

AAC: Austrian Academy Corpus

Text Encoding Initiative (TEI)

Stephen Miller

Texts & Files: Using Corpora and Digital
Archives / Editions

Institut für Germanistik, Universität Wien

SS 2009

Oxford Text Archive

- Personal initiative of Lou Burnard (OUCS)
- Initial (and continuing) policy of accession: the “dustbin” approach
- Most texts tagged in OCP COCOA format derived (“diverted”) from print publication projects
- Creation of etexts *incidental*

SGML ISO 8879 (1986)

- Development driven by
 - AAP: American Association of Publishers
 - Pentagon (CALIS project)
- Background in problems and issues of large-scale document provision and maintenance in US corporate enterprises

Keystone Concept

- Descriptive Markup

Steven De Rose, "Markup Systems and the Future of Scholarly Text Processing," *Communications of the ACM*, 30 (1987): 933-47.

TEI Sponsors

- Association for Computers and the Humanities
- Association for Computational Linguistics
- Association for Literary and Linguistic Computing

TEI Funding

- US National Endowment for the Humanities
- Directorate General XIII of the Commission of the European Communities
- Andrew W. Mellon Foundation
- Social Science and Humanities Research Council of Canada

“Poughkeepsie Principles” Vassar College, NY (1987)

- Guidelines principles:
 1. Suffice to represent the textual features needed for research;
 2. Be simple, clear, and concrete;
 3. Be easy for researchers to use without special-purpose software;
 4. Allow the rigorous definition and efficient processing of texts;
 5. Provide for user-defined extensions;
 6. Conform to existing and emergent standards.

TEI Guidelines (P3)

- TEI reported in 1994 with:
 - Guidelines for Electronic Text Encoding and Interchange* (TEI P3) (Chicago & Oxford: ACH-ALLC-ACL Text Encoding Initiative, 1994)
 - Set of DTDs
 - “Chicago Pizza Model”

TEI Guidelines (P5)

- Released 1 November 2007
- Complete rewrite of P4
- XML-based
- <http://www.tei-c.org/release/doc/tei-p5-doc/en/html/index.html>

TEI Homepage

- <http://www.tei-c.org>
- Teach Yourself TEI
<http://www.tei-c.org/Support/Learn/tutorials.xml>