

TEIガイドライン日本語版プロジェクト

大矢一志
鶴見大学

Christian Wittern
京都大学

日本におけるマークアップ資料の作成を支援するために、TEIガイドライン日本語版を作る計画があり、現在、P4を土台に翻訳が進められている。目標としては、P5の日本語版を作成する予定である。計画の概要と進行状況を紹介し、参加者を募りたい。

Japanese translation project of the TEI Guidelines

OHYA Kazushi
Tsurumi University
Christian Wittern
Kyoto University

In order to provide support to the encoding of texts in Japanese, a project to translate the TEI Guidelines into Japanese has started. At the moment, a rough draft version of large parts of the P4 version has been prepared. This will be enhanced and refined, with the ultimate goal of preparing a Japanese version of P5, as soon as that becomes stable enough to be translated. This poster will present the current state of the work and hopes to attract more collaborators in this project.

1 日本における人文科学資料のマークアップ活動の活性化

本会議が日本で開催され、日本語による討論が行われたことは、今後、日本におけるマークアップ活動や TEI の利用が活発になる、重要な契機になるだろう。但し、日本では、電子テキスト論すら、十分に行われ、コンセンサスが出来上がっているように感じられない。本会議の土屋俊先生の発表にもあるように、文字コード、異字体の認定、字形の論議に関心が偏り、一般的な電子テキスト論や、それに関連したマークアップ付データの論議は、主流として行われてこなかった。結果として、マークアップ活動を検討する基礎資料が、日本語で用意される機会は、大変に限られてきた。一方、欧米では、電子テキスト論や、電子ジャーナルの出版に関する編集技術やビジネスモデル論が活発に論議され、また、電子図書館プロジェクトやコーパスプロジェクトではマークアップ言語が積極的に採用され、実績を上げてきた。とりわけ、TEI の活動は、傍目からは、上手く世代交代を行いながら、論議が継続され、現在では、XML の高度なスキーム操作を含む、新しいモジュールの取り扱いにも論議を広げている。TEI 活動の初期に、マークアップ言語として SGML を採用する際に行われてきた、マークアップ言語一般を論議する姿勢は、現在も、現行 XML 規格の限界を超えた記述手法を模索する活動に見ることができる。日本においても、電子テキスト論や、とりわけ、人文科学資料のマークアップ活動を一般的なものにするには、まず、この溝を埋めるために、先人の論議を独学できる環境を整える必要があると考えている。

この一環として、TEI ガイドライン日本語版の作成を計画している。日本で TEI を実際に使用する場合、日本語資料に合った拡張タグ集合を、独自に開発する必要がある。これは、資料の内容を分析するだけではなく、マークアップ言語の高度な知識も要求され、かつ、TEI スキームの理解も必要となる、大変に難しい作業になる。これを実行するには、少しでも、現場の論議に集中できるような環境を整える必要があり、その第一の資料は、日本語で読める TEI ガイドラインであろう。TEI Lite による TEI 全体像の把握や、TEI を使った事例の解説、TEI そのものの解説よりも、まず、TEI ガイドラインを用意することが重要であると考えている。

TEI ガイドライン日本語版が用意されることにより、マークアップ言語の論議の場が生まれ、資料を実際にマークアップし、その記述 (実験) が発表され、評価される環境が生まれてくると期待している。

2 TEI ガイドライン日本語版作成の進行状況

当初、大矢は、数年前から進めていた P4 の翻訳を公開することを検討していたが、TEI の重要メンバーである京都大学の C.Wittern 先生のアドバイスにより、P5 の日本語版を作成することにした。翻訳の対象は、P5 の 1 章から 31 章を対象としている。そのうち重要な 1 章から 20 章までを、第一期分として計画している。現在、2(P5), 3(P5), 6(P4), 7(P4), 8(P4), 9(P4), 10(P4), 11(P4), 12(P4), 13(P4), 14(P4) 章の下訳がほぼ済んでいる。但し、現在、P5 は draft 段階で、かつ、精力的に修正作業が進められており、この 1 年でも、大幅な修正が行われている (e.g. 3 章 P5)。幾つかの下訳は、完成していたが、使えなくなってしまう¹。

現在は、P5 の編集状況を眺めながら、対訳語彙集の整理を目的とした新規下訳を進め、同時に公開方法を検討している²。P5 の策定作業は、まだ終わりそうになく³、現時点では、本格的な活動は、2007 年以降になると予想される。本プロジェクトに関心のある方は、ohya(at)tei.infon.org までご連絡頂きたい。

¹P4 の 13 章 "Terminological Databases" は廃止され、代わって、"Manuscript Description" が新規の 13 章として用意された。

²現在、TEI は、ODD という、D.Knuth 氏の WEB システムに似た発想の、スキーム定義、解説、出力形式を同時に示す手法を検討しているため、単純に文書のみとするか、ODD 関連の解説文も用意して公開するかといった、公開手法自体も検討しなくてはならない。

³予定としては、2006 年中に正式版を公開するとある。