Spanish Research in Knowledge Organization (2002-2010)

Mercedes de la Moneda Corrochano*, María J. López-Huertas**, and Evaristo Jiménez-Contreras***

*/**/*** Universidad de Granada. Departamento de Biblioteconomía y Documentación, *<dlmoneda@ugr.es>, ** <mjlopez@ugr.es>, *** <evaristo@ugr.es>

Prof. Evaristo Jiménez-Contreras is Ph.D. in Documentation (1993) and professor in Bibliometrics in the Department of Information and Documentation of the University of Granada. He coordinates the research group Evaluation of the Scientific Information and Communication. His main research lines are methodologies for the evaluation of science in the academic institutions.

Mercedes de la Moneda Corrochano is Ph.D. in Documentation at the University of Granada. She teaches in the Department of Information and Documentation of the same University. Her publications are on the evaluation of science. She is part of the research group of the University of Granada EC3 which carries out several research project inside the National Research Plan I+D. She is in charge of the Impact Index of Social Science Journals (INRECS)

María J. López-Huertas is Ph.D. in Roman Linguistics at the University of Granada. She teaches Knowledge Organization at the Department of Information and Documentation of the University of Granada at Graduate and Ph.D. level. Her research interests are knowledge organization design and methods, interdisciplinary knowledge organization, and evaluation of interdisciplinary knowledge. She is the former president and current vice president of ISKO.

De la Moneda Corrochano, Mercedes, López-Huertas, María J., and Jiménez-Contreras, Evaristo. **Spanish Research in Knowledge Organization (2002-2010).** *Knowledge Organization. 40(1)*, 28-41. 9 references.

ABSTRACT: This study analyzes Spanish research on Knowledge Organization from 2002 to 2010. The first stage involved extraction of records from national and international databases that were interrogated. After getting the pertinent records, they were normalized and processed according to the usual bibliometric procedure. The results point to a mature specialty following the path of the past decade. There is a remarkable increase of male vs. female authors per publication, although the gender gap is not big. It is also evident that there is a remarkable internationalization in publication and that the content map of the specialty is more varied than in the previous decade.

Received 2 October 2012; Revised 27 October 2012; Accepted 5 November 2012

1.0 Introduction

In a previous study (López-Huertas and Jiménez-Contreras 2004), scientific output in the area of knowledge organization was first analyzed for the period 1992-2001. Since then, there have been great changes in the field itself, as well as in the university setting where most of this research and these authors are rooted. The present study attempts to reflect the state of knowledge organization research in Spain during the period 2002 to 2010, and compare it with the results described for the previous decade of 1992-2001.

With the focus on knowledge organization (KO), the trend detected in 1992-2001 was one of positive evolution and expansion, with overflow into corporate settings or the workplace in general, into decision-making and so-called competitive intelligence (2012). At any rate, there is still no overall consensus among specialists as to whether the aforementioned contexts pertain to KO or not (Hjørland 2008; Smiraglia 2005). The present study is limited to the realm of KO in a strict sense, centering on information retrieval systems. This focus will be justified later on. Moreover, as affirmed in the paper published in 2004 and cited above, the difficulty of drawing conceptual boundaries for KO and its epistemological weakness or lack of theoretical coherence have been stressed by previous authors (Hjørland 2002).

Very few contributions about KO studies have come to light in recent years. One publication partly regarding the subject has a more limited temporal coverage than our study (Oliveira, Grácio, and Silva 2010) or limited to a source (Alves et al. 2011). Other studies have a similar coverage of time (Moneda, López-Huertas, and Jiménez-Contreras 2011) or they consider a longer period of time (Travieso 2011).

1.1. Justification and objectives

Since the aforementioned paper, published in 2004, hardly anyone has conducted research into this subject area, regarding Spanish contributions to KO, giving us good reason to undertake a review of the state of the art. Furthermore, this second endeavour comes to complement the perspective traced in 2004 while allowing us to follow the evolution of the field and pinpoint possible changes in the (roughly) two decades analyzed. The time spam covered by the present study is of nine years instead of the ten covered by the 2004 publication. The reason is that we presented a short paper on this topic in the 10th ISKO-Spain Conference held in Ferrol in 2011. The conference topic was the evolution of KO in Spain, and it seemed to us interesting to study the Spanish research on KO from the period 2002-2010, since the presentation would be in 2011, following up on the 2004 study.

As conceptual limits, in order to produce homogeneous results that would permit comparison of the two periods involved (1992-2001 and 2002-2010), we adopted the same bases as in the previous study. That is, we restricted the concept of KO to systems that approach the subject area from a linguistic-conceptual perspective, fundamentally; although other approaches clearly focused on Knowledge Organization are also included. All research into specialized conceptual structures or encyclopaedias was taken into account, regardless of whether the approach was theoretical, methodological, practical, or professional. Therefore, content analysis and indexing per se were not considered. Accordingly, only publications by researchers born in Spain and by naturalized citizens of Spain were included.

2.0 Material and methods

Considering that the specialized area chosen has diversified output, in a number of different formats, our study embraced all of them—monographs, theses, conference papers (national or international), and articles of any extension published in all the journals indexed by the databases specified below.

2.1 Databases consulted

In general, the same patterns as in the previous study were followed, for the sake of consistency. Notwithstanding, some changes were necessary due to the appearance of new databases, such as Dialnet, which, in turn, led us to disregard the databases of Teseo, Rebiun, and Rueca, given that they refer to the same documental type, Ph.D. theses and monographs, and they offer a similar degree of coverage in their collections. Therefore, the databases finally consulted were ISI, LISA, Dialnet, and ISOC. In addition, the publications of the International ISKO acts were incorporated manually, as they are not included in the databases consulted; those of ISKO-Spain were included in view of the importance they have in the context of our study.

