AN OVERVIEW OF EUROPEAN SUPPLY OF SOCIAL SCIENCE INFORMATION *

Paper presented in ECSSID VI Conference, Canterbury, 1991

Jean Meyriat*
Adelaida Román Román**

Resumen: Tomando como fuente de información para esta comunicación el Repertorio EUROSPES de Bases de Datos en Ciencias Sociales editado en 1990, se analiza la oferta de Información a través de las Bases de Datos creadas en los países europeos de habla latina. El objetivo del trabajo que se presenta es conocer, de manera detallada, los rasgos que definen esta oferta en cada país y en el conjunto de países estudiados. Estructurado en dos partes, en la primera se ofrecen, en primer lugar, datos sobre las características de la información contenida en las bases de datos: naturaleza de la información, temas sobre los que versa, cobertura temporal, técnicas documentales utilizadas, normas de descripción bibliográfica, productores de esas Bases de Datos. En segundo lugar, se estudian los aspectos relativos a la distribución de la información: distribuidores, paquetes de gestión de bases de datos, instrumentos de ayuda al usuario para las consultas, tipo de acceso que se ofrece. La segunda parte se dedica al análisis de los datos aportados en la primera: distribución de la producción de bases de datos por temas, países y años, estudio de las áreas temáticas no cubiertas, comparación de las técnicas documentales empleadas, compatibilidad de los paquetes de gestión y de los formatos informáticos, análisis de la calidad de los servicios de acceso, precios, formación ofrecida, etc. Se finaliza con unas conclusiones orientadas a proponer los desarrollos deseables para el conjunto de países así como las posibilidades de cooperación con vistas a mejorar la oferta de información automatizada en Ciencias Sociales.

Palabras clave: Ciencias sociales, bases de datos, España, Francia, Italia, Portugal, Bélgica, análisis comparativo.

Abstract: On the basis of the EUROSPES Directory of Social Sciences Data Bases, issued in 1990, this paper aims to analyze the information supply throughout the data bases created in the latin language European countries. It is divided into two parts. In the first part the study shows the main traits of the information contained in the data bases: kind of information offered, subjects covered, information technologies used, bibliographic rules, etc. The information diffusion is also studied: vendors, software, helping tools for users, services offered for access. The second part is devoted to the analysis of the data described in the first part: a comparative study on the availability of information by country is given: this comparative analysis goes over softwares compatibility, exchange formats, quality of access services, prices, training offered, etc. It finishes with some conclusions about the foreseen developments and also about the possibility of cooperation in order to a more complete and efficient supply of automatic information on Social Sciences.

Keywords: Social sciences, databases, Spain, France, Italy, Portugal, Belgium, comparative analysis.

An overview of European supply of social science information

EUROSPES is an informal grouping of the main institutions dealing with social science information and documentation in a number of countries of Western Europe. These countries are those using mainly romance languages: Belgium, France, Italy, Portugal and Spain. One of the main aims of EUROSPES is to promote the electronic delivery of social science information.

For this purpose, it compiled and published a *Directory of Social Science Data Bases*, (1) issued in 1990, after a first preliminary edition dated 1987. The Directory describes 173 bases, comprising 24 produced in Belgium, 65 in France, 41 in Italy, 14 in Portugal, and 24 in Spain. In addition 3 bases produced in Luxembourg are also described (but they are not national, originating within the ECE), and 2 produced in Switzerland (but they are not fully representative of the country). In this report, reference will only be made to the five major producing countries.

The Directory intends to cover all types of data bases: bibliographical, textual, factual and numerical. In principle only bases available to the general public by electronic means are covered, but it was found convenient to include some bases only planning to give this access during the next few months. All data were, when possible, provided and thereafter checked by the producers themselves; many letters were sent and received for this purpose during 1988 and 1989. It is a matter of course that such a process does not guarantee that all information is updated at the same moment, as the whole set of data has to be retained until the answers from the slowest correspondents have been received. However, it may be said that the Directory gives a fair representation of the situation prevailing at the end of 1989. Everybody knows that any situation, in this field is rapidly changing, and we do not deal with that of March 1991. Even for 1989, we accept that the Directory may be incomplete and sometimes inaccurate.

