts*[3,4]) have pointed out that a citation is the result of a decision made by an individual, or at best, by a small group of individuals. However, the overall view provided by the references in papers from a given discipline ensures that this image of its structure has some measure of objectivity.

Several attempts to identify zones of high citation density have been described. One of the most frequently used approaches was developed simultaneously in the early 1970s in the United States (*Small*[5]) and the former Soviet Union (*Marshakova*[6]). Both studies were based on the principle that the number of times two papers are cocited correlates directly with likelihood that they share affinities in their subject matter. This idea later formed the basis of three approaches to citation analysis: paper, author, and journal cocitation. The bibliometric model of cocitation groups authors, articles or journals on the basis of subject similarities, and also makes it possible to measure the interaction between different fronts, thus identifying a hierarchy of interrelated specialties (*Franklin*[7]).

The first author cocitation studies from the early 1980s (*White*[8,9]) suggested that author cocitation analysis (ACA) could shed light on the intellectual structure of a given discipline. As in previous studies of article cocitation, ACA is used to draw maps of the discipline that show groupings of authors at different fronts. These maps are based on the principle that the number of cocitations between any two authors is inversely related to the distance (similarity) between these authors in a two-dimensional display. The inverse proportion is subjected to mathematical treatments to calculate the coordinates of each author on the map from the relative distances between different authors, and to illustrate groupings of authors at different specialty fronts. Although such cocitation maps contain a degree of mathematical artifact, these artifacts are none

230

the less considered useful (*Rip*[10]). Cocitation maps have been shown to be an effective tool for calculating the dependent variable in studies designed to identify the factors that influence the increase in knowledge or the structure of scientific communities in a specific field, and to determine the impact of science policies on the dynamics of scientific development.

Each cluster of authors on a cocitation map consists of a visible and an invisible component. The former comprises authors that are intensely cocited in a specific research area; the more they cocite each other, the closer together they appear. The latter comprises citing authors who form part of the front to the degree that they use the members of the visible group as the intellectual base for their own research. Thus the intellectual structure of a given discipline appears on the cocitation map as all participating authors, although as we will describe below, the representation provided by the map is an aid to further analyses that needs to be interpreted in the light of characteristics peculiar to a given discipline. In some cases cocitation analysis has been used to corroborate previous individual analyses (*White*[11]).

Object

In this study we analyzed publications and author cocitations in library and information sciences (LIS) in Spain during the period from 1985 to 1994 as a measure of the structure, specificity and composition of research fronts in this country. In a previous paper we analyzed the author productivity in the Spanish LIS research (Jiménez[12]).

Material and methods

Material

In any cocitation analysis, the procedure used to select and treat the data conditions the results. The way in which the authors are chosen can introduce biases that are later untraceable; hence the decisions affecting methodological procedures need to be reported in detail (*Leydesdorf*[13]). The starting point for multivariate analysis is the cocitation matrix, which represents the intensity of cocitation between any two authors. The more often two authors are cited together, the closer together the topics they research appear in the eyes of the scientific community that cites (or declines to cite) them. Construction of the matrix is thus a key step in cocitation studies.

To construct the matrices, the most frequently cited authors must first be identified. Therefore a minimum number of citations received per author must be chosen; authors receiving fewer citations do not appear on the maps. The authors should be chosen in such a way that the resulting matrix contains as few zeroes as possible (*White*[14]). In some disciplines the cocitation relationship between authors is so strong that a second selection step is needed to keep the matrix from becoming too large.

We used a database constructed specifically for this study using as source material specialized Spanish journals and the proceedings of major national library and information science (LIS) conferences. Articles and their references were recorded for the scientific production in LIS fields during the period from 1985 to 1994. The specific journals and conferences that formed the source material are listed in the Appendix. The final database contained 1500 articles and 10000 references. Among the 2250 authors of citable items, we selected those who were most frequently cited and those who were most frequently cocited. Because many authors each published only a small number of the citable items, cocitation groupings were infrequent. We none the less were able to identify 53 authors who fulfilled two criteria: they received more than 15 citations in the database (preselection), and more than 60 cocitations with other authors who also satisfied the preselection criterion. These two cutoff points were chosen because they indicated the threshold below which the number of relationships that scored zero in the matrix increased considerably.

Earlier studies that used ACA were based on the Institute of Scientific Information (ISI) database. In the present study we used a database that included all authors in all references cited in the source items, instead of only the first author or the first three authors. This procedure allowed us to detect cocitations that would have been overlooked with databases that use *et al.* in the reference lists.

