**Francia. Forschungen zur westeuropäischen Geschichte**
Herausgegeben vom Deutschen Historischen Institut Paris
(Institut historique allemand)
Band 12 (1984)

DOI: 10.11588/fr.1984.0.51468

Geirr I. Leistad – Ferdinand Linthoe Næshagen
Per-Axel Wiktorsson

# ONLINE PROSOPOGRAPHY: THE PLAN FOR NORDIC MEDIEVAL DATA BASES

## 1. The plan for Nordic medieval prosopographic data bases

In the autumn 1982 the Nordic Council for Scientific Information and Research Libraries (NORDINFO) granted a sum of 75 000 Finnish marks to a team consisting of editor Geirr Leistad of the Norwegian Centre for Informatics, senior archivist Ferdinand Næshagen of the Norwegian Institute for Historical Sources and associate professor Per-Axel Wiktorsson of the Diplomatarium Suecanum, formerly of the »Sveriges Medeltida personnamn« (Sweden's medieval person names). The aim of this ad hoc team was to work out and test a plan for Nordic medieval prosopographic data bases.

The plan is – that ought to be underlined – for prosopographical data bases, not for prosopographies. What could turn them into prosopographies, identification, is such a tricky and time-consuming task that it had better be left aside for the present[1]. With careful and thorough computer programming the prospective users, professional and amateur researchers, need not, however, lose much, for computer search procedures can help them a long way. In fact the prosopography can most practically be developed from something very like the data base, and one might therefore say that the user is given access to the first stage of a prosopography instead of having to wait until the task is completed.

The point of departure must necessarily be the needs of the user, and these can be charted by asking, »what use do researchers make of prosopographical data?«. It seems to the authors that this is a point at which many projects, and not just computer projects, have gone wrong, and we believe most readers can call to mind instances where false notions of economy or lack of imagination have made projects miss important opportunities which could be had for a small extra effort.

There was never any doubt that a modern prosopographical data base had to be a computer data base[2]. Not only is the computer an indispensable tool for efficient search procedures, but by this means information can also be piped through to prospective users, even if they should sit before the screen in faraway Tromsø. Nor was there any doubt that the system chosen had to be an online one, allowing for interaction between user and data base, and desiderata like flexibility and user-friendliness were given high priority.

1 See E. A. Wrigley (ed.), Identifying People in the Past, London (Edward Arnold) 1973.
2 See V. Stibic, Tools of the Mind, Amsterdam (North-Holland) 1982.

## 2. The concept of data bases

Any collection of data on alphanumerical form may be called a data base. Stored in computer it is of course a computer data base. Moreover, if it can be read and searched by someone connected to the computer via communication lines, it is also an online data base. Modern computer techniques allow for several persons using the computer and accessing the data base simultaneously. An online data base may thus be available in a computer network and used by persons in widely separated geographical locations, usually through public telephone and data communication networks.

To facilitate the use of the online data bases the data have to be organized in a way conducive to their processing and retrieval. Usually the data base will be organized in records or documents, each regarded as conceptually identical, but giving unique information about each subject entity. In the case of prosopographical data bases we take the records to represent information on persons extracted from historical and demographical sources. The subject entities will be any singular instance of occurrence of persons mentioned in the sources.

The records in the data base may be organized to consist of data fields, each field covering semantically distinguishable information relating to the singular occurrence of persons in the source. The number of data fields is partly dictated by practical reasons, partly by the nature of the source material. The practical reasons depend on the chosen information system, that is the set of software used in generating, storing and searching the data base. Compromises have to be made, as always, but there will usually be no difficulties in setting a field format for the data base in such a way that both the information content of the source and the search needs of the user are taken care of. Each data field is given a unique field code (fig. 1), which regularly is

⟩ FIELDS

| No | Fieldname | Code | No | Fieldname | Code |
|---|---|---|---|---|---|
| 01 | Serial number | NR | 21 | Language | LA |
| 02 | Updating code | UC | 22 | Function | FN |
| 03 | Identification number | ID | 23 | Title | TI |
| 04 | First name | NF | 24 | Social position | SP |
| 05 | Parent name | NP | 25 | Occupation | OC |
| 06 | By-name | NB | 26 | Office | OF |
| 07 | Localisation name | NL | 27 | Office location | OL |
| 08 | Original name | NO | 28 | Scope note | SN |
| 09 | Normalized name | NN | 29 | Father | RF |
| 10 | Sex | SE | 30 | Mother | RM |
| 11 | Date | DA | 31 | Husband | RH |
| 12 | Year | YR | 32 | Wife | RW |
| 13 | Place of issue | LI | 33 | Children | RC |
| 14 | Locality | LO | 34 | Siblings | RS |
| 15 | County | LF | 35 | Other relatives | RO |
| 16 | Country | LC | 36 | Relatives by marriage | RL |
| 17 | District | LD | 37 | Seal | SL |
| 18 | Reference | BR | 38 | Seal inscription | SI |
| 19 | Source code | SC | 39 | Seal source | SS |
| 20 | Source type | ST | | | |

Fig. 1 Data fields and codes

attached to the information the field contains. This information will be in the form of data elements, the smallest semantically meaningful piece of information that the record contains. Several data elements in the same field may be organized into one or more data blocks. The data elements themselves will always consist of alphanumerical character strings.

Information will be retrieved from the data base by the user simply asking the computer for the wanted data element by entering it together with the appropriate field code. The computer will then respond by giving the number of records containing the data element; each record may then be printed or shown on the screen of the user's data terminal. The user may also combine the wanted data elements, called search words or keywords, by logical operators, whether the elements are contained in the same field or not. An example of searching in this mode is shown in fig. 2.

```
⟩BERGEN(LO) OR HO(LF)
     1    29    BERGEN (LO)
     3    30    HO (LF)
     2    30     1 + 2

⟩F(SE) AND 14; (YR)
     4    19    F (SE)
     5    30    14:
     6    11     4 ★ 5

⟩3 AND 6
     7     6     3 ★ 6

⟩ADEL(SP) OR FRU(TI) OR
HUSTRU(TI)
     8     7    ADEL (SP)
     9     7    FRU (TI)
    10     3    HUSTRU (TI)
    11    10     8 + 9 + 10

⟩NOT 11
    12     3     7 − 11
```

Fig. 2 Searching. The search queries with keywords, codes and operators are shown tagged ⟩. The first column is the set number, the second the number of instances, the third the corresponding search set. Set no. 12 with 3 instances ist the wanted one.

## 3. Creating data bases for historical research

Data bases have long been common in the natural sciences[3] and not only as bibliographic data bases, containing information on literature, but also as data bases giving qualitative and quantitative information relating to the subject matter of the disciplines in question. The star catalogues of astronomy and compound catalogues of chemistry should exemplify wellknown and indispensable data bases in these fields.

The application of computers to the processing of data bases has been a quantum leap in easing the task of collecting and handling scientific information. The application of computer data bases in the social sciences is now rapidly spreading.

A large part of the historian's task consists of collecting, sifting and collating information contained in the sources. Although the sources may exist in excellent editions, their use may nevertheless be extremely time-consuming. Hence the use of computerized data bases, especially in online versions, would increase not only the effectiveness of the research, but also the reliability of the results, simply by allowing the historian a far better control and overview of the source material, due to the possibility of using the data base directly as input for quantitative and statistical investigations of historical problems.

3 See Derek J. de Solla Price, Little Science, Big Science, New York and London (Columbia University Press) 1963.

With the advent of prosopography as an important tool in most branches of early history, it is quite natural that we did consider the possibility of creating online prosopographical data bases. This resulted in the aforementioned project for developing Nordic medieval prosopographic data bases.

The first phase of this project was a study of the system demands and techniques necessary for an online prosopographic data base which could be used to advantage by historians and social scientists whether familiar with computers or not. This we regard as a *sine qua non*. – By the system we here understand both the software, i. e. the computer programs, and the hardware, computers, terminals etc. used.

## 4. Online data bases: System demands and techniques[4]

A computer does not only compute, it may also be used for storing large amounts of information more or less indefinitely. By the use of appropriate software, this information can be organized in various ways, moved about, sorted, searched and retrieved. The information may also be fed from the data base to special software for performing computations, statistical analyses or preparing it for computer typesetting.

Conceptually, the computer's handling of information does not differ much from that of humans. But there are, at least presently, important differences. The computer cannot think and it cannot create. One should not expect »intrinsic newness« from a computer. It acts on the instructions fed into it by the operator and is not able to simulate the powers of associative thinking shown by the human mind. It may be programmed to learn from experience and at the time being this is the most advanced form of intelligence the computer is able to show[4a]. It can only handle information in the form of sequences of alternating physical states, usually as binary-valued signals. It is by reading, counting, storing and writing such states the computer is able to do its work. Evolutionarily speaking, the computer is an imbecile compared to the human brain. But the electronic computer has a tremendous advantage over the brain. It is able to handle enormous amounts of information at a very high speed. Thus, the same properties that enable the computer to perform millions of multiplications in a second, also make it eminently suitable for handling large online data bases.

