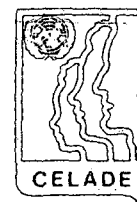


c. 2812.00 - 6907

pasas a vickie

Centro Latinoamericano de Demografía



Conference Documents

COMPATIBILITY AMONG POPULATION DOCUMENTATION SYSTEMS TO MEET THE NEEDS OF THE LATIN AMERICAN REGION

Arthur M. Conning

Betty Johnson de Vodanovic

DC/17
June, 1978
30

Paper presented to the
Technical Meeting on
Population Documentation
System Compatibility,
CELADE, Santiago,
3-7 July, 1978

I N D E X

	<u>Page</u>
I. INTRODUCTION	1
II. THE WORLD POPULATION LITERATURE AND LATIN AMERICAN REQUIREMENTS	2
Characteristics and needs of Latin American users	2
Providing Latin America with access to the World Population Literature	3
III. THE OPERATION OF DOCPAL WITH COMPATIBLE DATABASES FROM OTHER SYSTEMS	4
Inclusion of Selected World Literature in the Abstract Journal	4
Literature on Latin America Produced in the Developed Countries	4
IV. LEVELS AND ELEMENTS OF COMPATIBILITY	5
Levels of compatibility	5
Database elements of compatibility	5
Computer Program Compatibility	6
V. ACHIEVING AND MAINTAINING COMPATIBILITY	6
VI. RECOMMENDATIONS	7
BIBLIOGRAPHY	8

I. INTRODUCTION

Compatibility among population documentation systems could bring advantages to their users and to the internal operations of these systems. Maximizing these benefits, while minimizing the disadvantages of compatibility, is of particular concern to population documentation systems which have been designed to serve the needs of particular groups or specific national and/or regional users. This is particularly true of a system like the Latin American Population Documentation System (DOCPAL), which, as a component of the Latin American Regional Population Programme^{1/} of CEPAL, must serve the requirements of the Latin American Region.^{2/}

Consequently, in working toward compatibility of such systems, the basic criterion for decision making should be that the results of any measure to increase compatibility must enable the individual systems to better meet the needs of their major intended users, i.e., those for whom the system was designed. The application of this principle will help each system to determine the levels of compatibility that would be most beneficial -- for example, at the level of the database or at the level of products and services as well -- and, given agreement on the level(s), the elements that should be made compatible.

A less fundamental criterion, but one that must also be considered, is the effect of the move toward compatibility, and of compatibility, itself, on the internal operation of the individual systems.

To apply the basic criterion stated above it is necessary to specify the characteristics and needs of the major users toward which each system is directed. Although many of the national and regional users of other systems may be similar to those of Latin America, CELADE can legitimately give its views only on the needs and characteristics of its actual and potential Latin American users and on the circumstances under which compatibility would enable DOCPAL to better serve them.

The objectives of this paper are to analyze the characteristics and requirements of Latin America that are relevant to the discussion of compatibility and to make recommendations from the point of view of these needs and of the DOCPAL System.

Not considered in this paper are the possible benefits of having compatible systems to global agencies, such as foundations and international agencies operating at the world level. Their welfare is important, but the immediate benefits to users in the developing countries must be given priority from the point of view of a regional agency like CELADE.

^{1/} CELADE is the agency responsible for the execution of this programme.

^{2/} The "Latin American Region" is taken to include all countries which are members of the UN Economic Commission for Latin America (CEPAL), that is, the countries normally considered part of Latin America plus the countries of the Caribbean, many of which speak English.

II. THE WORLD POPULATION LITERATURE AND LATIN AMERICAN REQUIREMENTS

Characteristics and needs of Latin American users

From the very beginning of DOCPAL, an attempt has been made to base the overall design on the characteristics and needs of the Latin American users. The original proposal submitted to IDRC in 1975 for funding the creation of DOCPAL and its operation for the first two years (now known as the DOCPAL Foundation Building Phase, or Phase One) treated the matter in terms of potential users in the "well-endowed" and "poorly-endowed" centers of the Region (CELADE, 1975:3-8a).^{1/} This distinction is useful for the present purposes, even though it must be remembered that there is a continuum between the extremes.

