Documentation and Electronic Data Processing

How can a large company cope with the problem of the rapid increase of bibliographic data and the growing interest in technical literature? The following article describes a practical solution for the dissemination and classification of literature abstracts based on electronic data processing.

A. M. BECKER

Sulzer Bros. Ltd.
Wintherthur, Switzerland

• 1. Introduction

The Sulzer Company has recently worked out a system for the dissemination and classification of literature abstracts, also for the automatic supplementation of the common catalogs in the technical library. The system is named "SULIS," an abbreviation of *Sulzer Literaturverteilung und Sortierung* (Sulzer literature dissemination and classification).

The UDC number (number of the universal decimal classification system) is used as the criterion for the distribution of literature abstracts among the subscribers of the company.

For the processing of the SULIS system we use a CDC 3300 computer, a Xerox copying machine, and an IBM sorter for punched cards.

• 2. Definition of the Problem and Aim

The constant increase of bibliographic data and the growing interest of the employees in technical literature resulted in the technical library being flooded with a lot of literature abstracts, which the normal amount of personnel could no longer work up. By "work up" we mean punctually delivering literature abstracts to the right persons and keeping the different card indexes and catalogs of the technical library up to date.

To increase the number of employees for this purpose would be very difficult and, last but not least, uneconomical. Thus we defined our aim as follows:

To work out a system for automatic dissemination of literature abstracts, their classification in different card indices, and maintenance of updated catalogs in the technical libraries.

• 3. The UDC Number as the Criterion of Distribution

The distribution of literature abstracts is based on the UDC number. This Decimal Classification System is very common in Europe. Most of the scientific libraries classify their books and newspaper articles with the aid of UDC numbers. Other systems of distribution—using keywords or microfilm systems, for example—would require extensive preparation of data and long processing times on the computer. Such expense would be unjustified for an engineering works.

The logical structure of the UDC number is another advantage that justifies its use as a criterion for the distribution of literature abstracts. In order to be able to analyze this system, we give here some basic information about the UDC number:

The universal decimal classification method is based on the decimal numerical system. The whole range of human knowledge is divided into 10 main sections, each of which is subdivided into 10 sections of a second order, etc. The 10 main sections of the UDC numbers are:

0. General, bibliography, library matters
1. Philosophy
2. Religion, theology
3. Sociology, law, administration
4. Philology, linguistics
5. Mathematics, natural science
6. Applied sciences, medicine, technology
7. Art, applied arts, games, sports
8. Literature
9. Geography, history

The 10 subsections of the second decimal number of group 6 (applied sciences, medicine, technology—which is the most interesting group for an engineering works) are:

60. Applied sciences in general
61. Medicine

62. Engineering
63. Agriculture
64. Domestic economy
65. Trade and commerce
66. Technical chemistry
67. Diverse industries
68. Diverse industries continued
69. Building materials, building trade

Through continual subdivision, each section of a higher order can be further differentiated as required. The following examples concern some of our branches:

621	Mechanical engineering
621.4	Heat engines (without steam engines)
621.43	Internal combustion engines
621.436	Diesel engines
621.436.12	Four-stroke diesel engines
621.436.14	Two-stroke diesel engines

All UDC numbers applied in our system contain 12 digits. UDC numbers with less than 12 digits (without special marks) are "filled up" with zeros up to 12 digits.

THE CONCEPTION OF THE SULIS SYSTEM (see Fig. 1)

We divide the system into five main phases, which repeat in a 14-day cycle.

1. Preparation of data and comparison decks
2. Data processing, output of punched cards and magnetic tape
3. Lettering and reproduction
4. Updating of the magnetic tape and printing of supplements to new edition of literature catalogs
5. Separating, sorting, distribution, and classification of the punched information cards

• 4. Preparation of Data and Comparison Decks

4.1 PREPARATION OF DATA

Literature abstracts from books and journals prepared in the firm or provided from outside are typed on cards that have the dimensions DIN-A-6 (104 × 147 mm) or less and contain: UDC numbers, name of author(s), name of journal, year of publication, and a short summary of the book or journal article. Each card is provided manually with a serial number (see Fig. 2).

For each literature abstract two punched cards are prepared. They contain the serial number, standardized UDC numbers (12 digits per number), authors or special designations, and the name of the journal.

Eight UDC numbers are at the moment available for covering the content; however, in 95% of all cases we need from four to six UDC numbers for the description of the subjects. This pack of punched cards represents the Information Deck and provides the input for the SULIS program.

4.2 COMPARISON DECKS

Four registers (card decks) give the basic information to the SULIS system about who in the firm is interested in what.

