

# *Cahiers* **GUT** *enberg*

∞ TAILORED DATABASE PUBLISHING WITH  
T<sub>E</sub>X

© Steen LARSEN, Arne Flemming JENSEN

*Cahiers GUTenberg*, n° 10-11 (1991), p. 117-134.

<[http://cahiers.gutenberg.eu.org/fitem?id=CG\\_1991\\_\\_10-11\\_117\\_0](http://cahiers.gutenberg.eu.org/fitem?id=CG_1991__10-11_117_0)>

© Association GUTenberg, 1991, tous droits réservés.

L'accès aux articles des *Cahiers GUTenberg*

(<http://cahiers.gutenberg.eu.org/>),

implique l'accord avec les conditions générales

d'utilisation (<http://cahiers.gutenberg.eu.org/legal.html>).

Toute utilisation commerciale ou impression systématique

est constitutive d'une infraction pénale. Toute copie ou impression

de ce fichier doit contenir la présente mention de copyright.



# Tailored database publishing with T<sub>E</sub>X

---

Steen LARSEN and Arne Flemming JENSEN

*UNI-C (Danish Computing Centre for Research and Education)  
Olof Palmes Allé 38, DK-8200 Aarhus N, Denmark*

**Abstract.** T<sub>E</sub>X is well suited to produce inventories such as bibliographies or dictionaries. Such publications are characterized by a large number of entries, a high uniformity of structure, typographical variation, and high demands to line and page breaking. Furthermore, sorting of entries and compiling of indexes will often be necessary. In the necessary T<sub>E</sub>X input files there will be a large percentage of control sequences.

Producing inventories based on text editors present numerous difficulties as regards for example T<sub>E</sub>X syntax control, data validation, and sorting. Producing them via a standard database system gives better data control, but forces the user to accept the limitations of the system's user interface.

This paper presents an approach chosen when establishing a tailored T<sub>E</sub>X-based database publishing system for the bibliography Nordic Archaeological Abstracts. The solution was implemented by combining three different systems: an interface management system, a database management system, and T<sub>E</sub>X. The system is described and compared to the previous editor-based production, and future possibilities are briefly sketched.

**Résumé.** T<sub>E</sub>X est bien adapté pour produire des listes, telles qu'une bibliographie ou un dictionnaire. De telles publications ont pour caractéristiques un grand nombre d'entrées, une structure uniforme, et des problèmes de mise en page liés aux coupures de lignes et de pages. De plus, les entrées sont généralement triées, et un index est nécessaire. Un tel texte T<sub>E</sub>X contient beaucoup de commandes.

Produire de tels inventaires avec seulement un traitement de textes est difficile, en raison de la syntaxe de l'éditeur, de l'impossibilité de contrôler la validité des entrées, et de l'absence de tri. Une telle production est simplifiée par l'utilisation d'outils de gestion de base de données, mais la qualité de résultat est moindre.

Nous présentons ici une approche consistant à habiller une base de données avec T<sub>E</sub>X. Une application réelle a été réalisée sur une base de données bibliographique (Nordic Archaeological Abstracts). Notre approche combine trois systèmes : une interface, une gestion de bases de données, et T<sub>E</sub>X. Nous décrivons notre système et le comparons à une solution fondée sur un éditeur. Les futures évolutions sont évoquées.

**Key words:** database publishing, interface management, bibliography.

## 1. Background

### 1.1. Nordic Archaeological Abstracts

The bibliography Nordic Archaeological Abstracts (NAA) is published once a year. It covers the Nordic archaeological literature mainly from the preceding year.

A volume of NAA extends to about 350 pages. It contains

- a few short introductory text sections;
- 800-1000 abstracts of selected Nordic archaeological articles;
- 5 indexes (authors, sites, subjects, periodicals and collective works).

The abstracts are short and concise summaries which, together with the indexes, provide an easy access to the published articles.

NAA is produced by a large editorial staff from all over the Nordic countries. After a long period of Danish editorship, the general editors are now based in Norway.

