

CENTRO LATINOAMERICANO DE DEMOGRAFIA

7146.00

ISIS: 16759

c. 2

INFOPAL

THE LATIN AMERICAN  
POPULATION INFORMATION PROGRAM  
OF CELADE

Arthur M. Conning  
October 1977

CONTENTS

page

Abstract	
I. INTRODUCTION .....	1
Population Information and the Priorities of the Latin American Countries	
Types of Population Information	
Users of Population Information in Latin America	
The Process of Using Information	
II. OBJECTIVES OF THE INFOPAL PROGRAM .....	6
General Objectives	
Specific Objectives	
Storage and Retrieval of Latin American Population Information	
Computer Processing of Population Information	
III. THE COMPONENTS OF THE PROGRAM .....	9
Population Information Storage and Retrieval	
Data Bank	
DOCPAL	
Population Information Processing and Systems	
IV. AN INTEGRATED APPROACH .....	14

ABSTRACT

Following the recommendations of the World Population Plan of Action and the Second Latin American Meeting on Population, the CELADE Latin American Population Information Program (INFOPAL) endeavours to assist the countries and CELADE personnel and its students in locating, obtaining and processing population information with special emphasis given to priority areas, defined by the governments, involving the formation, implementation and evaluation of development policies incorporating population parameters.

Distinguishing primary information -- micro-data and documents (including tabulations) -- from secondary information, which describes the primary information to facilitate its retrieval, the Program assists various classes of users by creating and processing secondary information, supplying the primary micro-data and documents, and by making available the most suitable techniques for processing population information by computer. The Program only undertakes regional level activities and those which cannot be done now by the countries and which, hence, must be carried out by CELADE, the specialized agency executing the Regional Population Program of the United Nations.

The Program is concerned with the storage and retrieval of population information and its processing. The former involves two component elements, the Data Bank and DOCPAL, treating the storage and retrieval of micro-data and documents, respectively. The Data Bank will eventually be placed in database form making it easier to process census, survey and other micro-data interactively. DOCPAL is concerned with the creation and provision of secondary information on all population documents written in or about Latin America, the preparation of specialized bibliographies via on-line searches with a computer terminal, the provision of full texts and technical assistance for using DOCPAL procedures in national Centers.

The Population Information Processing and Systems component of the Program provides technical assistance in the processing of population information to the countries, processes population data for CELADE staff and students, and operates the computer systems for the storage and retrieval aspects of the Program. It also provides specialized training in the processing of survey and census data for country investigators and

## CENTRO LATINOAMERICANO DE DEMOGRAFIA

programmers and designs systems for population information processing that meet specific country needs and limitations. Major activities in the past, and continuing into the future, have involved the processing of population census data and the World Fertility Survey being carried out in around eleven Latin American countries.

The INFOPAL Program represents an attempt to integrate population information activities within a single program so as better to meet the needs of workers in the field of population. The ability of the Program to make major contributions to the countries derives in large measure from the knowledge, experience and committment gained from the integration of population information storage, retrieval and processing activities with the substantive technical assistance, training and research activities of CELADE.

I. INTRODUCTION

Population Information and the Priorities of the Latin American Countries<sup>1/</sup>

The countries of Latin America are increasingly becoming concerned with the formulation, implementation and evaluation of specific policies affecting population variables and with the incorporation of population parameters into their development plans. This is seen in the recommendations of the Second Latin American Meeting on Population (SLAMP, 1975)<sup>2/</sup> which give high priority to such policy relevant inputs as research projects aimed at establishing the interrelationships between population and development in specific historical contexts (SLAMP, 1975: 41) and to the improvement of census and other population statistics to meet the needs of the policy makers and their professional staffs (SLAMP, 1975: 40). Furthermore, there is the recognition that methodologies and analytical tools for incorporating demographic variables into the process of economic and social planning must be developed or adapted (SLAMP, 1975: 39), a task that requires both the review of existing procedures and experiences and the ability to manipulate and process the information as desired.

In order for the countries of the region and the international organizations which serve them to undertake activities in these priority areas, the required information must be identified, obtained and, when necessary, processed rapidly. This is more complex than in the past, when the focus was primarily on national level analyses, since the new orientation toward specific policies often requires more involved processing by computer so that empirical analyses and evaluations can be done on the specific population categories and geographic sub-units that are affected by the policies.