The Register of the Subscribers contains the names and internal addresses of those persons in the firm who are interested in literature information cards. Each subscriber has a serial number with which the computer program operates.

The Simple UDC Deck assigns to each subject (division of knowledge) the numbers of the interested subscribers. Thus each punched card contains a UDC number of 12 digits, followed by the numbers of the corresponding persons (according to the register of subscribers). For example: if a person is interested in the subject of diesel engines, his subscriber number will appear among other numbers at the corresponding UDC number for this subject as follows:

621436000000 ... ... 010 ... ...

The Combined UDC Number Deck gives a more exact definition of a division of knowledge and hence a more accurate dissemination of literature.

Each UDC number contains a combination of two UDC numbers with 6 digits each. For example: if subscriber 007 is interested in the subject of vibrations in crankshafts, but is not interested in vibrations and crankshafts as separate subjects, then he will appear as a subscriber in the combined UDC number deck:

Vibrations = UDC number 534.1
Crankshafts = UDC number 621.824

The corresponding combined UDC number 534100621824 is created by the consolidation of the first number 534.1 (special signs are omitted and two zeros are added in order to fill up the first 6 digits) and the second number 621.824 (special signs are omitted). Thus we get our combined UDC number deck for subscriber 007:

534100621824 ... ... 007 ... ...

The logic of the combined UDC number enables us also to allow for the special appended numbers of the UDC classification system. Special or general divisions of knowledge can be brought into direct relation with a second subject. For example:

621.4–5	control of heat engines only
621–5	control in mechanical engineering (including the subgroups 621.–621.9)

The Blocking Deck excludes whole divisions of knowledge from the simple and combined UDC deck that a certain subscriber is not interested in. Experience has demonstrated that several divisions of knowledge defined by our two previous comparison decks are much too general and that the subscribers are often fed with information cards that are not of great interest to them. The blocking deck provides a filter for unnecessary information. For example:

subscriber 218 is interested in
658700000000 = sales general (quoted in this simple UDC deck).

But he is not interested in

658770000000 = receipt of goods and
658780000000 = storage

which are both subdivisions of general sales (658700000000) according to the UDC classification system.