The search strategies were likewise repeated, as detailed in Tables 1 and 2. This terminological approach to the databases made it necessary to perform several searches so that the combined sum of all would guarantee exhaustive retrieval of our subject area, even though that implied that the results would have some duplication, given that the use of Knowledge Organization as the only term could have silenced numerous relevant documents. The duplicated references were detected and eliminated after loading all the search items into the Procite database.

With respect to the structure of the consulted databases, we should acknowledge the lack of homogeneity and standardization in the formats for data retrieval, and above all the lack of information in fields that are of great important for bibliometric studies, such as author affiliation, something that was found

to be common in Dialnet, and which led to considerable manual labor afterwards. We should also mention that Dialnet does not offer search syntax, meaning that the retrieval of documents with terms is difficult and time-consuming. Finally, we underline the renowned lack of normalization of author names, above all in international data bases (Ruiz, Delgado, and Jiménez 2002), which, along with the all-too-frequent appearance of first names shown by initials, can severely affect analyses concerned with the study of author gender. The bibliographic processor Procite was used to process data.

2.2 Obtaining and processing data

Thematic searches were carried out for the selection of documents, mainly searches by terms and in some cases by classification codes, depending on the database.

The international results were obtained by consulting the databases of ISI and LISA. In the case of ISI, the query was made with the list of terms shown in Table 1. Please note that, in addition to the use of the field "topic," where the terms were stored, a further refined search was conducted by place (Spain). No refined search was based on the specialized areas arising from each search in order to obtain a more pertinent retrieval, although this called for a posterior manual filtering to eliminate any irrelevant documents. In this way, we were able to include 42 publications that were not included under the tag of Library and Information Science.

The search conducted in the LISA database involved a list of a priori terms shown in Table 1.

The national results were extracted from the ISOC and Dialnet databases. For ISOC, the search strategy was twofold, using the classification codes of the database (which led to a search of low precision and wide scope) and a manual selection of pertinent documents for our study. The codes used are shown in Table 3. Aside from this search, another was carried out with terms to cross the results of the previous search. The terms used to retrieve information from ISOC and Dialnet are indicated in Table 2.

The list of terms used in English and in Spanish is basically the same as the one used in the study published in 2004. However, we added new expressions that are considered necessary given the appearance of new topics or the increasingly generalized use of some terms over the past decade. Such is the case of ontologies, taxonomies, folksonomies, and systems for knowledge organization, respectively (see Table 2).

The result of these searches had to be filtered by the revision of the results obtained in order to ensure the relevance of the results in any case.

Once the references had been selected, they were exported to a bibliographic processor to process the information obtained. The duplications were eliminated, and authority control was exercised to correct

CLASSIFICAT* LANGUAG*
CLASSIFICAT* SYSTEM*
DESCRIPTOR LANGUAG*
DOCUMENT* CLASIFICATION
DOCUMENT* ORGANIZATION*
DOCUMENT* LANGUAG*
FOLKSONOM*
INDEX* LANGUAG*
KNOWLEDGE ORGANIZATION
KOWLEDGE ORGANIZATION SYSTEMS
KNOWLEDGE STRUCTUR*
LIBRAR* CLASSIFICAT*
ONTOLOG*
SUBJECT HEAD*
TAXONOM*
THESAUR*

Table 1. Search terms in ISI and LISA

LENGUAJE* DOCUMENT*
LENGUAJE* DE INDIZ*
LINGÜÍST* DOCUMENT*
CLASIFICAC* BIBLIOGRAF*
SISTEMA*DE CLASIFICAC*
CLASIFICACION* DE BIBLIOTECA*
CLASIFICACIÓN* DE LIBRO*
INDIZACIÓN
CLASIFICACIÓN
ESTRUCTURA* CLASIFICAT*
ESTRUCTURA* CONCEPTUAL*
ENCABEZAMIENT* DE MATERIA*
FOLKSONOM*
LISTA* DE ENCABEZAMIENTO*
ONTOLOG*
ORGANIZACIÓN DEL CONOCIMIENTO
SISTEMAS DE ORGANIZACIÓN DEL CONOCIMIENTO
TAXONOM*
THESAUR*
TESAUR*

Table 2. Search terms in ISOC and Dialnet

Códigos	Contenido genérico	Búsqueda completa o parcial
200100	Scientific output	200104 (Basic and applied research)
200200	Documentation and Information Policies	200200 al 200299
200300	Information Resources	200300 al 200399
200400	Information Analysis	200400 al 200499
200500	Information Management, Information storage and Information retrieval	200500 al 200599
200600	Información Indistry and Tecnology development	200604 (Automatic Indexing)
200700	Library System	200700 al 200799
200800	Archives and Museums Documentation	200802 (Archives Management)
200900	Information Management	200900 (Information Management);
		200901 (Services Planning)
201001	Docencia	Restricted to Knowledge Organization

Table 3. Search codes in ISOC

and normalize the names of authors, which, in many cases, called for consulting alternative sources such as personal webpages.

Statistical treatment of the data was trivial and will not be specified here, except to clarify that, in the recount of authors associated with institutions, we used fractioned recount; that is, the portion resulting from each institution resulted from operating with each document was expressed as 1/n, n being the number of authors in question.

3.0 Results and discussion

The results obtained respond to the following research questions:

What are the characteristics of the population of publishing authors? Is there equality in terms of author gender? How much research is actually printed and divulged? How has output evolved over time? Where is Spanish research published, and how many studies have come to light in the period of study here?

Heterogeneity in the identified documents made it necessary to group them into three types: articles, monographs, and dissertations. Each group has its own characteristics, in terms of structure and objectives, as well as in the data identifying the authors. Thus, we first proceeded to perform a sectorial analysis to eventually arrive at an analysis of the data set as a whole, which allowed us to reflect the conduct and the dynamics of Spanish research in the field of Knowledge Organization.

The figures obtained were recounted after eliminating irrelevant documents or duplications in the databases consulted. These generally presented the aforementioned problems of little visibility of KO researchers in bibliometric studies, which may have to do with inadequate categorization of the subject matters included under the specialty, and the inclusion of specific categories within other more general categories, which makes it difficult to identify them while furthermore producing noise in the retrieval process.