The poorly-endowed institutions have very limited financial, infrastructure and human resources. Their staff members are usually poorly trained, few have studied abroad, particularly in the developed countries or major centers of the Region, and many have learned on-the-job. They have little contact even with other centers in their own countries, particularly with relevant non-population institutions; and they normally have little or no contact with major international centers. Library facilities are very limited and there is no direct experience with computation for data processing in their field although they may have had some data processed for them.

On the other hand, the well-endowed institutions tend to have specialists who have been well trained, often abroad in major universities or training centers; an increasing number have doctoral degrees from major developed country universities. The staff maintain contacts with other Latin American institutions and often have strong links with one or more international centers outside the Region. Library facilities may be fairly good and data processing by computer is fairly routine although most have difficulties in this area and do not have their own computer facilities.

While the situation will change over time, at present the majority of the Latin American institutions working in population or likely to use DOCPAL services tend to be poorly-endowed; this is true of the majority of governmental agencies, particularly in the poorer and smaller countries and certainly of most of the provincial level centers. The well-endowed centers are far fewer although their influence, quality and ability to attract funds is out of proportion to their number or size.

The benefits resulting from making DOCPAL compatible at some level with other systems will be felt differently in these two types of Latin American user institutions. Again keeping in mind that any real institution may not have exactly the same features described here, Table 1 outlines the characteristics and needs of the users likely to be found in each type of institution and relevant to the compatibility issue. Although there are still no empirical studies on this matter, the analysis here is consistent with similar evaluations made when designing the DOCPAL Journal (Conning, 1977), analyzing the POPINS Thesaurus (Acuña, Conning, Johnson, 1978), and planning the DOCPAL Country Population Documentation Development Second Phase activities (CELADE, 1978).

^{1/} See Bibliography at the end.

Table 1

CHARACTERISTICS AND NEEDS, RELEVANT TO SYSTEM COMPATIBILITY, OF PERSONAL OF LATIN AMERICAN INSTITUTIONS ACCORDING TO THEIR LEVEL OF RESOURCES

(See text for the definitions of the types of Institutions)

Poorly-endowed Institutions

1. Primarily interested in information on their own country or similar countries of the Region.
2. Little immediate need for, or interest in, documents from other Regions, except when they contain directly useful techniques (major studies by recognized authorities on the field).
3. Some persons beginning to see population in wider development-oriented context, but because of limited training, are not often able to use material that goes much beyond their traditional orientation.
4. The final users of the documents or abstracts do not read very easily in other languages. In the Spanish-speaking countries, some may be able and willing to struggle with English (and fewer in French and almost none in other languages if the document is of major relevance for their work.
5. Often do not have librarians or documentalists to assist in searching for information. Those available are often poorly trained in their field and do not read foreign languages well, if at all.
6. Limited access to major international journals relevant to population and development, and very limited access to books from outside the Region (in part because of limited budgets). Have almost no access to unpublished documents from other Regions.
7. No possibility of searching database tapes because little or no experience with computer data processing and no access to computers.

Well-endowed Institutions

1. While mainly interested in their own country or similar countries of the Region, have a strong interest in theoretical and methodological work done in other countries of the Region. Interested in keeping up to date on all major work produced in their fields.
2. Strong interest in technical, methodological or theoretical works from other Regions and in major empirical studies from other Regions. Prestige in keeping up on work done in major world centers. Little interest in information of other regions for comparative analysis.
3. View population in broad development-oriented context. Actively working in this direction and therefore need access to the social-economic development literature in their daily work in the field of population.
4. The final users of the documents or abstracts usually can read in English, although not always with great facility, and they are more likely to browse with Spanish abstracts (if their native language is Spanish).
5. Usually have librarians or documentalists to assist in searching for information. They read in English, but not with great facility.
6. Access to most major international journals relevant to population and development, and reasonable access to major books by well-known authors in the field. Very limited, or no access, to unpublished documents from other Regions.
7. Only a few well-endowed centers could now consider searching database tapes on a computer to which they have access.