In the Blocking Deck we excluded the divisions from

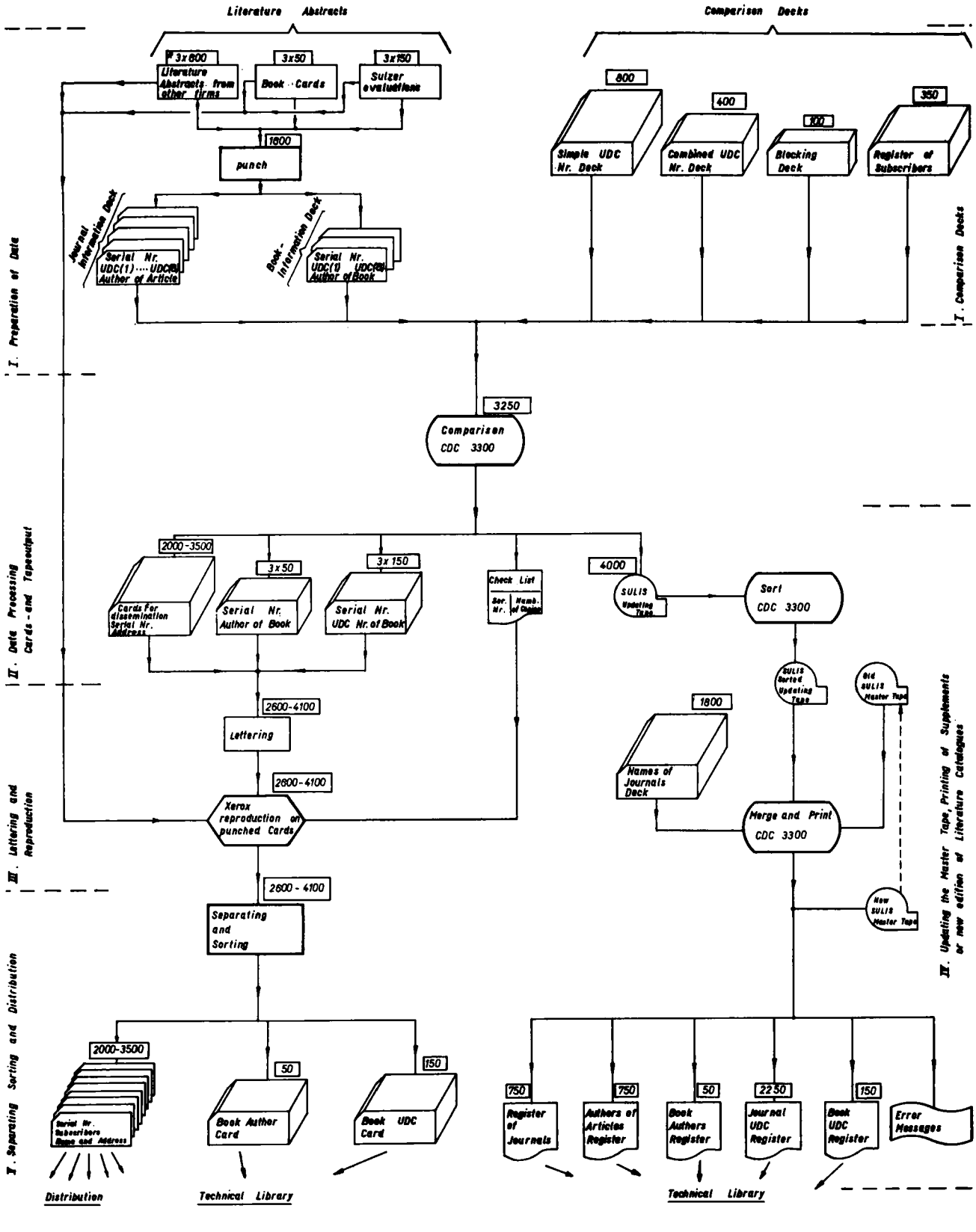

FIG. 1. Flow chart of the SULIS system. (*The numbers in the rectangles express cards resp. records per 14 days.)

ROBERT MAXWELL & CO. LTD. DOCUMENTATION & SUPPLY CENTRE

ANNOUNCES A NEW/FORTHCOMING PUBLICATION OF SPECIAL INTEREST AND IMPORTANCE:-

665.5

BLAND, W. F.; DAVIDSON, R. L.
Petroleum Processing Handbook
1000 pp. approx. £14.0.6d./F 252.45

This book provides a collection of facts and information about the technology of converting crude oil into finished petroleum products such as gasoline, kerosine, fuel oils, diesel oil, and lubricants. Also included is information on the various types of equipment used in processing plants, their maintenance and construction. BF/25302/L

Supplied on Firm Account only/Pas en communication

PUBLICATION DATE: Approx. October 1967, London, McGraw-Hill

FOR SPEEDY SERVICE PLEASE POST THIS CARD TODAY! QUOTING YOUR ORDER NUMBER AND WHERE REQUIRED SEND WITH YOUR OFFICIAL ORDER FORM →

LEVEL * A

PLEASE SUPPLY ☐ ON APPROVAL ☐ FIRM ACCOUNT

* LEVEL AT WHICH PUBLICATION IS WRITTEN

A ADVANCED G GENERAL H HIGHLY ADVANCED
I INTRODUCTORY O INTERMEDIATE R REFERENCE

REFERENCE NUMBER 0 0 0 0 (76 77 78 79 80)
1 1 1 1 1
2 2 2 2
3 3 3 3 3
4 4 4 4 4
5 5 5 5 5
6 6 . . 6
7 7 7 7 7
8 8 8 8 8
9 9 9 9 9 (76 77 78 79 80)

207-1653-027

DK 531.71 621.91/.95-52

DK 5268 (Best.-Nr.)

VDMA 052.3 061.1 066 211

Messen und Prüfen während der Bearbeitung
(In-Process Measurement and Size Control)
8 B.

Tipton, H.

Metalworking Production
110 (1966) 49, S. 77-84

Nach Betrachtungen über das Messen der Werkstückdimensionen während der Bearbeitung und die Steuerung der Werkzeugmaschinen nach diesen Ergebnissen wird die Anwendung der Meßsteuerung bei numerisch gesteuerten Werkzeugmaschinen erörtert. Überblick über die gebräuchlichen mechanischen, pneumatischen, elektrischen und optischen Einrichtungen und Bericht über bereits entwickelte, aber noch nicht ausreichend erprobte Verfahren: u. a. Mikrowellen-Interferometer, Holographie, Ultraschall, Servoübertrager, Geschwindigkeitsmessung. Beispiele für die Anwendung von Meßsteuerungen beim Schleifen, Drehen und Honen.
sef

SKA
427.552

TH Per.410 918

Brunee, Y. and J. Finn. When will a Nuclear Merchant Ship be Competitive? 15 figs., 10 tabs., 14 ref.
ACEC Rev., -(1962)No.1, p. 12-31.
Safety conditions required - choice of the PWR - economical analysis of operating conditions of nuclear and conventional ships - comparative study of different charges that constitute cost price per ton of cargo - nuclear propulsion will cross before long the threshold of competitivity.
1009 ETH

ETH ZH 25.6.62.

