TEI Information Sheet

Obtaining TEI P2 and Using TEI-L

TEI EDJ8

May 4, 1993 (14:14:34)

Draft May 4, 1993 (14:14:34)

Abstract

The Text Encoding Initiative (TEI) is an international research project funded by the US National Endowment for the Humanities, the European Community (DG XIII), the Andrew W. Mellon Foundation and the Canadian Social Science and Humanities Research Council. The TEI is sponsored by the Association for Computing in the Humanities, the Association for Literary and Linguistic Computing and the Association for Computational Linguistics. Since 1988, it has been working towards the definition of a suite of extensible Guidelines and Recommendations for use when encoding all kinds of text in machine readable form for all kinds of research purposes. Its initial proposals, derived from extensive consultation in the research community represented by its three sponsoring organizations, appeared as an initial report in November 1990. That document (TEI P1) recommended the adoption of a standard based on the Standard Generalized Markup Language (SGML, ISO 8879) and made very detailed proposals for document type definitions covering a large range of document types, including tagsets for basic prose, dictionaries, lexical and syntactic analyses and textual criticism amongst others. These proposals have since been further refined and extended by a number of specialist working groups, and are now being published in fascicle form, as further described below.

1

Addresses

The TEI is managed by a steering committee composed of representatives of the three sponsoring organization (see TEI J16 for further information on the TEI). The editorial work is co-ordinated by two editors, whose addresses follow:

C. M. Sperberg-McQueen, (editor in chief)
University of Illinois at Chicago (M/C 135)
Computer Center
1940 W. Taylor St.
Chicago, Illinois 60612-7352 US
U35395@uicvm.uic.edu
+1 (312) 413-0317 Fax: +1 (312) 996-6834

Lou Burnard, (European Editor)
Oxford University Computing Services
13 Banbury Road, Oxford OX26NN, UK
lou@vax.ox.ac.uk
+44 (865) 273200 Fax: +44 (865) 273275

2

Contents of TEI P2

Version 2 of the *Guidelines for the Encoding and Interchange of Machine-Readable Texts (TEI P2)* is being published initially in electronic form. Rather than issuing TEI P2 as a complete, comprehensive, single volume work, the TEI will publish each section of TEI P2 as soon as it has been approved by the drafting committee and the editors. A copy of the current table of contents accompanies each fascicle and is also attached to this document.

This serial mode of publication has the major benefit of simplifying the review of the Guidelines for the reader. Not everyone has the time or inclination to comment on a volume the size of TEI P1 when it lands on the desk all of a piece, while individual chapters on specific topics can perhaps be more readily and speedily considered. We hope that this means the quality and quantity of public comment will increase: you can help by commenting yourself or by making sure that other interested persons in your organization who may not have access to electronic mail are made aware of the drafts as they come.

Each fascicle is also accompanied by a User Response Form. The TEI is very anxious to receive detailed technical comments on any aspects of its recommendations, and this form is simply a convenient way of soliciting them. Comments are of course welcome in whatever format is most convenient: they may also be directed to TEI-L, as discussed below.

3

File Formats

Each file containing a draft fascicle has a name in the form P2xx.yyy where xx is a two letter code (earlier drafts used a number) for the chapter name and yyy a code indicating the file format or filetype. If you have plenty of file space and access to a PostScript printer, we recommend you to download files with the filetype PS. If you just want to read the text, on screen or with screen-like formatting, download files with the filetype DOC. If you have access to the TeX text formatting system, download files with the filetype TEX (but note that very few files are as yet available in this format). Or if you want to see the drafts in their true SGMLshape, then download the files with filetypes P2X and REF.