If the statements given in Table 1 are reasonably close to reality, the 1976 POPINS Technical Task Force Report (TTF) is correct in its statement that the "first need in developing countries is probably adequate inventories of research, bibliographies, and abstract and full-text services at the national level" (Par.146; their underlining). The Second Phase of DOCPAL will attempt to meet these needs through technical assistance, training and the provision of special materials, but recognizes that for some time to come many of the problems outlined by the TTF have to be met through the regional-level services provided by DOCPAL to the country institutions because it is impossible to bring the infrastructure of each country of the Region up to an adequate level.

It is also clear that providing access to the world literature would be very useful to many centers in the Region. From the Table, the following specific conclusions relevant to Latin America emerge:

- a) It is desirable to have information on and access to the documents from other regions that define new technologies or that effectively demonstrate their use.
- b) It is desirable to have information on and access to important studies in other regions, particularly when they involve new theories, models, or methodology; there is a growing interest and a particular need for studies done anywhere that show how to approach development planning taking into account population variables.
- c) There is little demand, on the other hand, for inter-regional comparative materials (this is also noted by the TTF (POPINS, 1976; Par.147)).
- d) For the majority of countries of Latin America, whether the users are located in well- or poorly-endowed institutions, it is always highly desirable to have informative abstracts written in Spanish, since even users able to read in English or French are more likely to browse in the system if the abstracts and descriptors are in Spanish. This also would permit librarians with limited knowledge of English to be better able to prepare specialized bibliographies from the journal.
- e) Copies of the documents must be available, even of published materials, since most poorly-endowed centers do not always have access to major international journals and even less to important books.
- f) Documents of interest may be in English or French as well as in Spanish or Portuguese. There is very limited ability to read documents in other languages although those documents referring to new techniques might be of value if Spanish-language abstracts (with formulae included) were available.

Providing Latin America with Access to the World Population Literature

On the basis of the previous analysis it appears that Latin American institutions, whether well- or poorly-endowed, would be significantly benefited from increased access to the world literature if the material were organized to meet the needs of the Region.

Since the great majority of actual and potential Latin American users can not read foreign languages easily and those who are most isolated from the literature can read only in Spanish (if

Spanish speakers), it also follows that information on the world literature has to be provided through Spanish language abstracts and indexing. To keep them up to date on this literature on a regular basis and enable them to find specific information rapidly and conveniently, most of Latin American users would have to obtain the information through a Spanish language journal tailored to their needs. Logically the DOCPAL journal could be used for this purpose. Demand searches also must be available, but it will be some time before they become the predominant way in the Region to obtain specific information.

III. THE OPERATION OF DOCPAL WITH COMPARATIVE DATABASES FROM OTHER COUNTRIES

It is instructive to imagine, in broad outline, how DOCPAL (or other similar systems) might operate if there were sufficient database compatibility to permit exchange tapes to be processed by ISIS as a DOCPAL file, perhaps after passing the donor tape through a transformation program.

Inclusion of Selected World Literature in the Abstract Journal

If the donor system had flagged the items of major interest to users in other regions (see Table 1, points 2 and 3), these might be transferred automatically to the DOCPAL file for eventual inclusion in the abstract journal. However, it would be necessary to acquire the documents for the clearinghouse and to produce the abstracts. Because of the costs involved in the inclusion of the abstract, only works of likely value to persons in Latin America would be flagged.

Literature on Latin America Produced in the Developed Countries

The TTF noted that a great deal of published work produced on the developing countries by scholars of the industrialized countries, is available only in institutions of the developed countries (POPINS, 1976: Par.158). This is one of the reasons why Population Index can locate the literature published on Latin America in the United States and Europe more completely and easily than can DOCPAL.

Considering that Population Index does have apparent control over the literature on Latin American population produced in the industrialized countries, it is possible to see how compatibility would enable the Latin American Region to maximize the use of that information.

If the two databases were sufficiently compatible (see below), DOCPAL could obtain the Latin American materials published outside Latin America from the Population Index tape. Once this were properly organized, DOCPAL would rely entirely on Population Index for the location of such materials and, therefore, would not even have to check whether any given document published outside Latin America was already in DOCPAL.

Similarly, when Population Index is convinced that DOCPAL can locate materials produced in Latin America more completely and easily than they can, they, in turn, might decide to rely entirely on the DOCPAL datatape for such materials; as noted in the first example, here, it might be desirable for DOCPAL, now the donor system, to flag items of sufficient general interest for inclusion in Population Index.