FIG. 2. Examples of literature abstracts for dissemination and classification

658770000000 to 658789999999 for subscriber 218. Subscriber 218 will now have the following simple UDC number assigned: 658700000000

	Subscriber	*From*	*To*
Blocking:	No. 218	658770000000	658789999999

• 5. Data Processing and Output of Punched Cards and Magnetic Tape

The computer compares the UDC numbers of the Information Decks with those of the Comparison Decks and decides, according to a certain complex of criteria (see 5.1—logical processing), who shall get which information. If the computer finds a fitting subscriber, it punches a card containing the serial number of the information card, the name of the subscriber, and his address. If the information card concerns a new book in the technical library, the computer also punches cards of the authors and subject registers in the library. Those cards contain the serial number, the author's name or UDC numbers.

All new entries of literature abstracts into the technical library are analyzed and ordered in the five common catalogs kept by the technical library. These records are written on a special magnetic output tape.

At the end of the processing the computer prints a check-list showing how many copies of each original from the literature abstracts are desired.

5.1 The Logic of Processing

In order to decide if and to whom an Information Card is to be assigned, each UDC number from this card is compared with the numbers of the simple UDC Number Deck. The comparison does not test for 100% agreement between the two UDC numbers; instead it is based on the structure of the UDC number.

If a subscriber is interested in the UDC number 621400000000 he will get all information corresponding to the first four digits of this number, independent of the residual digits (from 5 to 12). However, if one of the first four digits of the UDC number from the examined Information Card does not correspond to the defined UDC number of the subscriber, no assignment will take place. Table 1 may help to explain this principle.

Table 1. Assignment of information cards to subscribers by simple UDC number deck

Subject of subscriber	Information card	Assignment
621400000000	621430000000	yes
	621438762000	yes
	621348700000	no
	621400001000	yes
621436120000	621436120000	yes
	621436122100	yes
	621436110000	no
	621400000000	no
	621436020000	no

The combined UDC Number Deck has a similar logic to the simple UDC Number Decks for dissemination of literature. Instead of one, two comparative tests are carried out, one for the first 6 digits and the other for the last 6 digits of the combined UDC number. Only when both tests give a positive answer does an assignment of the Information Card take place. The example in Table 2 illustrates the logic of the distribution.

Table 2. Assignment of information cards to subscribers by combined UDC number deck

Subject of subscriber	Information card	Assignment
534100621824	534000000000 621824000000	no
	534120800000 621824600000	yes
	534120000000 621820000000	no

• 6. Lettering and Reproduction

The output of the punched cards is interpreted by an IBM 557, so that names and addresses of the interested subscribers and the serial numbers of the literature abstract appear on their front side. In the next step the originals of the literature abstract are reproduced with the help of a Xerox copying machine 914 on the reverse side of the punched cards (see Fig. 3). The identification of original literature abstracts and the corresponding punched cards is done with the help of the serial numbers, which appear on both of them. The checklist helps the operator to check his reproduction work.

• 7. Updating of the Magnetic Tape and Printing of Supplements to New Edition of Literature Catalogs

In the next step the magnetic output tape is sorted and merged on the CDC 3300 with the old stock of the catalogs. With the aid of a register of journal names common in our library, the new supplements to our five catalogs are prepared. We print new supplements every 14 days and once in 6 months we completely revise the catalogs of the last 2 years.

• 8. Separating, Sorting, Distribution, and Classification of the Punched Information Cards

The imprinted cards are now separated (according to special codes) and sorted by an IBM 084 sorter.

The punched cards that contain the names and addresses of the subscribers are sorted according to the addresses and distributed by the internal post. The punched cards for the book registers of the technical library are mechanically ordered into the current "yearly stock." At the end of every second year the two last stocks are mixed together.