The bibliography is financially supported by the humanistic research councils in Denmark, Finland, Norway and Sweden.

### 1.2. UNI-C

UNI-C (Danish Computing Centre for Research and Education) is a public company affiliated to the Danish Ministry of Education. It offers its services to the Danish academic world as well as to private companies. UNI-C provides central computing on vector processors and mainframes, hosts a national network with access to international networks, and develops special computer systems in cooperation with other institutions and companies. It also offers consulting and courses in a broad range of applications.

As regards  $\text{\TeX}$ , UNI-C provides various  $\text{\TeX}$  implementations on mainframes and personal computers,  $\text{\TeX}$  and  $\text{\LaTeX}$  courses, and  $\text{\TeX}$  consulting. UNI-C offers to take care of the entire production of a publication or to help the user build his or her own production facilities based on  $\text{\TeX}$ .

### 1.3. The NAASYS project

For the past four years UNI-C has been formatting and typesetting NAA on the basis of editorial data produced with a text editor on an IBM PC.

During this process it became evident that a database publishing system could give the editors a lot more control of their data in the editing process, remove the nuisance of coding, and save a lot of time.

A project for establishing such a system to run on a 386 IBM PS was sketched by UNI-C, and the necessary grant was given by the Nordic humanistic research councils in 1990. The project, called NAASYS, started in the autumn of 1990, and an almost final version was delivered to NAA in the late spring of 1991.

## 2. Data structure

### 2.1. Abstracts and references

A simple NAA abstract looks somewhat like this in the final publication:

---

7F 7H Sw NAA 1989/17  
**Neues zu Brakteaten und Anhängern in Schweden.** (News about  
bracteates and pendants from Sweden)  
  
Lamm, Jan Peder & Morten Axboe. *Frühmittelalterliche Studien* 23, 1989, pp  
453-477. 7 figs, pls XXI-XXXII, refs. Ger.  
  
Presentation of golden ring-cross pendants (cf NAA 1987/299) and three  
hitherto unpublished gold bracteates. A small number of bracteates without  
rim and loop, found in Gotlandic graves, were apparently specially made to  
be used as Charon's obols. (Au)

---

Figure 1. A simple NAA abstract

Abstracts are sorted according to

1. key ("7F" in the example, reflecting periods 1-11 and topics A-L);
2. author ("Lamm, Jan Peder" in the example);
3. title ("Neues zu Brakteaten ..." in the example).

The abstract number ("17" in the example) reflects the position of the abstract in the sorting order.

In the text paragraphs of the abstracts, text strings such as extra authors ("Morten Axboe" in the example) or periodicals ("*Frühmittelalterliche Studien*" in the example) function as ordinary text in the paragraphs, but must also be compiled to the relevant indexes. Such text strings in NAA are called *text references*. The *ordinary references* in NAA are text strings not represented in the text but only compiled to the indexes.

## 2.2. Indexes

The five indexes are more or less complicated. The following figure shows small excerpts from the simplest index (authors) and the most complex (subjects with up to five levels):

---

AUTHOR INDEX	SUBJECT INDEX
Avelle-Widhalm, Gloria 482.	ORNAMENTS Amber (Neo) 71,
Axboe, Morten 17, 341e, 344, 346	antler, ivory etc. (GerIA-Med) 862,
.	bells (Med-PM:Dan) 549e
.	.
Lait, Alexander 94	.
Lamb, Hubert H 109, 110, 144, 157,	- Brooches
166, 167, 188, 808	- BA: Gold (Norw) 245.
Lamm, Jan Peder 3, 17, 469, 803	- CeltIA: Ball brooches (Dan)
.	274, Kostrzewski M (Dan)
.	307, La Tène 269.
Petersen, Erik Brinch 108k, 111,	- Pendants ring-cross & gold
129, 137, 145, 153, 177, 798	bracteates (GerIA:Ger) 17, 23,
Petersen, Peter 17, 292, 341c	(GerIA:Pol) 25.
.	.

