

# Bases de datos bibliográficas y clasificación de revistas científicas: problemas de la interdisciplinariedad para la automatización de procesos

Beatriz Gutiérrez García<sup>1</sup>, Luis Rodríguez Yunta<sup>2</sup>, Adelaida Román Román<sup>3</sup>

Publicado en: *La interdisciplinariedad y la transdisciplinariedad en la organización del conocimiento científico. Actas del VIII Congreso ISKO-España, León, 18, 19 y 20 de Abril de 2007*. Universidad de León: Secretariado de Publicaciones, 2007: pp. 411-421.

<sup>1</sup>Consejo Superior de Investigaciones Científicas (CSIC), Centro de Información y Documentación Científica (CINDOC), Madrid, España, email: valrev@cindoc.csic.es

<sup>2</sup>CSIC, CINDOC, Madrid, España, email: luisry@cindoc.csic.es

<sup>3</sup>CSIC, CINDOC, Madrid, España, email: adelaida@cindoc.csic.es

## Resumen

En las bases de datos bibliográficas, los sistemas de trabajo tradicionales han recurrido mayoritariamente a la clasificación manual de cada artículo o documento individual. Sólo en casos excepcionales se recurría a la asignación de un epígrafe clasificatorio para la revista, afectando consecuentemente a todos sus artículos. Sin embargo, este modo de trabajo puede ser una alternativa más utilizada en el futuro inmediato ante la necesidad de automatizar los procesos de carga de datos. Sobre esta hipótesis, se plantea el problema de la interdisciplinariedad y el mantenimiento de perfiles disciplinares por parte de las publicaciones. Se realiza un análisis basado en las revistas españolas de las áreas de Humanidades. Se analizará la definición temática que realiza la propia publicación, su relación con el sistema de clasificación y la distribución temática porcentual de sus artículos en la base de datos ISOC. Se pretende extraer conclusiones prácticas sobre la viabilidad de un sistema de base de datos bibliográfica de actualización automática con asignación de clasificaciones a través del título de revista.

## Abstract

In the bibliographical databases, the traditional work systems have resorted mainly to the manual classification of each article or individual document. Only some database producers used to assign a classification heading to the journal, which affected consequently all its articles. Nevertheless, this way of work can be a more useful alternative in the immediate future because of the need of automating the data load processes. Taking this hypothesis as a starting point, the problem of the interdisciplinary nature of the research as well as the maintenance of disciplinary profiles by the journals are considered. An analysis based on the Spanish journals in the Humanities is made. The subject definition made by the publication, its relation with the classification system and the subject distribution by percentage of its articles in the ISOC database, will be analyzed. Practical conclusions will be drawn on the viability of a bibliographical database system automatically updated, with allocation of classification headings through the journal title.

**Palabras clave:** Bases de datos bibliográficas, revistas científicas, sistemas de clasificación, recuperación de información.

**Keyword:** Bibliographical databases, scientific journals, classification systems, information retrieval

## 1. Planteamiento y objetivo. Los sistemas de clasificación por revistas en las bases de datos bibliográficas

La producción de bases de datos bibliográficas se ha basado tradicionalmente en el análisis documental realizado artículo por artículo por especialistas (bien los propios autores, bien expertos o bien bibliotecarios o documentalistas especializados). La mayor parte de estos productos son herederos de los repertorios bibliográficos impresos, por lo cuál mantuvieron los mismos hábitos de trabajo en su elaboración. Aunque en algunos casos las bases de datos utilizan sistemas universales de clasificación, los sistemas predominantes son tablas de clasificación especializadas, realizadas a medida y asignadas de forma manual por los analistas a cada artículo. Sólo en casos muy singulares, la clasificación se ha asignado de forma genérica a cada revista, de modo que cada uno de sus artículos resulta encuadrado automáticamente en el mismo epígrafe.

El modelo tradicional, basado en un gran esfuerzo humano, será difícilmente sostenible. Sin embargo, la aplicación de técnicas de clasificación se considera indispensable para ofrecer a los usuarios un acceso sistemático y un limitador eficaz en la recuperación de información (GARCÍA MARCO, 1997), o para realizar evaluaciones de la actividad investigadora (URBANO, 2005). Otros autores lo consideran como una herramienta para desarrollar sistemas amigables de navegación o browsing (ESPELT, 1999). La utilización de estos sistemas a través de tablas de clasificación es aún escasa en los sistemas actuales de recuperación de información bibliográfica. Como señala Mari-Carmen Marcos, “tradicionalmente se ha menospreciado el browsing como técnica de búsqueda en las bibliotecas” (MARCOS, 2004)<sup>1</sup>. Sin embargo, se trata de un recurso de uso creciente en nuevos sistemas, como los archivos abiertos. Un ejemplo de ello puede verse en ArXiv<sup>2</sup>, en donde es posible navegar por categorías temáticas y visualizar directamente mediante un simple clic los últimos registros incorporados sobre un apartado concreto.