A secondary analysis of the data collected in this Directory allows us to report on social science information supplied at the end of 1989 via electronic media by romance-language speaking countries of Western Europe. This report will deal with three aspects: origin and coverage of the bases, technologies used and linguistic contents.

Figures will be given as percentages, unless otherwise stated. Admittedly, percentages may be meaningless when referring to so small totals as e.g. 14 bases for Portugal; but it is the only practical way to make comparable data from several countries. However, account has to be taken of the limited number of producers in some countries; when there is somewhere a body producing an important proportion of the national bases, generally using the same technologies, the global percentages are *ipso facto* biased. This is for instance the case of Italy, where the Corte Suprema di Cassazione produces 21 bases from a total of 41, or even of France, where the Institut de l'Information Scientifique et Technique produces 12 bases from a total of 65.

1. Origin and coverage

In the five countries, the bases are mostly recent, if we consider the dates of their computerization —which in many cases followed a sometimes large period of

manual collecting and processing of data. Only 17 % of the bases in Belgium were created before 1981, 47 % in France, 24 % in Italy, 7 % in Portugal and 12 % in Spain (the actual figures being 4 in Belgium, 3 in Spain and 1 in Portugal). France began to follow this path earlier: 9 % of its computerized bases were already in operation in 1970.

In four of these countries, producers are predominantly in the public sector, the exception being Belgium, where 63 % of the bases are produced by private companies. But 77 % of the French bases and 76 % of the Italian ones, come from the public sector. Only one producer in Portugal against 10 and one in Spain against 16, is private. It appears that in Latin countries, there is some slowness in establishing a profit-making information industry.

The majority of the bases remain bibliographic: 67 % in Belgium, 58 % in France, 54 % in Portugal, 71 % in Spain. The exception is here Italy, where 51 % of the bases are textual. This is due to the large number of bases produced by the Corte Suprema di Cassazione and the Camera dei Deputati, making available legal and judicial texts. Textual bases are only 17 % in Belgium, 23 % in France, 12 % in Spain and do not exist in Portugal.

An analysis of the subjects covered cannot go very far, for two reasons. First, the compilers of the Directory decided to include some important bases which are not devoted to the social sciences, but are more general and contain *inter alia* a wealth of social science data. This is e.g. the case of PORBASE (Portugal), BIBL (Spain), ORIADOC (France), etc. The second reason is that «social sciences» are understood with a loose meaning: a number of the bases have no connection with any science, but cover data produced by business, by financial activities, by daily life... which could be material for analysis by social scientists (2). Not surprisingly, 12 % of the bases in Belgium, in France as well as in Spain cannot be categorized otherwise than «multidisciplinary».

The two fields best covered, in the majority of the five countries, are: law and jurisprudence, by 49 % of the bases in Italy (for the specific reasons mentioned above), 25 % in Belgium, 21 % in Spain, 20 % in France; and economics and society, by 28 % of the bases in Portugal, 25 % in Spain, 25 % in Belgium, 17 % in Italy.

Other concentrations of interest are more specific to one or two countries, e.g. enterprises (28 % in Portugal, 10 % in Italy), labour and employment (12 % in Belgium), parliamentary and political life (15 % in France). Only in France a number of bases are specifically dedicated to one or other among the social sciences, such as sociology, ethnology, economics, educational sciences...

Taking an opposite viewpoint on the whole situation, one may conclude that in each of these countries, some important topics are not yet covered by electronic information in an appropriate way; room is left for development (3) (4).

2. Technical and functional aspects

There is a limited amount of standardization, not only at the global level of the five countries, but also within each country. This would make more difficult the organization of cooperation and the setting up of exchange networks between diverse bases.

Within the subset of bibliographical data bases, even the bibliographical description of the documents is poorly standardized. The number of bases purporting to follow the prescriptions of ISBD, or of national cataloguing standards (which themselves conform to ISBD) is limited: 19 % in Belgium, 53 % in France, 11 % in Italy, 67 % in Portugal, 29 % in Spain. Some bases refer to UNISIST guidelines, without more precision. All the others use internal rules.