This arrangement would seem to be easily arranged since the differences in languages and user needs between the two systems means that the same item appearing in both journals does not represent

a duplication, although the previously duplicated effort to include it would be eliminated. Clearly this could reduce the costs of location and selection of documents and increase the number of documents in each system.

Since DOCPAL must have a copy of the documents obtained from the Population Index database, if Population Index were interested in copies of non-conventional Latin American materials, an exchange of copies of the documents between the two systems might be feasible.

IV. LEVELS AND ELEMENTS OF COMPATIBILITY

Levels of Compatibility

As seen in the two examples presented in the previous section, the most evident means of providing the Latin American Region with relatively low-cost access to the world population literature through DOCPAL requires that there be some degree of compatibility among the population documentation systems at the level of the database. This would not require all fields to be defined identically but would require agreement on the substantive definitions of certain essential fields (see below).

Compatibility at other levels appears less compelling; flexibility is highly advantageous to systems that attempt to tailor their products and services to their users and that must take into account changing conditions. For example, at present there would seem to be relatively little benefit to be gained from formally standardizing journal formats and indices, when these may vary with the types of needs of the users. However, when the design of a product of one system does meet the needs of another, it would save development costs and time for the latter system to take advantage of the ideas of the former, perhaps making adaptations and improvements.

Similarly, since the present concern is with compatibility among systems, rather than the creation of a new super-system, there is no reason to standardize on system procedures and worksheet formats. Rather only the definitions of database compatible fields must be normalized. Since DOCPAL will concentrate much of its effort on assisting country institutions to improve their capabilities, in part by taking advantage of the DOCPAL procedures, it is important that the worksheet format and procedures be open to modification to meet local needs and characteristics.

Database Elements of Compatibility

Not all the individual data elements of each database should be compatible since each system has its own special needs. While this paper is not the place to discuss the list of fields that should be made compatible to achieve the types of benefits discussed in previous sections, it is important to identify, in general, the elements that should be made compatible. From the point of view of DOCPAL, it is essential or desirable that the following sets of data elements be compatible:

- Essential : Descriptive cataloging (UNISIST essential fields). Authority lists (corporate authors; journal short titles, etc.)
- Highly Desirable: Geographic and descriptor indexing (Use of POPINS Multilingual Thesaurus to permit automatic translation).
- Desireable : Availability of the abstract or annotation to assist the writing of the abstract in another language.

It is important to note that compatibility of data elements implies more, than, for example, equivalent definitions. Thus a system, which employs UNISIST definitions for the publisher, year of publication, page number and report number, but which puts all this information in the same field, will be able to convert a UNISIST exchange tape to its format, but will be unable to provide a UNISIST tape of its own database since it will be unable to separate out the individual parts of the composite field. For this and other reasons, DOCPAL places each UNISIST data element in its own field.

Standards of quality of the databases exchanged (i.e., of the compatible elements) would have to be established. It might be desirable to apply an edit program to check for all computer detectable errors before sending any tape from one system to another.

Computer Program Compatibility

If there is agreement on the substantive definitions of the data elements that are to be compatible (with one data element per field), in general, it should be possible to write a program to transform from one computer program format to another.^{1/} However, this will have a cost that could be significant if the databases are large and require complex transformations. Ideally, it would be preferable for each documentation system to use the "same" computer program. Then data tapes would require little or no routine transformation and participating documentation systems would be able to use the applications programs of one another, when suitable.

However, the matter of computer programs should be left flexible since many of the programs are not portable between computers of different manufacturers, or even, sometimes between computers of the same manufacturer. It might be noted that from the DOCPAL point of view, everyone should use the ISIS-DOS version it employs, since it is especially appropriate to the computer environment in which CELADE processes its data. In other environments, another version of ISIS might be preferable or it may not even be possible to utilize ISIS.

V. ACHIEVING AND MAINTAINING COMPATIBILITY

Since the achievement and maintenance of compatibility involves three or more separate systems, it would appear desirable that there be an organization responsible for a number of functions and acting as a focal point. These might be:

- a) Insure that each developing region is adequately covered by an appropriate population documentation system.
- b) General coordination of the move toward compatible databases.
- c) Maintenance of the common standards among the participating population documentation systems and with other relevant systems.