As CELADE is responsible for implementing the activities of the Latin American Regional Population Program (RPPLA, 1975: 284), which includes the above priorities in its long term objectives, one of CELADE's contributions

---

<sup>1/</sup> "Latin America" is taken to comprise all countries which are members of the Economic Commission for Latin America (CEPAL). This includes the English and other non-Spanish speaking countries of the Caribbean.

<sup>2/</sup> See the bibliography at the end of the text.

must be to improve the utilization of population information by the countries of the region. This can be fostered through the provision of technical assistance and instrumental services closely integrated with substantive population activities and training provided by CELADE. In addition, CELADE's staff must be enabled to make more efficient use of the information available in the center. To better integrate its ongoing activities in the field of population information <sup>1/</sup> and to identify unmet present or future needs and carry out activities to fill them, CELADE has created the Latin American Population Information Program (INFOPAL). In this way the recommendations of the World Population Plan of Action regarding population information are being carried out <sup>2/</sup>.

#### Types of Population Information

It is useful to distinguish "primary information", which is that actually employed by a user for an analysis, from "secondary information" which describes the primary information and thereby facilitates its location and retrieval. Secondary information sources are bibliographies, abstract journals, library catalogues, etc. The term "population information" will always refer to primary information unless otherwise noted.

It is convenient to subdivide the primary information into two general types, micro-data and documents, which are defined in Table 1. As noted in the table, a "tabulation" may be viewed as a special form of document. The table also shows the extent to which computer processing is likely to be involved in the use of each type of information in the field of population.

---

1/ The Interregional Consultative Group of Experts on the World Population Plan of Action said that "support should be given ... to the central establishment of documentation and data banks, and to the processing of data that can be promptly available for programme and operational purposes ... (UNFPA, 1976: 20) and that "Special effort should be made to facilitate rapid and wide dissemination of research results" (Ibid: 49).

2/ The relevant recommendations and statements are contained in paragraphs 15, 72, 73, 76, 77, 78(m), 80, 88, 100 and 102 of the World Population Plan of Action.

TABLE 1. DEFINITIONS AND EXAMPLES OF THE MAJOR TYPES OF PRIMARY INFORMATION TREATED IN THE CELADE LATIN AMERICAN POPULATION INFORMATION PROGRAM AND THE COMPUTER PROCESSING REQUIRED

TYPE OF PRIMARY INF.	DEFINITION	EXAMPLES	COMPUTER PROCESSING REQUIRED
I. MICRODATA	A SET OF CASES (QUESTIONNAIRES, ETC.); EACH INDIVIDUAL CASE HAS VARIOUS ITEMS (VARIABLES) DESCRIBING IT, USUALLY IN CODED FORM. NORMALLY A GIVEN SET OF MICRO-DATA CONTAINS A LARGE NUMBER OF CASES (FROM HUNDREDS TO MANY MILLIONS) AND A FEW TO MANY VARIABLES (PERHAPS UP TO 1000) DESCRIBING EACH CASE.	CENSUSES; SURVEYS; FAMILY PLANNING SERVICE STATISTICS ON INDIVIDUALS; ETC.	ALWAYS (OR ALMOST ALWAYS). INDIVIDUAL RECORDS ARE SELDOM OF INTEREST; RATHER THE MICRO-DATA ARE PROCESSED TO OBTAIN STATISTICAL ANALYSES ("TABULATIONS"). OFTEN THE TABULATIONS PRODUCED FROM THE MICRO-DATA REQUIRE FURTHER PROCESSING.
II. DOCUMENTS (WRITTEN)	MATERIALS, NORMALLY WRITTEN ALL OR PART IN NATURAL LANGUAGE, PUBLISHED OR NOT. OFTEN TAKEN TO INCLUDE AUDIO-VISUAL EQUIPMENT SUCH AS FILM STRIPS AND RECORDS; THESE ARE EXCLUDED HERE SINCE THEY WILL NOT BE IN THE CELADE INFOPAL PROGRAM.	BOOKS, ARTICLES, REPORTS, THESES, DICTIONARIES, ETC.;	SELDOM (TABULAR NUMERICAL DATA MAY REQUIRE PROCESSING; ON RARE OCCASIONS CONTENT ANALYSES, ETC., MAY ALSO REQUIRE PROCESSING).
(TABULATIONS)	FOR PURPOSES HERE, ANY <u>EXISTING</u> SYSTEMATIZED NUMERICAL INFORMATION PRESENTED IN WRITTEN FORM (ALTHOUGH SOMETIMES STORED IN MACHINE READABLE FORM). NORMALLY OBTAINED BY PROCESSING MICRO-DATA. CONSIDERED HERE AS A FORM OF DOCUMENT.	CONTINGENCY TABLES, REGRESSIONS, FACTOR ANALYSES, PROJECTIONS, ETC.	FREQUENTLY FURTHER PROCESSING IS REQUIRED TO PRODUCE A STATISTICAL ANALYSIS -- FOR EXAMPLE, AN ANALYSIS OF VARIANCE IS CARRIED OUT ON A TABULATION, OR TABULAR FERTILITY INFORMATION IS USED TO ESTIMATE FERTILITY RATES VIA A BRASS TECHNIQUE.