Diversity is even greater in computer exchange formats. Most of the bases are unable or unwilling to specify what format they rely on. When they mention a standardized format, one finds 2 different formats in Belgium, 10 in France, 3 in Italy, 4 in Potugal, 7 in Spain. Among them, only STAIRS is broadly represented, with the exception of France: 9 times in Belgium, 4 in Italy, 2 in Portugal, 2 in Spain. Some bases mention MARC, but apparently they refer to different members of the MARC family; only UNIMARC is present in two countries (1 base in France, 3 in Portugal). Some others mention ISO 2709, but this is rather a general frame hosting several formats. MISTRAL is used by 4 bases in France, 2 in Spain.

Some formats are used in only one country: BASIS by 12 bases in Spain, DSO by 10 bases hosted in France by L'Européenne de données, TEXTO by 5 bases in France. In this same country, 12 bases produced by l'Institut de l'Information scientifique et technique use the same format, called 6-74, which has not been exported.

Due to the diversity of the data concerned, it is difficult to present general comments on the documentary processing they receive. This difficulty had also to be met by the compilers of the Directory who could not offer completely homogeneous information on the topic (maybe because there is no general agreement on the use of technical terms as «keyword», «index», «classification»...)

Very few of the bases are satisfied with only cataloguing and classifying the data. As a matter of fact they are most often indexed in one way or another; more often than not, indexing relies on a controlled vocabulary, which is frequently a thesaurus. A majority of the bases add summaries or abstracts to the keywords or descriptors, but the Directory does not inform on the types of summaries.

Access to the data is made available either directly by the producer or by a computing center which hosts the base and sells the data. A limited number of bases use both methods. Some are hosted by two or even three different vendors. For this reason, the sum of calculated percentages may exceed 100.

There are significant differences between countries under this respect. In Belgium, Portugal and Spain, most producers directly give access to their data. In Belgium there is only one vendor (exception being made for two hosts of the European Communities, ECHO and EUROBASES): BELINDIS, set up by the Ministry of Economic Affairs, which hosts one base produced by the same Ministry plus five from other origins. In Portugal, there are 10 computing centers for 14 bases; in Spain, 13 for 24 bases (7 of them being distributed by the same Center for data processing of the Ministry of Education and Science).

In Italy, there are only 6 vendors, 5 of them being at the same time the most important producers: only one base is hosted by an independent vendor. This explains, in that country, an unusual uniformity of the means of access to the data (see below).

In France, 23 % of the bases directly give access to their data, or are hosted by a specialized computing center. But most frequently use is made of the services of a commercial vendor. 52 % of all the bases are hosted by L'Européenne de données; 26% by Télésystèmes-Questel; 11 % by SUNIST (a vendor set up by the Ministry of Education). In addition, two bases aiming at an international market are also hosted by ESA-IRS, and one by DIALOG. Here also we may expect some degree of uniformity in the means of access.

The softwares used to retrieve data from the bases depend upon the above-mentioned conditions. 8 % of the bases, in Belgium as in France, rely on home-tailored softwares; the figure is 17 % in Italy as in Spain; it is so high as 50 % in Portugal. Some bases do not clarify what software they use. In all other cases retrieval is made by a standard software. But only one, STAIRS, is used in all countries excepting France: by 46 % of the Belgian bases, 32% of the Italian, 21 % of the Portuguese, 12 % of the Spanish. MISTRAL is used in Belgium, France and Spain, but only by 2 % from the total number of the bases.

Other softwares, well known internationally, are scarcely used: BASIS only by 50 % of Spanish bases, GEAC and DBASE III each by two bases in Portugal, CDS/ISIS by three bases also in Portugal, DOBIS-LIBIS by only one base in Spain. Some softwares are merely national: ITALGIURE-FIND, used in Italy by 49 % of the bases; and in France QUESTEL + (46 %), DSO of DATASTAR (37 %), TEXTO (9 %). This is often due to their installation at a major vendor. The same occurs with UNIDAS in Spain (29 % of the bases).