The contribution in the number of documents of each data base used in this study is shown in Table 4.

Databases	Documents
ISI	96
LISA	66
ISOC	145
DIALNET	226
TOTAL	533

Table 4. Documents in the databases consulted

After the final filter, the number of articles consulted was just 357, a figure slightly below that of the previous period 1992-2001 (399 documents). Yet we must emphasize that, in the latter case, one more year of study was included. Indeed, if we extrapolate the data gathered in this study to a ten year period, the number of publications would be around 497 hypothetical articles. Hence, we stress the numerical difference observed with respect to the decade 1992-2001, as documented in international databases and, in particular, the ISI, where a great increase in publications indexed in the period 2002-2010 is witnessed, as a to-

tal of 497 works were published. This stands in remarkable contrast to the previous decades and their publications indexed in the ISI: the rise in publication reached as much as 18% of total output. Meanwhile, the ISI publications identified in the previous decade represented only 4.2% of the entire set of documents. In LISA, there is also an increase, but it is not as surprising as the case just described. This finding has very interesting implications, as it suggests that Spanish research in KO has greater impact internationally than at the national level.

3.1. Quantification of author output

Because the collection of documents obtained was irregular from the documental standpoint, as commented earlier, we expound the results in three groups: authors with articles and presentations, authors of monographs, and authors of Ph.D. theses.

3.1.1 Authors of articles and their output

Authors	Number of papers published
Garcia Marco, F. J.	11
López-Huertas, M.J.	10
Moreiro González, Jose A.	10
Sorly Rojo, A.	7
López Alonso, M.A.	6
Morato Lara, Jorge	6
San Segundo, R.	6
Sánchez Cuadrado, S.	6
Sicilia Urban, M.A.	6
Ureña López, L.A.	6
Eito Brun, R.	5
Granados, M.	5
Montejo Raez, A.	5
Sánchez Jiménez, R.	5
Caldera Serrano, J.	4
Caro Castro, C.	4
García Barriocanal, E.	4
Llorens Morillo, J.B	4
Martínez Méndez, F.J.	4
Pastor Sánchez, J.A.	4
Pérez Agnera, J.R.	4
Rodriguez Bravo, B.	4
Sánchez Alonso, S.	4
Authors with 3 publications	12
Authors with 2 publications	59
Authors with 1 publication	395

Table 5. Publications in journals and proceedings of conferences, by author

Under this heading, we describe both the articles published in journals and those printed as acts of national or international conferences/congresses. We identified 489 authors, who produced 298 papers in periodicals (179 journal articles plus 119 conference communications). A summary of the most productive ones is offered in Table 5. Accordingly, there were 23 authors behind a total of 489 papers published, who may therefore be considered "productive" in the development and diffusion of Knowledge Organization. They represent 4.7% of total authors. The output by this particular group is 130 articles, which stands as 43% of overall KO publication.

It is seen that, according to the model put forth by A. J. Lotka, and corroborated in the previous decade studied, a small percentage of authors does in fact produce a high percentage of publications, in this case 43%.

If we compare these results with those of the previous decade, a period for which 201 authors were identified, we find that, between 2001 and 2010, the number of authors increased to 395. However, the production in this decade is not greater than the previous one, during which a total of 330 articles came to press. That is, the number of productive authors is on the rise, but productivity *per se* is not, showing a somewhat disappointing harvest of 298 articles. Hence, we must conclude that the increase in author ranks is related with the number of undersigning authors: 32.5% of the documents analyzed were signed by three or more authors, and 15% were co-authored by four to six researchers. Table 7 reflects these figures, taking all the document types into account.

No. Authors	No. Works
1	159
2	81
3	57
4	35
5	16
6	5
7	1
9	1
12	1

Table 6. Number of authors per work published

It is important to point out that many journals chosen by the cited authors to publish their research are not LIS journals. The total of the articles published in

these journals comprise 42.65% of the total titles and incorporate 26.78 % of the articles. The main areas of knowledge of these journals are: informatics (11.48% of the articles), economy, and enterprises (7.10% of the articles). Journals devoted to health sciences, psychology, translation, etc. follow with less representation. It is also remarkable the lack of collaboration between areas LIS/non-LIS in these publications, where almost all authors do not belong to the LIS area of knowledge. Considering the articles published in LIS journals, authors coming from areas out of LIS represent only the 11%. Table 7 shows the number and percentage of articles published in LIS and non LIS journals.

Knowledge Areas	Articles	Journals
LIS	134	39
Informatics	21	6
Economy-Enterprise	13	10
Health Sciences	5	5
Translation	4	2
Social sciences	3	3
Psychology	1	1
Architecture	1	1
Museums	1	1
TOTAL	183	68
LIS	134	39
Non LIS	49	29
% of Non LIS	26,78	42,65
% of LIS	73,22	57,35

Table 7. Knowledge areas of journals of selected publications

3.1.2 Authors of Monographs and Their Output

In this group, we look at complete monographs (13) and book chapters (23), giving a total of 36 publications. This collection amounts to 8.4% of all the works referenced. They were signed by 43 authors, who represent just 8.26% of all authors identified for all the document types published in the period 2002-2011. Here, unlike the case of articles, co-authorship is very low. At the very most, we can encounter three authors. In terms of productivity, we again see that a small number of authors (15) produce 41.5% of all the monographs. In contrast to the period 1992-2001, here the collaborative authors are few, and therefore were not analyzed separately. Results are given in Table 8.

Authors	No. of publications
Gil Urdiaciain, Blanca	3
López-Huertas, María J.	3
Moreiro González, José A.	3
Agustín Lacruz, M. del Carmen	2
Caro Castro, Carmen	2
Torres Ramírez, Isabel	2
38 authors	1

Table 8. Authors of monographic works and their productivity

Comparison of these results with those from the previous decade make evident a sharp decline in publication. In this period, there were 141 authors who produced 278 monographic works. One possible explanation is the fact that the institutions or organizations that undertook publishing tasks in the past—largely involving thesauri or material headings—have since become less active in this area of activity.

