The UKTeX Archive at the University of Aston

PETER ABBOTT

The Computing Service
University of Aston
Aston Triangle
Birmingham B4 7ET
England
AbbottP@Aston.Ac.Uk

ABSTRACT

A viable archive should contain material which is useful, accessible and easily implemented on the target system. The growth of electronic communication, electronic mail in particular, has brought into being computer-based archives, removing the restrictions of geographical or political boundaries.

This paper describes the growth of the Aston Archive, the Archivists and how we manage the archive. I shall describe our approach to make implementation across a wide range of systems consistent and how we answer questions such as "Does the archive contains fonts for device Y" or "How do I use printer X".

I hope to answer the question "Is the Archive Useful?"

1. Definition of an Archive

If the proverbial "man in the street" were asked the question "What is an archive?", the most likely reply would be "A collection of books or records". A popularly held view is of the 19th-century room filled with books collecting dust, or records of births, deaths and marriages. In the United Kingdom for many centuries, parish records formed the basis of archives and were usually kept in the Parish Church — a situation which no longer applies, I might add. Chamber's Student Dictionary (Revised edition 1980) defines archives as "the place in which government records are kept: public records". The University is mainly funded by government monies and the archive does contain public records, so the definition seems relevant.

I suppose that the modern concept of an archive was created with the advent of the digital computer, and it was only a matter of time before "collections" of computer programs and similar material acquired the title ARCHIVE. In fact, the common term used when preserving the filestore of a computer against corruption for future use is archiving, and a number of products on the open market make use of the word. The major growth in electronic communication in the last few years has opened up the concept of a central archive to serve large areas of a country or even the globe.

There are a number of well-known archives in the world, and several which serve the TeX community. I shall, of course, be concentrating on the one located at Aston University in the United Kingdom.

2. The System at Aston

Aston University is located in the City of Birmingham on a 35-acre city-centre campus. Birmingham has two universities, and the older one (carrying the title "Birmingham University") owes much to Chamberlain and is classed as a "red-brick" University.[1] The Aston Campus does not conform to the standard city-centre mould, and its campus (albeit small) would not disgrace a country-park approach. The Computing Service which hosts the archive is located on the fourth floor of the Main Building,

---

[1] 'Red-brick' — The Victorians were fond of using red terra cotta bricks for buildings and universities founded in Victorian times have acquired the name "red-brick".

which was designed in 1933 and built after the second World War. As a result, the accommodation arrangements are generous for the hardware and acceptable for the staff and users.

To all intents and purposes, the equipment used to support the archive is immaterial, other than for the unusual problems that arise; the present computer system is a VAX processor running the VMS operating system.

3. Archive History

In the summer of 1987, a number of interested users in the United Kingdom discussed (via e-mail) the idea of a UK-based version of TEXhax, whereby local problems could be aired and informal contact made between like users. At that time, Aston were interested in text publishing and the Vice Chancellor expressed an interest in TEX (having spent some years at Stanford).

I agreed to set up a distribution list and to collate and distribute the queries. The first few issues were experimental, with a number of hiccups and problems, but eventually a pattern emerged of a weekly digest normally posted on a Friday afternoon beginning at 1700 hours local time.

As a result of the digest software, a large number of questions started arriving at Aston; these mainly asked if 'x' was available, but also frequently offered items which might be useful to others. Like Topsy, it grew and grew and grew.

For many months, the digest was from the UK for the UK, but as it became more widely known, requests for subscription arrived from many areas of the world including the North American continent. This brought more material for the archive and requests for distribution both via e-mail and by more conventional means (such as magnetic tape). The archive and digest are now considered as originating in the United Kingdom with almost universal distribution.

All of this activity consumes a large part of my "free" time and I became increasing worried that the archive was becoming disorganised with duplicate material. At a meeting in November 1988 at Nottingham, I appealed for help in preparing an article for *University Computing*, a publication of the Inter University Computer Committee (IUCC); this has since developed into a book (Clark and Abbott 1989; I have brought copies of part of the book with me to the meeting). I also approached a number of people with a view to assistance with the archive.

3.1 Archivists

Adrian Clark

Adrian (who prefers the alias *Alien*) is a researcher in the image-processing field, being especially interested in algorithmic and software aspects. His interests in TEX are hence mainly concerned with graphics — particularly the inclusion of grey-level pictures ("halftones") in documents. He is the author of a number of articles in *TUGboat* concerning halftone output and aspects of the implementation of TEX under VAX/VMS. Adrian is responsible for several areas of the archive, including contributed LATEX and BIBTEX styles, VMS implementations, and halftones. He also developed the Aston mail-server.

Malcolm Clark

Malcolm really needs no introduction but he has been producing books in TEX since 1984; is editor of *TEXline*, a newsletter of the TEX community, since 1985; organised TEX88, the 3rd European TEX Conference at Exeter University; is the editor of the TEX88 Conference Proceedings (which we are still awaiting). He is currently TUG's European Co-ordinator and also Secretary of the British Computer Society's Electronic Publishing SIG; teaches TUG courses in the US and Europe; carefully watches SGML, PostScript and ODA; currently completing *A plain TEX Primer* — an introduction to TEX, and *HyperTEX* — a Hypertext guide to TEX and LATEX. Responsible for a tiny portion of the Aston Archive.