A number of searches were run to identify all cocitations received by the 53 selected authors, and a cocitation matrix was prepared with the results. Because the number of cocitations is directly proportional to the similarity between research topics, the value yielded by a given author's cocitation of him- or herself would be expected to be the highest cocitation value for this author. To obviate this problem, we assigned to this cell a value higher than any of those actually yielded by the matrix. This step is necessary only for analyses that require a symmetrical matrix. Different systems have been used to estimate the cocitation value for a given author with him- or herself (*Culnan*;[15] *McCain*;[16] *White*[8]). An appropriate way to determine this value is to add the three highest values for that author and divide the resulting sum by two. The resulting value is usually higher than the highest value from the matrix, and provides some assurance that a given author is more similar to him- or herself than to any other.

The next step was to transform the cocitation matrix into a correlation matrix of Pearson r correlations, which provided the entry values used in subsequent multivariate analyses. This transformation (White[8,14,11]) was used because of the need for normalized values in the matrix of similarities between authors in order to obviate the effects of scale caused by differences in the number of citations received, and because of the need to reduce the number of zero relationships between authors that results when there are no associated cocitations. Otherwise. the usefulness of the number of cocitations as a similarity factor is compromised, as this number is the same at a given degree of relationship for all pairs of authors who might be located at different levels because of their relations with others.

The similarity given by Pearson's r ranges from -1 (minimum) to $+1$ (maximum), and expresses the dependence between variables. The resulting matrix thus expresses the positive or negative similarity between each pair of authors. The values in diagonal cells are equal to $+1$, and express the highest possible similarity. Despite the apparent incompatibility between an r value of 1 in association with nonidentical values for the variables to be correlated (Bayer[17]), the use of Pearson's correlation coefficient in the present analysis is justified by our intention to relate authors on the basis of relationships between these authors and other citing authors, rather than between the two components of the pair.

Methods

Because our objective was to represent and analyze matrix data, we used multivariate analysis to reduce the number of dimensions and construct a geometric representation of the variables that provided a graphic image of similarities, differences and clusters (Cuadras[18]). The input data were obtained from the correlation matrix, and each variable (i.e., each author) was represented on the basis of his or her relations with other authors. A natural representation of these relationships would have been possible only in an n-dimensional space, where n is the number of authors in the map. To obviate this problem we generated cocitation maps with several different techniques commonly used in ACA to represent objects belonging to a many-dimensional space in a two- or three-dimensional space (Egghe;[19] Kinnucan[20]). These techniques, to the extent that they make it possible to use a reduced space, simplify the display and facilitate subsequent measurements and analyses. The distances between objects can be measured before and after reduction, and a comparison of the two values can be used as an indicator of the validity of the representation in reduced space (goodness of fit).

Statistical analyses

Cluster analysis. This technique is used to create a two-dimensional display (dendrogram) of clusters or classes of different objects whose relationships are represented by the matrix values. This type of automatic classification, also known as numerical taxonomy, currently comprises more than 150 different techniques that are grouped in families according to shared procedures. In general, the LIS discipline involves polythetic clustering hierarchies to create trees illustrating the hierarchy of relationships among authors on the basis of individual characteristics, with groups being defined on the basis of the entire set of characteristics of all individual authors (*Willet*[21]). These agglomerative techniques require at least two elements: the distance function and clustering rules. These two factors make it possible to group individual objects (each of which initially represents a separate class) into classes regardless of the initial distance between them. An iterative process is used to group the objects that are closest to each other, and to move to higher levels of the hierarchy to produce a branching display of the relationships. In the present study the similarity matrix was transformed into a distance matrix (1–*r*), thus the distance function required by the classification process is the complement of *r*. As the clustering rule we used the Ward method, as recommended by *Griffiths*,[22] *McCain*[16] and *Zitt*.[23] This method uses analysis of variance to evaluate the distance between clusters, and is intended to reduce the sum of the squares of the distances between the clusters formed in each iteration. This approach can be assumed to create small clusters, and to give rise to a number of finely branching bushes rather than a single tree with a few main branches.

Our purpose in using this approach was to create clusters of authors who represent different research fronts that could be further analyzed with the procedures described below.

Principal components analysis. The basic premise of principal components analysis (PCA) is that the linear relation between any two variables is best summarized by a regression line. In other words, the variable that represents the regression line as a point cloud contains essential information about both variables. The two variables are thus combined into a single factor. This mechanism can be used to reduce pairs of variables to single dimensions in order to simplify the graphic display of the authors included in the matrix. This process initially includes as many factors as are needed for an optimal representation of the set of variables; the actual number of factors is determined by their internal variance. The most widely used stopping procedure for determining how many factors to include is to calculate the eigenvalue of each factor. The group of all factors with an eigenvalue greater than 1 accounts for most of the total variance

(*Boyce*[24]). If a two-dimensional plot is needed, the plane that maximizes internal variance of a projection of the variables as points in space, or any parallel plane that includes the origin of *n* dimensional space, needs to be determined (*Boyce*;[24] *Egghe*[19]).