3.1.3 Global analysis of authors of articles and monographs

Finally, we prepared a joint list of all the most productive authors of articles or monographs, so as to derive an integral notion of the group dynamics and assess productivity overall. The results are shown in Table 9.

In this context, we should point out that 12 of the 25 authors from the Table of 1992-2001 are seen to be active a decade later. On the other hand, 2002-2010 is witness to 35 new authors who, due to their low productivity in most cases, are not referenced by name in Table 9.

3.1.4 Authors of PhD theses and their output

The number of Ph.D. dissertations published comes to 23, a higher figure than the 15 of the previous period. Thus, we can speak of a moderately heightened activity if we moreover bear in mind that the second period of analysis is one year longer than the first. Ph.D. theses were generated in eleven Spanish universities, listed in order of importance: Universidad de Valencia, Universidad Politécnica de Valencia, Universidad Carlos III of Madrid, Universidad de Alcalá de Henares, Universidad de Murcia, Universidad Complutense de Madrid, Universidad de la Coruña, Universidad de Granada, Universidad de Málaga, Oberta de Cataluña, and the universities of León and Salamanca.

Moreiro Gonzalez, Jose Antonio	14
Lopez- Huertas, Maria Jose	13
Garcia Marco, Francisco Javier	12
Garcia Jimenez, Antonio	8
Mochon Bezares, Jose Angel	8
Morato Lara, Jorge	8
Sorli Rojo, Angela	8
Lopez Alonso, Miguel Angel	7
San Segundo Rosa.	7
Sanchez Cuadrado, Sonia	7
Ureña Lopez, Luis Alfonso	7
Caro Castro, Carmen	6
Sicilia Urban, Miguel Angeles	6
Agustin Lacruz, Maria del Carmen	5
Eito Brun, Ricardo	5
Granados, Mariangels	5
Montejo Raez, Arturo	5
Pastor Sanchez, Juan Antonio	5
Sanchez Jimeno, Rodrigo	5
Vicedo, Jose Luis	5
Authors with 4 works	12
Authors with 3 works	15
Authors with 2 works	67
Authors with 1 work	422

Table 9. Most productive authors of articles and monographs

3.1.5 Most cited authors

Along the lines of the previous methodology (and Jiménez 2004), we located citations of the works recorded in the ISI. Of the 96 publications identified, 41 were cited. Thus, we can say that the international visibility is greater, as we are speaking of 96 ISI papers as opposed to 17 in the previous decade. Similarly, we observed that the repercussions as measured in the number of citations and, in absolute terms, was also greater than in the previous period, since total citations received was 135 versus seven from the previous decade. An explanation of the increased international visibility of the Spanish research could be the fact that ISI has introduced Conference Proceedings in its database and, specially, the inclusion of two Spanish journals: *El Profesional de la Información*, which published 19 ISI selected articles which received 12 citations, and the *Revista Española de Documentación Científica*, with nine works selected and three citations thereof. The existence of these two journal might makes it easy for authors the publication process. An external cause could also be responsible for the increase: The Spanish Agency of Evaluation for Universities is more and more considering that international publications are a must for promotion. All, taken together, may explain this phenomenon.

Altogether, the number of citations received is distributed as shown in the Table below. We should underline that 60% of the citations (81) are concentrated in five papers; and it is also noteworthy that the most productive authors in the area of KO are not precisely the ones showing a greater number of citations of their work.

Regarding the geographic origin of the citing authors (Table 11), a considerable degree of diversification is seen, though the order of the first two positions is maintained with respect to the previous period. Logically, the top spot is occupied by Spain,

No. of publications	Citations received	No auto-citation	Identified works
1. (Moya, F. et al. A new technique for building maps of large scientific domains based on the cocitation of classes and categories	39	19	4
1. García-Berrocal, E. et al. Usability evaluation of ontology editors	15	14	2
1. Díaz, I. et al. A specification pattern for use cases	11	7	1
1. Zazo, N.F. et al. Reformulation of queries using similarity thesauri	10	8	3
1. Sánchez-Alonso, S. et al. Making use of upper ontologies to foster interoperability between SKOS concept schemes	6	4	4
1. Guerrero, V.P. Automatic extraction of relationships between terms by means of Kohonen's algorithm	4	4	2
6 works	3		
5 works	2		
24 works	1		
Total	135		

Table 10. Distribution of citations received

Origin of the First Author of the Citing Works	No. of publications
Spain	61
USA	23
Brazil	6
UK	5
Canada	4
China	4
South Korea	4
Germany	3
Mexico	3
Argentina	2
Australia	2
Belgium	2
Cuba	2
France	2
Italy	2
Taiwan	2
Colombia	1
Croatia	1
Ecuador	1
Slovenia	1
Finland	1
Holland	1
Jordan	1
Poland	1
Total	135

Table 11. Origin of the citations received

with 45.19% of the citing works; in second place, we find USA with 17.045% of the citing works. Grouping the citations by geopolitical areas shows Europe to be the first citing region (59.26% of works), followed by North America (20%), Iberoamerica (11%), the Far East (7.41%), and other countries, with a representation of 2.22%.

3.2 Institutional affiliation of authors

Spatial and institutional distribution of the authors, as summed up in Table 12, reflects the spatial and corporative geography of Spanish research in Knowledge Organization.