In summary, the following filetypes are currently used:

PS: Postscript (not readable on screen)
P2X, REF, DTD [all SGML]: (readable on screen, but full of SGML tags)
DOC, LIST: Formatted file (readable on screen, no tags)
TEX: Suitable for input to the LaTeX or TeX processor

As you might expect, TEI P2 is being drafted using SGML, and all drafts are available in this form. Files of type DTD contain DTD fragments; files of type P2X and REF contain distinct parts of the SGML document. As their names suggest, P2X uses an extended version of the TEI P2 DTDs themselves, while REF files contain the formal definitions of TEI elements, parameter entities and element classes which will make up the alphabetical reference section for TEI P2. The current drafts use a preliminary version of this DTD, which we are not making generally available since it is not yet fully documented nor likely to be of general interest; SGML hackers with a burning desire to know more should contact the editors.

The draft DTD files contain the Document Type Declaration (DTD) for the material covered in each fascicle and these files will be made available on the server for retrieval by anyone interested in detailed study of the TEI proposals. Please note however, that these DTD fragments are unlikely to be usable as a whole for some time, and may be subject to substantial revision.

4
The TEI-L Listserver

TEI-L is the name of a publicly-accessible ListServ maintained at the University of Illinois at Chicago, to which anyone interested in the TEI should be subscribed. This ListServ is the primary communications channel between the TEI and the research community it serves. Subscribers to TEI-L can exchange messages, discussion and comment with each other and with other participants in the TEI. They are also automatically informed, by electronic mail, whenever a new fascicle of the TEI Recommendations is available and how it may be downloaded.

To subscribe to TEI-L, send an electronic mail message to the address

```
listserv@uicvm.uic.edu
```

The exact form of this address may vary on different computer systems; for systems on Bitnet, for example, it is is simply listserv@uicvm. You should consult your local support staff for advice on which format to use.

The text of the mail message you send should contain a line like the following:

```
subscribe TEI-L J. Q. Public
```

substituting your real name [not your network address!] for J. Q. Public. The ListServ program is clever enough to work out your network address by inspecting the envelope of your message, but needs to know your real name for its membership list.

When you are registered as a subscriber to TEI-L you will receive an introductory package of materials giving further details of services available, TEI publications etc. The following briefly summarizes the most commonly used commands. Like the subscribe command above, each one should be sent as the body of a mail message to the address

```
listserv@uicvm.uic.edu
```

To get a list of files available from the TEI-L fileserver:

```
index TEI-L
```

To get a copy of the specific file called FOO BAR (filenames on this system consist of two words separated by a space):

```
get FOO BAR
```

To get a list of subscribers to TEI-L and their addresses:

```
review TEI-L
```

To suspend messages from TEI-L temporarily (for example because you are going on holiday):

```
set TEI-L NOMAIL
```

To resume your subscription (Caution! you will never see any messages that came while you were away!):

```
set TEI-L MAIL
```

Whichever message you send, the ListServ will always reply to you by electronic mail. If you are requesting a very large file from it, therefore, you should be sure not only that you have enough disk space but also that your mail system can cope with large messages. If this is likely to cause problems, you might like to consider using FTP instead (see further below)

Finally, if you want to make a comment on TEI P2 or enter into a general discussion with other subscribers, you can send an ordinary mail message to the address

```
TEI-L@uicvm.uic.edu
```

Messages sent to this address are automatically forwarded to all subscribers to the list. Please bear this in mind when sending your message! Not every subscriber is interested in any difficulty you may be having downloading files, for example; messages on such topics should be sent to the TEI secretariat at u49127@uicvm.uic.edu. However, general comments on the content of the TEI Recommendations or specific questions and criticism about them are very welcome indeed.

5

Other ListServ sites

Identical facilities to those listed above are available from a ListServ maintained in Germany at the University of Göttingen, which distributes copies of TEI drafts electronically and also hosts a discussion list. The name of the list is MARKUP-L, and its address is Listserv@ibm.gwdg.de (or, for EARN sites, Listserv@dgogwdg1)

To use this ListServ, use exactly the same instructions as those above, substituting MARKUP-L for TEI-L, and ibm.gwdg.de for uicvm.uic.edu throughout.