Users of Population Information  
in Latin America

The Program is directed toward satisfying, in varying degrees, the actual and potential information needs of the following types of users in the countries and in CELADE and other regional and international organizations:

- 1.- Researchers, technicians, university professors and other professionals highly specialized in their own fields;
- 2.- Governmental and other policy makers, development planners and their professional staffs and advisors;
- 3.- Administrators and action programs and their staffs;
- 4.- Channelers of information (librarians, documentalists, etc.);
- 5.- Non-specialized professors and university students;
- 6.- General public (reporters, secondary students, etc.)

The CELADE program concentrates most of its effort on the first three types, which are the users with whom CELADE works on most of its substantive activities and the major users actively requiring population information and its processing. However, channelers are also of major importance because they make information available to the first three types of users. To a lesser extent the interests of the last two groups for non-technical literature are also taken into account with regard to documentation.

The Process of Using Information

For a given purpose, which may vary from a theoretical research project to the administration of an action program to the presentation of alternatives for a new policy, the users of population information tend to follow a series of broad general steps which are independent of the substantive matters involved:

- a) Determine the type and content of the primary information required and the analysis to be conducted;
- b) Identify the relevant existing information;
- c) Obtain the existing information required;
- d) Collect new empirical information (if necessary);
- e) Process the information;
- f) Make analyses and draw conclusions;



Whatever the matter being considered, it is desirable that there be an interactive relation between the user and the information since it is likely that hypotheses will be suggested or new approaches found which should be followed up; hence, some or all the steps may have to be repeated. Ideally, this interaction between the user and the information should be possible even in national centers in which human, technological and financial resources are very limited.

A population information system can make major contributions to its users by facilitating step b, through the creation and provision of suitable secondary information to make easier the retrieval of the required primary information; step c, through a data bank for population micro-data and the provision of full texts of documents and tabulations; and, step e, through making available the most suitable processing techniques for a given user's needs and resources.

#### The Role of CELADE

Since CELADE is a regional organization while most users of population information are in the countries, one must ask what services and assistance CELADE should provide and which activities are or should be done in the countries. To permit country users to interact with their micro-data it seems clear that the countries must be able to do their own processing and should have their own data banks. CELADE's role must be to provide technical assistance and to adapt or create technology that meets the processing needs of those countries which do not have the human resources to do so themselves; only when a given institution is unable to obtain timely results should CELADE actually process data on request.

To identify and retrieve desired documents from the large mass of published and unpublished documents, including tabulations, it is necessary to create and process secondary information describing each item. Since most Latin American countries presently do not have national population information centers capable of doing this and few will be able to do so on their own in the intermediate future, CELADE now must obtain and process the documents in CELADE and make the primary and secondary information available to all potential users. Given that documentary information of one country often is of value to another country in the region with similar problems in similar contexts, a documentation system located in CELADE also can facilitate this interchange of information. As national population information centers are developed, CELADE must provide

technical assistance and training in its procedures to avoid duplication of effort and to maintain international compatibility.