It is evident that the vast majority of authors are affiliated with Spanish universities (80.10%). This collective is followed closely by authors who work in non-university archives and libraries. Further behind stands the CSIC (Spain´s Scientific Research Council), foundations, and hospitals. The trend for universities to generate more publications appears as a constant, as in the previous decade of study, it showed 80% of output as well. Table 13 sums up these results. The degree of productivity of the universities reflects some changes with respect to the period 1992-2001, when the ranking was: Zaragoza, Carlos III, Murcia, Granada, Salamanca, Sevilla, Autónoma de Madrid, Valencia, and Barcelona. At any rate, we find that only five universities (Zaragoza, Carlos III, Murcia, Granada and Salamanca) were among the most productive in both periods.

INSTITUTIONS	% OF OUTOUT
UNIVERSITY	80.10
SERVICES (administration, archives, non-university libraries)	7.10
CSIC (*Consejo Superior de Investigaciones Científicas*)	2..96
PUBLIC AND PRIVATE ENTERPRISES	2.84
FOUNDATIONS	2.38
HOSPITALS	0.49
NOT LOCATED	4.00
Total	99.87

Table 12. Most productive institutions

In the ranking of the most productive universities (2002-2012), we have to point the entrance of the Universities of Jaén and Alicante, which do not have LIS in their curricula. The authors of the selected papers are working in the Department of Informatics in the former case and the Department of Languages and Informatic Systems in the latter one. In both cases, the articles written by them do not include an author from LIS Departments. This is another reason that let us consider, on top of the journals where they publish, the interest toward knowledge organization from other specialties, although their productivity in KO is much lower. In the same line, we could count more than 15 articles from authors belonging to Departments of Economy and Management of Enterprises. We also found authors belonging to the Departments of Informatics, Health Sciences, Translation, Psychology, or Architecture, in which production is lower than 10 papers. Nevertheless, it evidences the interest for KO from other specialties.

Likewise noteworthy is the presence of non-university entities, which generated 15.77% of total output. At the same time, we see some diversification

of the institutions involved, especially libraries and archives, but also film archives and press documentation centers. Moreover, there are centers that did not appear in the previous period, including diverse enterprises, foundations, and hospitals, responsible for 31 publications. The broadening area of interest in knowledge organization suggests greater social sensitivity regarding the benefits that it may hold for private companies or institutions, in special business and hospitals together with its use in informatics, social research, museums, etc. As stated above, these facts might mean recognition of the usefulness of knowledge organization in contexts not only linked to information retrieval.

3.3 Evolution of output over time

In general, output is seen to be more or less stable over the nine years studied here (Figure 1), ranking between the 29 publications of 2010 to as many as 53 in 2007. There are peaks of greater production in the years 2002, 2003, and 2007, coinciding with the publication of the ISKO Proceedings. These congresses also show the greatest volume of activity in the 1990's. We can therefore speak of a definite impact of ISKO events on the volume of Spain's scientific output in the area of knowledge organization.

3.4 Analysis of scientific output by gender

Without a doubt, approaching this type of analysis is of general interest, but it gains extra interest in a specialized field where women are present at all levels, whether as professionals, students, or teachers. We hoped to determine whether this reality was reflected in the scientific output. Yet it was impossible to determine the first name of some authors (and therefore their gender), since only first initials were used in some records. As an average value, we found that, for 41% of the studied publications, at least half of

SOURCE	No. of contributions	% Contributions
Universidad Carlos III de Madrid	40.83	11.44
Universidad de Granada	25.94	7.27
Universidad de Zaragoza	20.75	5.81
Universidad Complutense de Madrid	20	5.60
Universidad de Extremadura	14.33	4.01
Universidad de Murcia	14.25	3.99
Universidad de Salamanca	12.25	3.43
Universidad de Jaén	10	2.80
Universidad de Alcalá de Henares	9.91	2.78
Universidad de Alicante	9.40	2.63
Other Universities (39)	108.26	30.33

Table 13. Most productive universities 2002-2012

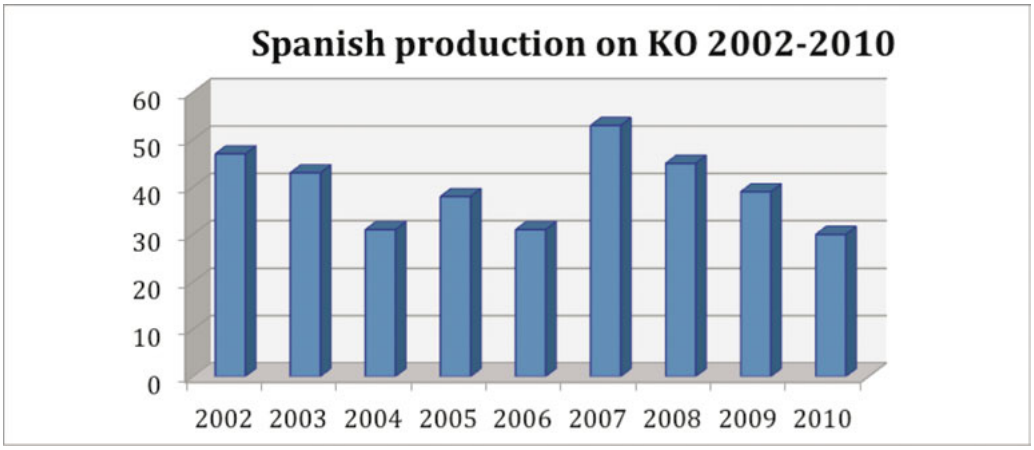

Figure 1. Output by year

Number of authors in papers	Without female authorship	1 female authorship	2 female authorships	3 female authorships	4 female authorships	5 or more female authorships	Total number of papers	% of papers signed by women	Papers with less than 50% of female authorship
1 author	93	66					159	58,49	58,49
2 authors	29	41	8				78	37,18	37,18
3 authors	26	19	11	1			57	45,61	45,61
4 authors	12	11	5	4	2		34	35,29	67,65
5 authors	2	8	4	2	0	0	16	12,50	62,50
6 or more authors	3	3	1	1	1	0	9	33,3	88,88
Without identification	4								
Totals	169	148	29	8	3	0	353	46,74	58,93

Table 14. Distribution of authors by gender

the undersigning authors were women. This trend depends on whether there was one author or more. Among the published works with just one author, women represent 42% of the total, but as the number of authors increases, so does the proportion of female authorship. Overall, we found that the more the co-authors, the higher the participation of women. At any rate, however, the differences were slight and do not point to any significant gender gap.