Por otra parte, la expansión de la edición electrónica y la aplicación de protocolos de transmisión o captura de metadatos, facilita la automatización de los procesos de actualización de bases de datos bibliográficas. ¿Qué metodología se utilizará de forma mayoritaria para la clasificación de los documentos en este nuevo entorno? ¿Es posible automatizar la asignación de clasificaciones? Se plantea como hipótesis el desarrollo de tres posibles vías:

- La asignación de clasificaciones en origen por los autores o editores, como parte de los metadatos que acompañan a los propios documentos en las plataformas de los editores y agregadores de publicaciones científicas. En esta opción, diferentes sistemas se alimentarán de forma automática a partir de una fuente única, pero la interoperabilidad de este proceso sólo parece aplicable a sistemas universales de clasificación.
- La utilización de sistemas expertos de clasificación automática por procesamiento de lenguaje natural, en las bases de datos y repositorios de documentos a texto completo. Aunque se trata de una opción ya aplicada en algunos servicios, es aún un campo experimental de compleja generalización, ya que exige diseños a medida.
- La asignación de epígrafes clasificatorios aplicados de forma genérica a las fuentes (revistas, congresos, etc). Esta posibilidad parece idónea para sistemas que mantengan clasificaciones realizadas a medida.

En esta comunicación se pretende valorar la viabilidad de esta última hipótesis, en el caso de las revistas españolas de Humanidades. No se trata de una opción novedosa, ya que ha sido aplicada tradicionalmente por el Institute of Scientific Information (ISI)<sup>3</sup>, uno de los productores de bases de datos bibliográficas de mayor influencia internacional. Su modelo de trabajo se ha basado en la utilización de los datos contenidos en las propias publicaciones primarias (sumarios y citas bibliográficas). En este sistema, las referencias bibliográficas de cada artículo no son clasificadas de forma individual, sino que quedan encuadradas en los epígrafes que se asignen a la revista en la que aparecen publicadas. En la base de datos *Current Contents* es posible utilizar las categorías temáticas como criterio directo de búsqueda. Por el

<sup>1</sup> Marcos, Mari-Carmen, 2004, p. 192-199.

<sup>2</sup> ArXiv.org e-Print archive: <http://arxiv.org/>

<sup>3</sup> ISI – Thomson Scientific: <http://www.isinet.com/>

contrario, en *Arts&Humanities Citation Index*, se aplican dentro de las opciones de refinar y analizar resultados (Web of Science 7.8 Workshop, 2006).

Un producto similar a los índices de citas del ISI es la base de datos *Scopus*<sup>4</sup>, que realiza así mismo una clasificación de las publicaciones, en categorías y subcategorías. Las bases de datos de sumarios elaboradas en España, también han procedido a una categorización por revistas. Sin embargo, no se aprovecha esta característica para ofrecer al usuario una posible limitación de sus búsquedas por artículos. *Dialnet*<sup>5</sup> no permite utilizar la clasificación temática para localizar documentos sino tan sólo en la búsqueda de revistas, al igual que la *Base de Dades de Sumaris*<sup>6</sup>. El portal *PubMed*<sup>7</sup>, sin duda uno de los servicios bibliográficos más utilizados internacionalmente, sí incluye, entre sus opciones de limitación de una búsqueda, la posibilidad de reducir la consulta a algunos subconjuntos de revistas.

En consecuencia, la aplicación de sistemas de asignación de epígrafes clasificatorios por revistas es una opción actualmente se puede considerar como secundaria en la presentación de productos bibliográficos. Sin embargo, resulta de fácil implementación como proceso automático en la alimentación de sistemas de información, y aporta un recurso de gran utilidad para la usabilidad de las interfaces de usuario. En esta comunicación se pretende analizar las consecuencias prácticas de este modelo, en el caso de las revistas españolas de Humanidades.