As CELADE is the specialized regional United Nations organization treating population, the functions of CELADE's INFOPAL Program cannot be taken over by more general statistical or other agencies since such organizations do not have the specialized technical competence in population, cannot give population the priority that the topic requires, and do not have the recognition in the field of population that CELADE has, a major requirement for operating a population information system. Of course, this means that CELADE and the agencies concerned with other developmental activities must cooperate to permit users to draw on more than one area of information. 1/

---

1/ The development of separate information systems for specialized topics, with compatibility secured by following international standards, is not unique to population. The planners of DEVSIS (Developmental Science Information System) envision that sectoral information such as population will be handled by specialized agencies (DEVSIS, 1975: 62).

## II. OBJECTIVES OF THE INFOPAL PROGRAM

### General Objectives

The central objective of the INFOPAL Program is to facilitate the utilization of population information by the various types of users in Latin American countries and regional and international organizations through the provision of integrated technical assistance training and information activities relating to the storage, retrieval and processing of population micro-data and documents (including tabulations). In designing the component elements comprising the Program, particular attention was given to those aspects of population work which the countries identified as being of high priority at the Second Latin American Meeting on Population and which CELADE has declared as its general or specific objectives (RPPLA, 1975: 285).

### Specific Objectives

The specific objectives to be carried out in the next few years are conveniently divided into those involving the storage and retrieval of population information and those concerning the processing of this information.

#### Storage and Retrieval of Latin American Population Information

##### A. Micro-data

1. Maintain the CELADE Data Bank, acquiring data sets of wide interest (mainly censuses and the World Fertility Survey) or specifically required by CELADE researchers.
2. Provide data sets on request to institutions and persons, in or outside the region, after the required releases have been obtained from the original producer of the data.
3. Place the datasets in database form, thereby creating the Latin American Population Data System (DATOPAL) permitting the use of a database survey and census tabulations system (see point 3 of Computer Processing below).

B. Documents

1. Work toward attaining bibliographic control of Latin American population books, articles and other documents written since 1970; obtain the documents and create secondary information for them.
2. Store general interest, unpublished, tabulations produced by CELADE from the Data Bank information (since the emphasis in the INFOPAL Program is on interactive use of data by investigators, only a limited number of tabulations, mainly from census data, will be stored.
3. Produce the abstract journal, DOCPAL Latin American Population Abstracts twice a year to reach the widest number of users in and outside the region.
4. Produce specialized bibliographies, on request, based on all the documents under control.
5. Make the documents (and tabulations) available to users in and outside the region through a clearinghouse.
6. Assist the development of national population information centers and services through technical assistance, training and informational activities.

Computer Processing of Population Information

1. Provide technical assistance and training to the countries, transferring technology created in CELADE or adapted by it to meet the Latin American population processing needs.
2. Assist the countries in processing population data, with efforts directed toward their becoming more self-sufficient and able to work interactively with data (special attention to censuses, demographic surveys, the World Fertility Survey and family planning statistics).
3. Develop software for a database census and survey tabulation system (DATOPAL) which will facilitate the "interactive" manipulation and use of census, survey and other micro-data at CELADE; adapt the system for use by the countries which require it.

4. Carry out complex computer processing for the CELADE staff and students and train them in the employment of user-oriented systems.
5. Maintain the systems for the storage and retrieval of primary and secondary information of other components of the INFOPAL Program and provide technical assistance and training to the countries in the use of those systems for population information.

## III. THE COMPONENTS OF THE PROGRAM

Population Information Storage and Retrieval

The Data Bank and DOCPAL are the two component elements of the INFOPAL Program concerned with the storage and retrieval of Latin American population micro-data and documents, respectively.

Data Bank (since 1970)

The CELADE Data Bank stores and makes available, in standardized format, samples of the 1960 and 1970 Latin American censuses. It also has various fertility, abortion, migration, demographic and other surveys in which CELADE participated or which are required for investigation, training or technical assistance.

As many of the countries are beginning to organize population data banks that will be used for development and population policy research and planning with the emphasis on the production of the special tabulations required for such purposes, various national institutions will want to store their micro-data in a form to permit easy manipulation and rapid and low cost processing. At present, CELADE's own Data Bank is organized in linear form making it costly to perform complex manipulations or work "interactively" with the data. To overcome these limitations and to be able to assist the countries, CELADE is developing the Latin American Population Data System (DATOPAL) <sup>1/</sup> which will utilize a direct access data base approach suitably linked with a general survey and census processing system. It is hoped that the system can employ Canadian technology adapted and modified by the CELADE Population Information Processing and Systems group. When the DATOPAL system has been completed -- in large part a substantive matter which must take into account the many possible uses of the micro-data by the countries, CELADE and other organizations -- the present CELADE Data Bank holdings, as well as new micro-data sets as they are required, will be placed in the DATOPAL database(s).

