

Proposition de formation TEI

Lou Burnard Consulting

20140217

1 Objectifs de la formation

En ce qui concerne la pratique de création des corpus numériques les propositions de la Text Encoding Initiative (TEI) restent incontournables depuis une trentaine d'années. Actuellement basée sur le standard informatique XML, la TEI permet la réutilisation l'intégration et l'échange de ces corpus parmi les scientifiques de toutes les disciplines des sciences humaines et sociales. Un grand communauté mondial reste responsable de son évolution et de sa mise à jour, un communauté surtout scientifique et ouvert.

La TEI est majoritairement utilisée pour les éditions numériques de sources anciennes (par exemple de textes littéraires, de manuscrits du moyen age, de brouillons d'auteurs modernes ...) ou archivistiques (chartes, registres, inscriptions de l'antiquité etc.); elle est essentielle dans la construction des corpus linguistiques et des dictionnaires informatisés, et peut bien s'appliquer également à la construction des corpus oraux ou visuels.

Cette formation propose de mettre en evidence la variété de ce que la TEI rend possible dans la création des ressources numérisés, et de fournir aux participants des expériences pratiques de son usage. Voici une liste prévisionnel des sujets qui nous espérons inclure :

- Initiation au balisage et à XML-TEI. Théorie du balisage. TP sur l'usage d'un éditeur XML.
- L'essentiel de la TEI. Création et validation d'un document minimal TEI.
- Metadonnées et indexage. TP sur la creation d'un entête (TEI Header) simple.
- Metadonnées spécialisés pour les manuscrits. TP sur la création d'un msDesc enrichie.
- Intégration des facsimilés et des transcriptions. TP transcription d'une source manuscrite ou tapuscrite
- Balisage sémantique: traitement de noms de personnes et de lieux. TP sur annotation d'une texte
- Personnalisation de la TEI. TP sur creation et usage d'un systeme spécialisé pour les cartes postales
- Présentation de quelques outils facilitant la dissémination et édition des ressources TEI XML

2 Démarche proposée

La formation se déroulera sur 3 journées dans une salle informatique équipée de postes de travail sous Windows, Mac, ou Linux.

La formation sera composée d'un mélange de discussions générales, de présentations détaillées, et de travaux pratiques. Pour ces derniers, on peut se servir de n'importe quelle éditeur XML, mais oXygen est recommandé parce qu'il est d'habitude installé avec un système TEI complète par défaut.

Tous les matériaux préparés pour la formation (les diapos, les scripts de TP, les fichiers exemplaires etc.) seront mis à la disposition de l'INHA pour réutilisation ou redistribution ultérieure sous licence CC.

3 Considerations pratiques et budgétaires

public Un public restreint est envisagé (10-15 personnes) de participants informés sur les essentiels de l'XML, mais sans connaissance approfondie. Des compétences informatiques générales (par exemple, manipulation de fichiers, mode d'emploi d'un éditeur etc. sont prérequis.

formateurs La formation sera assurée par Lou Burnard.

budget Un budget total pour 3 journées autour de 2500 euros est envisagé, pour suffir aux frais de déplacement (transports Oxford-Paris, 3 nuitées d'hôtel) et d'heures de formation @ 40 euro/heure. Un devis sera fourni