^{1/} This statement would seem to be true if each document record is treated separately in the computer program, such as is the case with the ISIS-DOS used by CELADE. If the document records are nested and linked, this may no longer be so. In the case of ISIS, it has been shown that it is possible to transform ISIS-DOCPAL records to the UNISIST exchange record format (see Packer and Conning, 1978).

- d) Maintenance of the multilingual thesaurus.
- e) Monitoring the exchange of data tapes among the participating systems.
- f) Promotion of the better use of the existing systems.

VI. RECOMMENDATIONS

The previous sections can be summarized in a number of general recommendations that are stated here only for the purposes of discussion since they are seen almost exclusively from the point of Latin America and DOCPAL and any move to compatibility will involve trade-offs between all the systems involved.

- 1). The decisions concerning compatibility should always consider the relative advantages and disadvantages to the actual and potential users of the participating national and regional systems. That is, the individual systems must not lose sight of the fact that their first responsibility is to provide services tailored to the needs of their nation or region.^{1/}
- 2). The participating systems should aim at compatibility at the database level. Other aspects of the systems should be left flexible although participants should take advantage of the developmental work and new applications of other systems.
- 3). Documents included for exchange should be pre-selected by the donor system to be of general interest to other systems. The documents of inter-regional interest normally will be those with new techniques, models, methodologies or theories. Of particular interest also are those illustrating relationships between development planning and population variables.
- 4). The database should have compatible descriptive cataloging (preferably following the UNISIST Reference Manual) and it is highly desirable that a multilingual thesaurus be utilized to permit automatic translation. Agreement should be reached on corporate author and other authority lists.
- 5). Standards of quality and methods of assuring it should be established.
- 6). Since ISIS was developed by the United Nations and is very flexible and low cost, it should be the preferred system.
- 7). A Central organization, preferably located in the United Nations, should be responsible for general coordination of activities and the maintenance of agreed-upon compatibility.

^{1/} This emphasis on the existing systems meeting the needs of the users in their nations or regions, while seeking to establish intersystem links, is very much in accord with the Draft Resolution II of the 19th Session of the UN Population Commission. This "Affirms ... the need to develop at the national level appropriate institutional arrangements and supporting programs to promote the sharing and exchange of information; the importance of establishing or strengthening regional capacities in the field of population information, and the desirability of establishing appropriate co-ordination and linkages among the already existing information activities and services in the field of population". (Populatio Commission, 1977: p.2).

BIBLIOGRAPHY

- Acuña, S., Conning, A., and Johnson, B., 1978. The POPINS Multilingual Thesaurus for Latin America: An analysis of the draft prepared in March 1978. Paper presented to the Third Meeting of the POPINS Multilingual Thesaurus Working Group, United Nations, N.Y., 2-4 May 1978. CELADE-Santiago, DC/16.
- CELADE, 1975. An Outline of the Latin American Population Documentation Service (DOCPAL): A Request to IDRC, Ottawa. CELADE-Santiago.
- CELADE, 1978. DOCPAL Country Population Documentation Development: A Request to IDRC, Ottawa. CELADE-Santiago.
- Conning, A., 1977. The Journal "DOCPAL Latin American Population Abstracts": The Design of User-Oriented Features. Paper prepared for the Session "Information Systems in Demography" of the IUSSP General Conference, Mexico City, 8-13 August 1977. CELADE-Santiago, DC/13.
- Packer, Abel and Conning, Arthur M., 1978. The Utilization of the UNISIST Reference Manual in the DOCPAL System and the Computer Conversion From the DOCPAL ISIS Format to the UNISIST Format. CELADE-Santiago, In press.
- POPINS, 1976. Study Prepared for the POPINS Interim Steering Committee by the Technical Task Force. Pp.1-122d in Towards a World-Wide Information System of the Field of Population. (United Nations, N.Y.).
- Population Commission, 1977. Report on the 19th Session 10-21 January 1977. Economic and Social Council. Official Records: 62 session, Supplement N°4, United Nations, N.Y. (E/5913; E/CN.9/334).

