The ACH-ACL-ALLC Text Encoding Initiative: A Brief Overview

1

What is the TEI?

The Text Encoding Initiative (TEI) is a major international project, sponsored jointly by the Association for Computers and the Humanities (ACH), the Association for Computational Linguistics (ACL), and the Association for Literary and Linguistic Computing (ALLC), with funding from the National Endowment for the Humanities, the Commission of the European Communities and the Andrew W. Mellon Foundation. It arose out of a planning conference convened by ACH at Vassar College, Poughkeepsie, New York on 12-13 November 1987, at which three overall objectives were defined for the project:

1. To specify a common interchange format for machine readable texts.
2. To provide a set of recommendations for encoding new textual materials. The recommendations would specify both what features are to be encoded and how those features are to be represented.
3. To document the major existing encoding schemes, and develop a metalanguage in which to describe them.

2

Organization

The TEI is managed by a Steering Committee consisting of two representatives from each of the sponsoring organizations. Fifteen scholarly organizations participate in the TEI by representation on the project's Advisory Board, which approved the plan of work and will be asked to endorse the Guidelines, once they are complete. Two editors, one European and one North American, coordinate the work and are responsible for drafting the TEI guidelines.

3

Publication of Draft Guidelines

At the end of the first two-year development cycle in July 1990, the TEI published the first draft of the Guidelines (TEI P1), a document of almost 300 pages which incorporates the recommendations of four Working Committees, drawn from a broad cross section of the scholarly community, which addressed the topics of Text Documentation, Text Representation, Text Analysis and Interpretation, and Metalanguage and Syntax Issues.[1] The draft Guidelines consist of recommendations of SGML tags for core features common to many text types and to a few specific discipline areas.

---

[1] C.M. Sperberg-McQueen and Lou Burnard, ed., *Guidelines For the Encoding and Interchange of Machine-Readable Texts* (Chicago and Oxford: Text Encoding Initiative, 1990).

4

Work Groups and Affiliated Projects

The major objectives during the second two-year funding cycle are to test the guidelines and extend their scope and coverage in the light of user comments.[2] For the second cycle the TEI has set up a number of small but tightly-focused work groups to tackle specific areas in more detail. So far, about a dozen such groups have been set up, including: character sets, text criticism, hypertext and hypermedia, mathematical formulae and tables, language corpora, physical description of manuscripts, analytic bibliography, general linguistics, spoken texts, literary studies (including prose, drama, and verse), historical studies, machine-readable dictionaries, computational lexica, and terminological data.

The TEI has also set up arrangements with a number of affiliated projects which are testing the Guidelines and reporting back on any problems encountered and how they have been overcome.

5

Bulletin Board

The TEI maintains an electronic bulletin board on which news about the TEI is posted and which provides a forum for detailed discussion of the TEI Guidelines. To subscribe, send an electronic mail containing only the line SUBSCRIBE TEI-L to LISTSERV@UICVM.BITNET or LISTSERV@UICVM.UIC.EDU.

6

The Future

Once the technical specification of the TEI Guidelines has been established, the focus of the TEI will turn more to usability, user education and general promotion of the Guidelines. There is a need to consider the issues of maintenance, further publications, training materials, software, as well as encouragement for existing archives to convert to TEI SGML. The success of the TEI will depend on the ultimate acceptance of the Guidelines by the user community. Involving users as much as possible in their development has been the first stage in this effort. Suggestions for the future are very welcome.

For further information beyond this document contact the editors:

In Europe: Lou Burnard, Oxford University Computing Service, 13 Banbury Rd, Oxford OX2 6NN, England. E-mail: lou@vax.ox.ac.uk. Phone: (+44 865) 273238. Fax :(+44 865) 273275

Elsewhere: C. M. Sperberg-McQueen, Computer Center (M/C 135), University of Illinois at Chicago, 1940 W. Taylor St., Room 124, Chicago IL 60612-7352, USA. Bitnet: u35395@uicvm Internet: u35395@uicvm.uic.edu. Phone: (+1 312) 413-0317 Fax: (+1 312) 996-6834.

---

[2] The second cycle has been extended through June, 1993.