

TEI研究会2024-12研究会の記録

開催日時

2024年6月17日(月)12:30-14:30

[TEI研究会HP](#)

前回: [TEI研究会2024-11研究会の記録](#)

参加者

氏名(所属等)

王一凡(人文情報学研究所／国立国語研究所)

小風綾乃(東京大学史料編纂所)

永崎研宣(慶應義塾大学／人文情報学研究所)

岡田一祐(慶應義塾大学)

中川奈津子(九州大学)

石田友梨(岡山大学)

横山恵理(大阪工業大学)

蛭沼芽衣(九州大学)

高須賀萌

主な話題

1. 高須賀さんの報告

- 黄表紙本のTEI化
- 誰が何を話しているのか、絵と対応させたい
- 画像をどのように表示させるか
 - 最初からある箇所にzoomさせておいてほしいか、全体像が表示されてほしいか→タグで指定する
- 著者IDとしては、国書データベースの著者IDを使っても良い。
 - 国書DBのURLはTEI化などを想定されているように見える。
 - そうとばかりも言い切れない...(のでNIJL側の整理が必要)
 - <https://kokusho.nijl.ac.jp/author/165899> ← 最後の6桁数字がID
 - VIAFとリンクされているとなお良い。
- 画像・文字をどのように体系的に切り出すか？
 - 以下の対応関係をどのように記述するか？
 - 切り出された画像と文字
 - 台詞と話者

- 人物名と人物画像(これはすでにリンクされている)
- 話者がわからない場合は「話者不明ID」とリンクする。
- 特に画像のXML:IDの付け方は工夫の余地があるかも？
 - 人物名_ページ数、人物名_登場回数0など。
 - 人物名_20などにしておくと、20と30の間にもう1個見つけた場合も、人物名_25などに対応できる。
 - id="f01r_emma" f = folio (丁), rl = right/left などとするのが良さそう。
- page boundary にも <pb n="f02r"/> などと番号を振るとよいのかも？
 - 丁をまたいだ台詞や本文が数多くある。
 - pbタグは付けないという選択肢もある。
 - 本文だけはpbをつけた方が場面転換などを記述できて便利かもしれない。
- 現行は場面転換のときには <div type="scene"> を利用。

2. 翻訳会

- 前回: [TEI研究会2024-09研究会の記録](#)
- 翻訳中: [3 すべてのTEI文書で利用可能な要素](#)
- 元: [3 Elements Available in All TEI Documents - The TEI Guidelines](#)
- 3.6.4 最後の例については、正規化された…に関する一文から再開

3.6.4

- normalized...正規化

3.6.5

- abbreviation...省略形(これまでのこの単語の訳を省略形に統一するか否か)
- expansion...展開形
- copy text...底本
- special processing...個別の処理
- 符号化...打ち込みの時点で符号化という(Unicodeなど)
 - 符号化しないということは打ち込みもしないということに？
 - タグ付けをしないことを明確にする。

❖ 次回、@type属性は、機能によって…の一文から。