# TECHNICAL STANDARDS FOR MAGNETIC TAPE EXCHANGE BETWEEN DATA ORGANIZATIONS

Karsten Boye Rasmussen

Danish Data Archives

## INTRODUCTION

At the IASSIST-sessions in Uppsala August 17 - 18, 1978, there was some discussion of the technical aspects of data exchange. In recent years, the Danish Data Archives (DDA) has had considerable experience concerning exchange of data files on magnetic tape -- internationally between data organizations as well as inside Denmark between computer installations (notably IBM, CDC and UNIVAC). The DA therefore accepted the invitation to write a note on these problems proposing some usable standards. Although the DDA as a DO (Data Orgnization) emphasizes the importance of documentation, this note deals with the technical aspects of data transfer on magnetic tape in general.

The correct procedure for exchange of magnetic tapes naturally includes preparation of a complete technical description of the tapes. However, the standards proposed in this paper, while they may not always be the most effective ones or the easiest ones to use, should assure the possibility of reading the tape at other locations, even in cases where the tape description for some reason is lacking. The philosophy behind the standards is that of "simplication". You should never let a "data maniac" convince you to ship a multi-reel, machine and operating system dependent SPSS-file on seven track tape; that is, unless you are on really bad terms with the receiver of the file.

## TAPE LABELLING (NL, no label)

In principle, all products from the major computer firms should be capable of reading ANSI labels (American National Standard labels). However, different systems may produce slightly different ANSI-labels (due to frequent change of operating systems), and some systems may not process ANSI labels correctly. For these reasons, we would propose that all exchange tapes be written without labels (NL, no label), as NL-tapes can definitely be processed by any computer center. Leading tapemarks should be avoided.

## TRACKS (9-track, 1600 BPI, PE)

At most installations, 9-track tape drives are available; normally this is true even for computing centers where the system uses 7-track tape drives. At present, the most commonly used density is 1600 BPI (bits per inch) with phase encoded (PE) magnetization.

## DATA FORMAT (blocked, fixed record length, 80 char., ASCII)

Most system and/or machine dependent files should be abandoned, as even a very good

programmer would have to spend months converting them to the local file format. Some installations support conversion programs, but these programs may not exist at the receiving computer center, and even if they do, the program 'level' may be different. It is advisable to ship data files in the most conservative and simple format possible, i.e. card image character format. Given the fact that many computers will have trouble handling large blocks, the blocksize should not exceed 2048 bytes. Finally, we would recommend using the ASCII character set, which can normally be converted by standard software (or the operating system).

## CONCLUSION

For exchange purposes, a data file in magnetic tape should have the following technical characteristics: the tape should be without labels (NL), 9-track, 1600 BPI, PE. the file should have 80 byte records, be in character format (ASCII) and have a blocksize not exceeding 2048 bytes. Most installations will have software (supplied by the manufacturer or supported by the computing center) for reading and converting such files. (This fact is reflected in the Data Organization Registry Form, completed by the members of the International Federation of Data Organizations). Should this not be the case, files of this kind will present only a minor problem to a capable programmer.

As mentioned above, this note does not distinguish between data and documentation files, as this distinction is irrelevant for purposes of generating and reading the tape. Indeed, documentation files from the most commonly used packages (OSIRIS: the codebook; SPSS: the file produced by WRITE FILE-INFO) already follow the standards outlined above as far as data format is concerned; and if, in the future, a standard for documentation files is defined -- e.g. the "data interchange file" as proposed by Richard Roistacher -- the standard file will almost certainly be in character card image format too.

## REFERENCES

CDC. DMS-170, CYBER RECORD MANAGER VERSION 1. Reference Manual, No. 60495700, 1976.

ICPSR. OSIRIS III, VOL. 1. Ann Arbor: The University of Michigan, 1975.

IBM OS UTILITIES (IEBGENER), 16TH ED. (GC28-6586-15). April, 1973.

Nie, Norman, ET AL. SPSS SECOND EDITION. New York: McGraw-Hill, 1975.

Roistacher, Richard C. THE DATA INTERCHANGE FILE. CAC doc. No. 257. University of Illinois, February, 1978.

System supported software for generating and reading "standard" files:

IBM: OS UTILITIES (IEBGENER) Order No. GC28-6586-15, 16th ed. April 1973

CDC: DMS:170, Cyber Record Manager Version1, Reference Manual, No. 60495700, 1976