Introduction

Jean-Baptiste Camps & Simon Gabay

Qu'est-ce qu'une édition électronique?

Principes et fondamentaux

fondamentau Comment représes

Pourquoi opter pou

Outils et

La modélisatio comme outil

Les outils numériques aux différentes étapes de la réalisation

Introduction

Philologie numérique et édition électronique

Jean-Baptiste Camps & Simon Gabay

Univ. de Neuchâtel

Formation en philologie numérique : encoder, exploiter, diffuser 12-16 février 2018

Philologie numérique

Introduction

Jean-Baptiste Camps & Simon Gabay

Qu'est-ce qu'une éditior électronique?

Principes et fondamentaux

Comment représent un texte?

une édition électronique?

Outils et méthodes

d'analyse
Les outils numériques
aux différentes étapes
de la réalisation
d'une édition

digital approach to philology [Andrews, 2012]

L'emploi de méthodes computationnelles à toutes les étapes du travail philologique, de la *production* des données (recension, transcription, collation, ...) à leur *analyse*.

Philologie tournée vers les données

une approche qui mette les données au centre :

en amont tirer profit des méthodes computationnelles pour produire des données dans des quantités ou granularités jusque là inenvisageables.

en aval partir des données plutôt que de leur plaquer des hypothèses préexistantes.

au centre les *données* elles-mêmes, formalisées selon un *modèle* représentant les faits étudiés, et accessibles de manière *transparente*.

Que sont les données pour des philologues?

Introduction

Jean-Baptiste Camps & Simon Gabay

Qu'est-ce qu'une édition électronique?

Principes et fondamentaux

un texte?

Pourquoi opter pour une édition électronique?

Outils et méthodes

comme outil d'analyse Les outils numérique aux différentes étape de la réalisation d'uns édicion

- des textes ou, plus souvent, les témoins de textes, conservés notamment dans des manuscrits;
- ayant fait l'objet d'un travail philologique d'édition;
- enrichis par l'expertise ou l'interprétation philologique (annotation, interprétation,...)

Problèmes

- Comment rendre ce travail (interprétatif) philologique le plus *transparent* possible?
- Comment rendre ces données réutilisables par d'autres?
- Comment favoriser les nouveaux questionnements, les nouvelles enquêtes, les nouvelles utilisations de nos données?

Philologie numérique, science des données, science ouverte

Introduction

Stanford Digital Humanities, 2014,

digital-humanities-and-data-science.

https://digitalhumanities.stanford.edu/

Hacking Math & Stats Skills Machine Knowledge Learning Grad School Data Science Officemate Not that Outside dangerous, in Committee NSA Member retrospect That Guy Who Thesis Stole Your Advisor Identity Online Substantive James Bond Evil Expertise Villain

Source: Elijah Meeks, « Digital Humanities and Data Science ».

AVANTAGES

- nouvelles données, nouveaux questionnements;
- cumulabilité ;
- reproductibilité;
- réfutabilité.

FAIR Data

Introduction

Jean-Baptiste Camps & Simon Gabay

Qu'est-ce qu'une édition électronique ?

Principes et fondamentaux

un texte?

Pourquoi opter poune édition électronique?

Methodes

La modélisation comme outil d'analyse Les outils numérique aux différentes étape de la réalisation d'une édition



(Img. SangyaPundir CC BY-SA 4.0)

- Findable: données doivent être faciles à trouver en ligne, y compris sur le long terme (archivage pérenne);
- Accessible : format ouvert, libre, documenté et compréhensible;
- Interoperable: pas enfermées dans un logiciel propriétaire; standards, intégrité des données, documentation;
- Reusable: données ouvertes, réutilisables (légalement, techniquement,...).

Plan

Introduction

Jean-Baptisto Camps & Simon Gabay

Qu'est-ce qu'une édition électronique ?