The primary requirements of such data bases concern generation, storage, searching and retrieval, which in its turn means considering the appropriate hardware and software. A data base corresponding to PROL, for instance, may contain several hundred thousand records. A fairly large and fast minicomputer would then be required, viz. the VAX or the NORD 500. The data base will typically consist of several hundred megabytes, each byte representing a single character. This again means that a mass storage system is needed, usually in the form of a disk station, capable of storing from 300 Mbyte to 1 Gbyte. The disk station functions as an external memory to the computer; any particular information on the disk station may be accessed and transferred to the computer in milliseconds. The computer must be able to handle several external users of the online data base at the same time, i. e. it has to have full timesharing and multiple communications facilities. The computer must moreover be able to treat queries rapidly. Online users tend to get spoiled quickly, hating to wait more than 5 seconds for information it would have taken them weeks to search manually.

Data will be fed into the computer, preferably by skilled terminal operators. The terminals consist of a video display unit and a keyboard. The intellectual preparation of the data that comes before entering them into the computer is better described in connection with the

4 See F. W. LANCASTER, Information Retrieval Systems, New York (Wiley) 1979.

4a But see F. L. NÆSHAGEN, Statistics and Historical Research, in: Francia 12 (1984), in particular note 13.

| | | |
|---|---|---|
| 01 | NR: | |
| 02 | UC: | |
| 03 | ID: | |
| 04 | NF: | Ragnhild |
| 05 | NP: | Mattisdatter |
| 06 | NB: | |
| 07 | NL: | |
| 08 | NO: | Ragnildh Matisdotter |
| 09 | NN: | Ragnhild Mattisdatter |
| 10 | SE: | F |
| 11 | DA: | 1486-07-10 |
| 12 | YR: | 1486 |
| 13 | LI: | Kongsgården |
| 14 | LO: | Bergen |
| 15 | LF: | Ho |
| 16 | LC: | NO |
| 17 | LD: | |
| 18 | RR: | DN 01-0949 |
| 19 | ST: | Fullmaktsbrev |
| 20 | SC: | DN |
| 21 | LA: | NOR |
| 22 | FN: | |
| 23 | TI: | |
| 24 | SP: | |
| 25 | OC: | |
| 26 | OF: | |
| 27 | OL: | |

Fig. 3 Filled-in and edited registration form of an experimental version of PRONOMA

| | | |
|---|---|---|
| 28 | SN: | Hennes eiendom gårdene Dæli på Ringsaker og Bjørke og Skjennum i Nannestad har ved arv kommet til søsterdøtrene Margreta Hansdatter og Elin |
| 29 | RF: | |
| 30 | RM: | |
| 31 | RH: | |
| 32 | RW: | |
| 33 | RC: | |
| 34 | RS: | |
| 35 | RO: | Margreta Hansdatter/søsterdatter<br>Elin/søsterdatter |
| 36 | RL | |
| 37 | SL: | |
| 38 | SI: | |
| 39 | SS: | |

description of the data base proper[5], here it is enough to say that the terminal operators usually will work with prepared registration forms (Fig. 3), making use of the computer's data entry system.

Lastly, we come to question of choosing the appropriate software, namely the set of computer programs that generate the data bases and allow the user to search and retrieve the wanted information.

Many such data base software systems exist and are commercially available. While the cost of developing such systems may run into hundred thousands of US dollars, even as much as a million, they can be bought or leased for a few thousand dollars – rarely more than ten or twenty thousand. One should never attempt to create one's own unique set of software as long as commercial packages are cheaply available on the software market, where development cost can be shared by hundreds or thousands of users of the same product.

The only thing we should require of our data base software is that it should be capable of handling prosopographical information as furnished by the sources. A fairly large number of the existing data base systems were designed for handling bibliographic information. Expe-rience has shown that with suitable search options, these systems are well tailored for prosopographical purposes.

The fact that programmers so often are engaged to make new programs for historians (and probably other humanists as well), seems to indicate some defect in the communication between historians and computer experts: In fact, when one has occasion to compare what historians do to what is done by other groups of users, for instance natural scientists, one will often find important structural similarities. For this reason existing, commercial software will often serve as well for historians as they have done for these other groups, though perhaps with some minor adaptations. For obvious reasons the team chose the POLYDOC system of the Norwegian Centre for Informatics which, among other things, is used for processing the Centre's engineering data base, the Artikkel-Indeks/AID. With this software then, prospective users will be able to get information on persons mentioned by name or persons with specific characteristics in an operation which in all essential respects is similar to what the natural scientist does when searching for information on articles dealing with, for instance, the use of vanadium steel in oil well drilling. (A bit more information on the POLYDOC system is given in the appendix.)

## 5. A prosopographical data base for medieval Norway

### a. *Scope of the project*

Up to now the number of person instances (Belege) or persons mentioned in the Norwegian medieval sources has been unknown. Moreover, the prosopographical material in these sources have been made use of to no great extent, reflecting that the Norwegian medievalist has primarily been occupied with political and economic history.

There are indications, however, that the climate is changing and that there is a growing interest in the social history of Norway of that period. We think that prosopography is likely to play an important part in the Norwegian historical research of the future, either directly as an

5 On the intellectual preparation of data, consisting of abstracting, indexing and editing, see Harold BORKO and Charles L. BERNIER, Abstracting Concepts and Methods, New York (Academic Press) 1975, and ID., Indexing Concepts and Methods, New York (Academic Press) 1978, as well as KRIPPENDORF (note 8 below).

element of social history proper or as a tool for most other aspects of medieval history. Printed sources are, to be sure, much easier to deal with than unprinted ones, but without further processing they are not very well adapted to sociologically oriented methods of historical research, especially where there are large amounts of data or a review of all relevant sources is needed. Only a computerized, online data base containing all prosopographical information in the sources would answer the need for an efficient research tool. Our primary task was then to evaluate the different sources, estimate the number of records and formatting the records so as to capture all the relevant data in easily accessible and comprehensible form. Secondly, a trial data base was generated to perform as an illustrative example and test out concepts of formatting, registration and retrieval. Lastly, the total work demand and costs of the complete data base were estimated, together with an evaluation of its application in medieval historical research. Henceforth, this data base will be called by its acronym PRONOMA.

### b. *Sources and the number of records*

Something, surely, ought to be said about the size of the task for which this project first was conceived. To take the case of PRONOMA, a prosopographical data base going right from the beginnings to the year 1600 seemed likely to be of a reasonable size, and besides, by going beyond 1500, to serve as a bridge to the far better documented seventeenth century. The data for such a data base could most practically be taken from printed sources only: since most Norwegian sources previous to the year 1600 are printed, and the rest probably will follow within two or three generations (apart from what inevitably will crop up in little known archives abroad and at home), collecting data from unprinted sources would seem impractical. Finally it was decided to include all persons mentioned by name, regardless of social position, as well as those unnamed persons who somehow were singled out and might be identified from other data. Such cases would include »the archbishop of Nidaros« who certainly might be identified, and »Olav Galle's messenger to the archbishop« who possibly might be so.

Having decided on these guidelines the size of the task could also be determined: In the case of minor printed sources the number of instances was found by a simple count of names in the index. Thus an entry like »Sigurd Hallsteinsson 78, 157, 211–13«, would be counted as three references, which spot tests showed to be close enough. Apart from these minor sources, containing from less than a hundred names to ten thousand, three major ones stood out, the »Norske Regnskaber og Jordebøger« (Norwegian Accounts and Land Registers), the »Norske Lensrekneskapsbøker« (Norwegian Feudal Accounts) and the »Diplomatarium Norvegicum«. For these estimates were made, giving for the first 40 000, for the second 37 000 and for the third 133 000 instances. In the case of the latter the estimate was made from a sample of 200 charters which showed an average of 7,42 instances per charter and a standard deviation of 7,43 %. This means that there is a 95 percent probability that the true value lies between 114 000 and 153 000 (two standard deviations in each direction). For 100 diplomas a closer study was made showing that 86 percent of the instances were nominal and 14 percent were non-nominal (most of them of the »archbishop of Nidaros« kind and easy to identify). 7 percent of the nominal instances concerned women, and 14 percent of the non-nominal ones, which means that by including non-nominal instances the number of women is increased by one third.

