

TEIはなぜ日本で知られなかった、知られていないか、知られるようになるか

Why was and is TEI unknown in Japan and will it become better known?

土屋俊(千葉大学)

Syun Tutiya, Chiba University

2006年5月17日

歴史的反省

- 1988年のPoughkeepsie会議へ参加(京都大学 中村順一氏(故人))
- P2、P3作成へ土屋参加(1994年にP3公表)
 - ここまでのキーパーソンは長尾真氏
- P3の国内配布
- Susan Hockey来日
- Michael Sperberg-McQueen, Lou Burnard来日
- 「人文科学とコンピュータ -コンピュータ支援による人文科学研究の推進」(1995-99)に乗らず

いくつかの論点(1)

- プレインテキスト至上主義
 - 1980年代になって、日本語の入力・表示が容易になったために、さらにタグをつけるということについての抵抗が強かった？
 - 文字コードに関する議論がやかましく、文書の構造化の話までいかなかった？
 - ウェブの普及が遅かったので、1990年代ではHTMLすら十分に普及していなかったため、タグをつけることの有益さについて認識できていなかった？

いくつかの論点(2)

- いや、問題はもっと本質的だ。日本語の文書は「構造化」されていない(人文系の意見?)
 - 連綿と書き継ぐスタイルは、構造化になじまない
 - 日本語文書は、むしろ「様式」のような構造化が一般的である。「様式」は二次元の配置によるものであって、階層構造の存在を想定する構造化文書に考え方はなじまない
 - 現実には、HTMLよりPDFが好まれる
 - (いや、返り点、振り仮名など注記によってテキストを可読化する伝統はむしろ日本的である)

いくつかの論点(3)

- ウェブのようなインタラクティブなコミュニケーションがインターネットの本質であるならば、構造化文書標準化のような企図はインターネットでは不自然である
 - たしかにSGMLは使いにくかった
 - たしかに、画像、映像、音声などがインターネット配信されることのほうが話題にはなる
 - メタデータに対する関心は一般に低い

では、将来は？

- 楽観論

- このままやっていけば、いずれみんなタグをつけるようになる
- そのときに存在する標準をつかえばよいじゃないか

- 悲観論

- 日本における「構造化文書嫌い」は本格的であって、簡単には考え方は変わらない
- 標準は構築のための作業が重要であり、あとから勉強して使うのでは無理がある