## 2. Metodología

Se analizan en primer lugar los datos sobre la adscripción disciplinar que realizan las propias publicaciones en sus páginas de presentación. Cuando no existe una formulación expresa del área temática de los trabajos que pueden publicarse, se ha completado este dato a partir de la categorización presente en el subtítulo o en la adscripción de la institución responsable de la selección de contenidos.

Esta autoclasificación se compara con la distribución disciplinar real de sus artículos, evaluada a partir de los registros de la base de datos ISOC para los números editados del año 2000 en adelante. Por ello, se ha tenido en cuenta en ambas fases el sistema de áreas temáticas definido en este producto.

Las bases de datos ISOC (ALCAÍN, 2001), elaboradas y distribuidas por el CINDOC, aplican un sistema de clasificación de elaboración propia. En sus inicios, estas tablas de clasificación fueron utilizadas para la presentación de las referencias bibliográficas en los repertorios bibliográficos: Índice Español de Humanidades e Índice Español de Ciencias Sociales. Estas publicaciones se dividieron en series por disciplinas, de modo que el primer nivel jerárquico de la clasificación marcaba la ubicación que tendría cada referencia que se introducía en el sistema. Esta división se aplicó también al acceso en línea a la base de datos, estableciendo subproductos temáticos, que se han mantenido, una vez desaparecidos los repertorios impresos.

La clasificación ISOC es un sistema de epígrafes constituidos por 6 dígitos a semejanza de los códigos de la UNESCO. Los dos primeros epígrafes definen la disciplina y son utilizados para la división en sub-bases para el acceso en línea. Se dividen en 16 conjuntos, de los cuáles ocho corresponden a Ciencias Humanas: Antropología y Etnología, Arqueología y Prehistoria, Bellas Artes, Filosofía, Geografía, Historia, Lingüística y Literatura. A estos apartados hay que añadir dos subconjuntos más de carácter transversal: estudios interdisciplinarios y estudios sobre América Latina. Este sistema de división por áreas temáticas es aplicado por el CINDOC tanto a la categorización de las revistas como a la clasificación de los artículos. En este segundo caso, determina la presencia de los documentos en cada una de las sub-bases que conforman el sistema ISOC

Para este análisis se han seleccionado las revistas vivas de Ciencias Humanas, categorizadas como de mayor calidad (grupos A y B), a fin de limitar el estudio a aquellas publicaciones cuyo valor científico y académico es más relevante. Este conjunto es resultado de un complejo proceso de categorización de revistas, que se llevó a cabo con el objetivo de establecer unos criterios mínimos para la inclusión de las revistas en la base de datos ISOC. Se establecen tres categorías, A, B y C. La asignación a alguna de estas categorías está en función del cumplimiento de diferentes parámetros que tienen en cuenta el nivel de

<sup>4</sup> Scopus, Elsevier B.V.: <http://www.info.scopus.com/>

<sup>5</sup> Dialnet, Universidad de La Rioja: <http://dialnet.unirioja.es/>

<sup>6</sup> Base de Dades de Sumaris, Consorci de Biblioteques Universitàries de Catalunya: <http://sumaris.cbuc.es/>

<sup>7</sup> PubMed, National Library of Medicine: <http://www.ncbi.nlm.nih.gov/entrez/>

normalización, la apertura en la composición de los órganos de gestión y en las autorías de las contribuciones, un sistema de revisores externos para la selección de originales y la puntuación asignada por el profesorado universitario en función de la calidad de contenidos (ROMAN, 2006). Los datos recopilados se refieren a la situación de las bases de datos ISOC a fecha 5 de diciembre de 2006.

### 3. Revistas españolas de Humanidades: autodefinición disciplinar

El CINDOC recoge un total de 1234 publicaciones de Ciencias Humanas en el *Directorio de Revistas de Ciencias Sociales y Humanidades*<sup>8</sup>. De ellas, hay 897 revistas vivas, que se distribuyen de la siguiente manera, en las ocho áreas temáticas de Humanidades:

Disciplinas	Categoría A	Categoría B	Otras
América Latina	3	2	24
Antropología y etnología	0	4	27
Arqueología y prehistoria	4	17	63
Bellas Artes	1	6	85
Filosofía	0	9	79
Geografía	4	10	41
Historia	12	30	161
Lingüística y Literatura <sup>9</sup>	9	28	149
Interdisciplinares	3	4	149
Total <sup>10</sup>	36	101	897

Tabla I: Número de revistas vivas, por categorías. Fuente: Base de datos en línea del Directorio de Revistas de Ciencias Sociales y Humanidades. 5-XII-2006.