All told, charters and letters seem likely to yield about 140 000 instances, legal sources 20 000, cadasters 6000, accounts and land registers 89 000, diaries etc. 6000 and varia 7000, which gives a total of 270 000 of which 37,5 percent concern the period before 1500. What has not yet been reviewed is the sagas, the epigraphical sources, the runic inscriptions, unprinted feudal accounts, unprinted letters and charters, mine accounts, a *lagtingsbok* (court protocol) from Bergen and some copybooks and collections of regests. This would at a guess give about 130 000 references, but the guess might easily be wrong by some twenty percent. (Table 1)

Tab. 1 *Numbers of records and different persons*

| Source | Records, 1000 | | Persons, 1000[6] | |
|---|---|---|---|---|
| | Total | Before 1500 | Total | Before 1500 |
| Letters & charters | 142 | 87 | – | – |
| Legal sources | 20 | 5 | – | – |
| Cadasters | 6 | 5 | – | – |
| Accounts & land registers | 89 | 0 | – | – |
| Diaries etc. | 6 | 0 | – | – |
| Varia | 7 | 4 | – | – |
| Sum sources reviewed | 270 | 101 | 150 | 55 |
| Other sources (estimate)[7] | 130 | 40 | 62 | 11 |
| All sources | 400 | 141 | 212 | 66 |

### c. *Data collection and formatting*

The main task in creating PRONOMA will be the collection of data. These are extracted directly from the source and transferred to a registration form. This registration form reflects the chosen format of the record or document, so the process of data collection and formatting is conflated into one.

The data for the data base can be collected either as excerpts or as more or less processed data. A number of considerations indicate that the second strategy is preferable: Assume for instance that the user is searching for *lagmenn* (judges). If the computer were to search for the word *lagmann* in unprocessed excerpts, the computer would come up with every single case where the word was mentioned, like sons of *lagmenn*, lower courts referring cases to *lagmenn* et cetera, while it would leave out all implicit mentions of *lagmenn*. To avoid this noise and distortion a logical and semantic processing similar to the content analysis[8] of the younger social sciences is necessary.

This kind of data collection presupposes a form, similar to the social science questionnaire. The data fields, which are the counterpart of the questionnaire's questions, must be chosen with regard to the needs of the users and the characteristics of the sources. Fortunately the maximum of fifty data fields allowed by the Polydoc system was more than enough, and the form eventually ended up with thirty-nine, covering those relevant person characteristics which can be found in the sources and be used as search criteria. Thus there are data fields for first name, patronymic (or metronymic as sometimes is found), family name or nickname and localisation name (like the German *von* or *zu* –). There are also fields for the most important social characteristics like social status, occupation and office, and for family relationships, that is for names of father, mother, husband or wife, children, siblings, other relations and in-laws. Finally there are fields for the source context in which this information is found, reference, the type of source as well as language and seal. Relevant information which does not come under

6 Spot test indicate between 1,5 and 2 references per person. The average is somewhat higher before 1500 than after, probably due to the preponderance of new types of sources in the 16th century.

7 A rough guess, based upon the characteristics of the sources and their volume, most likely a bit on the low side.

8 See Klaus KRIPPENDORF, Content Analysis, an Introduction to its Methodology, Beverly Hills/London (Sage Publications) 1980.

any of these headings can be included in the scope note, which is similar to the historian's regest, and there is also a data field for identification number to be used when developing the data base into a prosopography[9].

It should not be thought that all of these data fields will be used when filling in a form for an instance, the unit envisaged. What with the information the sources usually give, a typical case would be that of Lidvard Aslaksson in »Diplomatarium Norvegicum« vol. I no. 961 (where he in true Telemark fashion kills his man[10], and like a true Norwegian in a quarrel about language). For this reference we can fill in data fields 1, 2, 4, 5, 8–22 and 28–29, while there is no information on title (23), social status (24), occupation (25), office (26 and 27), kin apart from his father (30–36) and seal (37–39). The name Lidvard Aslaksson is, however, mentioned elsewhere in the »Diplomatarium«, and in due time we shall fill in other forms on that name. We can then in one case fill in field 28 about some criminal act for which the man needed *kongsdag* (X no. 267), in another case field 24 about his position as a landowner (VII no. 730) and in a third case field 24 and 34 about his landowning and brothers (IX no. 520). Seeing that all these references come from Upper Telemark, that the name Lidvard is not very common, and that all these references come from the period between 1488 and 1538, it is likely but not certain that they all refer to the same man. When developing the data base into a prosopography, field 3, identification number, will be used, and these four references will then be given the same number if we take them to refer to the same man. An example of the computer generated prosopography of Lidvard Aslaksson is shown in fig. 4 (p. 708–709).

### d. *Editing and data processing*

Filled-in forms will be handled by competent editors. Their main task will be the correction of errors; in particular great care must be taken to provide consistency between the various terms used[11]. Most of the terms in the forms will later, after data processing, turn up as index terms and these constitute the keyword or search terms which the user will apply in searching the PRONOMA.

The records of PRONOMA are made up of no less than 39 separate data fields, 38 of which are searchable. The attempt to make the index terms consistent will therefore involve most of the data, and editing shall very likely prove a time-consuming task. To some extent this can be avoided by applying codes for the most frequent and unique terms, like geographical names, administration districts, source references, offices etc. Checking and correcting can then be done by using a special interactive editing program. The editors may then be able to perform much of their job by the computer terminal.

Feeding the data to the computer is followed by further checking and proofreading. Finally, the data base itself is generated or, if it already exists, updated by adding new records. The number of records and the frequency of updating can be chosen at will, but often there are system generated constraints on the amount of data that may be updated in a single batch. The data processing aspects of PRONOMA is to a large extent related to the technicalities of the chosen data base software, viz. the POLYDOC system, and further elucidation of these matters is therefore relegated to the appendix.

9 More could be said about the question of record linkage (identification), which our team at present is working on. The problem is a complicated one, but it is obvious that the linking of records can be automated. This would involve the use of statistical methods, but must take as its point of departure a sound empirical basis, relating to the information content and the structure of the records.

10 This propensity to manslaughter was noted by bishop Øystein of Oslo in 1395 (»Norges gamle Love« 2 I no. 161) and the reverend Peder Claussøn Friis around 1600 (»Samlede Skrifter«) (Kristiania 1881) p. 300. It is also borne out by the unusually large number of documents from that period relating to manslaughter.

11 See Borko and Bernier (1975), note 5 above.

Fig. 4 The prosopography of Lidvard Aslaksson, landowner in Telemark ca. 1500. In reality the records would have the same identification number, not shown here.

⟩LIDVARD ASLAKSSON

13 4 LIDVARD ASLAKSSON

– 1 –
Normalized name: Lidvard Aslaksson
Original name: Liduord Aslakson
Sex: M
Date: 1489-05-06
Place of issue: Skien
Locality: Skien
County: Te
Country: NO
District: Skien
Reference: DN 01-0961
Source type: Letter of evidence
Scope note: Has killed Arne Torleivsson in quarrel about language
Serial number: 0016

– 2 –
Normalized name: Lidvard Aslaksson
Original name: Lyduard Aslaksyni
Sex: M
Date: 1523-xx-xx
Place of issue: Dalen? i Eidsborg? parish
Locality: Lårdal
County: Te
Country: NO
Dixtrict: Eidsborg parish
Reference: DN 09-0520
Source type: Deed of conveyance
Social position: Landowner
Scope note: Sells together with his brothers Hage and Håvard Aslakssons 7 »markebol« in the eastern farm of Dalen in Eidsborg parish to Asmund Gregardsson
Serial number: 0015
Siblings: Hage Aslaksson/bror; Håvard Aslaksson/bror

– 3 –
Normalized name: Lidvard Aslaksson
Original name: Lewordh Aslackson
Sex: M
Date: 1488-09-15
Place of issue: Natadal i Flatdal parish
Locality: Seljord
County: Te
Country: NO
Reference: DN 10-0267
Source type: Withdrawal of charges; Settlement
Scope note: Arrested by Harald Toresson and others although he was granted the king's peace
Serial number: 0014

– 4 –
Normalized name: Lidvard Aslaksson
Original name: Lidword Aslackson
Sex: M
Date: 1538-02-24
Place of issue: Vinje
Locality: Vinje
County: Te
Country: NO
Reference: DN 07-0730
Source type: Deed of conveyance
Social position: Landowner
Scope note: Sells 1 »markebol« in Midbø in Abø in Vinje parish for 9 »kyrlag« to Orm Øyvindsson
Serial number: 0013

### e. *Using* PRONOMA

In its final, projected form PRONOMA will probably contain about 300000 records or more as hitherto unprinted sources are published. These 300000 records constitute the prosopographies of about 165000 persons. Some of these, eventually a large proportion, will be furnished with identification numbers and the set of records with a given identification number will then be the prosopography of a particular person.