In ACA studies that use PCA, authors are assumed to contribute to different degrees to each factor. Their contribution can be measured, and the composition of each factor is interpreted on the basis of those authors whose contribution is higher than ±0.7 (*McCain*[16]). A given author normally appears only once on the map.

In Table 1, the numerical value of a given author's contribution to each factor gives an idea of his or her relative influence within a scientific field. The scientific content of each factor can also be judged from an analysis of the authors' specialties (Table 2).

Principal components analysis yields a clearer representation of the variables (authors) within a smaller space. This advantage, together with the results of the clustering process, provides information on the composition and content of research fronts, as well as on the relationships between fronts.

Multidimensional scaling. This multivariate analysis technique is used to identify the dimensions that best explain similarities and differences between variables. Because the purpose of multidimensional scaling (MDS) is to generate a map of objects, this approach can be considered an alternative to PCA.

Multidimensional scaling procedures use a matrix of similarities or dissimilarities to calculate the coordinates of a two- or three-dimensional space such that the resulting dissimilarities or similarities resemble as closely as possible those in the matrix. In ACA studies, the correlation matrix is used to calculate the author's position in the reduced space. The analysis identifies positions that represent the smallest differences possible between the observed and the calculated distances.

To ensure the best possible fit between distances we used the so-called stress statistic, which measures the goodness of fit between observed and calculated similarities:

$$\varphi = \Sigma \, (d_{ij} - \delta_{ij})^2$$

where *d* is the observed distance and δ the calculated distance in the reduced space. This variable, which is calculated in each iteration, is used as a stopping criterion. When a minimal value is obtained for the stress variable, we can assume that the coordinates for a given author on the map are those which best, illustrate, in the reduced space, the relation between the coordinates in the correlation matrix (*Boyce*[24]). In other words, when the difference between the observed and the calculated distances are reduced to a minimum, the resulting map of authors accurately reflects how authors in a given discipline see authors who are used as the intellectual structure for this discipline.

Thus the resulting map illustrates the intellectual space of a scientific discipline, and reflects how the members of this discipline view the set of contributions produced during a given period. The distribution of points on the map thus illustrates the location of individual authors, the location and extent of clusters (research fronts), and the overall structure of the discipline.

Results and discussion

Figure 1 shows the classification of authors in the correlation matrix. The dendrogram illustrates two clusters of authors separated by a distance of $y=3.5$. Two subclusters appeared in each main cluster; the separation between subclusters was $y>2$ in the left group, and $y>1$ in the right. At lower levels of the hierarchy, smaller clusters are separated by much shorter distances. These clusters represent authors who share affinities, as determined by the clustering algorithm. The first subcluster on the left of the diagram comprises the first 13 authors, including those from Spanish universities and national research council centers who have published in bibliometrics. Foreign authors in the first subcluster are Price, Garfield, Line, Cronin, and Lancaster; their presence corroborates that members of this subcluster have published mainly in bibliometrics and informetrics. Although not all publications by these authors belong to these two disciplines, all authors share a publication record in these specialties.

The second cluster from the left, comprising 16 authors, represents a research front in the area of librarianship. A larger percentage of these Spanish and foreign authors are professional librarians; a few are university professors whose most frequently cited publications have appeared in the area of library management.

The third and largest front (19 authors) comprises mainly university researchers. In contrast with the first two clusters, there appears to be no unifying area of research interest. Members of this front have apparently published in a variety of fields, but share a common theoretical approach to research generalities, as well as a concern for professional training.

The smallest cluster (4 authors) consists of university researchers from a single geographical area (Catalonia) whose interests are similar to those held by members of the preceding cluster. Because these clusters are separated by a distance of only $y=1.2$, the "Catalonian" cluster can be considered a subcluster of the large front of university researchers described above.
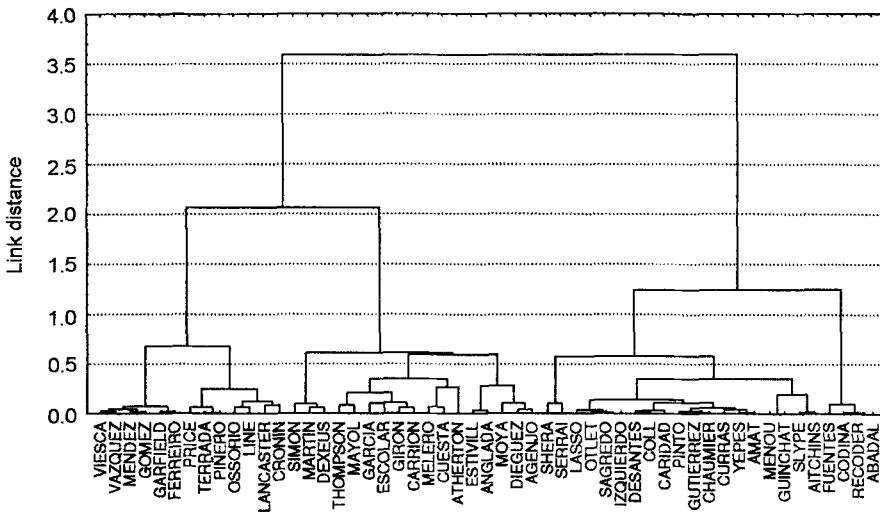
Fig. 1. Classification of authors with the Ward method