---

Figure 2. Reference representation in author index and subject index

An index entry points back to the relevant abstract by means of the abstract number ("17" in the example). In the author index you can see entries created as text references ("Axboe, Morten") and as ordinary references ("Petersen, Peter"). In the subject index only entries created as ordinary references occur.

### 3. Editor-based production

#### 3.1. Input

In the previous editor-based production process our typical NAA abstract would be typed somewhat like this in the *TEX* input file:

```
\key{7F} {7H} {Sw}

\title{Neues zu Brakteaten und Anh\angern in Schweden.} (News about
bracteates and pendants from Sweden)

\author{Lamm, Jan Peder} & \coauthor{Morten Axboe}. \periodical
{Fr\"uhmittelalterliche Studien} {23, 1989}, pp 453-477. 7 figs, pls
XXI-XXXII, refs. Ger.

Presentation of golden ring-cross pendants (cf NAA 1987/299) and three
hitherto unpublished gold bracteates. A small number of bracteates
without rim and loop, found in Gotlandic graves, were apparently
specially made to be used as Charon's obols. (Au)

\exauthor{Peter Petersen}
\site{Barsaldershed} {}
\site{Bohusl\"an} {GerIA}
\site{Broa/Broe} {}
\subject{ORNAMENTS}{Pendants}{}{ring-cross & gold bracteates}{{GerIA:Sw}}
\subject{PAGAN RELIGION}{Funerary ritual}{}{Charon's obol}{{GerIA:Sw}}
\subject{RELIGIOUS OBJECTS}{}{}{{GerIA:Sw}}
```

Following the text paragraphs of the abstracts are the ordinary references, in this case for the author, site, and subject indexes.

#### 3.2. Process

The abstracts and subsequent references were typed and edited on 11 separate text files, one for each archaeological period. During the entire editing process, the abstracts were unsorted. When all abstracts had been typed the text files were cross-checked and sorted by a separate Pascal program, resulting in 11 sorted text files. The sorted text files were then formatted by *TEX*, producing the formatted abstracts plus 11 new text files containing the unsorted index files. At this stage of production it was *not* possible to include new abstracts because that would change the abstract numbers used in the indexes. The index files were merged, sorted, and formatted. Finally the introductory text sections were formatted.

The following figure shows a simplified model of the editor-based production process:

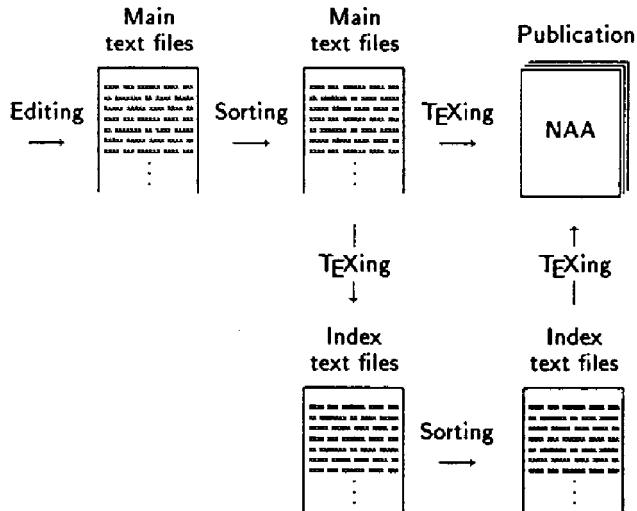


Figure 3. Simplified model of editor-based production process

Not represented in the figure above are all the arrows showing the correction processes following the proof readings.

A production process like this has numerous disadvantages. Some of the main problems were:

- endless possibilities of TeX syntax errors;
- no instant validation of data (spelling, format etc.);
- no sorting before all abstracts were included;
- no possible check of indexes before they were generated by TeX.

The main cause for most of the problems were:

- poor control of vital data;
- non-dynamic relation between abstracts and indexes.