En total, el conjunto de revistas que se analizan en este trabajo, publicaciones vivas de las categorías A o B<sup>11</sup>, está formado por 137 títulos. De estas revistas se ha procedido a recoger los datos de presentación, en los que debe establecerse las líneas temáticas, como información dirigida a lectores y posibles autores. De las 137 revistas tan sólo 97 incluyen una definición de ámbito temático. En la mayor parte de los casos, se trata de categorizaciones genéricas de lo que puede publicarse en la revista. Pero algunas revistas incluyen enumeraciones más prolijas de temas de interés o de áreas geográficas concretas (aspectos que a menudo quedan al margen de los contenidos descritos en los sistemas de clasificación)<sup>12</sup>.

Otro aspecto analizado ha consistido en ojear los sumarios para comprobar si las publicaciones contenían subdivisiones por secciones temáticas. De las 137 revistas seleccionadas, sólo 8 utilizaban este recurso, que puede interpretarse como una clasificación de sus artículos en origen. En otros casos se establecen apartados para ubicar dossieres monográficos, pero no constituyen secciones fijas.

A través de este proceso se estableció la categoría temática de cada una de las publicaciones. En ausencia de definición expresa, se completó este dato a partir de la categorización presente en el subtítulo o en la adscripción de la institución responsable de la selección de contenidos. Como resultado final de la categorización de las publicaciones de Humanidades, se han considerado tres casos diferentes, debido a que plantean condiciones peculiaridades para la hipótesis de este trabajo:

- Revistas que pueden clasificarse en una única área temática: 80 casos (58,4%). Incluye 2 de Antropología-Etnología, 12 de arqueología, 5 de Bellas Artes, 8 de Filosofía, 13 de Geografía, 33 de Historia, 5 de Literatura y 3 de Lingüística. En algunas de estas publicaciones la especialización temática es más precisa, pero también de difícil sistematización. Por ejemplo en Historia, el grupo de mayor tamaño, se distinguen revistas especializadas por épocas (Antigua, Medieval, Moderna o

<sup>8</sup> El Directorio está accesible en línea de forma gratuita en la dirección <http://bddoc.csic.es:8080/>

<sup>9</sup> En el Directorio de revistas las áreas de Lingüística y Literatura se consideran de forma conjunta, dado que el número de publicaciones conjuntas es mayor que el de aquellas que se especializan sólo en una de ellas.

<sup>10</sup> Algunas publicaciones se hayan incluidas en más de un área temática, por lo cuál la suma total de revistas es algo inferior a la mera suma aritmética de las cifras de cada apartado.

<sup>11</sup> Véase la lista completa de publicaciones en anexo que se incluirá de este documento en su versión depositada por los autores en el archivo abierto E-LIS <http://eprints.rclis.org>.

<sup>12</sup> Por ejemplo, materias presentes en la definición de algunas revistas, como Historia oral, Arqueología mediterránea, Historia de Al-Andalus, Filología francesa o Estudio Cervantinos, no quedan reflejados de forma precisa en las tablas de clasificación utilizadas en las bases de datos ISOC.

Contemporánea) mientras que otras se definen por líneas transversales a cualquier época (Historia Agraria, Económica, de la Ciencia, de las Instituciones, de las Religiones).

- Revistas que comparten dos disciplinas: 38 casos (27,7%). El grupo principal de estas características lo constituyen las 24 publicaciones de Literatura y Lingüística. Se trata en muchos casos de revistas especializadas en una rama de la filología: griega, latina, alemana, francesa, inglesa, española,... En la división disciplinar utilizada en la base de datos ISOC es preciso considerarlas en dos áreas temáticas, al igual que otras publicaciones que se definen directamente de esta manera.
- Revistas multidisciplinares: 19 casos (13,9%), que se definen como Multidisciplinares de Humanidades o que abarcan en su categorización más de 2 disciplinas. Se trata en muchos casos de publicaciones especializadas en un área cultural (estudios árabes, hebreos, clásicos), y que aceptan aportaciones desde cualquier disciplina.

#### 4. Grado de interdisciplinariedad, análisis de los artículos vaciados en la base de datos ISOC

Para este trabajo se han analizado los epígrafes clasificatorios, por áreas temáticas, asignados a cada uno de los artículos de las revistas seleccionadas, en las referencias bibliográficas introducidas en la base de datos ISOC. En este sistema de información, el campo de clasificación es obligatorio y multiocurrente, de modo que un mismo documento puede recibir hasta 4 epígrafes de clasificación diferentes.