Table 1 shows the results of PCA applied to the matrix of Pearson correlations. The variables appear in groups that reflect the principal components of the group of authors studied here. These components reflect the differences in the contributions of each author depending on his or her relationship with each of the other authors. The differences represent the factor loading for each author on each other factor. Authors above the threshold for minimum loading can be considered the most representative of a given factor. In subsequent analyses we used a minimum factor loading to classify authors and identify structural relations between variables. Similarities in the areas of research dealt with by the authors who were selected for each factor corroborate the results obtained with cluster analysis.

As recommended by *McCain*[16] and *White*,[9] we used a threshold factor loading for the variable "author" of 0.4. Variables for which the loading was greater than 0.7 were considered to be the most representative of a given factor, and were used to determine the relationships between a given factor and all other factors. Although our procedure included orthogonal (varimax) rotation to ensure that the variable "author" yielded a high loading on only one factor, a given author could have different loadings on different factors. This means that an author might be prominent in research in different

specialities within a discipline. Some earlier studies considered that when an author appeared to be influential in several factors, this was objective evidence that the author made a significant contribution to the discipline.

Table 1
Distribution of the authors studied according to their loading factor

Factor 1	9,17	Factor 2	5,35	Factor 3	4,3	Factor 4	2,95	Factor 5	2,7
SAGREDO	0,88	VIESCA	0,87	ABADAL	0,93	SLYPE	0,85	MENOU	0,85
IZQUIERDO	0,87	VAZQUEZ	0,82	RECODER	0,93	AITCHINS	0,81	GUINCHAT	0,84
YEPES	0,81	GOMEZ	0,78	CODINA	0,91	CHAUMIER	0,54	ATHERTON	0,50
OTLET	0,80	MENDEZ	0,76	FUENTES	0,56	LANCASTE	0,45	OSSORIO	0,47
PINTO	0,77	FERREIRO	0,53					AMAT	0,40
GUTIERREZ	0,68	OSSORIO	0,52						
COLL	0,65	GARFIELD	0,45						
CHAUMIER	0,64	**Factor 6**	**2,45**	**Factor 7**	**2,1**	**Factor 8**	**1,88**	**Factor 9**	**1,65**
DESANTES	0,64	AGENJO	0,88	ESTIVILL	0,84	SHERA	0,75	CUESTA	0,88
AMAT	0,60	DIEGUEZ	0,84	ANGLADA	0,81	SERRAI	0,74	MELERO	0,68
CURRAS	0,60	DEXEUS	0,63						
LASSO	0,59	MOYA	0,59						
CARIDAD	0,52	MELERO	0,44						
TERRADA	0,40								
Factor 10	**1,59**	**Factor 11**	**1,24**	**Factor 12**	**1,2**	**Factor 13**	**1,11**	**Factor 14**	**1,08**
PINERO	0,83	CRONIN	0,77	MAYOL	0,73	CARRION	0,76	LINE	0,77
TERRADA	0,74	LANCASTER	0,53	CARIDAD	0,60	GIRON	0,73		
PRICE	0,69			FUENTES	0,59	ESCOLAR	0,58		
GARFIELD	0,46					MARTIN	0,45		

Table 1 shows the distribution of the authors we studied according to their loadings on each factor. None of the authors in the present analysis were representative (factor loading > 0.7, bold names on the table) of more than one factor, and clusters of authors appeared to be well segregated. In terms of cocitation frequencies, these results suggest that citing authors who cite the most representative authors for a given factor rarely cite authors who are predominant for other factors. This was corroborated by the fact that none of the correlation coefficients for any combination of the 14 factors we considered was greater than 0.1. The results of this analysis are also compatible with the large distances that separated the four main clusters obtained with the Ward classification.

The most representative authors for each of the 14 factors shown in Table 1 were clearly related with the four author clusters. The factors in the table are given in decreasing order of internal variance; the higher the variance, the greater the likelihood that an author associated with this factor appeared in association with at least one other factor. Table 2 shows the relationships between the four author clusters and the 14 factors shown in Table 1. The greatest percentage of variance was accounted for by factors 1 to 3. All factors represent a proportion of the variance greater than the mean (eigenvalue > 1, mean variance = 1).

