IDRC
CRDI
CIID

CANADA

NEW HORIZONS IN AGRICULTURAL INFORMATION MANAGEMENT

PROCEEDINGS

OF AN INTERNATIONAL SYMPOSIUM

MARCH 13-16, 1991

BEIJING, CHINA

The International Development Research Centre is a public corporation created by the Parliament of Canada in 1970 to support research designed to adapt science and technology to the needs of developing countries. The Centre's activity is concentrated in six sectors: agriculture, food and nutrition sciences; health sciences; information sciences; social sciences; earth and engineering sciences; and communications. IDRC is financed solely by the Parliament of Canada; its policies, however, are set by an international Board of Governors. The Centre's headquarters are in Ottawa, Canada. Regional offices are located in Africa, Asia, Latin America, and the Middle East.

Le Centre de recherches pour le développement international, société publique créée en 1970 par une loi du Parlement canadien, a pour mission d'appuyer des recherches visant à adapter la science et la technologie aux besoins des pays en développement; il concentre son activité dans six secteurs : agriculture, alimentation et nutrition; information; santé; sciences sociales; sciences de la terre et du génie et communications. Le CRDI est financé entièrement par le Parlement canadien, mais c'est un Conseil des gouverneurs international qui en détermine l'orientation et les politiques. Établi à Ottawa (Canada), il a des bureaux régionaux en Afrique, en Asie, en Amérique latine et au Moyen-Orient.

El Centro Internacional de Investigaciones para el Desarrollo es una corporación pública creada en 1970 por el Parlamento de Canadá con el objeto de apoyar la investigación destinada a adaptar la ciencia y la tecnología a las necesidades de los países en desarrollo. Su actividad se concentra en seis sectores: ciencias agrícolas, alimentos y nutrición; ciencias de la salud; ciencias de la información; ciencias sociales; ciencias de la tierra e ingeniería; y comunicaciones. El Centro es financiado exclusivamente por el Parlamento de Canadá; sin embargo, sus políticas son trazadas por un Consejo de Gobernadores de carácter internacional. La sede del Centro está en Ottawa, Canadá, y sus oficinas regionales en América Latina, Africa, Asia y el Medio Oriente.

This series includes meeting documents, internal reports, and preliminary technical documents that may later form the basis of a formal publication. A Manuscript Report is given a small distribution to a highly specialized audience.

La présente série est réservée aux documents issus de colloques, aux rapports internes et aux documents techniques susceptibles d'être publiés plus tard dans une série de publications plus soignées. D'un tirage restreint, le rapport manuscrit est destiné à un public très spécialisé.

Esta serie incluye ponencias de reuniones, informes internos y documentos técnicos que pueden posteriormente conformar la base de una publicación formal. El informe recibe distribución limitada entre una audiencia altamente especializada.

IDRC-MR293e
May 1991

New Horizons in Agricultural Information Management

Proceedings of an International Symposium,

March 13-16, 1991, Beijing, China

Compiled and Edited by

Gary K. McCone

IDRC
CANADA

SDIC

Sponsored by

International Development Research Centre

Organized by

Scientech Documentation and Information Centre

Chinese Academy of Agricultural Sciences

Organizing Committee

Main Organizing Committee

Chairman	WANG Xianfu
Vice Chairman	JIAO Bin
	HE Chunpei
Secretary	MIAO Zhuoran
Members	HAN Ling
	JIA Shangang
	HUANG Xuegao
	GUO Dianrui
	ZHAO Huaying
	PAN Shuchun
	LI Kaiyang

Secretariat

Chairman	MIAO Zhuoran
Members	CHEN Junying
	HU Jia
	YU Fenghui
	TAI Weidong
	FANG Baoqin