Se analizan de forma independiente los tres casos establecidos en el apartado anterior:

##### 4.1. Revistas que se definen en una única área temática:

Se localizaron en ISOC 80 publicaciones de categoría A o B que se asignaban con claridad a una única área temática. En estas publicaciones se han recogido (a 5-XII-2006), un total de 7695 artículos editados desde el año 2000. Se analizó la proporción de documentos cuya adscripción disciplinar coincidía con la de la revista, con los siguientes resultados globales:

Disciplina (nº revistas)	Nº reg <sup>13</sup>	Misma Clas	Id. + otras	Otras clas.
Antropología (2)	104	103	40	1
Arqueología (12)	873	846	92	27
Bellas Artes (5)	692	678	22	14
Filosofía (8)	811	809	27	2
Geografía (13)	1390	1389	19	1
Historia (33)	3162	3059	117	103
Literatura (5)	356	326	12	30
Lingüística (3)	166	166	4	0
Total (80)	7554 (100%)	7376 (97,6%)	333 (4,4%)	178 (2,4%)

Tabla II: Distribución temática de los artículos incluidos en revistas especializadas en una única disciplina. Fuente: Base de datos ISOC. 5-XII-2006, revistas vivas de Humanidades, categorías A o B.

A través de estos datos se observa que, en las revistas de estas características, si los epígrafes de clasificación de los artículos se hubiesen asignado de forma automática a partir del área de la revista, en lugar de haber sido asignados uno a uno por un analista, los resultados habrían sido muy similares:

- El porcentaje de ruido (epígrafes incorrectos que no se corresponden con el área temática de la revista), se limita a un 2,4%. En el peor de los casos, se sitúa en el 8,4% en el área de Literatura.
- El porcentaje de silencio (epígrafes pertinentes de áreas temáticas ajenas a la de la revista), supondría tan sólo un 6,8%. Sin embargo, si se sitúa en cotas superiores al 10% en algunas disciplinas: 39,4% en Antropología, 13,6% en Arqueología y 11,8% en Literatura.

<sup>13</sup> El número de registros se corresponde con la suma de las columnas 3 (misma clas) y 5 (otras clas). Debe tenerse en cuenta que las cifras de la cuarta columna (id.+ otras) ya están incluidas en la tercera (misma clas.), ya que se refiere a los artículos que han sido clasificados en más de una disciplina, siendo una de ellas la propia de la revista. Se ha establecido esta diferenciación para localizar el número de documentos con epígrafes diferentes a la clasificación de la revista.

#### 4.2. Revistas que comparten dos disciplinas:

Se localizaron en ISOC 37 publicaciones de categoría A o B que se encuadraban a caballo entre dos áreas temáticas. En estas publicaciones se han recogido (a 5-XII-2006), un total de 3716 artículos editados desde el año 2000. Se analizó la proporción de documentos cuya adscripción disciplinar coincidía con estas dos áreas de la revista, con los siguientes resultados globales:

Disciplinas (nº revistas)	Nº reg <sup>14</sup>	Area A	Area B	A y B	A o B + otras	Otras no A ni B
Antropología+Lingüística (1)	150	146	14	12	75	2
Arqueología+Historia (4)	356	291	77	23	26	11
Arqueología+Antropología (1)	36	36	0	0	1	0
Bellas Artes+Educación (1)	214	100	149	41	1	6
Filosofía+Derecho (1)	138	63	132	57	0	0
Historia+Antropología (2)	287	225	87	39	43	14
Historia+Derecho (1)	75	74	3	2	2	0
Historia+Geografía (1)	60	53	51	44	0	0
Historia+Economía (2)	198	195	86	83	1	0
Literatura+Lingüística (24)	2476	1343	1049	80	29	164
Total (38)	3990	2519	1655	381	178	197
	(100%)	(63,1%)	(41,5%)	(9,5%)	(4,5%)	(4,9%)

Tabla III: Distribución temática de los artículos incluidos en revistas que se encuadran en dos disciplinas. Fuente: Base de datos ISOC. 5-XII-2006, revistas vivas de Humanidades, categorías A o B.