Table 2
Factors related with each author cocitation cluster identified in Spanish
library and information science items published between 1985 and 1994

Cluster 1: Informetrics	Factors 2,10,11,14
Cluster 2: Libraries	Factors 6,7,9,13
Cluster 3: University 1	Factors 1,4,5,8,12
Cluster 4: University 2	Factor 3

If the factors with the highest eigenvalues best represent the authors' cocitation behavior, its factor loadings also can be used to create a map that illustrates the relationships between authors. When we plotted the loadings of each of the 53 authors on factor 1 (x axis), factor 2 (y axis) and factor 3 (z axis), authors with large loadings on one factor and small loadings on another appeared as widely separate from each other (Fig. 2). This three-dimensional plot does not reflect the complexity of the relations between authors with regard to all 14 factors, but does provide an overall view. In combination with other analyses, these results shed light on the intellectual structure of the discipline.

The map of authors produced with MDS once again identified the same groups as were found with cluster analysis and PCA. The results of MDS provided insights into the significance of the spatial relationships between clusters, and made it possible to analyze the both relationships between authors within a given cluster and the positions of the clusters and authors in each (Fig. 3) (*White*[9]). These four points are shown in the MDS map, which depicts authors at distances as similar as possible to those in the original matrix. The resulting stress for the two-dimensional plot shown in Fig. 3 was 0.22, a value compatible with those obtained by other authors for similar representations (*McCain;*[16] *White*[9]). Although *Boyce*[24] found that only those stress values lower than 0.15 were useful, in the present case this would yield a confusing picture because of the large number of variables and their relative positions in three-dimensional displays.
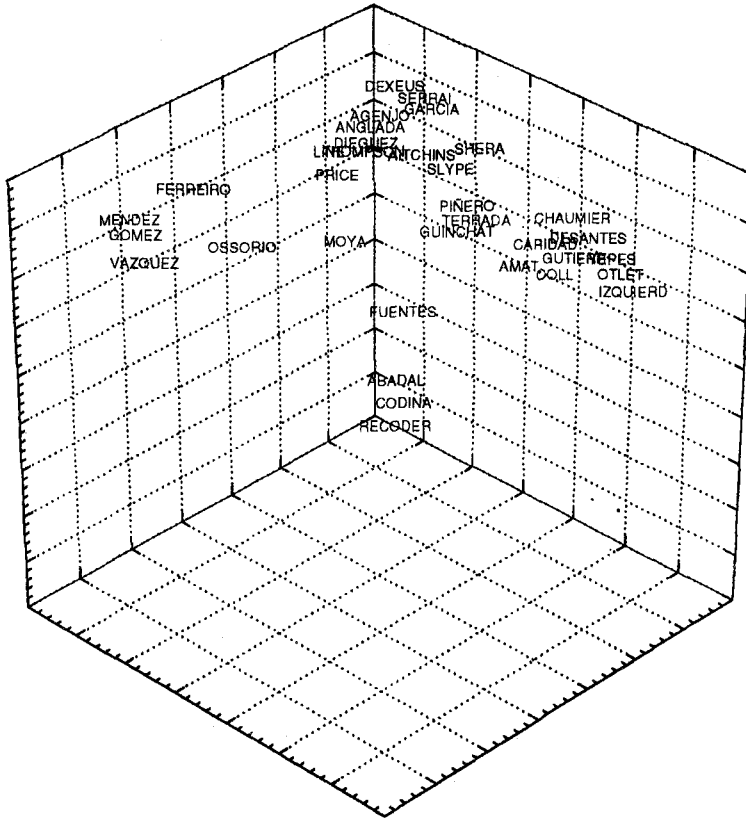
Fig. 2. Map showing the relationships between authors on the basis of the three factors with the greatest variance

Figure 3 shows four author clusters that closely match those identified in the initial Ward classification and in the PCA. All clusters appear near the periphery of the map, suggesting that authors belonging to different clusters were rarely cocited. The most noteworthy feature of the map is the absence of authors near the center, a finding that illustrates the few, if any authors or publications were considered common to or fundamental to research in any of the areas in which these authors have published. This result, a rare occurrence in similar studies, suggests that authors at the four research fronts we identified do not use a common knowledge base.
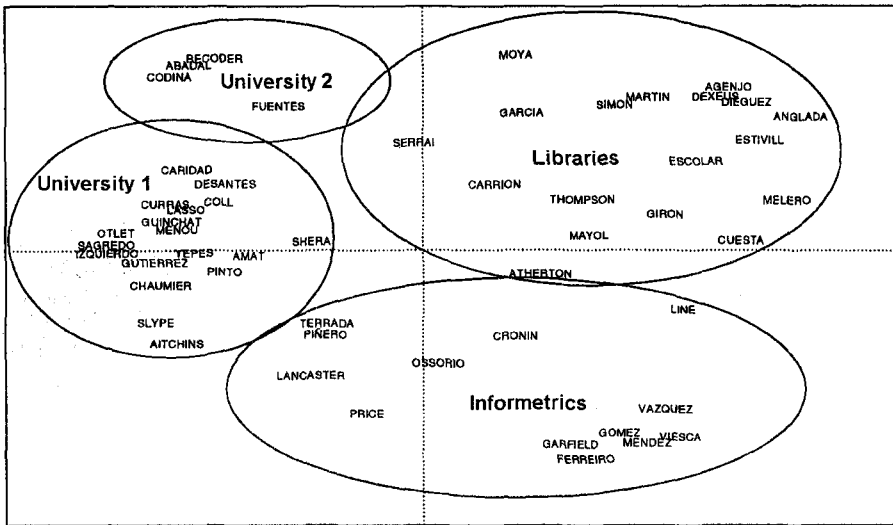
Fig. 3. Author cocitation map generated with multidimensional scaling