Accomodation & Transport

Chairman	HUANG Xuegao
Members	ZHAO Huaying
	QIN Juanjuan

Conference Site

Chairman	HE Chunpei
Members	PAN Shuchun
	ZHANG Rongchang

Papers

Chairman	JIA Shangang
Members	LI Kaiyang
	LIANG Suzhen
	WANG Zhenjiang
	GUO Jian

Table of Contents

Keynote Address

Session I: Management and Development of National Agro-Information Systems

Session II: Information Management and New Technology Application

Session III: Management and Development of Regional Agro-Information Systems

Session IV: Scientech Information and Productivity

Session V: Development and Utilization of Agro-Information Resources

Appendix 1: Supporting Papers

Appendix 2: Symposium Participants

Appendix 3: Author Index

Establishment of the Chinese Agriculture Abstracts Database

GUO Jian

Scientech Documentation and Information Centre
Chinese Academy of Agricultural Sciences
Beijing, China

Abstract

This article describes the procedure of establishing the database of Chinese Agriculture Abstracts with MINISIS software on minicomputer and Micro CDS/ISIS software on microcomputer, including a connected sequence of building the database, typesetting the journal publication, and producing the subject index by computer.

I. Information Source

Chinese Agriculture Abstracts (*CAA*) is a set of journals which publish comprehensive abstracts in agriculture. The publication of these journals was started in 1980 by the Scientech Documentation and Information Centre of the Chinese Academy of Agricultural Sciences. These journals are bimonthly and provide about 11,000 abstracts, including a subject index at the end of each year. It includes six individual journals: *Soil and Fertilizer*, *Veterinary Medicine*, *Animal Science*, *Horticulture*, *Plant Protection*, and *Grain and Economic Crops*. At present, about 1,000 different periodicals are available for abstracting by the editorial board. Of these, 400 are considered to be key periodicals. CAA concentrates on collecting and reporting the literature and documents of agriculture as well as those that are closely related to agriculture in our country. The abstracts are made from articles in periodicals, journals, trade magazines, monographs, conference proceedings, etc. The statistics show that articles from periodicals, trade magazines and journals are the main source of information.

II. The Main Concept of the Database

The computerized database of *CAA* is a Chinese character abstracts database built on an HP 3000/70 computer with MINISIS DBMS (the software was kindly provided by IDRC). The Database of Chinese Agriculture Abstracts (DB CAA) is in fact a comprehensive computerized information processing system which possesses the capabilities of compiling indexes and database retrieval.

The establishment of DB CAA began in 1988. At the beginning, we directly input the data on an HP 3000/37 computer, also all the data were saved in the HP 3000/37 computer every day. After working in this way for a period of time, we found that the input speed of the Chinese characters was slow, the computer system's expense was high and that Chinese character processing in personal computers had improved.

Because of a personal computer's fast input speed, simple operation and convenient management, we determined to switch to using a personal computer as the input device.

The DBMS software we used in the personal computer to build the database is Micro CDS/ISIS which was developed by UNESCO. It allows us to build and manage structured non-numerical databases. Its functions are similar to and compatible with MINISIS software. It provides a number of functions for the establishment of a database: definition, modification, input and output, etc. On the basis of such conditions, we have developed a flowchart of the system. Figure 1 shows the flowchart of the operation of this system.