The search system of POLYDOC, together with the index terms and the field codes are the main tools at disposal of the user of PRONOMA. It is worth noting that some of the fields can be searched as full text, i. e. every word (character string) is taken as an index term, and others as consisting of subfields, each subfield, whether containing one or several data elements, being treated as a separate index term. The concept of subfield searching and full text searching may be applied to the same field. For instance, in the case of a person's full name being *Olav Gunnarsson Bagge*, he may be searched as *Olav Gunnarsson Bagge*, or as *Olav* or *Gunnarsson* or *Bagge*, each of which may be combined by logical operators. The first case exemplifies subfield searching with the field identical to its own subfield; the last three instances full text searching. Fig. 5 shows the fields and field codes of PRONOMA, together with search options and standard print formats.

Any keyword from any field may be combined with one or several others into search statements. This is done by the use of the logical operators AND, OR and NOT. How these operators function is best illustrated by Venn diagrams as shown in fig. 6 (p. 711).

The search statements are treated as queries by the search system, which in response generates search sets, corresponding to all records with the properties specified by the search statements. Search sets may be combined further by the logical operators. One important restriction applies; namely that only one particular operator is allowed in a single search statement. The combination of AND, OR or NOT must therefore be done by combining different search statements. (This restriction could be modified by using parentheses in the search statements, thus creating full Boolean expressions, but presently POLYDOC does not allow this.)

Numbers and dates may be searched in the same way, but it is also possible for them to be searched by the clauses GREATER THAN or GT, implying after; LESS THAN or LT, implying before and FROM...TO, implying the closed interval. Unknown numbers are assigned an *x*, which is always sorted as coming after a digit. (Likewise, anonyms are denoted by an A.) The search procedures do not discriminate between small and large letters. Examples concerning searching and sorting numbers are shown in table 2.

⟩FIELDS

| No | Fieldname | Code | Searchable | | P | P TRIAL | P FULL |
|---|---|---|---|---|---|---|---|
| 01 | Serial number | NR | | ★ | | | ★ |
| 02 | Updating code | UC | | ★ | | | |
| 03 | Identification number | ID | | ★ | ★ | ★ | ★ |
| 04 | First name | NF | | ★ | | | |
| 05 | Parent name | NP | | ★ | | | |
| 06 | By-name | NB | | ★ | | | |
| 07 | Localisation name | NL | ★ | ★ | | | ★ |
| 08 | Original name | NO | | ★ | ★ | ★ | ★ |
| 09 | Normalized name | NN | D | D | ★ | ★ | ★ |
| 10 | Sex | SE | | ★ | ★ | ★ | ★ |
| 11 | Date | DA | | ★ | ★ | ★ | ★ |
| 12 | Year | YR | | D | | | |
| 13 | Place of issue | LI | ★ | ★ | ★ | ★ | ★ |
| 14 | Locality | LO | | ★ | ★ | ★ | ★ |
| 15 | County | LF | | ★ | | | ★ |
| 16 | Country | LC | | ★ | | | ★ |
| 17 | District | LD | | ★ | | | ★ |
| 18 | Reference | BR | | ★ | ★ | | ★ |
| 19 | Source code | SC | | ★ | | | |
| 20 | Source type | ST | | ★ | | | ★ |
| 21 | Language | LA | | ★ | | | |
| 22 | Function | FN | | ★ | | | ★ |
| 23 | Title | TI | | ★ | ★ | | ★ |
| 24 | Social position | SP | | ★ | | | ★ |
| 25 | Occupation | OC | | ★ | | | ★ |
| 26 | Office | OF | | ★ | ★ | | ★ |
| 27 | Office location | OL | | ★ | ★ | | ★ |
| 28 | Scope note | SN | ★ | | ★ | | ★ |
| 29 | Father | RF | ★ | ★ | | ★ | |
| 30 | Mother | RM | ★ | ★ | | ★ | |
| 31 | Husband | RH | ★ | ★ | | ★ | |
| 32 | Wife | RW | ★ | ★ | | ★ | |
| 33 | Children | RC | ★ | ★ | | ★ | |
| 34 | Siblings | RS | ★ | ★ | | ★ | |
| 35 | Other relatives | RO | ★ | ★ | | ★ | |
| 36 | Relatives by marriage | RL | ★ | ★ | | ★ | |
| 37 | Seal | SL | | ★ | | ★ | ★ |
| 38 | Seal inscription | SI | | | | ★ | ★ |
| 39 | Seal source | SS | ★ | | ★ | | ★ |

Fig. 5 Field format of PRONOMA. The first column under Searchable denotes full text searchability

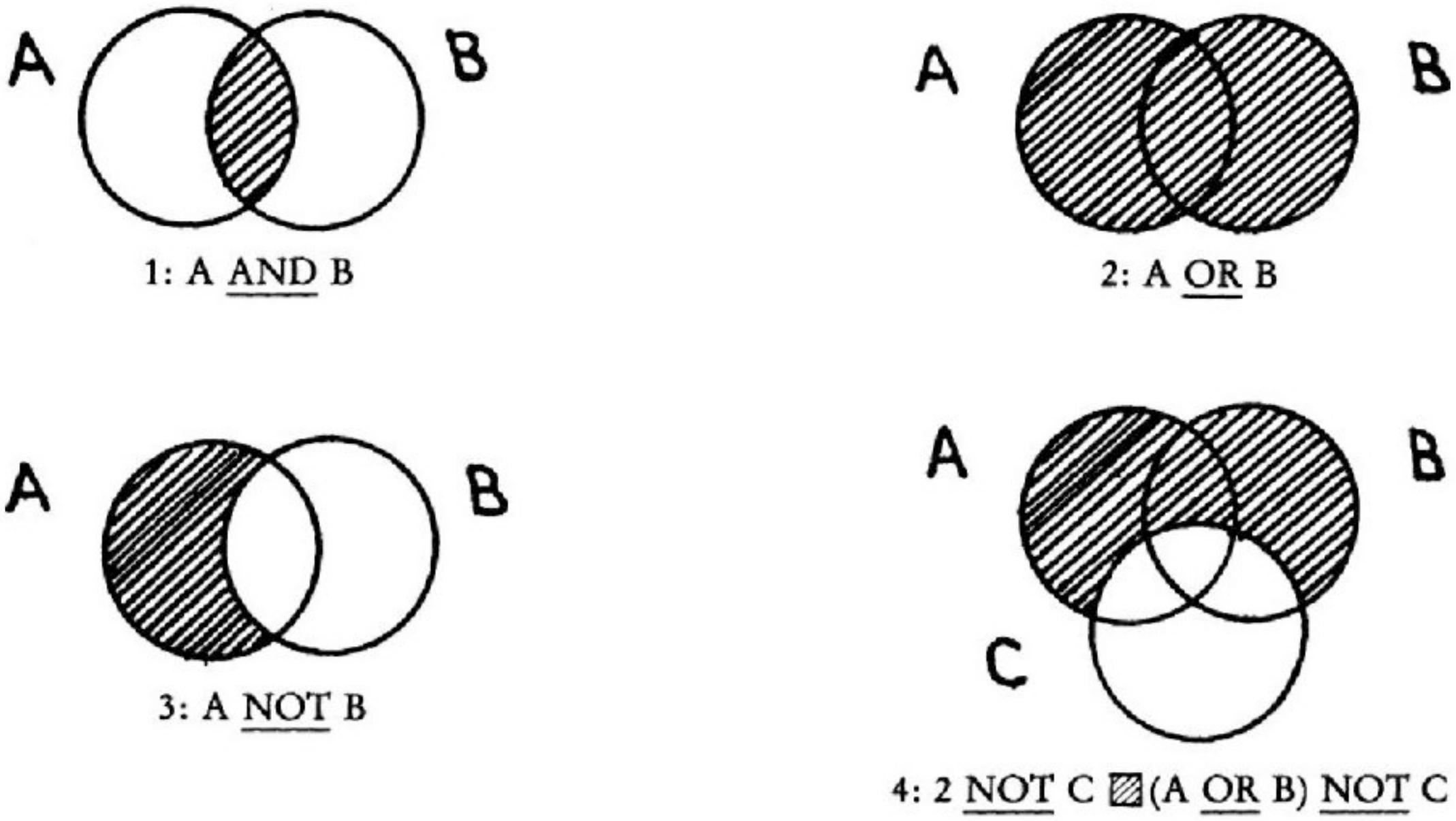


Fig. 6 Venn diagrams, exemplifying search sets, where A may be *lagrettemann* (office), B *Romerike* (jurisdiction)

Tab. 2 *Searching and sorting numbers/dates*

| Search statement | Meaning |
|---|---|
| 13: | All numbers beginning with 13 |
| 139: | All numbers beginning with 139 |
| GT 1390: | All numbers from and including 1391 |
| LT 1390: | All numbers up to and including 1389 |
| FROM 1390: TO 1399: | All numbers in the closed interval (1390, 1399) which also includes 1390 and 1399 |
| **Sorted data element** | **Sorted** |
| 1390-12-xx | Unknown day in December, sorted after 1390-12-31 |
| 1390-xx-xx | Unknown day in year 1390, sorted after 1390-12-xx |
| 139x | Unknown year in the period 1390–1399, sorted after 1399 |
| 13xx | Unknown year in the period 1300–1399, sorted after 139x |

The numbers in table 2 can be thought of as chronological data. The ability to search and sort on chronological criteria is of course of prime importance in any prosopographical data base.