Although there may be some relation between certain clusters, we identified no author or authors that served as a common referent for the entire discipline. Although the centers of gravity of the different clusters are widely separated, some authors are located at "bridge" positions between fronts. For example, Piñero and Terrada mapped halfway between the "informetrics" and "university 1" clusters; Serrai was located halfway between the "libraries" and "university 1" clusters; and Atherton, Line and Cronin appeared halfway between the "informetrics" and "libraries" clusters. These authors have apparently been used as an intellectual reference by citing authors in different fronts.

We noted a single exception to the large separations between fronts: the shortest distance was that between "university 1" and "university 2." On the basis of the methods we used, this distance is sufficient to consider these as independent fronts; however, we could find no differences in the content of the articles published by authors at these fronts. We suspect that this finding reflects widespread social and professional ties between authors at different universities. The left half of the ACA map for Spanish LIS was occupied predominantly by university researchers; although the density of the points in these quadrants suggests that this is a large front, we cannot

determine which specific scientific contents are associated with this author cluster. We can only say that 1) their publications have some features in common, 2) many of the publications cited are monographs containing a large proportion of theoretical and introductory material, and 3) citing authors belong to the same type of institution (university, library, or national research council) as cited authors.

The distribution of authors within the three largest fronts differed strikingly. Author density was greatest in the "university" front, and a single author (Yepes) was located near the center of the cluster. In contrast, the "informetrics" front contained a clear nucleus near the periphery of the map, and consisted of authors who published only in this research area (Gómez, Méndez, Viesca, Vázquez, Ferreiro and Garfield), surrounded by a belt of authors who appeared closer to other fronts (Line, Cronin, Ossorio, Price, Terrada, Piñero and Lancaster). The "libraries" cluster was the most disperse of the four, with authors spread out across a large surface. This reflects the variety of topics these authors have investigated, and the correspondingly lower frequency of cocitations in comparison with other clusters. This cluster also contained fewer non-Spanish authors than the other two.

Conclusions

Our results show that during the period from 1985 to 1994, three research fronts existed in Spanish LIS. The distances between the fronts on the ACA map suggest that the knowledge bases used by authors at each front were clearly different.

The front designated here "informetrics" comprised publications that centered on specific topics, and appeared to be similar to subject area fronts in earlier studies (Persson;[25] White[9]). In fact. this was the only similarity between the knowledge base used in LIS in Spain during this period and the findings of similar studies in other populations. This result emphasizes the peculiarities of Spanish LIS research. Certain authors in the "informetrics" from (e.g., Terrada and Piñero, among others) acted as a bridge to the "university" front. This may have resulted from the bibliometric studies produced by invisible authors in the "university" front. However, because such studies and authors are scarce, we suspect that social and professional ties account for the appearance of "bridges" between fronts, as the authors in bridge positions were the only members of the "informetrics" cluster with a university affiliation. In any case, the "informetrics" front contained few authors (visible or invisible), and most publication in this front appeared in a single journal (Revista Española de Documentación Científica).

The "libraries" front was the most disperse of those identified in this study, as a result of the variety of topics dealt with in these publications. Moreover, the appearance in recent years of studies in the field of information technology has widened the scope of this research area. The visible part of this front consisted mainly of librarians who published predominantly applied studies in the journal Boletín de la ANABAD (Asociación Nacional de Archiveros, Bibliotecarios, Arqueólogos y Documentalistas). This front was tenuously linked to others by non-Spanish authors who have published studies that can be considered interdisciplinary in nature (e.g., Line and Atherton). Thus there appears to be little common intellectual ground between this front and the others; citing authors from different fronts most likely use these sources for quite different reasons.

The "university" front was the largest and most compact, comprising the greatest number of cocitations. On the other hand, it was the least well defined in terms of research areas. Because of the absence of common research topics investigated by visible authors, we conclude that the publications in this front had in common a more general approach to research problems (i.e., handbooks, manuals and introductory texts). Unlike the other fronts, monographs were frequently cited (27%). In this front most studies that comprise the knowledge base are monographs, which are more likely than other types of publications to be general overviews, manuals, and texts dealing with specific specialities within the discipline (63%).