Principes et fondamentaux

un texte ?
Pourquoi opter poulune édition électronique ?

méthode

d'analyse
Les outils numérique
aux différentes étape
de la réalisation

- 1 Qu'est-ce qu'une édition électronique?
- Principes et fondamentaux de l'édition électronique
 - Comment représenter un texte?
 - Pourquoi opter pour une édition électronique?
- 3 Outils et méthodes de l'édition électronique
 - La modélisation comme outil d'analyse : construire son propre modèle TEI
 - Les outils numériques aux différentes étapes de la réalisation d'une édition

Plan

Introduction

Jean-Baptiste Camps & Simon Gabay

Qu'est-ce qu'une édition électronique?

Principes et fondamentaux

Comment représer

Pourquoi opter poi une édition

Outils et méthodes

La modélisation comme outil d'analyse

Les outils numériques aux différentes étapes de la réalisation d'une édition

- 1 Qu'est-ce qu'une édition électronique?
- Principes et fondamentaux de l'édition électronique
 - Comment représenter un texte?
 - Pourquoi opter pour une édition électronique?
- Outils et méthodes de l'édition électronique
 - La modélisation comme outil d'analyse : construire son propre modèle TEI
 - Les outils numériques aux différentes étapes de la réalisation d'une édition

Intuition

Introduction

Jean-Baptiste Camps & Simon Gabay

Qu'est-ce qu'une édition électronique?

Principes et fondamentaux

un texte?

Pourquoi opter pour une édition

Outils et méthodes

La modélisation comme outil d'analyse

Les outils numériques aux différentes étapes de la réalisation d'une édition

- Un ensemble de fichiers électroniques,
- encodés de manière sémantique, en XML (XML/TEI),
- publiés sous forme de page web,
- et plutôt sur un site institutionnel.



Page d'accueil des Éditions en ligne de l'École des chartes (ÉLEC)

Définition minimaliste

Introduction

Jean-Baptiste Camps & Simon Gabay

Qu'est-ce qu'une édition électronique?

Principes et

tondamental Comment renrése

Pourquoi opter poi

une édition électronique ?

Outils et méthodes

La modélisatio comme outil d'analyse

Les outils numériques aux différentes étapes de la réalisation

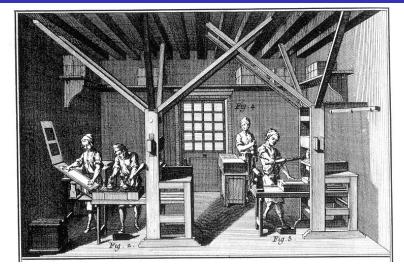
Édition électronique

Une édition qui fait appel, à un stade ou un autre de sa réalisation (établissement du texte et des index, analyse historique, linguistique ou littéraire et commentaire, publication, diffusion...) au medium électronique.

Qu'est-ce qu'une édition qui *n'est pas* électronique?

Introduction

Qu'est-ce qu'une édition électronique?



Encyclopédie, art. «Imprimerie»

Nouvelle définition

Introduction

Jean-Baptisto Camps & Simon Gabay

Qu'est-ce qu'une édition électronique?

Principes et fondamentaux

un texte?

Pourquoi opter pour une édition

Outils et méthodes

La modélisation comme outil d'analyse Les outils numérique aux différentes étape de la réalisation d'une édition

Édition électronique

Une édition dont la forme finale est constituée de fichiers électroniques, et qui est diffusée sous cette forme, soit par un réseau (Internet,...), soit sur un support de stockage (CD, ...).

- Une page html statique est-elle une édition électronique?
- Un fichier pdf ou Open Document constitue-t-il une édition électronique?
- Un ensemble de fichiers de texte mis à disposition constitue-t-il une édition électronique?

Édition numérisée ou numérique?

Introduction

Jean-Baptiste Camps & Simon Gabay

Qu'est-ce qu'une édition électronique?

Principes et fondamentaux

un texte?

Pourquoi opter pou une édition électronique?