The methods and possibilities of searching an online data base is a large, although not complicated subject. Let us rather see how this will work for the researcher looking for data, by taking the following imaginary case:

Having at his disposal thirty-nine data fields and three logical operators, AND, OR and NOT, the researcher has in theory $3^{39}$ possible search alternatives, a very large number indeed. In practice, however, only a few data fields will be relevant for a given research task, and even so experience shows that researchers often use these opportunities in a fairly primitive way. Let us assume now that a researcher sits in front of a terminal, wanting to study the *lagrettemenn*

(permanent jury members chosen from well off peasants) of Romerike in the fifteenth century. If so, he can find them through field 26, office: *lagrettemann* and field 27, jurisdiction: Romerike and field 12, year: 14:. This will give him all instances concerning men who are called *lagrettemenn*, having Romerike for jurisdiction and are mentioned between 1400 and 1499, inclusive. Now these *lagrettemenn* may obviously be mentioned on other occasions, and then without the title, for instance as litigants in civil cases. To find these other instances, those he already has, must be turned into a search file. At this point the shortcomings of a data base versus a prosopography can be seen, for with a prosopography the other instances of the persons could be found through field 3, identification number. With a data base the names (field 4–7) must be used, and if he wants to play safe, he will ask for all persons with similar names and then sort out those who could not possibly be identical. If he, on the other hand, wants to save trouble and considers he can afford a certain attrition of his population, he will ask for the names but set limits, for instance in field 12, year: 1350–1550, and field 15, county: Akershus (which includes Romerike), Østfold, Hedmark, Oppland, Buskerud, Vestfold, Oslo, Sarpsborg and Tønsberg (the neighbouring counties and cities) and then start sorting.

It is to be hoped that this example can bring home to the user two important lessons: In the first place the researcher's relationship with the online data base is an interactive one; the researcher will ask, be answered and ask again until he has what he wants. In the second place the shortcomings of a data base versus a prosopography can to a considerable extent be overcome given a moderate inventiveness on the part of the researcher.

```
⟩NBR RAGNHILD

Search word                                    Items

 1  OTTE ..... 1
 2  OTTE MATTIASSON ..... 1
 3  PER ..... 1
 4  PER ANDERSSON ORKNØYING ..... 1
 5  RAGNHILD ..... 1
 6  RAGNHILD MATTISDATTER ..... 1
 7  RASMUS ..... 1
 8  RAVALD ..... 1
 9  RAVALD NILSSON ..... 1
10  REIDAR ..... 1
11  RUSTUNG ..... 1
12  SIGURDSSON ..... 2
13  SKINNER ..... 2
14  SKREDDER ..... 2
15  SVEIN ..... 1

– UP n or DOWN n (max 20), pick search term or
give a command

⟩6

        14    1    RAGNHILD MATTISDATTER
```

Fig. 7 Online search for Ragnhild Mattisdatter, using the neighbour command

```
⟩P
- 1 -
NN  : Ragnhild Mattisdatter
NO  : Ragnildh Matisdotter
SE  : F
DA  : 1486-07-10
LI  : Kongsgården
LO  : Bergen
BR  : DN 01-0949
SN  : Her property, the farms Dæli on Ringsaker and Bjørke and Skjennum in Nannestad,
      has by inheritance become the property of her sister-daughters Margarete Hansdatter
      and Elin

⟩P TRIAL
- 1 -
NN  : Ragnhild Mattisdatter
NO  : Ragnildh Matisdotter
SE  : F
DA  : 1486-07-10
LI  : Kongsgården
LO  : Bergen
RO  : Margarete Hansdatter/sister; Elin/daughter

⟩P FULL
- 1 -
NN  : Ragnhild Mattisdatter
NO  : Ragnildh Matisdotter
SE  : F
DA  : 1486-07-10
LI  : Kongsgården
LO  : Bergen
LF  : Ho
LC  : NO
BR  : DN 01-0949
ST  : Power of attorney
SN  : Her property, the farms Dæli on Ringsaker and Bjørke and Skjennum in Nannestad,
      has by inheritance become the property of her sister-daughters Margarete Hansdatter
      and Elin
NR  : 0008
⟩STOP
- Come back soon
```

Fig. 8 The standard print formats, using the result of the search shown in fig. 7 as an example

The information the researcher finds, is usually first seen on the screen, but it can of course also be »copied«, as printouts, for computer typesetting or as data for statistical analyses. In the latter case a `mode program` will make it possible to group the information on each variable (for that is what the data fields really are) into values and then for instance do a multivariate analysis.

Online prints may be »written« on the data terminal's screen, one record at a time. If the user also has a printer at his disposal a hardcopy on paper may be obtained. The online prints are formatted by the print format chosen, either from 3 standard formats or according to the user's wishes simply by selecting the wanted data fields. Examples of a search and various print formats are shown in fig. 7 and 8.

### f. *Work demand and costs*

PRONOMA is believed eventually to contain approximately 300000 or more records. The whole project could be finished in about ten years' time. If new sources are published, they will of course be added to the data base. Manpower and budget requirements are shown in table 3.

Tab. 3

| Type of work | Manyears | Cost. NOK 1000 at '84 prices |
|---|---|---|
| Abstracting & indexing | 30 | 5000 |
| Data entry | 10 | 1700 |
| Proof-reading | 10 | 1700 |
| Management & editing | 15 | 4200 |
| System development | 5 | 1400 |
| Sum | 70 | 12850 |
| Data processing & hardware | | 3000 |
| Total costs | | 15850 |

An extra NOK 1,15 million is allotted for special projects, bringing the grand total to NOK 17 million or about 2 million US dollars. This is less than 0,004 % of Norway's yearly gross national product at current prices.

### g. *Applications*

It is easy to see what a valuable tool PRONOMA would be for almost all branches of Norwegian historical research, and not just for professional historians, but also for the amateurs who are very active within the flourishing field of local history in Norway.

The data base will be stored permanently on a central host computer, preferably in connection with Norwegian National Archives. It may be accessed over public data networks or telephone lines, both in Norway or from abroad. Data terminals for searching online are now available at libraries, archives and educational establishments all over the country. Home users may employ the telephone as their online connection, by hooking the terminal to their receiver set. Local users who have access to a computer or possess their own personal computers, may download parts of PRONOMA for their own research, thus creating a local data base.

The use of the data base should be free, although one of course must pay a communication charge on using the network. That is why the downloading and creation of special data bases may sound as an attractive proposition, above all in local research environments.

We foresee immediate and far-reaching effects from online searching of the PRONOMA. In particular PRONOMA will be of great value for studies of Norwegian medieval social history, name research, genealogy and statistical and demographical investigations of early Norwegian society.

An easily attainable goal would be that old pipe-dream of a renowned Norwegian historian, namely a statskalender (Government yearbook, Staatshandbuch) of Norway in the Middle Ages.

Eventually, the possibilities of PRONOMA will only be limited by a lack of fantasy of the historian's part.

## 6. Swedish university students in the Middle Ages

In the matter of prosopography Sweden may have a more favourable starting point than the other Nordic countries in the institute »Sveriges medeltida personnamn« (Sweden's medieval personal names). In the institute's collections, which consist of five separate series[12], are found approximately 775 000 personal names of the period before ca 1570. These series which are excerpted from manuscript sources as well as printed ones[13] would in the least be a valuable help in prosopographical research.

One such project, undertaken by one of the authors (Per-Axel Wiktorsson) by means of the software described above, concerns Swedish university students in the Middle Ages.

In the study of a particular group as this, as compared to a nationwide prosopographical data base, some problems may become more salient and new problems may arise. The main problems in this case concern identification, the information needed for the purposes of the study, and the compatibility of these data with those of a nationwide prosopography.

As regards identifying persons mentioned in foreign university matricles, it is obviously next to impossible to say from which country a *Nicolaus Benedicti* comes, unless further information is available. With, however, the help of a nationwide online prosopographical data base, or, in the absence of that, data bases such as those of »Sveriges medeltida personnamn«, some guidance may be had by looking for a Nils Bengtsson who has an academic title such as *doctor* or *magister* (though such titles are seldom used in Swedish medieval sources) or has an office for which an academic past must have been desirable, such as bishop, dean, archdeacon, precentor and some others, as well as secretaries to kings or bishops. And, of course, even if a man has an obviously Nordic name, such as *Bero Haquini*, it would be a great deal easier to establish his nationality if other Nordic data bases, similar to »Sveriges medeltida personnamn« were available.