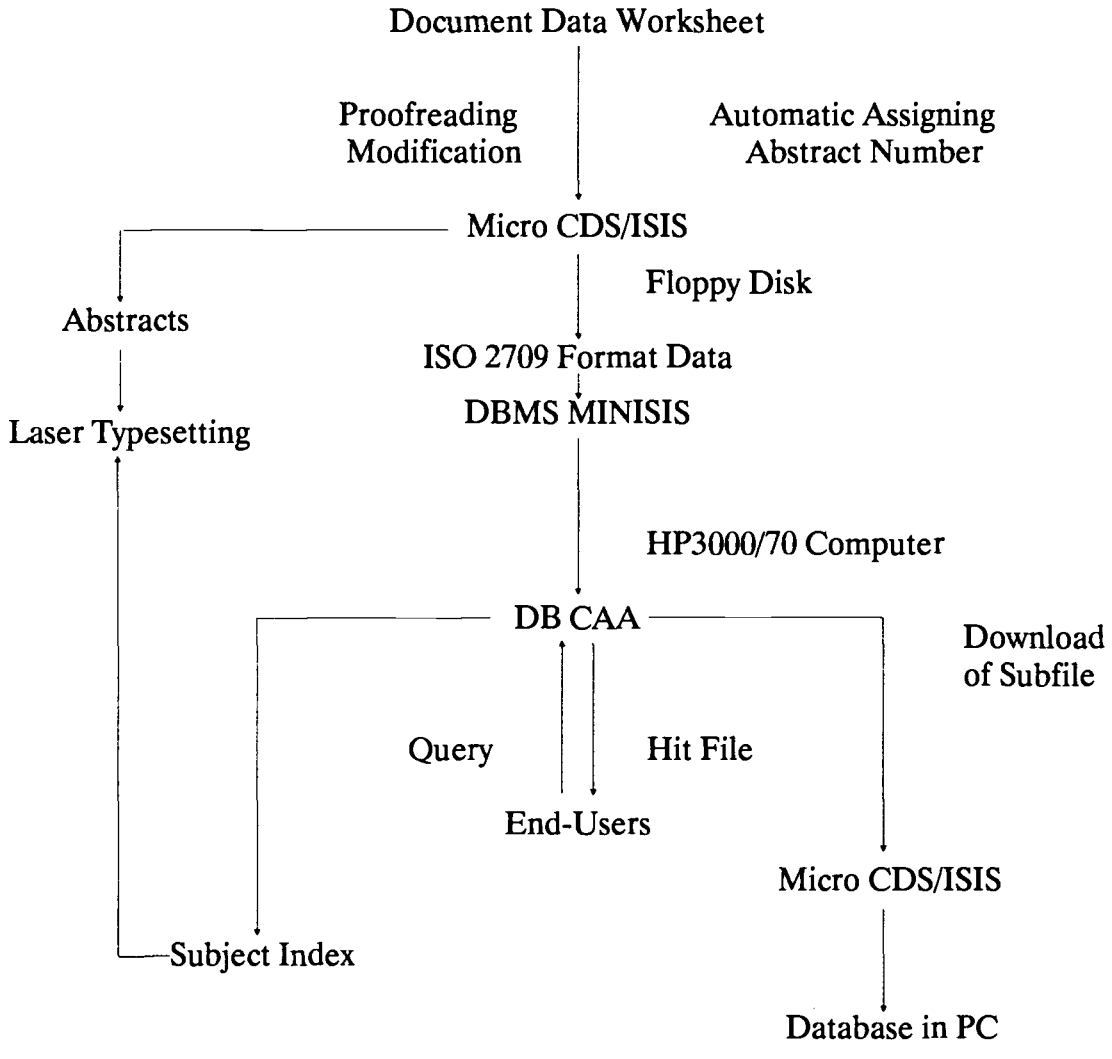

Figure 1. System Flowchart Diagram of DB CAA

III. Realization Of The Establishment Of DB CAA

Each record of the DB CAA is designed to include 25 fields. These fields include the contents of the abstracts. It is searched by classification number, author, keywords in title, descriptors and feed terms. During retrieval, users can set up Boolean expressions including "logical OR, logical AND, logical NOT, left-truncation and right-truncation" to retrieve information from the database.

During the building of DB CAA, we wrote a program which can assign abstract numbers by computer according to the classification number of each record. In this way, we can not only reduce the work burden of editors, but also avoid errors. This skill is a simple one. See Figure 2.