6

FTP access

FTP (File Transfer Protocol) is an alternative means of transferring files from one computer to another now very widely used on the Internet. If your computer supports this system, you may find it more convenient than ListServ, since it allows for files to be transferred directly from one machine to another rather than as electronic mail messages. On the other hand, you will not be advised of the availability of new fascicles automatically unless you are subscribed to a ListServ.

Your local communications support staff should be able to advise you on the use of anonymous FTP, which is more or less standard across a very wide range of computer systems. The example below assumes that you are accessing the FTP server maintained by the SGML Project at the University of Exeter in the UK, which has agreed to make TEI drafts generally available in this way, but the same principles apply to many other sites.

To connect to the file server:

At the operating system prompt, type

```
ftp sgml1.ex.ac.uk
```

or

```
ftp 144.173.6.61
```

You will be prompted for an account or user name (the exact form of the prompt will be different on different machines), to which you should simply respond

```
ftp
```

or

```
anonymous
```

There may be some delay following a "Connected" message before you are asked for a password. You should then supply your full e-mail address.

For full information on available FTP commands, ask your local support staff. Some examples of the most commonly used commands follow:

To get the file readme, renaming it as TEI.ReadMe:

```
get readme TEI.ReadMe
```

To get all the files of type doc in the current directory:

```
mget *.doc
```

To change the current directory to tei/drafts:

```
cd tei/drafts
```

To list the names of the files in the current directory:

```
ls
```

To disconnect from the server:

```
bye
```

Remember that most FTP servers run under UNIX or use UNIX style file naming conventions and structures, in which (for example) upper and lower case letters are regarded as distinct.

For information about the structure and organization of the TEI files held at the Exeter file server, your first command after getting connected should be

```
get tei/readme
```

7
Other FTP sites

7.1 Japan

The TEI P2 fascicles may also be obtained by anonymous FTP from the following two sites in Japan. Those who can reach these sites are strongly recommended to get the files from either of them.

```
pine.kuee.kyoto-u.ac.jp
ftp.hitachi-sk.co.jp
```

PINE is for those accessing from the western part of the country and the HITACHI site is for those accessing from the eastern part.

Standard anonymous ftp login procedure and restrictions apply at each site. Each fascicle is stored in compressed tar format in a file called P2xx.tar.Z (where xx is the identifying code for the fascicle), containing all released formats. These files are held in directories

```
pub/TEI (at pine.kuee.kyoto-u.ac.jp)
pub/doc/TEI (at ftp.hitachi-sk.co.jp)
```

Like other .tar.Z files, these files are not suitable for transfer to non-Unix machines, and should be transferred in Binary Mode. In each directory, an additional README file is available containing updated information on its contents.

TEI enthusiasts in the Far East are requested to contact Professor Syun Tutiya of Chiba University, who has responsibility for East Asian distribution of TEI materials on behalf of the TEI Japan Committee (an independent organization with the goal of ensuring that East Asian needs, particularly but not exclusively character-set related, are met by the TEI Guidelines). Contact address:

Syun Tutiya
Department of Philosophy
Chiba University
1-33 Yayoi-cho
Inage
Chiba Chiba 263 JAPAN
email: tutiya@culle.l.chiba-u.ac.jp
fax: +81-43-256-7032

The TEI Japan Committee is also interested in formulating a common set of comments from Asian points of view; those interested in having input to this process should contact Prof. Tutiya. A Japanese TEI discussion group -- JTEI-L -- is also in operation. To subscribe to JTEI-L, send a message containing your name, professional background, and interests to tei-l-request@culle.l.chiba-u.ac.jp

7.2 Norway

The SGML Archive maintained by Erik Naggum at the University of Oslo in Norway also shadows TEI drafts on its anonymous FTP server, the address of which is

```
ifi.uio.no
```

You can download copies of all current published drafts from this server, in the directory SGML/TEI. Filenaming conventions are the same as those used at the TEI-L fileserver (uppercase only!) Discussion on the TEI-L fileserver is also archived at this site, in monthly batches.