Within this front we distinguished two smaller subfronts. One consisted of Catalonian university authors cited by other Catalonian researchers. The larger subfront comprised members of the so-called Madrid school of researchers affiliated with universities in this city, along with colleagues who have intellectual or personal (mentor-mentee) ties (e.g., Yepes, Sagredo, Caridad, Gutiérrez, Izquierdo, Desantes, Lasso, Pinto and Currás). Cocitation behaviors in this front was especially persistent; we note that pairs of authors were often cocited despite the lack of similarity in the content of their respective research publications. This situation arises frequently when citing authors belong to the same university, and is favored by the general, theoretical nature of the citing and cited works. Cocitation patterns in the Spanish population we investigated may thus represent the behavior designated as "hat-tipping" (*Thorne*[26]).

*

Appendix

Journal & Proceeding titles:
- Revista española de documentación científica
- Boletín de la Anabad
- Boletín de la Asociación Andaluza de Bibliotecarios
- Item
- Documentación de las ciencias de la información
- Revista general de información y documentación
- Jornadas españolas de documentación automatizada
- Jornadas bibliotecarias de Andalucía
- Jornades catalanes de documentació

Authors	Citations	Self-citations	Cit-Self.	Co-citations
LOPEZ YEPES, JOSE	84	4	80	87
AGENJO BULLON, XAVIER	64	23	41	158
PEREZ ALVAREZ OSSORIO, JOSE RAMON	63	10	53	96
CURRAS PUENTE, EMILIA	60	31	29	324
AMAT NOGUERA, NURIA	56	1	55	94
LINE, MAURICE B.	48	0	48	81
CARRION GUTIEZ, MANUEL	46	0	46	154
LANCASTER, FREDERICK WILFRID	38	0	38	313
LOPEZ PIÑERO, JOSE MARIA	37	0	37	113
GARCIA GUTIERREZ, ANTONIO LUIS	36	5	31	62
GARFIELD, EUGENE	36	0	36	163
TERRADA FERRANDIS, MARIA LUZ	35	7	28	61
FUENTES I PUJOL, M. EULALIA	34	9	25	218
GARCIA MELERO, LUIS ANGEL	34	6	28	167
DESANTES GUANTER, JOSE MARIA	33	9	24	151
PRICE, DEREK JOHN DE SOLLA	33	0	33	91
COLL-VINENT, ROBERTO	32	0	32	66
MENDEZ MIAJA, AIDA	31	2	29	93
VAZQUEZ VALERO, MANUELA	31	9	22	90
CHAUMIER, JACQUES	30	0	30	70
ANGLADA I DE FERRER, LLUIS M.	28	3	25	121
CARIDAD SEBASTIAN, MERCEDES	28	0	28	148
SLYPE, GEORGES VAN	27	0	27	183

CODINA BONILLA, LLUIS	26	8	18	127
SAGREDO FERNANDEZ, FELIX	26	1	25	96
IZQUIERDO ARROYO, JOSE MARIA	24	8	16	74
MAYOL I FERNANDEZ, CARMEN	24	4	20	141
PINTO MOLINA, MARIA	23	3	20	144
GOMEZ CARIDAD, M. ISABEL	22	8	14	113
SIMON DIAZ, JOSE	22	0	22	142
CRONIN, BLAISE	21	3	18	189
ESCOLAR SOBRINO, HIPOLITO	21	0	21	145
FERREIRO ALAEZ, LUIS	21	7	14	346
GIRON GARCIA, ALICIA	21	0	21	61
LASSO DE LA VEGA, JAVIER	21	0	21	153
AITCHINSON JEAN	20	0	20	135
ESTIVILL I RIUS, ASSUMPCIO	20	1	19	161
OTLET, PAUL	20	0	20	83
SHERA, JESSE HAUK	20	0	20	116
THOMPSON, JAMES	20	0	20	124
GARCIA EJARQUE, LUIS	19	3	16	228
VIESCA ESPINOSA DE LOS MONTEROS, ROSA DE LA	19	2	17	149
RECODER SELLARES, M. JOSE	18	6	12	120
CUESTA ESCUDERO, M. JESUS	17	0	17	154
DEXEUS MAYOL, MERCEDES	17	2	15	81
MARTIN ABAD, JULIAN	17	0	17	71
ABADAL FALGUERAS, ERNEST	16	4	12	101
ATHERTON, PAULINE	16	0	16	68
DIEGUEZ, FRANCISCO	16	1	15	98
GUINCHAT, CLAIRE	16	0	16	84
MENOU, MICHEL	16	0	16	161
MOYA ANEGON, FELIX DE	16	3	13	146
SERRAI, ALFREDO	16	0	16	110