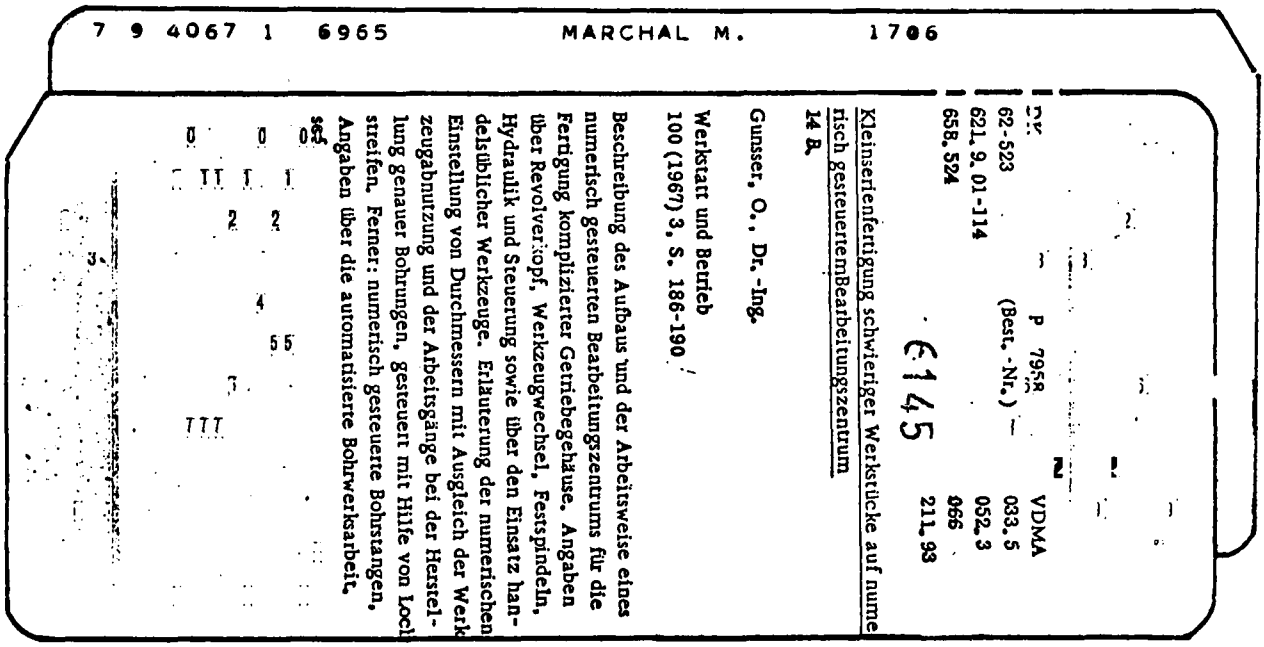

FIG. 3. Serial number, name, and address of the subscriber. Reverse: Reproduction of the original literature abstract.

• 9. Retrieval

Literature researches can be done with the aid of the five catalogs ordered according to:

authors of books
subjects of books
names of journals
authors of publications in journals
subjects of publications in journals

The catalogs refer to the serial numbers of the originals. The originals—containing the abstracts of books or of publications in newspapers—are kept in a card index according to their serial numbers. A special program is drawn up for complicated researches in which the computer has to search for a definite subject on the magnetic tapes containing the stocks of our library. As a result a list is printed containing authors, and names of journals or books and their serial number in the SULIS system concerning the subject in question.

• 10. Some Figures on the SULIS *System*

The following figures may provide a picture of the amount of material expected to be processed during a year:

processing	about	20,000	literature abstracts a year
subscribers	about	400	
distribution	about	75,000	information cards a year
ordering in card registers	about	5,000	book cards a year
supplementations in catalogues	about	100,000	records a year

• 11. Computer and Program Language

The SULIS program is written in COMPASS for the CDC 3300 computer: 32 K words, 24 bits a word. The facility of dynamic core assignment was extensively explored; thus the utilization of the core is near to the optimum. The features required from the system are: card reader, card puncher, printer, and three magnetic tape units.

• 12. Final Remarks

The success of such a system, once it is completed, depends, among others, on at least five main points:

1. Accurate Comparison Decks for which the Technical Library is responsible
2. Good maintenance of the Comparison Decks for which the Technical Library and the Computer Center are responsible
3. Feedback by the subscriber to the Technical Library in case he receives irrelevant information cards
4. Provision of suitable literature abstracts
5. Correctly defined UDC numbers on the provided literature abstracts

The system has been running since March 1967 to the satisfaction of most of our 400 subscribers; the Comparison Decks are continuously updated. A rough calculation shows that the expenses of this system are 25% lower than the manual method and that the output of the Technical Library has increased by about 40% with the SULIS method.

Acknowledgments

Mr. M. Marchal, former Manager of the Computer Center of Sulzer Brothers Ltd., took part in the conception of the system. Mr. B. Stüdeli, Manager of the Technical Library, gave us detailed information about the structure of the UDC numbers and gathered the various data for the system. Mr. V. Hausherr wrote the program of the SULIS System.