Si se aplicase un sistema de asignación automática a partir del área de la revista en las publicaciones de estas características, lo más habitual es considerar indistintamente que todos los documentos comparten estas dos disciplinas en todos los artículos. Por tanto, en este caso, las diferencias con la asignación de epígrafes por un analista, si son importantes:

- El porcentaje de ruido (epígrafes incorrectos que no se corresponden con el área temática de la revista), supondría un valor medio entre el 36,9% (en la primera área temática) y el 58,5% (en la segunda). Pero, en el caso más extremo, llega al 100% de los documentos que se asignarían al área de Antropología en una revista que se autodefine como de Arqueología y Antropología, pero que no ha publicado ningún artículo de esta segunda disciplina a partir del 2000.
- El porcentaje de silencio (epígrafes pertinentes de áreas temáticas ajenas a las de la revista), supondría tan sólo un 9,4%. Sin embargo, sí se sitúa en cotas superiores al 10% en algunos casos particulares: 51,3% en el caso de Antropología + Lingüística, 19,8% en Historia + Antropología y 10,4% en Arqueología + Historia.

#### 4.3. Revistas multidisciplinares

En este grupo se encuadran un total de 19 publicaciones que se definen como Multidisciplinares de Humanidades o que abarcan más de 2 disciplinas. La dispersión en áreas temáticas de sus artículos es muy variable.

Revistas multidisciplinares (19)	Valor mínimo	Valor máximo	Valor medio
Porcentaje de la disciplina mayoritaria	32,1%	97,4%	53,3%
Disciplinas diferentes	5	11	6,8
Disciplinas con más del 20%	1	3	2,1

Tabla IV: Distribución temática de los artículos incluidos en revistas multidisciplinares de Humanidades. Fuente: Base de datos ISOC. 5-XII-2006, revistas vivas de Humanidades, categorías A o B.

<sup>14</sup> En esta tabla, el número de registros se corresponde con la suma de las columnas 3 (Área A), 4 (Área B) y 7 (Otras no A ni B), pero restando luego la columna 5 (A y B), para evitar contabilizar dos veces los documentos que pertenecen a ambas áreas a la vez. Debe tenerse en cuenta que las cifras de la columna 6 (A o B+otras) ya están incluidas en la 3 (Área A) o en la 4 (Área B), ya que se refiere a los artículos que han sido clasificados en más de una disciplina. Las de la columna 5 (AyB) están presentes tanto en la 3 (Área A) como en la 4 (Área B).

Las publicaciones especializadas en estudios clásicos (7 revistas) y estudios árabes (5 revistas), representan los casos particulares predominantes en la selección de revistas realizada. En la siguiente tabla, se analiza el grado de presencia de las diferentes áreas temáticas, en ambos subconjuntos:

<b>Disciplina</b>	<b>Estudios clásicos</b>	<b>Estudios árabes</b>
Antropología	15	4
Arqueología	180	9
Bellas Artes	30	13
Filosofía	25	10
Geografía	6	4
Historia	330	147
Lingüística	89	55
Literatura	215	71

Tabla V: Distribución temática de los artículos incluidos en revistas multidisciplinares de Estudios Clásicos y Estudios Árabes. Fuente: Base de datos ISOC. 5-XII-2006, revistas vivas de Humanidades, categorías A o B.

Para aplicar un sistema de asignación automática a partir del área de la revista en este conjunto de publicaciones, cabrían diferentes opciones:

- Incluir cada artículo en todas las áreas temáticas. El porcentaje de ruido sería muy elevado, aunque variable según el caso, pero llegaría al 100% en aquellas disciplinas que no estuvieran presentes en determinadas publicaciones.
- Asignar a los artículos sólo las áreas temáticas con presencia superior al 20% (porcentaje que puede establecer mediante un sondeo de los números editados en un periodo de dos años). Con esta opción se reduciría el ruido, pero aumentaría el silencio, ya que los documentos de disciplinas de menor presencia, dejarían de aparecer en sus respectivos grupos de clasificación.
- Crear un epígrafe independiente en las áreas temáticas para las publicaciones multidisciplinares. Con ello, la decisión queda en manos del usuario que puede incluir o no estas publicaciones al realizar su agrupación disciplinar en las búsquedas. Con esta opción, siempre que se limite la recuperación a una disciplina concreta, los artículos encuadrados en revistas multidisciplinares quedarán condenados al silencio documental.