To access the data stored by the computers, diverse means are offered. Comparisons are meaningful only for four countries; the case of France is different and has to be separately dealt with. Few of the bases give only on line access via the public telephone network: 25 % in Belgium, 7 % in Italy, 14 % in Portugal, 17 % in Spain. Much more often the telephone is supplemented by other means, preferably a data transmission network. DCS is used by 54 % of the bases in Belgium, ITAPAC by 93 % in Italy, IBERPARC by 54 % in Spain, TELEPAC by 57 % in Portugal. Telephonic connection via a rented line is used only in Belgium by 33 % of the bases and in Portugal by 43 %.

Videotex is still very unfrequent: three bases use it in Belgium, one in Portugal, one in Spain. Spain alone resorts to CD-ROM to disseminate data: 29 % from its bases do so. Additionaly, two producers in Spain and one in Belgium sell magnetic tapes.

On the contrary, the great majority of the bases in France (92 %) resort to videotex (Minitel), jointly with one or several other means. The data network TRANSPAC gives access to 91 % of the bases, and the telex to 54 %. In addition, the public telephone network is used by 5 %, and rented lines by another 5 %. 15 % also sell magnetic tapes; 5 % only (i.e. 3 producers) publish CD-ROMs.

Most of the producers and/or vendors provide assistance to users and train them to an appropriate use of their products; there is no significant difference in this respect among the five countries. More than half of them issue for this purpose handbooks or guides. Surprisingly enough, few among them put forward the availability of the thesauri or lexicons they use for indexing the data: 22 % from the French bases, but only one Belgian, 2 Italian, 2 Portuguese and 2 Spanish.

There are also significant differences in the charges paid by the users. In the four countries where online access is as yet less developed, a number of bases are open free of charge: 29 % in Belgium, 54 % in Italy, 86 % in Portugal, 67 % in Spain. This situation, which in the Mediterranean countries is clearly related to the public status of most of the bases, and to the will to stimulate the demand and create new needs, does not exist any more in France.

When there is a charge, it takes often in Italy (41 %) and sometimes in Spain (12 %) the form of an annual subscription. Belgian and French bases rather charge on the basis of the duration of connections, either alone (17 % in Belgium, 29 % in France) or preferably supplemented by a charge for each data retrieved (33 % of the bases in Belgium, 62 % in France). 17 % of the Spanish bases also charge for the time of connection. Another 17 % of the Spanish bases and 5 % of the French charge only for the number of data retrieved.

3. Linguistic aspects

Given the significance for European cooperation of due consideration to the diversity of languages prevailing in the region, an analysis was made of the languages covered and used by the bases described in this Directory.

As a matter of fact, analysis was limited to information published in the Directory, which is not available for all bases. We could include all textual bases and most of the bibliographic and factual ones (inventories). Statistical bases and some factual ones are not taken here into consideration, due to their non-verbal nature. So we shall thereafter deal with only 21 bases in Belgium, 59 in France, 33 in Italy, 11 in Portugal and 22 in Spain. Percentages, when given, will refer to these figures.

Belgium is here a special case. It is a bilingual country, and «national language» has to be understood either as French, Flemish or both. Furthermore, it is the headquarters of European institutions, and several among the bases produced in Belgium (and labelled in the Directory as Belgian) are in fact international bases, so experiencing specific linguistic conditions. It may be added that at least one from the French bases is in the same situation.

As far as coverage is concerned, most of the bases cover exclusively or mainly (more than 80 %) materials written in the national language (or languages in Spain): this is the case for 73 % of the bases in France, 88 % in Italy, 36 % in Portugal, 82 % in Spain. In addition, 7 % of the bases in France, 9 % in Italy, 36 % in Portugal and 9 % in Spain contain between 50 % and 80 % documents in the national language. When these ones are less than half the total, they leave room to other European languages (3 bases in France, 1 in Italy, 3 in Portugal, 2 in Spain). In France, 6 bases contain a fair proportion of documents in English, and 3 others even contain a majority of documents in English.

From the 21 bases produced in Belgium, only 3 cover documents in French, 1 in Flemish, 2 in both languages. But 10 are multilingual, covering documents from various countries of the European Communities, and five use at the same time national language(s) and English: in 3 of them, more than half of the documents are in English, and one base translates all documents into English.