## 4. System development tools

The NAASYS system was created with three different tools:

- an interface management system;
- a database management system;
- a formatting system.

### 4.1. User interface

The interface was designed and created with the interface management system C-scape version 3.2 from Oakland Group [1]. As the name suggests, C-scape is based on the programming language C. It contains a library of C routines and the program Look & Feel for interactive screen design. C-scape operates under various operation systems including DOS and UNIX.

C-scape is used to design screens that can contain other screens, menus, fixed data fields, and text editor fields in many combinations. The screens can be manipulated by mouse clicks or key combinations assigned by the programmer. There are various menu types as for example pull-down and pop-up menus. The text editor fields support for example scroll, word wrap and different foreground and background colour.

### 4.2. Database

The database was created with the database management system db\_VISTA version 3.1 from Raima Corporation [2]. The system is based on C, the network database model and B-tree indexing. It operates under various operating systems including DOS, UNIX and VMS.

### 4.3. Formatting

The format was created with the formatting system T<sub>E</sub>X [3], the state-of-the-art formatting system and typographical programming language. A T<sub>E</sub>X format for NAA has been in use for the latest four volumes of NAA. For historical reasons the format was made in plain T<sub>E</sub>X but will perhaps be made with the standard macro package L<sup>A</sup>T<sub>E</sub>X [4] in the future.

## 5. System and production model

### 5.1. Goals

The main ambition for the project was to build a system with the following characteristics:

- interface reflecting the structure of the final publication and the way the users look upon the data structure;
- menus and functions reflecting the working routines of the users;
- capability to handle various data types (ordinary text, text references floating around in the text, references outside ordinary text);
- instant (not only automatic) sorting of abstracts;
- instant (not only automatic) generating of indexes.

### 5.2. Process

A simplified model of the system looks somewhat like this (seen with the eyes of the users):

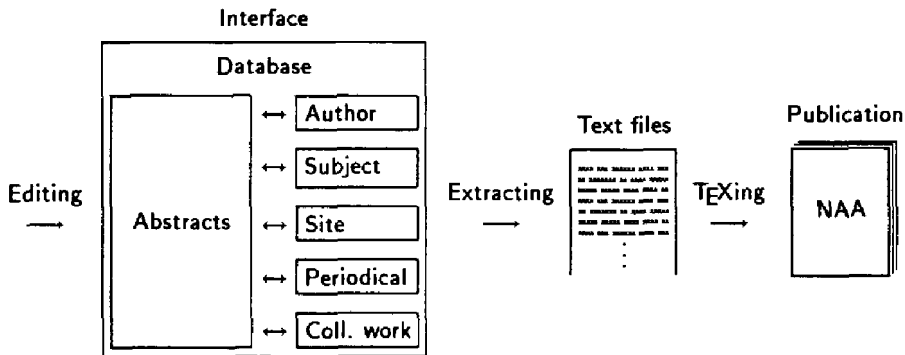


Figure 4. Simplified model of the database publishing solution

All user operations are handled by the interface which “takes care of the communication” to the database. The interface is responsible for the following operations:

- handles menus and functions;
- provides the editing facilities;
- validates all data put into the system;
- prepares transfer of all data into the relevant part of the database;
- formulates all requests to the database (selecting, preparing of index entries etc.).

After having received data or requests from the interface the database is responsible for the following operations:

- stores all data;
- takes care of the dynamic connection between the abstract and the indexes;
- sorts abstracts, references, and indexes;

At the final stage, all data are extracted from the database. The data are written on ASCII-files automatically including the proper T<sub>E</sub>X control sequences.

After having received the data extracted from the database T<sub>E</sub>X and its assistant program DVIPS take care of the final steps in the process:

- formats the files according to the rules laid down in the T<sub>E</sub>X format and produces a description in T<sub>E</sub>X's DVI-format;
- translates the DVI-files to PostScript format.

The PostScript files are finally printed on a PostScript laser printer (for proof reading) or typeset on a Linotronic 300 typesetter (final version).