<sup>1/</sup> The development of the DATOPAL system is being financed by funds from CIDA (Canada) to CELADE.

The Data Bank Bulletin is published periodically with a listing of all tapes and standardized tabulations available and the procedures for requesting these or new tabulations.

DOCPAL<sup>1/</sup> (since April 1976)

The Latin American Population Documentation System (DOCPAL) is working toward achieving bibliographic control over Latin American population materials by attempting to identify and acquire "all" books, articles, reports and other documents, published, or not, treating population topics and written in or about Latin America since 1970. To provide users in the countries and in CELADE with secondary information on each document, a detailed informative abstract is produced, key words are assigned from a controlled vocabulary (thesaurus) and a full bibliographic description of the item is written following international norms. This secondary information is stored and manipulated by a computer program<sup>2/</sup> to permit the compilation of specialized bibliographies on request via interactive searches by video terminal.

To insure that users in the countries with limited resources obtain full advantage from DOCPAL, the Journal DOCPAL Latin American Population Abstracts is published twice a year.<sup>3/</sup> The journal has a Literature Review Index designed to keep population specialists up to date on new work in their specialities and a detailed Subject Index based on keywords that enables specialists as well as other users to locate documents meeting specific needs. (The user-oriented features of the journal are described by Conning, 1977).

---

<sup>1/</sup> For a detailed description, see Conning, 1976a and 1976b. DOCPAL is being developed and operated during its first two years with a grant from IDRC of Canada.

<sup>2/</sup> ISIS

<sup>3/</sup> The first issue produced entirely by computer in camera-ready form was published in June 1977.

Although the DOCPAL system operates a computerized system, the procedures for cataloguing, abstracting and indexing via the thesaurus are applicable in manual systems. Hence, DOCPAL can provide technical assistance to national centres to set up manual or computerized population information systems to meet specific needs, thereby avoiding the need for the national centres to waste resources creating their own procedures which are not compatible with other systems or the regional system DOCPAL. Some of these country centers also will be able to act as national input centers to DOCPAL. In addition, DOCPAL will provide specialized training seminars in the use of its procedures, particularly in those involving indexing and searching.

The DOCPAL system in CELADE also maintains a "Complementary File" of books, journals and other documents necessary for the work of CELADE staff and students but not explicitly treating Latin American population. Specialized texts on economics, sociology, etc., are not included in this file since they are available in other institutions in Santiago. At present the Complementary File is not computerized.

#### Population Information Processing and Systems

The Population Information Processing and Systems group provides a wide variety of technical assistance, training and specialized services in the computer processing of population information to the countries. It also works with CELADE substantive units in the design of studies and their processing and maintains the systems for storing and retrieving population micro-data and documentary information. As a result of its experience the Computer group has found it necessary on occasions to engage in the development of new software systems to suit Latin American population needs, computer and human resources. The most ambitious effort to date has been CONCOR, a user-oriented language to detect and correct consistency and other errors in censuses and surveys (see CELADE, 1976, for a full description).



Although adequate computer hardware for processing population micro-data is now available in the countries or will be by the 1980 round of censuses, many national centers are not able to produce timely results of the population surveys, censuses and program statistics because of difficulties in one or more steps in which computer personnel should or must intervene. There are at least seven stages of an empirical population study or project evaluation in which computer knowledge or processing are required: (a). planning (optional, but highly desirable to avoid unnecessary delays and costs during the data elaboration); (b). data entry; (c). data consistency checking and correction (optional, but increasingly accepted as necessary); (d). data manipulation; (e). initial tabulations; (f). further data manipulation: tabulations and other statistical procedures for specific analysis; (g). data storage and retrieval (optional, but highly desirable).

In part, the problems occurring in one or more of these stages are the result of the turnover of programming staff attracted to industry, and in part due to the lack of adequate training and the difficulty of keeping up with new developments or the attempt to apply them when they are inappropriate. Whatever the reason, a census or population survey that is not processed soon after the fieldwork, is not useful for policy decisions that require knowledge of the present state of the population.