It is of interest also to know which languages give access to the data. Here again there is a difference between Belgium and the other four countries. From the

18 bases in Belgium which explicitely state the use of indexing terms (descriptors, keywords or terms from a nomenclature), 8 only use French while 10 add English and/or other European languages. On the contrary, Italian, Portuguese and Spanish bases index only with their national language(s). The same is true in France for 96 % of the bases; among the two exceptions which index also in English and a third language, one is an international base produced in France but hosted by an international vendor.

In the same way, from 11 bases in Belgium stating the availability of summaries, 2 write them in French, 2 in French and English, 7 in various European languages. In France, only one base provides multilanguage summaries, while 85 % summarize only in French, 12 % adding English. Summaries are also monolingual in the three other countries, with only one exception in Portugal (summaries in several languages).

At the end of this analysis, it is necessary to outline again that it is only based on one source, the Eurospes Directory, and therefore has to be cautiously interpreted. However, it throws light on the situation of on-line social science information in five European countries which do not mainly use English for their international communication. It tries to elucidate and explain some of the main differences between these countries (5). It provides some data which may be used when planning for broader exchange and cooperation between all countries and sub-regions within Europe.

Bibliografía

1. *Repertorio Eurospes de Bases de Datos en Ciencias Sociales,* Madrid/París, ISOC/FNSP, 1990.
2. MEYRIAT, J. Use of Information in Science and Research: The Social Science, *International Forum Information and Documentation,* 9 (3), 10-12, 1984.
3. LANCASHIRE, J.; McCARTHY, W. *The Humanities Computing Yearbook,* Clarendom Press, 1988.
4. MOBERG, F. *Databases in the Humanities and Social Sciences 1985,* Pardigim Press, 1987.
5. STERN, P. Online in the Humanities: Problems and possibilities, *The Journal of Academic Librarianship,* 14 (3), 1988, pp.161-164.

DE LA INVISIBILIDAD A LA VISIBILIDAD: REFLEXIONES SOBRE LA FORMACION DE ESPECIALISTAS Y USUARIOS DE LAS TECNOLOGIAS DE LA INFORMACION

Montserrat Sebastiá Salat*

Resumen: La actual evolución e impacto de las tecnologías de la información en las prácticas profesionales ligadas a las Ciencias de la Documentación e Información ha determinado la falta de formación de profesionales especialistas y de formación de los usuarios. Ambos problemas se analizan mediante un estado de la cuestión en el ámbito europeo.
Palabras clave: tecnologías de la información, formación de profesionales, formación de especialistas, formación de usuarios.

Abstract: Current evolution and impact of information technologies in professional practices linked to Information/Documentation Sciences has yielded lack of training of specialists professionals and users training. Both problems have been analyzed through the «state-of-the-art» in European background.
Keywords: information technologies, professional training, specialists training, users training.

1. Introducción

La «International Federation of Library Associations and Institutions» (IFLA) ha publicado una segunda edición ampliada del «Medium-Term Programme 1986-1991», documento en el que la reconocida institución propone las directrices de actualización, que deberían presidir el desarrollo intelectual y profesional en las áreas de las Ciencias de la Documentación. Un análisis comparativo entre la primera y la segunda edición del mencionado documento nos permite observar que, desde 1985, fecha de la primera edición del «Medium-Term Programme...» (MTP), hasta 1988, fecha de la segunda edición del citado «MTP», una de las áreas más destacadas en la práctica profesional y en la investigación es la Tecnología de la Información (TI). Esta nueva disciplina es uno de los aspectos a enfatizar en los programas de formación de profesionales y especialistas de la Documentación e Información. Disciplina que preside las tareas de la «Section of Information Technology» desde la cual la IFLA propone las líneas de investigación a seguir en materia de Tecnología de la Información (1).

Este progresivo interés por la TI, también se observa en los restantes organismos profesionales como: la FID con su «FID Strategic Plan», el CIA con el programa «RAMP», y la UNESCO con el «Programa General de Información (PGI)». Todos ellos proponen a la Tecnología de la Información como factor básico para el desarrollo de múltiples líneas de acción, orientadas a la formación de profesionales, a la creación de la infraestructura en archivos, bibliotecas, centros de documentación y servicios de información (2). En este mismo sentido cabe destacar la importancia dada a la TI por parte de organismos políticos suprana-