All corrections following proof reading are done in the system similar to the original editing process. After the necessary corrections following a proof reading all data are extracted again, and a new formatting takes place.

## 6. System facilities

The following section contains some selected system facilities, showing only the main features of the system. The aspects focused on are

- the structure of the user interface;
- menus and functions;
- data validation.

### 6.1. User interface

When the system is started the user is presented with the main screen:

Abstract	Reference	Cross-Ref	Index	Text	List	Extract	System	Quit
KEY: 1A	PER/SUBJ:		LOC: Finn			NAA 1989/	1	
TITLE TEXT								
(Aarne Åyräpää. tutkija Opettaja Kansalainen. 1988-luvun näkökulmia.)								
AUTHOR TEXT								
{Var. authors}, ed by Matti Huure. Museoviresto. Esihistorian toimisto julkaisu(1989). 68 pp, ill, refs. Finn & Sw.								
ABSTRACT TEXT								
Papers presented on the centenary of Aarne Åyräpää 21.10.1987. (MS-L) {a.: Saatteeksi.} (Introduction). By {Torsten Edgren}. P 5.								
Create	Select	Edit	Delete	Write	First	Last	Previous	Next
							Main	

Figure 5. Main screen

The screen shows the first of the existing abstracts so far. It contains the main menu (top), the submenu (bottom) and the screen used for creating and editing the abstracts. This screen contains 3 fixed fields and 3 text editor fields, reflecting the structure in the final publication (see figure 1).

## 6.2. Menus and functions

The main menu and submenu can be activated by moving the cursor and pressing the Return key, by typing the highlighted letter, or by a mouse click.

The functions of the main menu reflect the necessary working routines for the users:

- creating and manipulating abstracts;
- creating and manipulating references in the current abstract;
- creating and manipulating cross-references for the indexes;
- printing a selected entry from one of the indexes;
- editing text sections;
- structuring and changing validation lists;
- extracting the text files for T<sub>E</sub>X;
- controlling system facilities such as backup and restore.

The main menu sometimes presents a supplementary pull-down menu:

In the submenu you can create abstracts, references and cross-references. You can also select an existing abstract by indicating key, title, author or current abstract number, or by paging up and down. When selected the abstract or reference can be edited, deleted, or written on the local printer.

## 6.3. Data types and data validation

If you want to create a new abstract, you activate the function **Abstract** in the main menu and subsequently the function **Create** in the submenu:

When the initial fixed fields have been entered the title text field is automatically activated. As the title is always the first item in this field, a special *assist screen* occurs:

This special screen contains the title (the part of the title text that is used for sorting). When the title has been entered the system returns to the main screen, and the rest of the title text can be typed. The title and the rest of the title text are represented this way:

Abstract	Reference	Cross-Ref	Index	Text	List	Extract	System	Quit
KEY: 1A	Author Site Subject Show all			LOC: Finn			MAA 1989/	1
TITLE TEXT								
(Aarne R. ja Opettaja Kansalainen. 1988-luvun näkökulmia.)								
AUTHOR TEXT								
(Var. authors), ed by Matti Huure. Museovirasto. Esihistorian toimisto julkaisu(1989). 68 pp, ill, refs. Finn & Sw.								
ABSTRACT TEXT								
Papers presented on the centenary of Aarne Åyräpää 21.10.1967. (MS-L) (a: Saatteeksi.) (Introduction). By (Torsten Edgren). P 5.								
Create Select Edit Delete Write First Last Previous Next Main								

Figure 6. Main screen

ABSTRACT								
KEY: 0	PER/SUBJ:			LOC:			MAA 1989/	0
TITLE TEXT								
[Empty text field]								
AUTHOR TEXT								
[Empty text field]								
ABSTRACT TEXT								
[Empty text field]								
Create Select Edit Delete Write First Last Previous Next Main								

Figure 7. Screen for creating an abstract

When the title text has been entered the system automatically jumps to the next editor field – the author text field. The first item in this field always