The following table displays the participation of women in authorship.

3.5 Distribution of output by subject

In order to carry out this part of the study, the content of each one of the published works was analyzed. We observed considerable thematic variety in the collection of publications, particularly in this second period. Furthermore, we found other topics that were not present in the previous period, which generated new terminology and the need for a certain internal restructuring of the subject matter. Although we attempted to maintain the thematic groups used previously, at times it was necessary to introduce changes due to the evolution of the area. In the first place, very general groups were drawn to provide a clearer view of the contents of the publications analyzed. Figure 2 shows the general subject areas and their percentage-wise distribution.

It is interesting to note that the terminology used to represent the contents of the output from 2002-2010, if compared with that of the previous decade, shows only one coincidence: *Knowledge Organization*—which here represents 11% of all output—stood for 13% in the previous period. Over the period 2002-2010, there were a number of terminological and conceptual changes, and only one of the groups of the previous period is still present, namely *Knowledge Organization Systems*. It is an expression rooted in the specialized area studied and which came to be largely to denominate what was once referred to as Documental Languages. For this reason, documental languages have been included within *Knowledge Organization Systems* together with all the specific types of systems: Classifications, Subject Headings, Thesauri, Taxonomies, Ontologies, etc., making it the most important group of the set, with 56% of the output.

The rest of the subjects that appear in the figure are novel, and we believe can be attributed to the fact that research into knowledge organization has become increasingly specialized and is now more focused on *searches*, with a presence of 4%, or *retrieval*, with 5%. Also new is *Knowledge Representation*, with 7% of the total, and which includes the study of any linguistic, conceptual, or algorithmic method used to represent the contents of documents in information systems. *Knowledge Processing* likewise appears for the first time in the realm of study, with 5% of output, but is

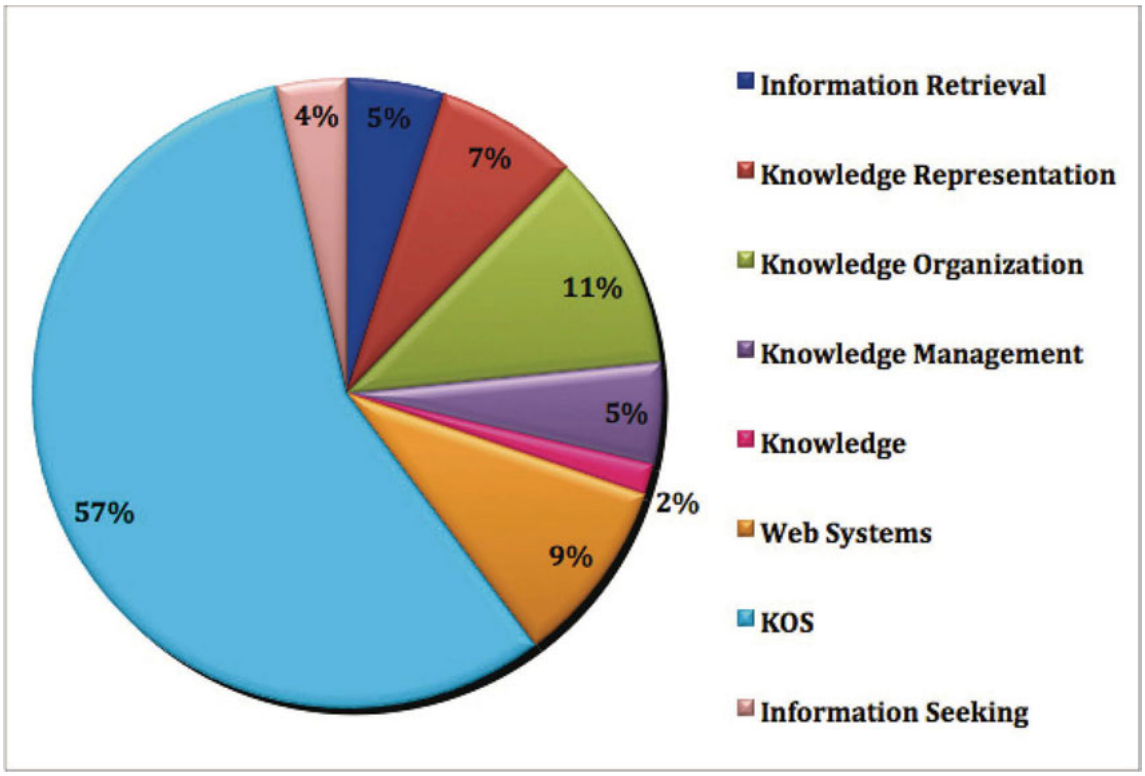

Figure 2. Display of percentage of general themes

oriented more towards managerial knowledge—that is for companies and organizations—still having strong connections with knowledge organization in a strict sense. For example, there are works describing the need to organize and prioritize this knowledge to later process it adequately, even in decision-making processes. Some authors focus on the study of *knowledge* in general, this minor group representing just 2% of output. *Web Systems* arise as a new subject area of interest in this second period, with a presence of 10%. The contents represented with this tag make mention of knowledge organization on the web, portals, social networks (Facebook, etc.) folksonomies, etc.

It is interesting to examine the internal composition of the most representative group, which is *Knowledge Organization Systems* and compare it, in turn, with the results of the period 1992-2001, as shown in the tables below.