Charles Curran

Charles has worked in the Systems Development section at Oxford University Computing Service for the last fourteen years. His main interests are operating systems and document architectures. He supports the UNIX services, which are provided on Suns and Convex supercomputers. He is currently chairman of the UK Convex Users Group. He first became familiar with TEX about eight years ago,

when he installed it on a local VMS VAX and also on the sadly maligned Perq. He has also spent a fair amount of time with METAFONT. It is with founts and things METAFONT that he is primarily concerned in his work on the Archive.

David Osborne

Dave joined the Cripps Computing Centre, University of Nottingham, in 1984 as systems programmer. His special interests are UNIX and text processing. He helped in setting up TeX on the Centre's VAX 11/750 some time in 1986, and has since worked on the Centre's TeX implementation on an ICL Clan 7 running UNIX System V. His areas of TeXnical interest are general, and mainly related to the implementation side of things, partly from a UNIX system manager's point of view. Dave joined TUG in 1988.

His responsibilities in Archive management include: the Beebe drivers, Andrew Trevorrow's software, Common TeX, and PC items.

Sebastian Rahtz

Sebastian is a lecturer in the Department of Electronics and Computer Science at the University of Southampton, where he has a special interest in humanities computing, teaching courses in text-processing for computer scientists and computing for archæology M.Sc. students. He originally got involved with computers for publication of an archæological report, and his research interests of type-setting, generic markup, hypertext and graphical databases continue to reflect a preoccupation with publication.

He has been using TeX since 1985, and has produced a number of books. The archæological aspect has produced a particular interest in the problems of graphical inclusion in TeX, and the integration of PostScript. Within the Aston Archive, he attempts to cover the UNIX implementations, some of the PostScript problems and some of the graphical systems.

Philip Taylor

Phil has been with the University of London for almost eighteen years, having migrated from Westfield College (University of London) to Bedford College (London), and thence to Royal Holloway and Bedford New College (RHBNC, University of London), where he is presently a member of the Computer Centre. His interest in TeX was evoked by a chance encounter at British Petroleum, where he first saw a sample of TeXset text, and was amazed to learn that it had been produced on a system identical to that at RHBNC. He rushed back to College with a tape, spent the next month pestering various colleagues at KQC for assistance with his naïve attempts to produce TeX output, and has never looked back since. He is currently visiting Canada (yet again) to collaborate with the Professors Gibson and Gibson, at the Universities of Guelph and Waterloo, in TeXsetting their book on nutrition. When not TeXsetting, he spends most of his time falling off horses.

Within the archive group, Phil is responsible for advice on VAX/VMS file aliasing, and for maintenance of the VAX/VMS list-server.

Areas of interest: Primitive and Plain TeX; do-it-yourself macro design; TeX esoteria and arcana. *Bêtes noirs*: LaTeX, UNIX & Wimps.

Niel Kempson and Brian Hamilton Kelly

Niel and Brian joined the archivists in May 1989 to take over the VMS part from Adrian, who was having difficulty in coping with it in addition to the mail server. Adrian still retains the mail server.

Brian (note the double-barrelled surname, *without* a hyphen) is a Senior Research Officer in the Software Engineering Group at the Royal Military College of Science, which is now part of Cranfield Institute of Technology: he is currently working for an M.Phil. connected with code generation derived directly from graphical representations of systems' designs (although *he spends more of his time on* TeXware!)

He has spent the whole of his working life in software, since 1964, the majority of it with the MoD (until the privatization of RMCS), but also worked for a short spell in 1966 at Imperial College. He has had an interest in computerized typesetting since 1968, and in type and typography going back to his schooldays. (Ah! The joys of an Adana hand press!) He has been using LaTeX since 1985, and his

first TeXware project was to rewrite Rose's original DVI driver for the DEC LN03 wholly in WEB, to overcome the bugs in the PL/I part of it. (His program, DVItoLN03 is not to be confused with Rose's later DVI2LN3 offering.)

He would like to see the authors of TeX- and METAFONT-ware pay much more attention to writing for ease of maintenance, and feels that every C- programmer should have three men with whips standing over him ensuring that good software engineering precepts are followed! Furthermore, he would like to see every TeX macro package re-written and documented through Frank Mittelbach's doc style — any volunteers!!?!

Niel is by profession a chartered electrical engineer, but is currently on sabbatical leave at RMCS Shrivenham, studying for a PhD. His academic interests are centered around the use of multivariate analysis to identify different types of communications signals.

As with most of the post-graduate students at RMCS, he soon discovered the joys of using LaTeX to produce reams of mathematical equations. He soon became involved with the maintenance of TeXware at RMCS, specializing in the Unix systems, DEC's language sensitive editor and translating BibTeX into C for use on the IBM PC.