**ABSTRACT**

KEY: 7F    PER/SUBJ: 7H    LOC: Sw    NAA 1989/ 6

**TITLE TEXT**

**TITLE**

AUT Neues zu Brakteaten und Anhängern in Schweden.▶

**ABSTRACT TEXT**

Create Select Edit Delete Write    First Last Previous Next Main

Figure 8. Assist screen for entering the title

**ABSTRACT**

KEY: 7F    PER/SUBJ: 7H    LOC: Sw    NAA 1989/ 6

**TITLE TEXT**

(Neues zu Brakteaten und Anhängern in Schweden.) (News about bracteates and pendants from Sweden)

**AUTHOR TEXT**

**ABSTRACT TEXT**

Create Select Edit Delete Write    First Last Previous Next Main

Figure 9. Title represented in the title text editor field

is the primary author, used for sorting as well as being a text reference for

the author index. Therefore, a new assist screen is automatically activated. But this time, it can be operated in connection with a *pick list*:

The screenshot shows a terminal window with a header bar labeled 'ABSTRACT'. Below the header, there are fields for 'KEY: 7F', 'PER/SUBJ: 7H', and 'LOC: 30'. The main area is divided into sections: 'TITLE TEXT', 'AUTHOR REFERENCE', and 'ABSTRACT TEXT'. The 'AUTHOR REFERENCE' section contains fields for 'FAMILY NAME:' and 'FIRST NAMES:'. A pick list is open on the right side, displaying a list of names: Königsson Lars-König, Lahtiperä Pirjo, Lamm Jan Peder, Lamm Kristina, Landell Petter, Landsverk Halvor, Lang James T, Lange Bernt C, Lange Kalle, Lannby M Golabiewski, Larsen Bent Bang, Larsen Carsten U, Larsen Erling Benner, Larsen Jan Henning, Larsen Lene, Larsson Gunilla, Larsson Lars, Larsson Mats, and Larsson Stefan. At the bottom of the window, there is a menu bar with options: Create, Select, Edit, Delete, Write, First, Last, Previous, Next, and Main.

Figure 10. Combination of assist screen and pick list

If you just *typed* the authors in all abstracts you would almost certainly end up with various representations (spellings) of the same name. Not only would that make proof reading more difficult, but it would certainly produce incorrect sortings of abstracts and indexes. And mistakes like these are difficult to track.

The pick list is activated by pressing the Alt-key in combination with the wanted initial (here L). Physically an author only exists in the list (the index). The author is placed in the abstract or reference simply by picking him or her from the list. If the author does not exist in the author list, you just type the name, and it is included in the list for future use.

This approach is used for authors, periods, localities, subjects, sites and a few other structures, not only minimizing spelling and sorting errors but also saving a lot of time.

Text references like extra authors or periodicals are created the same way, the only difference being that the user has to activate the assist screen by pressing a key combination like Ctrl-a (author) or Ctrl-p (periodical)



because text references can occur everywhere in the text. When they have been created they behave like ordinary text in the paragraph (wrap etc.).

After having typed the rest of the text and having created the necessary text references the abstract is complete:

ABSTRACT			
KEY: 7F	PER/SUBJ: 7H	LDC: Sw	NAA 1989/ 17
TITLE TEXT			
(Noues zu Brakteaten und Anhängern in Schweden.) (News about bracteates and pendants from Sweden)			
AUTHOR TEXT			
(Lamm, Jan Peder) & (Morten- Axboe). (Frühmittelalterliche Studien.123, 1989), pp 453-477. 7 figs, pls XXI-XXXII, refs. Ger.			
ABSTRACT TEXT			
Presentation of golden ring-cross pendants (cf NAA 1987/299) and three hitherto unpublished gold bracteates. A small number of bracteates without rim and loop, found in Gotlandic graves, were apparently specially made to be used as Charon's obols. (Au)			
Create Select Edit Delete Write First Last Previous Next Main			