Outils et méthode

comme outil d'analyse Les outils numérique aux différentes étape de la réalisation d'une édition

numérisée vs. numérique [Sahle, 2008]

distinction entre édition numérisée (imprimée à l'écran) et édition numérique, i.e., une édition qui ne peut être imprimée sans perte d'information, dont les fonctionnalités ne sont pas entièrement reproductibles par une version imprimée.

Digital scholarly editions are not just scholarly editions in digital media. I distinguish between digital and digitized. A digitized print edition is not a « digital edition » in the strict sense used here. A digital edition can not be printed without a loss of information and/or functionality. The digital edition is guided by a different paradigm. If the paradigm of an edition is limited to the two-dimensional space of the « page » and to typographic means of information representation, then it's not a digital edition

Édition électronique savante / critique

Introduction

Jean-Baptist Camps & Simon Gabay

Qu'est-ce qu'une édition électronique?

Principes et fondamentaux

un texte?
Pourquoi opter pou une édition électronique?

Outils et méthodes

La modélisation comme outil d'analyse

Les outils numériques aux différentes étapes de la réalisation

Édition électronique savante (Digital scholarly edition)

Une édition qui utilise les possibilités de l'outil informatique pour proposer

- une représentation critique ou enrichie sémantiquement d'une source;
- des fonctionnalités utiles à leur compréhension et leur analyse.

Plan

Introduction

Jean-Baptiste Camps & Simon Gabay

Qu'est-ce qu'une édition électronique ?

Principes et fondamentaux

Comment représe

Pourquoi opter po une édition

Outils et méthodes

La modélisation comme outil d'analyse

Les outils numérique aux différentes étape de la réalisation d'une édition

- 1 Qu'est-ce qu'une édition électronique
- Principes et fondamentaux de l'édition électronique
 - Comment représenter un texte?
 - Pourquoi opter pour une édition électronique?
- Outils et méthodes de l'édition électronique
 - La modélisation comme outil d'analyse : construire son propre modèle TEI
 - Les outils numériques aux différentes étapes de la réalisation d'une édition

Comment représenter un texte?

Introduction

Jean-Baptiste Camps & Simon Gabay

Qu'est-ce qu'une édition électronique?

Principes et

fondamentaux

Comment représenter

un texte?

Pourquoi opter po une édition

Outils et

La modélisation comme outil

Les outils numériques aux différentes étapes de la réalisation On emploie *a priori* les italiques pour les locutions et termes empruntés à d'autres langues.

Deux approches possibles

Introduction

Jean-Baptiste Camps & Simon Gabay

Qu'est-ce qu'une éditior électronique?

Principes et fondamentaux

Comment représenter un texte?

Pourquoi opter pour

une édition électronique?

Outils et méthodes

La modélisation comme outil d'analyse

Les outils numérique aux différentes étaps de la réalisation d'une édition On emploie *a priori* les italiques pour les locutions et termes empruntés à d'autres langues.

Procédural

On emploie

/ici commence un texte en italiques/ a priori /ici finit un texte en $% \left(1\right) =\left(1\right) =\left(1\right) +\left(1\right) =\left(1\right) =\left(1\right) +\left(1\right) =\left(1\right$

italiques/ les italiques...

Sémantique

On emploie

/ici commence une locution étrangère/ a priori /ici finit une locution

étrangère/les italiques

Solutions XML

Introduction

Jean-Baptiste Camps & Simon Gabay

Qu'est-ce qu'une édition électronique?

Principes et fondamentaux

Comment représenter

un texte? Pourquoi opter pour

Outils et

La modélisation comme outil

Les outils numériques aux différentes étapes de la réalisation d'une édition

Une solution en XML

On emploie <locutionÉtrangère>a priori</locutionÉtrangère> les italiques pour les locutions et termes empruntés à d'autres langues.

En TEI

On emploie <foreign xml:lang="la">a priori</foreign> les italiques...