The standard form for collecting data will also have to be adapted for this particular purpose, so as to make the most out of the data needed. Considering the purpose and the characteristics of the sources used, data fields will be needed for: former academic title, present university,

12 The collections are: The main collection, containing personal names from about 1100 to 1520, the collections from the *Kammararkiv* with personal names from manuscript sources of the years 1530 to 1571, the collection excerpted from sixteenth-century printed sources, the person register taken from publications about medieval persons and the etymological register also taken from scholarly publications.

13 Excerpts have been taken from manuscript sources: From the Swedish National Archives' series of parchment and paper letters, from Swedish letters in the Danish National Archives, from medieval letters in the Uppsala University Library, the Cathedral Libraries of Linköping, Skara and Strängnäs, the archives of Sävstaholm, Skokloster and Ericsberg, the various *landsarkiv* and in private archives, from medieval cadasters, accounts and copybooks, as well as from cadasters and tithe accounts in the Swedish National Archives (formerly in the *Kammararkiv*) and the Finnish National Archives. Excerpts have also been taken from printed sources: From the »Svenskt Diplomatarium«, from the collections of the »Fornskriftsällskapet« such as the *tänkeböcker* of the towns of Arboga and Kalmar, the Erik chronicle and other medieval rhymed chronicles, and Birgitta's and Katarina's canonisation processes, from »Finlands medeltidsurkunder«, from the copybook of the Åbo (Turku) cathedral, from the Stockholm *stadsböcker* and from »Jämtland och Härjedalens diplomatarium«.

present faculty, function (immatriculated/examined), other function, mentioned together with, formerly studied at, later studied at, later office and identified with, as well as fields for context. There are, however, no problems about this, just a matter of changing the form.

Although some of the data fields used in a study of this kind will differ from those of a nationwide prosopographical data base, there are no great difficulties in either using the two together or using both as separate units. In the first case data fields with the same content, for instance social status, may easily be numbered according to the nationwide standard, and other information, such as in the particular data fields mentioned above, may be summed up in a scope note. One may also, or instead, preserve the data collected for a study like this or for instance for local history as a separate data base which nevertheless can be searched together with the nationwide one, in one operation. To do so conversion software or a »save«-command is needed. In this way the work done on a particular study, such as medieval Swedish university students, does not have to be done again for the nationwide data base, as long as all information needed for its data fields is collected. Thus, the system is as flexible as one could wish for.

## 7. Future outlook

As far as we know an online prosopographical data base like PRONOMA is the very first of its kind conceived anywhere in the world.

The concept of data bases is now coming of age and its application to prosopography seems a natural step. However, transforming a data base from an information collection into a working research tool can best be done by making it not only computerized, but online and interactively searchable as well.

In our opinion historical research will benefit from the use of electronic data processing as much as any other scientific activity. The need for a sounder empirical basis in historical research has long made itself felt. The lack of continuous time series of data in history may be successfully countered by the intelligent application of statistical methods. In this context we foresee a widespread use of online data bases as an indispensable research tool, besides a generally greater reliance on computers and edp as well.

Not only will this make historical research more effective, it should also stimulate a methodological improvement by putting the historian's theories on a sounder basis, allowing for better models and more severe testing procedures.

Nor should the historian, either out of modesty or fear, abstain from the possibilities offered by new research tools of his trade. The importance of history as a branch of knowledge cannot be overestimated and no attempt that brings it little bit closer to the unattainable truth should be rejected.

## Appendix: The POLYDOC system

### a. *Description*

POLYDOC is a computer based documentation system developed at the Norwegian Centre for Informatics in Oslo. The system makes the indexing and retrieval of information of different types more effective. Originally, POLYDOC was developed with bibliographic data bases in mind, but has since successfully been applied to various information types, both in Scandinavian and other European countries. It has shown itself as a wellproven and flexible system, giving its users wide opportunities in creating and handling data bases. Moreover, it has shown itself to be highly functional in contexts ranging from libraries over business and engineering to archaeology. Compared to other systems offering the same capabilities, POLYDOC is at the low

end of the price range whether bought or leased. The user's total costs are of course also dependent on the computer system on which POLYDOC is run.

Presently POLYDOC can be run on all widely used brands of minicomputers and mainframes, like the Norwegian NORD, VAX, IBM, PRIME, UNIVAC etc.

The POLYDOC system consists of a set of computer programs, each devoted to specific information processing tasks. They are written in different versions of FORTRAN, depending on the computer applied. The program handles

- data entry
- generating databases
- updating of databases
- printout of indexes
- online searching

All these may be used in producing prosopographical data bases.

The data base proper consists of various data files, of which the most important are the DFILE or document file, consisting of the records that the data base contains and of the EFILE, which contains all the index terms or keywords used for searching the data base. The index terms are sorted in alphabetical or numerical sequences according to their field designations and codes. Each different index term is entered only once, but may be found in several records. This is done by the RFILE, which points from the index term in the EFILE to the AFILE, which consists of the addresses of all documents containing that particular index term. An overview of the file structure of POLYDOC is given in fig. 9 and fig. 10.

Searching the data base involves finding the wanted index term in the EFILE. The relevant records are then found in the DFILE and can be printed on paper or shown on a screen. Several temporary files are created to assist the searching. They are usually erased on terminating the search.