It is evident that there must be an integration of the data processing activities with substantive population work both in CELADE and in the countries. A function of CELADE, and particularly of the information processing group, is to stimulate this orientation in national centers. Hence, a major contribution of CELADE to substantive population activities in the countries is the provision of technical assistance<sup>1/</sup> and specialized training in the processing of population data. This is a necessary condition for the achievement of many of the goals of the World Population Plan of Action. Major efforts in the past that will continue in the future involve the processing of national censuses and of the World Fertility Surveys being carried out in at least eleven Latin American countries.

---

<sup>1/</sup> During 1977 a total of 22 separate technical assistance missions involving population data processing will have been carried out to more than 10 Latin American countries.

It should be noted that unlike the situation in some other fields, data elaboration of population studies and particularly the provision of technical assistance in this area require a familiarity with the substantive concepts and analytical procedures in order for them to be carried out rapidly and with a minimum of error. During a technical assistance mission time is limited and specific substantive problems of a given study must be grasped quickly so that the appropriate technology can be recommended and applied. For this reason, the wide experience of the CELADE programmers and system analysts in population data processing with CELADE investigators and the countries and in provision of technical assistance is an asset that cannot be duplicated elsewhere in institutions without the environment of substantive population work that exists in CELADE.

The integration of the activities of the Population Information Processing and Systems group with the substantive work of CELADE is complemented by the integration of the group with the micro-data and documentation storage and retrieval systems. As a result of this integration of information activities, it was possible, for example, within the space of little more than a year to design or adapt and make functional the data entry and retrieval systems that meet the special needs of DOCPAL and to produce the first issue of DOCPAL Latin American Population Abstracts (see the section on DOCPLA) completely by computer in camera-ready form and with many user-oriented features not possible without computer manipulation.

IV. AN INTEGRATED APPROACH

The INFOPAL Program integrates the population information storage, retrieval and processing activities of CELADE so as better to provide assistance, training and information to the countries and to CELADE's investigators and students. However, the present and future capacity and commitment of the INFOPAL Program to carry out the recommendations of the World Population Plan of Action and the Second Latin American Meeting on Population in the field of population information derives in large measure from the integration of the Program with the substantive technical assistance, training and research activities of CELADE.

The integration within the Program and with CELADE substantive activities must be further strengthened to meet the present country demands on the Program and to be ready for future needs that must be answered at the regional level.

BIBLIOGRAPHY

CELADE, 1976

CONCOR: Editing System. Manual in English prepared by the World Fertility Survey (London). Draft version (0). 25 Feb. 1976.

Conning, A., 1976a.

DOCPAL: Sistema de Documentación sobre Población en América Latina, Notas de Población. Año IV. Vol. 10.

\_\_\_\_\_, 1976b.

A Report on the Implementations of DOCPAL During the First Six Months of Operation; Background paper written for the Joint Meeting of the POPINS Interim Steering Committee and POPINS Technical Task Force, Washington 20-24 September 1976 (CELADE-Santiago DC/5).

\_\_\_\_\_, 1977

The Journal DOCPAL Latin American Population Abstracts: The Design of User-oriented Features. Prepared for the Informal Session of the IUSSP General Conference, Mexico City, 8-13 August 1977. (CELADE-Santiago DC/13).

DEVSI, 1975

The Preliminary Design of an International System for the Development Sciences. IDRC-065e. Ottawa, Canada

RPPLA, 1975

Regional Population Programme for Latin America 1976-1977 Capítulo F, United Nations Fund for Population Activities. pp. 276-343.

SLAMP, 1975

Conclusiones y Recomendaciones de la Segunda Reunión Latinoamericana de Población/Conclusions and Recommendations of the Second Latin American Meeting on Population. Mexico, 3-7 marzo 1975, ST/CEPAL/Conf. 54/L.9, Rev. 1.

UNFPA, 1976

Final Report of the Meeting. UNFPA/UN Interregional Consultative Group of Experts on the World Population Plan of Action (Geneve, 15-19 Sept. 1975). UNFPA/WPPA/20/Rev. 1, 18 Feb. 1976.