Classification	76
Subject Headings	47
Documental Languages	5
Thesauri	236
Total documents	364

Table 15. Representation of the subject areas of the Documents 1992-2001

Comparison of Tables 14 and 15 makes evident that the total number of documents in this group is greater in the previous period than in 2002-2010; therefore, interest in these subjects is on the decline. We also observe a diversification of subject matter, doubling in the second period, and, except for documental languages which increase their presence, the rest of the topics decrease, especially the thesaurus, which suffers a dramatic drop from 236 to 63, and subject headings go from 47 to 8. It appears that the migration of interest on the part of researchers toward new systems and new subject areas would explain the present situation. Among the latter, a growing interest is seen in ontologies, which are 15% of the group.

Classification	56
Subject Headings	8
Documental Languages	12
Tesauri	63
Conceptual Maps	7
Ontologies	48
Taxonomies	8
Systems for Knowledge Organization	47
Total documents	249

Table 16. Representation of the topics in no. of documents 2002-2010

Figure 3 below displays the percentages of the group Knowledge Organization Systems.

It reflects an evident interest in the study of documental languages, approached from a general perspective, at 24%, and by the now denominated Knowledge Organization Systems, with 15%, which, if added, give us 39% of the total. If we more closely analyze the groups deserving mention, *thesauri*, with 20%, mostly refer to thesauri of different specialized areas; the rest of the documents look into norms, theory, the state of the art, and methodology for constructing them; *classification*, presenting an internal composition from top to bottom, which includes: specialized classification, with 25 documents, theory and general aspects of classification, with 18 documents, and bibliographic classifications, with 9 documents, 7 of them corresponding to the UDC. Finally, *ontologies*, with 15%, are mostly ontologies built for specific subject areas, whereas the rest deal with aspects related with construction theory and methodology. The comparative evolution of the subjects configuring the group Knowledge Organization Systems in the two periods 2002-2010 can be viewed in Figure 4.

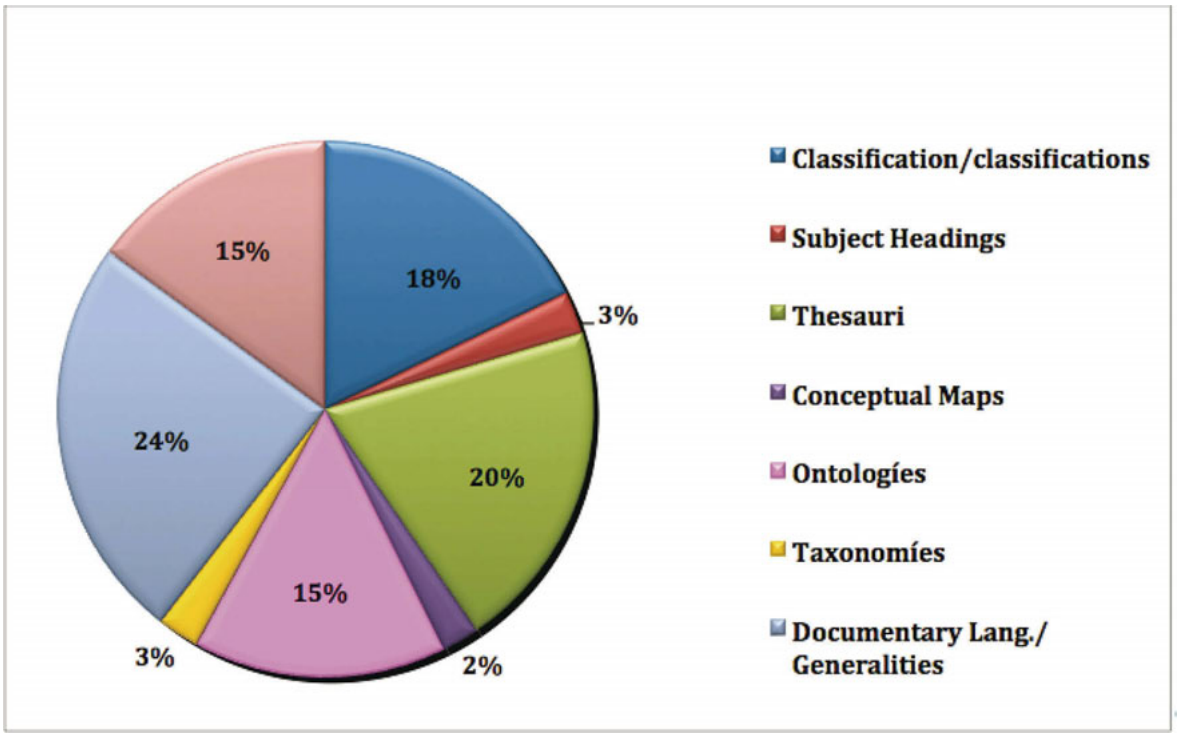

Figure 3. Percentage-wise distribution of the group Knowledge Organization Systems

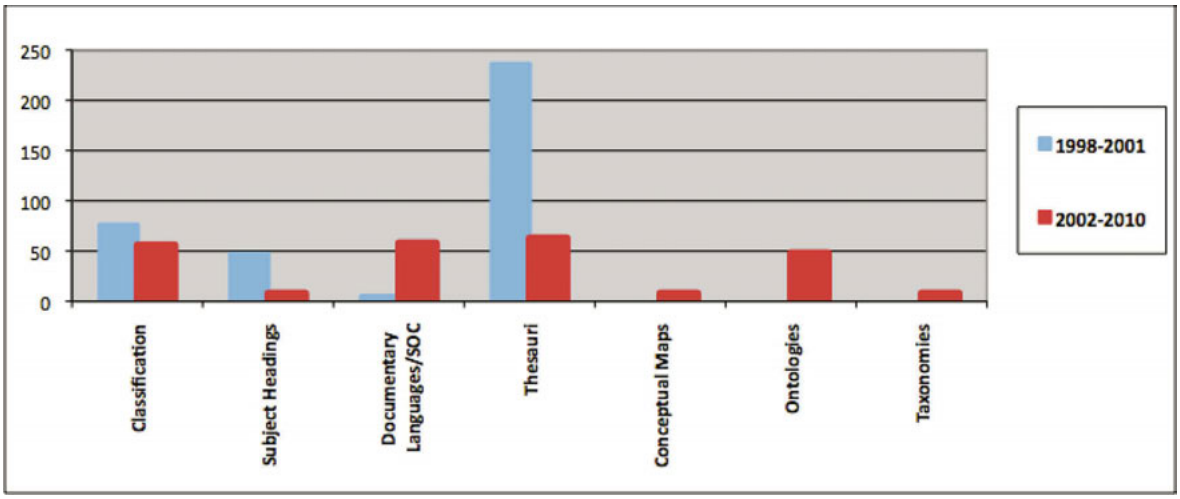

Figure 4. Comparative evolution of Knowledge Organization Systems (1992-2001 and 2001-2010)