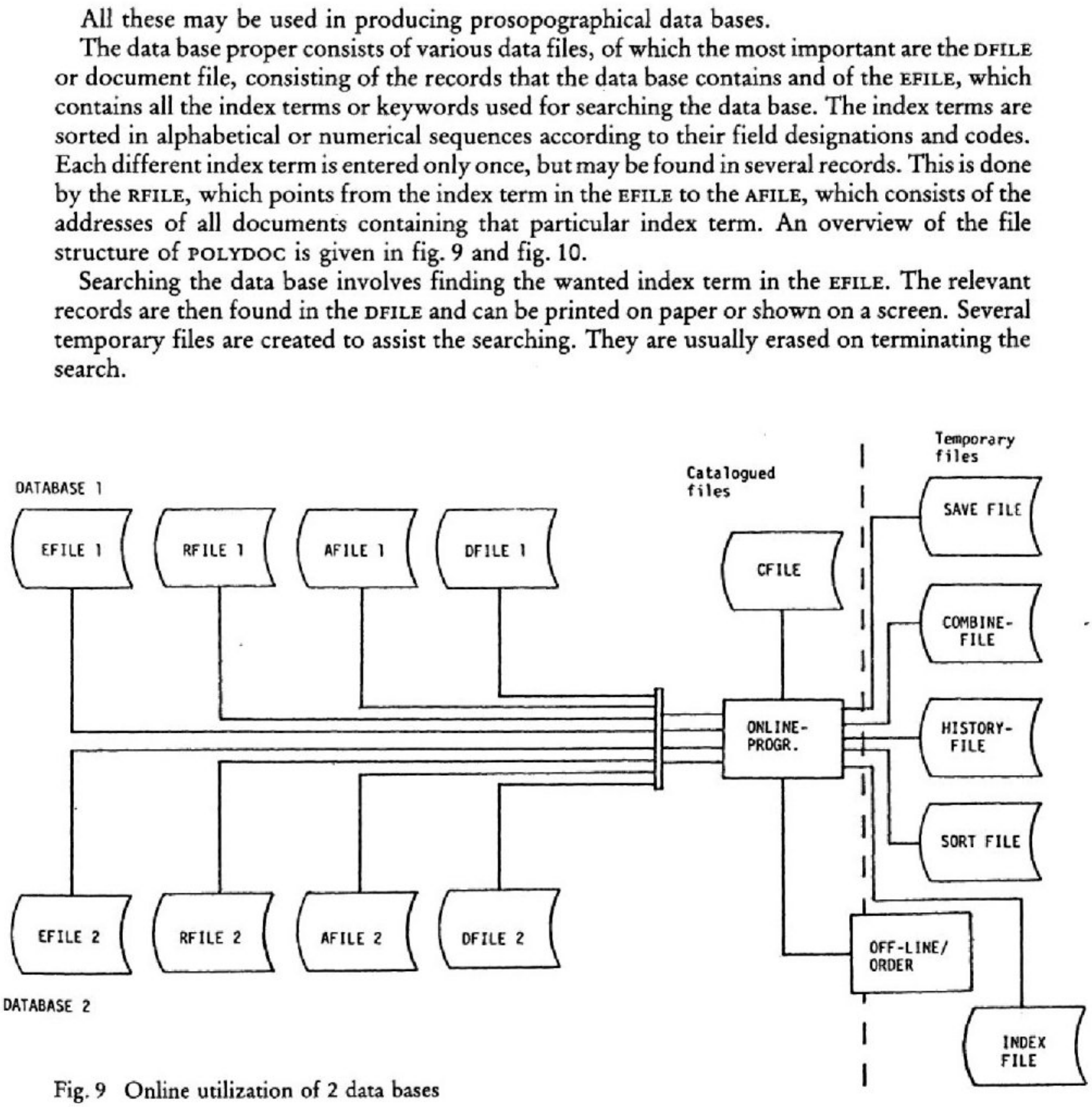


Fig. 9 Online utilization of 2 data bases

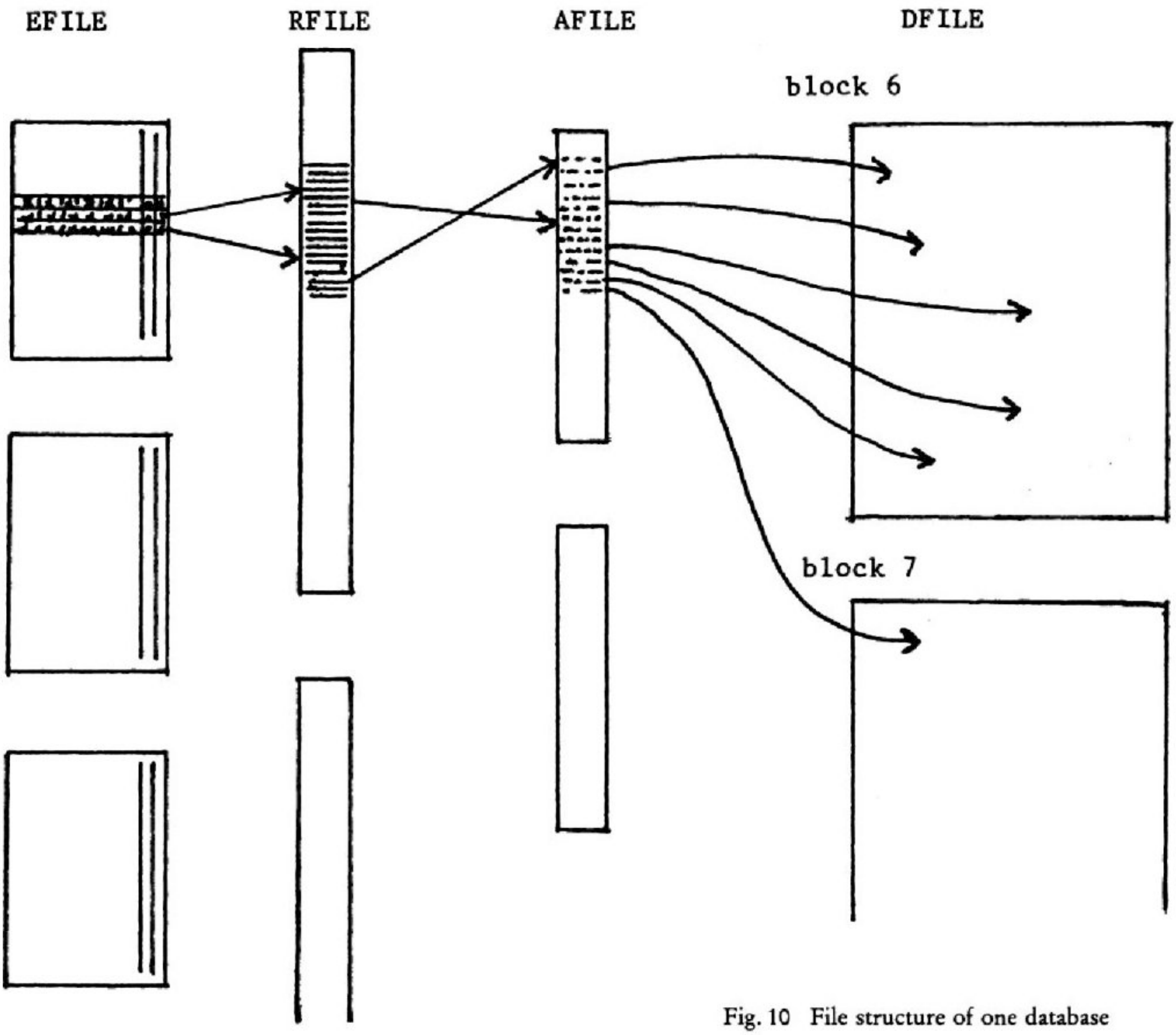


Fig. 10 File structure of one database

POLYDOC is able to handle several online data bases at the same time. Any data base may consist of up to 50 data fields and 4000 characters of text per record. These constraints may be modified if one should wish so. The number of documents in each data base depends in practice on the computer's storage facilities and the number of possible logical addresses which the computer offers. The last factor is usually related to the word length of the computer. For instance, a 16-bit computer may address $2^{16}$ documents or ca 65000, a 32-bit computer $2^{32}$ or more than $4 \times 10^9$ (four billion) documents. This should satisfy most users needs. The speed with which the computer performs a search, depends on many factors; the most important are the number of simultaneous users and the processing capability of the computer's central unit (CPU). A typical search usually does not take more than a few seconds of the computer's time, as measured in CPU time.

Lastly, it should be noted, while searching takes place online and interactively, updating and the printing of indexes take place as a batch job, often done overnight. A large updating may take several hours on a typical minicomputer (fig. 11).

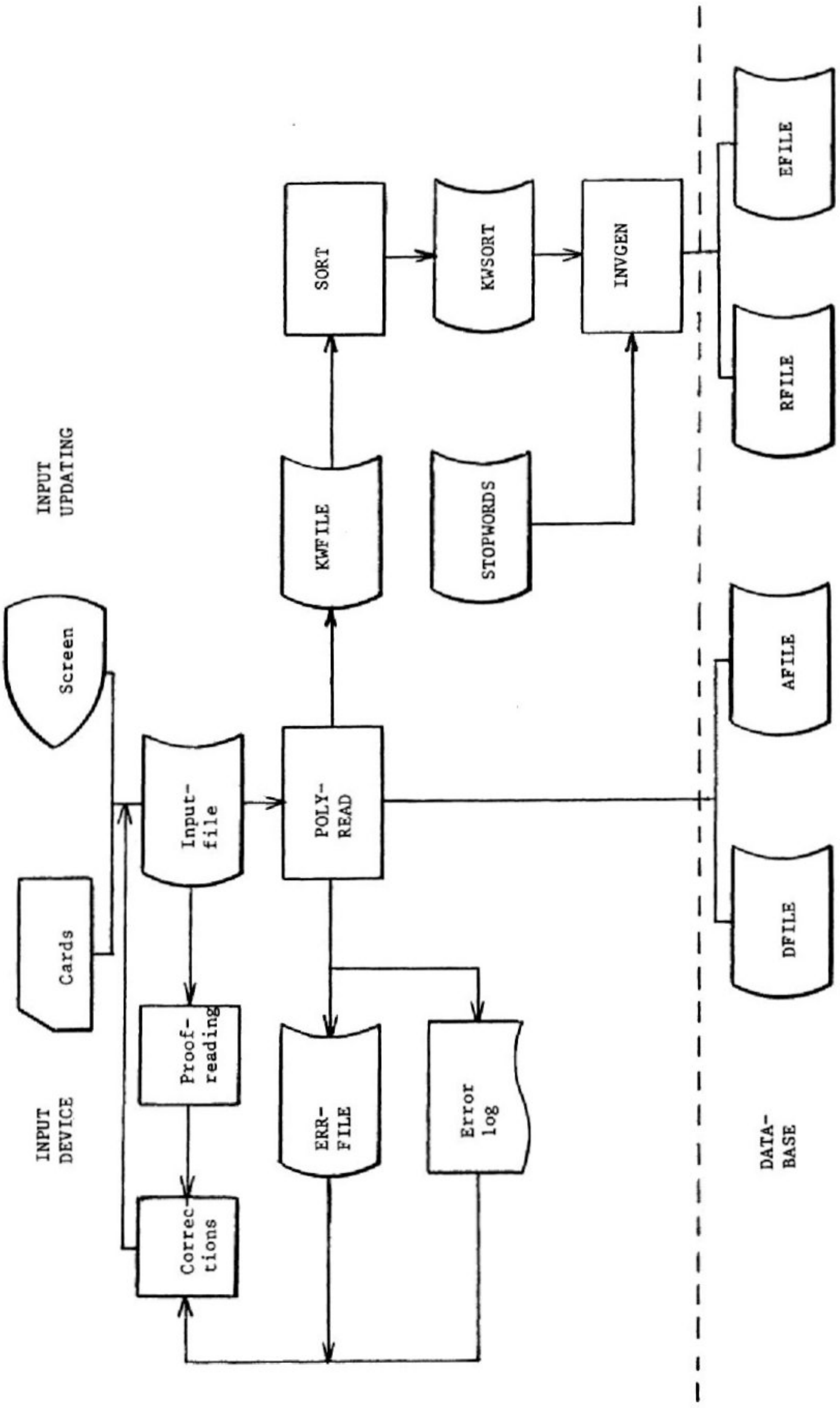


Fig. 11 Shows updating, using Polydoc

Also, it must be remarked that the POLYDOC data base system may be operated by anyone who has some knowledge of data processing and who has learned a bit about the system, typically by taking a 2-day course. This means that POLYDOC today is not only operated by computer scientists and engineers, but also by librarians and social scientists, who have come to see the value of online data bases in their fields of activity. »In the same way, by assimilating computers into their research tradition, historians may reap the same benefits as other scholars and scientists have.«

### b. *Printing of indexes*

In addition to ordinary online or offline printouts of the records in the data base, the operator or user may also want indexes of various types. The POLYDOC system offers several possibilities for the printing of indexes. Indexes are usually generated in batch oriented jobs, meaning that they must be ordered at and produced by the data base host, that is, by the computer on which the data base is stored.