## 5. Conclusiones

De los datos recogidos se desprende que la aplicación de un sistema automático de asignación de epígrafes clasificatorios a los artículos a partir de la categorización previa de las revistas es una opción válida solamente en determinados tipos de publicaciones, aquellas que se especializan en una única disciplina. En este caso, los niveles de ruido y silencio son muy reducidos (2,4% y 6,8% respectivamente). Esta circunstancia afecta al 58,4% de las revistas analizadas.

Por el contrario, el 41,6% de las publicaciones analizadas se encuadran en dos o más disciplinas. En estos casos, la recuperación de información se vería claramente condicionada por la presencia de niveles muy variables de ruido y silencio en las diferentes áreas temáticas.

El estudio se ha reducido a las publicaciones consideradas de mayor calidad editorial y reconocimiento por los propios investigadores. Con ello no se pretende extrapolar los datos y porcentajes presentados, al conjunto de revistas de Humanidades. Tan sólo se ha tomado como base para plantear el problema y establecer un panorama de las consecuencias que tendría un sistema de asignación de clasificaciones en base a la revista, de cara al ruido y silencio documental en la recuperación de información.

En consecuencia, no debería utilizarse un sistema de asignación automática de epígrafes clasificatorios en los documentos basado en la categorización de las revistas, sin aplicar mecanismos correctores. Una opción válida sería la presencia de secciones específicas en los sumarios de las propias publicaciones primarias de carácter multidisciplinar. Con ello, se limitaría considerablemente el problema del ruido por asignación errónea de epígrafes clasificatorios no pertinentes en los documentos. Sin embargo, se trata de una práctica actualmente muy poco extendida (un 5,8% de las publicaciones seleccionadas).

Del análisis realizado también se desprende la existencia de diferencias entre la clasificación empleada en una base de datos y la autoadscripción disciplinar de las publicaciones. En estudios posteriores debería profundizarse en cómo resolver el problema de encuadrar en una tabla de clasificación de revistas la riqueza de matices que muestran algunas revistas en su definición.

## 6. Bibliografía

ALCAÍN, Dolores; BARANDA, Piedad; RODRÍGUEZ YUNTA, Luis; ROMÁN, Adelaida; VILLAGRÁ, Angel. Evaluación de la base de datos ISOC a través de un estudio de usuarios. Homenaje a José María Sánchez Nistal. *Revista Española de Documentación Científica*, 2001, v. 24, n° 3, pp. 275-288  
[http://bddoc.csic.es:8080/basisbwdocs\\_rdisoc/rev0001/2001\\_vol24-3/2001\\_vol24-3\\_pp275-288.htm](http://bddoc.csic.es:8080/basisbwdocs_rdisoc/rev0001/2001_vol24-3/2001_vol24-3_pp275-288.htm)  
[Consulta: 5-12-2006]

ESPELT, Constança. Improving subject retrieval: user-friendly interfaces and effectiveness. *BiD: textos universitaris de biblioteconomia i documentació*, 1999, n° 1.  
<http://www.ub.es/biblio/bid/01espel1.htm> [Consulta: 10-12-2006].

GARCÍA MARCO, Francisco Javier. Clasificación y recuperación de información. En: PINTO, María (Ed.). *Manual de clasificación documental*. Madrid: Editorial Síntesis, 1997, pp. 247-285.

MARCOS, Mari-Carmen. *Interacción en interfaces de recuperación de información: conceptos, metáforas y visualización*. Gijón: Trea, 2004.

ROMÁN, Adelaida; ALCAÍN, María Dolores. Una apuesta por la calidad en la selección de las revistas españolas. *Clip. Boletín de la SEDIC*, 2006, n° 44.  
[http://www.sedic.es/p\\_boletinclip44\\_confirma.htm](http://www.sedic.es/p_boletinclip44_confirma.htm) [Consulta: 12-12-2006]

URBANO, Cristóbal; SEGUÍ, Rosa; BORREGO, Ángel. Clasificar el conocimiento para evaluar la generación de conocimiento. Clasificación y evaluación de la investigación en Ciencias Humanas y Sociales. En: *VII Congreso ISKO-España: La dimensión humana de la organización del conocimiento*, Barcelona, 2005, pp. 597-612.  
<http://bd.ub.es/isko2005/urbano.pdf> [Consulta: 22-11-2006]

*Web of Science 7.8 Workshop*. The Thompson Corporation: 2006  
<http://scientific.thomson.com/media/scpdf/WOSTRAIN78ES.pdf>, [Consulta: 4-12-2006].