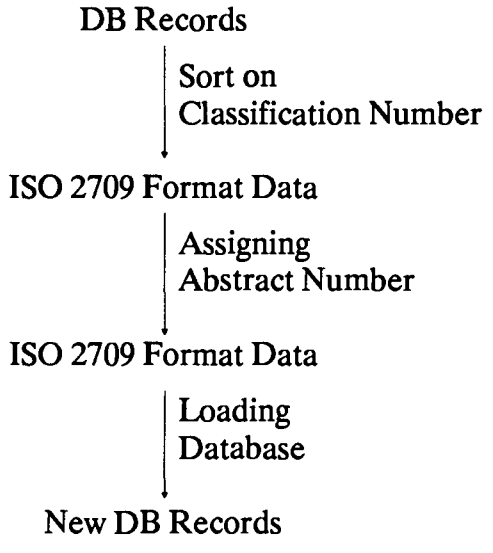

Figure 2. Flowchart Diagram on Assigning Abstract Number by Computer

Because of the complex structure of the database when using Micro CDS/ISIS software, we avoid calling database records directly. But we can call ISO-2709 formatted data as intermediate forms. It is a text file and easy to process. After having abstract numbers assigned by the program, the ISO-2709 format file becomes a new file which includes abstract numbers. It forms a new sequence of records in the database after the file is loaded into the master file in the minicomputer.

As subsidiary products, we also typeset the abstracts journal during the creation of DB CAA. We use the data from the database records to typeset the journal publication so as to avoid unnecessary repetition in data inputting. There are two kinds of typesetting software available in our country. One kind of software is used for Office Automation (OA), but it is not suitable for printing journals. Another kind of software is fit for

publishing. It includes batch processing, alternative type fonts and combinations of type. The software we chose is "KeYin" which is batch processing software. The procedure for typesetting the journals is described briefly as follows:

Using a formatting language with Micro CDS/ISIS software, we insert the functional symbols of typesetting software in the output file, e.g., size of the printed character, form of the printed character. After a file to be edited is produced by Micro CDS/ISIS, we artificially interpose some characters which are not inputted during data input, such as symbols of molecular formulae. When the file has been edited, the computer is used to automatically typeset the file and a typeset file is produced. At last, the typeset file is printed by HP LaserJet II printer. We produce a subject index of each abstract journal at the end of each year. Now the Chinese characters are classified into two levels. The first level is sorted according to phoneticism. The second level is sorted according to the basic structure of the Chinese characters, such as strokes, etc. Because of this, the output index file is not fit for publishing. We use the thesaurus management program of MINISIS software to solve this problem. Both first and second level Chinese characters (a total set of 6,763 Chinese characters) and other character symbols are sorted according to the sequence of the *XinHua Dictionary*. We then rebuild a new index comparison table on MINISIS. During sorting, the new table is used to replace the original table. Thus, we have basically solved the issue of sorting Chinese characters.

At present, the content of the database includes the following disciplines (corresponding to the printed abstract journals): soil and fertilizer, veterinary medicine, animal science, horticulture, plant protection, grain and economic crops. Through our efforts, DB CAA now contains about 5,400 records. We estimate that DB CAA will reach about 14,000 records at the end of this year.

IV. Problems

We have encountered many problems during the past two years. Some have been solved, others remain unsolved. The first problem is the shortage of Chinese characters in the computer. About sixty Chinese characters are missing in the DB CAA which has been loaded into HP 3000/70 computer. The second problem is the excessive artificial interpositions in the original typeset file during editing (on the basis of the contents of the abstract). The second problem refers mainly to symbols of molecular formulae and the italics of the Latin alphabet which the computer can hardly determine because they are in the text of the abstracts. The last problem is how to index a Chinese character which has multiple phoneticism. In a computer, the sequence of the phoneticism of a Chinese character is one-to-one. The computer can't determine how to phoneticize a character. These problems may be solved later when new software is available.

V. Conclusions

The establishment of the Database of Chinese Agriculture Abstracts has great significance because it is the first Chinese character agriculture abstract database in our country. We provide retrieval services while building this database. Users and readers

are welcome to query and test DB CAA. In so doing, we can promote both our work and the development of agriculture in China.