4.0 Conclusions

In view of the results obtained and described here, we may affirm that research into knowledge organization is well consolidated in Spain, and indeed shows growth and development with respect to the previous decade.

The ISKO conferences—in particular the national ones, but also the international ones—clearly contribute to the increase in Spain's research output.

Aside from a slight increase in the number of publications by some of the most prolific Spanish authors from the previous decade, we detected a fresh influx of newcomers: only 12 authors from the period 1992-2010 are also active in the period 2002-2010. Deserving mention is the appearance of 35 novel authors in this field of study, although their output is limited: José Antonio Moreiro co-signed 14 articles, María J. López-Huertas co-authored 13, and Javier García Marco produced or co-produced 12.

Interest in knowledge organization regarding areas beyond library and information science is evident due to the presence of papers coming from specialties others than that of LIS, in special from informatics and economy-business. This piqued interest is no doubt partly responsible for the appearance of new authors on the list of the most productive researchers, such as Ureña and Montejo Vicedo or Sánchez Alonso from the Informatics Department. Likewise, our analysis allows us to confirm that this field of study is increasingly interdisciplinary.

Despite a discrete overall increase in output during the period 2002-2010, there is a manifest drop in monographic publications. This points to a change in perspective on the part of researchers; we believe that they now tend towards social sciences as the realm of dissemination of research findings. There were nearly 50% more Ph.D. dissertations in the second period of study, indicating a greater degree of interest in knowledge organization on the part of students enrolled in LIS studies.

According to the ISI, there was a noteworthy increase in knowledge organization studies stemming from Spain. We highlight this finding as a sign of heightened quality in research output and greater international visibility of Spanish research efforts in the area of knowledge organization. It also suggests a change in publishing habits, perhaps due to Spain's overall scientific policy and decision-making procedures. Such development translates as an increased citation of Spanish authors, as recorded by the ISI database.

Results suggest that the gender gap has receded. Women were roughly half of the co-authors of 41.07% of the papers produced by Spanish research institutions. Notwithstanding, the fact that there are more women researchers active at present sheds some essential light on these data. Indeed, we found that the greater the number of undersigning authors, the greater the proportion of female authors.

In short, a considerable change is seen in the arena of knowledge organization output from Spanish institutions. Five topics are seen to emerge with vigor: knowledge representation, information search and retrieval, web systems, and knowledge management. Deserving special mention is the group we denote as knowledge organization systems (KOS), which incorporates documentary languages. Its internal composition reveals that specialized areas such as ontologies, conceptual maps, and taxonomies are gaining research interest. The new area known as folksonomies, generally included under web systems, is also a topic of growing interest. In short, the growth in output documented here reflects conceptual advances in knowledge organization on the part of Spanish researchers on the whole.

References

Alves, Bruno Henrique, Gracio, Maria Cláudia Cabrini and Oliveira, Ely Francina Tannuri. 2011. A produção científica da revista *Scire*: uma análise bibliométrica do período 2006/2010. Presented at I Congresso Brasileiro de Organização e Representação do Conhecimento (ISKO-Brasil), Brasília..

Hjørland, Birger. 2008. What is knowledge organization (KO)? *Knowledge organization* 35: 86-101.

Hjørland, Birger. 2002. The methodology of constructing classification schemes: A discussion of the state of the art. In: López-Huertas, María ed., *Challenges in knowledge representation and organization for the 21st century. Integration of knowledge across boundaries. Proceedings of the 7th International ISKO Conference, 10-13 July 2001, Granada, Spain*. Würzburg: Ergon Verlag, pp. 450-56.

López-Huertas, Marí.a José and Jiménez Contreras, Evaristo. 2004. Spanish research in knowledge organization (1992-2001). *Knowledge organization* 31: 136-50.

Moneda Corrochano, Mercedes de la, López- Huertas, María José and Jiménez Contreras, Evaristo. 2011. La investigación sobre organización del conocimiento en España (2002-2010). In: *Resúmenes X Congreso Capítulo Español de ISKO, Ferrol, 30 junio - 1 julio*, p. 74

Oliveira, Ely Francina Tannuri; Gracio, Maria Claudia; Cabrini Silva, Ana Claudia. 2010. Investigadores de mayor visibilidad en organización y representación del conocimiento: un estudio desde el análisis de cocitaciones. *Scire*, 16n2: 39-45.

Ruiz Pérez, Rafael, Delgado López-Cózar, Emilio and Jiménez-Contreras, Evaristo. 2002. Spanish personal name variations in national and international biomedical databases: implications for information retrieval and bibliometric studies. *Journal of the Medical Library Association* 90: 411-30.

Smiraglia, Richard. 2005. About knowledge organization: An editorial. *Knowledge organization*, 32: 139-40.

Travieso-Rodríguez, Críspulo, Lascurain-Sánchez, María Luisa, Sal-Agüero, Alberto and Sanz-Casado, Elías. 2011. La organización del conocimiento en España a partir del análisis bibliométrico de los Congresos ISKO-Capítulo Español. In: *Resúmenes X Congreso Capítulo Español de ISKO, Ferrol, 30 junio - 1 julio*, p. 73.