References

1. J. D. DE SOLLA PRICE, Networks of scientific papers, *Science*, 149 (1965) 510–515.
2. R. R. BRAAM, H. F. MOED, A. F. J. VAN RAAN, Mapping of science by combined co-citation and word analysis. I. Structural aspects, *Journal of the American Society for Information Science*, 42 (1991) No. 4, 233–251.
3. M. H. MACROBERTS, B. R. MACROBERTS, Problems of citation analysis, *Scientometrics*, 36 (1996) 435–444.
4. M. H. MACROBERTS, B. R. MACROBERTS, Problems of citation analysis: A critical review, *Journal of the American Society for Information Science*, 40 (1989) No. 5, 342–349.
5. H. SMALL, Co-citation in the scientific literature: a new measure of the relationship between two documents, *Journal of the American Society for Information Science*, 24 (1973) No. 4, 265–269.
6. V. MARSHAKOVA, System of document connections based on references, *Nauchno-Tekhnichescaya Informatisya*, Series II, 6 (1973), 3–8.
7. J. J. FRANKLIN, R. JOHNSTON, Co-citation bibliometric modeling as a tool for S&T policy and R&D management: issues applications and developments, In: A. F. J. VAN RAAN (Ed.), *Handbook of Quantitative Studies of Science and Technology*, North Holland, Amsterdam, 1988, 325–389.
8. H. D. WHITE, A cocitation map of the social indicators movement, *Journal of the American Society for Information Science*, 34 (1983) No. 5, 307–312.
9. H. D. WHITE, B. C. GRIFFITH, Author cocitation: a literature measure of intellectual structure, *Journal of the American Society for Information Science*, 32 (1981) No. 3, 163–171.
10. A RIPP, Mapping of science: Possibilities and limitations, In: A F. J. VAN RAAN (Ed.), *Handbook of Quantitative Studies of Science and Technology*, North Holland, Amsterdam, 1988, 253–273.
11. H. D. WHITE, Cocited author retrieval online: an experiment with the social indicators literature, *Journal of the American Society for Information Science*, 32 (1981) No. 1, 16–21.
12. JIMÉNEZ-CONTRERAS, F. D. MOYA ANEGÓN, Análisis de la autoría en revistas españolas de Biblioteconomía y Documentación, 1975–1995, *Revista Española de Documentación Científica*, 20 (1997) No. 3, 252–266.
13. L. LEYDESDORFF, Various methods for the mapping of science, *Scientometrics*, 11 (1987) 295–324.
14. H. D. WHITE, A cocitation map of the social indicators movement, *Journal of the American Society for Information Science*, 34 (1983) No. 5, 307–312.
15. M. J. CULNAN, C. A. O'REILLY, J. A. CHATMAN, Intellectual structure of research in organizational behavior, 1972–1984: a cocitation analysis, *Journal of the American Society for Information Science*, 41 (1990) No. 6, 453–458.
16. K. W. MCCAIN, Mapping authors in intellectual space: a technical overview, *Journal of the American Society for Information Science*, 41 (1990) No. 6, 433–443.
17. A. E. BAYER, J. C. SMART, G. W. MCLAUGHLIN, Mapping intellectual structure of a scientific through author cocitations, *Journal of the American Society for Information Science*, 41 (1990) No. 6, 444–452.
18. C. M. CUADRAS, *Métodos de análisis multivariante*, PPU, Barcelona, 1991.
19. L. EGGHE, R. ROUSSEAU, *Introduction to Informetrics. Quantitative methods in Library, Documentation and Information Science*, Elsevier Science Publishers, Amsterdam, New York, Oxford, Tokyo, 1990.
20. M. T. KINNUCAN, M. J. NELSON, B. L. ALLEN, Statistical methods in information science research, *Annual Review of Information Science and Technology*, 22 (1987) 147–178.
21. P. WILLET, Recent trends in hierarchic document clustering a critical review, *Information Processing & Management*, 24 (1988) No. 5, 577–597.
22. A. GRIFFITHS, L. A. ROBINSON, P. WILLET, Hierarchic agglomerative clustering methods for automatic document classification, *Journal of Documentation*, 40 (1984) No. 3, 175–205.
23. M. ZITT, E. BASSECOULARD, Reassessment of co-citation methods for science indicators: Effect of methods improving recall rates, *Scientometrics*, 37 (1996) 223–244.
24. B. BOYCE, C. T. MEADOWS, D. KRAFT, *Measurement in Information*, Academic Press, Londres, 1994.
25. O. PERSSON, The intellectual base and research fronts of JASIS 1986–1990, *Journal of the American Society for Information Science*, 45 (1994) No. 1, 31–38.
26. F. C. THORNE, The citation index: Another case of spurious validity, *Journal of Clinical Psychology*, 33 (1977) 1157–1161.