4. JANET

It is fair to say that the *only* way that we can work in maintaining and developing the archive and its facilities is by utilising modern technology. In this case I mean JANET — the (U.K.) Joint Academic Network. JANET is a private network operated by the Joint Network Team under a crown agency. The recurrent funding is met centrally and all universities have free and unfettered access. Polytechnics, research councils and other bodies deemed "suitable" are also connected (they generally have to pay a once-off connection charge).

The archivists have an account on the system at Aston and use JANET to carry out allotted tasks. Archive users are not permitted at present to access the archive interactively (archivists enjoy this perk), but Network Interface File Transfer Protocol (NIFTP) and e-mail access are both provided.

4.1 Gateways

Whilst this provides a service to the connected community, there are wider implications for its use. Gateways to other networks exist at various points in the United Kingdom, all linked via a backbone. The well-known gateways are:

Internet/Arpa/Nsfnet	Uk.Ac.Nsfnet-Relay
Earn/Bitnet/Netnorth	Uk.Ac.Earn-Relay
uucp	Uk.Ac.Ukc

plus others including commercial networks such as the British Telecom PSS service. The main trunks operate at 2 megabits and the Aston-Manchester link is currently 48K.

5. Structure of the Archive

We have decided to sub-divide the archive into the following main topic areas:

- tex
- metafont
- latex (subdirectory for slitex)
- amstex
- digests (texhax, uktex, texmag)
- bibtex
- utils
- fonts
- etc
- drivers
- docs
- langs
- tools (for use by the 'team', not the general public)

5.1 Areas Covered
There are two aspects to areas covered:

- What material should be accepted?
- What systems should it cover?

In dealing with the first area, I must mention the United Kingdom National Public Archive at Lancaster University. This archive contains public-domain software mainly for MS-DOS PCs and Apple Macs. At one time it carried some of the PC software related to TeX/LaTeX, but this has now been removed and all enquiries are re-directed to Aston.

The archivists are only interested in working with material directly related to TeX and related software, and a long debate has taken place over utilities such as the bootstrap for MS/DOS PCs. In the end, it was agreed that these would need to be stored at Aston as well as Lancaster, because not all subscribers (principally those on the other side of a gateway) could obtain the items from Lancaster.

We are working towards the situation whereby *all* needs for an implementation — from the smallest PC to the largest mainframe — can be obtained, preferably by e-mail, but in the interim supplemented by other media.

6. Who Uses the Archive?

6.1 Contributors
It might appear that we are working in a vacuum at Aston, but this is far from the truth — contributions are received from many countries and organisations. It is impossible to name the majority, but I must mention the following:

- Pierre MacKay (Univ. of Washington/Seattle), who regularly sends an update tape of the UNIX versions
- Michael DeCorte (Clarkson Univ., New York), who regular mails updates from his collection
- Max Calvani (International School for Advanced Studies (ISAS), Italy), who can supply material which avoids the Earn-Relay and the corruption that occurs there
- Andrew Trevorrow (freelance TeX programmer, Australia)
- Peter Flynn (Academic Projects Manager University College Cork, Ireland)
- Graham Toal (self-employed consultant, Edinburgh), who provided the early mail server to the Archive

plus, of course, all the others too numerous to mention.

6.2 Types of Access
As of August 1989 there are only two means of access: NIFTP and mail. Users on JANET can use NIFTP to transfer items from the archive to their local system. *All* users can access the archive via the mail server, and I am pleased to be able to report that the VAX/VMS problem of STREAM_LF files causing untold havoc has been overcome. There might be one or two files still to convert, but the end is in sight.

There are no interactive access facilities, although it is hoped to introduce a processor dedicated to the service in the near future. This will provide not only Kermit and Zmodem facilities, but also (hopefully) a secure environment for mail-boxes and bulletin-board services.

7. Digest Support
The digest is the major vehicle for information dissemination, and until access for external users becomes available, is also distributed via paper mail. This service is provided on a VAX 11/750 independent of the archive service, and it is intended to merge this with the archive as soon as the separate processor becomes available.

8. Non-Connected Users
There are many commercial users, most of whom do not have a connection or access to JANET, so for this group of users the advent of the independent service for the archive will open new opportunities.

The University has dial-in lines which will give access from any telephone, provided that the user has a suitable modem.

9. Is it Useful?
I think the following statistics will answer the question:

For the period November 1988 to end July 1989

No. of searches	26
No. of requests for a Directory listing	975
No. of requests for files	2,988
Files sent	3,986
Help sent	142
Italian help	6
Danish help	21
Invalid requests	195

Bibliography

Clark, Adrian, and Peter Abbott, eds. *The TEX Companion*. In preparation.

Clark, Malcolm. *HyperTEX*. Exeter and London: TEXpert Systems, 1989.

Clark, Malcolm. *A plainTEX Primer*. [Forthcoming].

Clark, Malcolm, ed. *TEX88 Conference Proceedings*. Chichester: Ellis Horwood, 1989.

Mittelbach, Frank. "The doc option." *TUGboat* 10(2):245–273, 1989.