Figure 11. Complete abstract

A more complicated assist screen is provided for making subject references. To create a subject reference you first activate the function Reference, then choose Subject from the pull-down menu, and finally the function Create from the submenu. Now you get the following screen where both the subject, the level 2 entry, the period and the locality have been picked from the proper pick lists:

#### 6.4. Other facilities

Other facilities include

- assist screens for creating and manipulating cross-references at different levels;
- screens for maintaining the contents and structures of the pick list;
- a separate text editor for editing the text sessions;

REFERENCE		Subject	
KEY: 7F	PER/SUBJ: 7H	LOC: Sw	NAA 1989/ 17
TITLE TEXT			
Menu and	SUBJECT REFERENCE		
	SUBJECT:	ORNAMENTS	
AUTHOR	2: ENTRY/PERIOD:	Pendants	
CLASS 1989	3: ENTRY/PERIOD:		
	4: ENTRY:	ring-cross & gold bracteates	
ABSTRACT	5: PERIODS:LOCALITY:	( GerIA : Sw )	
Press <b>Alt</b> <character> to insert lokality. <b>&lt;DEL&gt;</b> to clear field			
hitherto unpublished gold bracteates. A small number of bracteates without rim and loop, found in Gotlandic graves, were apparently specially made to be used as Charon's obols. (Au)			
Create Select Edit Delete Write		First Last Previous Next Main	

Figure 12. Creating a subject reference with pick lists

- system facilities such as backup and restore.

## 6.5. On-line help and documentation

C-scape makes it possible to include on-line help facilities at every level in the menus and functions. This has not been done due to lack of resource.

The system was (of course) documented with  $\text{\LaTeX}$ . We found it very useful to draw screens and parts of screen with the box facilities in  $\text{\LaTeX}$ , thus being able to focus on parts of a screen and saving a lot of time and disk space not having to represent them in Encapsulated PostScript format.

## 7. Possibilities and implications

### 7.1. Merits

Compared to the the editor-based production system NAASYS unquestionably gives the user

- better editing facilities;

- instant validation of vital data;
- instant sorting of abstracts;
- instant generating of indexes;
- better tools for cross-checking and search for vital data;
- better possibilities for editorial teamwork;
- considerably fewer working hours.

## **7.2. Possibilities and extensions**

We have already received a few requests from the users of the system and other interested persons. These requests indicate that the following facilities would be of interest when making similar systems:

- filter for text files from e.g. WordPerfect;
- possibility to merge data created by geographically separated users (merging two or more databases at a late stage of production);
- separate storing of indexes for each volume for later production of cumulative indexes;
- local output formatted by T<sub>E</sub>X;
- genuine multi-user systems in local networks;
- system running under Windows with a graphical user interface;
- output to other formats (SGML, Marc (a library standard));
- output to other media (on-line, CD-ROM etc.).

## **7.3. Implications for T<sub>E</sub>X**

T<sub>E</sub>X users who make reports, journals etc. are very well served with the excellent facilities provided by L<sup>A</sup>T<sub>E</sub>X. But potential T<sub>E</sub>X users who want to produce inventories with the characteristics described in the abstract of this article will need many of the facilities described in the article.

We see T<sub>E</sub>X as a valuable “engine” in tailored database publishing systems. Such systems would give some new and exciting applications for

TEX, making TEX even more widespread and making TEXperts look upon TEX's possibilities with new eyes. You should just have in mind that the publication in question must be so large or must be published so frequently that it is worthwhile to start such a project.

## References

- [1] *The C-scape Interface Management System*, Oakland Group, Inc., Cambridge, Massachusetts, 1990.
- [2] *db\_VISTA User's Guide*, Raima Corporation, Bellevue, WA, 1989.
- [3] Donald E. KNUTH, *The TEXbook*. Addison-Wesley, Reading, Massachusetts, 1984.
- [4] Leslie LAMPART, *LaTeX user's guide and reference manual*, Addison-Wesley, Reading, Massachusetts, 1986.