An index is a sorted list of specified data elements, taken from one or several data fields. If singular data elements are used as keywords, these are usually made to appear alphabetically. If several data elements from the same record are used as keywords, they can be rotated or permuted, so that each one appears alphabetically in its turn. It is of course also possible to print indexes sorted numerically.

Moreover it is possible to use several sorting concepts. For instance, in an index of first names these will be sorted and printed in alphabetical order. Identical names can then be sorted according to dates or geographical areas. An example of such an index is given in fig. 12. This is an index of first names or parent names, intended for name research. It is sorted alphabetically with either the first name or the parent name as a keyword and first sorting concept. Additional information on the instances of each name is added to the keywords. This information has been taken from the records and formatted according to the user's wishes. It has been sorted further by using the dates as a second sorting concept. The effect is an index printout containing all first and parent names in alphabetical order, with the instances of each name sorted chronologically. Furthermore the original spelling is given and the index also furnishes information about sources, titles, offices and geographical locations. In addition frequency lists may be generated (fig. 13, p. 722).

As mentioned it is possible to use the POLYDOC system for producing different types of indexes. These may be block indexes, rotation indexes and line indexes, as well as KWAC-indexes (Keyword And Context), KWOC-indexes (Keyword Out of Context) and KWIC-indexes (Keyword in Context).

The last three are especially suited to be accompanied by additional information, taken from the record text. This information may again be formatted according to the user's wishes, by using the field codes of the data base. The formatting allows for different graphical layouts, the adding of signs and standard explanatory texts.

### c. *Additional software*

For prosopographical purposes the user may want special software. The use of the data base as an input for printed editions, catalogues etc. may be done by applying standard computer typesetting software. The output is then transferred to magnetic tapes or disks and used directly as input in a photocomposer.

For online searching, it will be possible for the user to format and sort the printouts according to the need. As many as five different sorting concepts will be allowed. Such printouts must commonly be made in an offline mode. The user may enter printing, formatting, sorting and ordering instructions online from the terminal. The data base host then produces the offline print and mails it to the user.

ADALIS

Adaliz
1403-06-15 (DN 01-0588) Fru (Nøtterøy, Vf)

Adaliiz Ellinx dotter
1437-03-31 (DN 01-0760) Fru (Vågå, Op)

ALV

Alff Knwtzsyni
1463-09-29 (DN 01-0864) Herr (Bergen, Ho)

Alff Torgardsson
1463-09-29 (DN 01-0864) Herr, Electus – Prost (Bergen, Ho)

Alff Knwtson
1516-03-14 (DN 01-1048) Herr (Dovre, Op)

ALVSSON

Knwt Alfson
1516-03-14 (DN 01-1048) Herr (Dovre, Op)

AMUND

Amund Niclisson, Amundh Niclisson
1437-03-31 (DN 01-0760) (Vågå, Op)

A
1516-03-14 (DN 01-1048) (Dovre, Op)

AMUNDSSON

Guttorm Amundzsyne
1481-06-16 (DN 01-0930) Beskjedelig mann (Bergen?, Ho?)

Erick A⟨...⟩dsson
1486-07-10 (DN 01-0949) Lagmann (Bergen, Ho)

Nils Amundsson
1516-03-14 (DN 01-1048) (Dovre, Op)

ANDERSSON

Per Anderson Ortnisk
1570-01-01 (APB 156) (Bergen, Ho)

ARNBJØRG

Anbiorghe Tor⟨...⟩dotter
1437-03-31 (DN 01-0760) (Vågå, Op)

ARNESSON

Eskyld Arnesen
1520-xx-xx (NRJ 2-145) (Trondheim, ST)

ASBJØRN

Esbiorn
i Hallbjørtorp
1542-xx-xx (NRJ 5-121) (Inlands sødre, Bl)

ASKJELL

Eskyld Arnesen
1520-xx-xx (NRJ 2-145) (Trondheim, ST)

ASLAKSSON

Lewordh Aslackson
1488-09-15 (DN 10-1267) (Seljord, Te)

Liduord Aslakson
1489-05-06 (DN 01-0961) (Skien, Te)

Lyduard Aslaksyni
1523-xx-xx (DN 09-0520) (Lårdal, Te)

Lidword Aslackson
1538-02-24 (DN 07-0730) (Vinje, Te)

BENKTSSON

Ketil Beinsson
1481-06-16 (DN 01-0930) Bror, Generalkonfessor (Bergen?, Ho?)

Fig. 12 KWAC index of first and parent names

| | | | | | |
|---|---|---|---|---|---|
| A | 2 | GAUTE | 3 | MAGNHILD | 3 |
| ADALIS | 2 | GJEBLE | 1 | MAGNUS | 1 |
| ALV | 3 | GRIM | 1 | MARGARETE | 2 |
| AMUND | 2 | GUTTORM | 1 | MARINE | 2 |
| ARNBJØRG | 1 | HALLVARD | 1 | MATTIAS | 1 |
| ASBJØRN | 1 | HANS | 5 | NILS | 4 |
| ASKJELL | 1 | HARTVIG | 1 | OLAV | 8 |
| BERDOR | 1 | HOLTE | 1 | OTTE | 1 |
| BERTIL | 1 | HÅKON | 2 | PER | 1 |
| BIRGER | 1 | INGEBJØRG | 1 | RAGNHILD | 1 |
| BOTHILD | 1 | JOHAN | 1 | RASMUS | 1 |
| BOTOLV | 1 | JON | 2 | RAVALD | 1 |
| BÅRD | 1 | JØRGEN | 1 | REIDAR | 1 |
| DAVID | 1 | KJETIL | 1 | SVEIN | 1 |
| EINDRID | 1 | KNUT | 3 | TORALD | 1 |
| EIRIK | 2 | KRISTINA | 2 | TORBJØRN | 1 |
| ELIN | 1 | LASSE | 1 | TORSTEIN | 2 |
| ELSEBE | 1 | LIDVARD | 4 | TROND | 1 |
| ERIK | 2 | LODIN | 1 | ØYSTEIN | 1 |

---

| | | | | | |
|---|---|---|---|---|---|
| ALVSSON | 1 | GAUTESSON | 1 | NIKLISSON | 1 |
| AMUNDSSON | 3 | GUNNARSSON | 1 | NILSSON | 1 |
| ANDERSSON | 1 | HALLVARDSSON | 1 | ODDSDATTER | 2 |
| ARNESSON | 1 | HANSDATTER | 1 | OLAVSDATTER | 1 |
| ASLAKSSON | 4 | HENRIKSSON | 1 | OLAVSSON | 4 |
| BENKTSSON | 1 | IVARSSON | 2 | SIGURDSSON | 2 |
| BJØRNSSON | 2 | JENSSON | 1 | TOR⟨...⟩DATTER | 1 |
| EGGERTSDATTER | 1 | JONSSON | 2 | TORALDSDATTER | 1 |
| EINDRIDSSON | 6 | KNUTSSON | 2 | TORGARDSSON | 1 |
| EIRIKSSON | 4 | MATTIASSON | 1 | TORKJELLSSON | 1 |
| ERLINGSDATTER | 1 | MATTISDATTER | 1 | TORSTEINSDATTER | 1 |
| | | | | VIGFASTSSON | 1 |

Fig. 13 Frequency lists of first and parent names

Registration of words in Icelandic, Anglo-Saxon and Old Norse requires the use of diacritical and special signs. They are accounted for in the POLYDOC data entry system by the application of a properly placed accent. Alphabetization takes place according to modern Icelandic, while the correct character may only be inserted by the use of computer typesetting.

Creation of search statements may be performed automatically in online searching. This can be very helpful in special cases. Take for instance that we want to study, as mentioned in section 4c., the *lagrettemenn* of Romerike of the 15th century. These appear under several different names, may be as many as 100. It certainly will be a cumbersome process to create new search statements for each one, as we must if we want to study the social background of the *lagrettemenn*. They will of course appear in several records not referring to their position as such. These other records may therefore be searched by using a special instruction that automatically creates a new search set, consisting of all names employed by the known *lagrettemenn*.