Autres solutions XML

Introduction

Jean-Baptiste Camps & Simon Gabay

Qu'est-ce qu'une édition électronique?

Principes et fondamentaux

fondamentaux Comment représenter

un texte?

Pourquoi opter po

une édition électronique?

méthode

La modélisation comme outil d'analyse

Les outils numériques aux différentes étapes de la réalisation d'une édition

Utilisation pédagogique

```
<exemple>On emploie <cas>a priori</cas> les italiques
pour les locutions et termes empruntés
à d'autres langues.
```

Langue et morpho-syntaxe

```
<phrase>
  <mot type="pronom" subtype="personnel">On</mot>
  <mot type="verbe" subtype="conjugué">emploie</mot>
  <!-- etc. -->
</phrase>
```

eXtensible (2) Markup Language (1)

Introduction

Camps & Simon Gabay

Qu'est-ce qu'une édition électronique ?

Principes et fondamentaux

Comment représenter

Pourquoi opter p

électronique?

méthodes

La modélisation comme outil d'analyse

Les outils numériques aux différentes étapes de la réalisation

- 1 Le ML de XML : markup language, XML est un langage à balises;
- 2 le X de XML : XML est eXtensible (i.e., il ne propose pas un jeu prédéfini et fermé de balises, mais une syntaxe, et des règles sur ce que doit être un document XML bien formé, un document valide et sur l'écriture de grammaires spécifiques).

XML comme langage structuré (par des balises)

```
Introduction
```

Jean-Baptiste Camps & Simon Gabay

Qu'est-ce qu'une édition électronique?

Principes et fondamentaux

Comment représenter un texte?

Pourquoi opter pou une édition électronique?

Outils et méthodes

La modélisation comme outil d'analyse

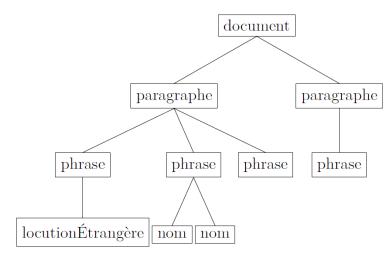
Les outils numérique aux différentes étape de la réalisation d'une édition

```
<document>
 <paragraphe>
    <phrase>
      On emploie <locutionÉtrangère>a priori</locutionÉtrangère>
      les italiques pour les locutions et termes empruntés
      à d'autres langues.
    </phrase>
    <phrase>
      On emploie souvent les petites capitales pour les noms propres
      comme <nom>Léopold Delisle</nom> ou <nom>Jules Quicherat</nom>
    </phrase>
    <phrase>
      On emploie en revanche généralement le gras pour des raisons
      coupables.
    </phrase>
 </paragraphe>
 <paragraphe>
    <phrase>
      Un second paragraphe...
    </phrase>
 </paragraphe>
</document>
```

Les documents XML comme arborescences

Introduction

Comment représenter



Arborescence de l'exemple précédent

Arborescence plus précise

Introduction

Jean-Baptiste
Camps &
Simon Gabay

Qu'est-ce qu'une édition électronique?

Principes et

Comment représenter

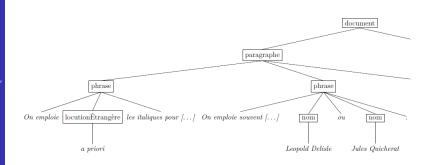
un texte?

une édition

Outils et méthodes

La modélisatio

Les outils numériqu aux différentes étap de la réalisation



Puissance et limitation de XML

Introduction

Jean-Baptiste Camps & Simon Gabay

Qu'est-ce qu'une éditior électronique?

Principes et fondamentaux Comment représenter

un texte?

Pourquoi opter pou

Outils et méthodes

La modélisatio comme outil d'analyse

Les outils numériques aux différentes étapes de la réalisation d'une édition

Bien formé

<auteur>

<nom>Dupont</nom>

</auteur>

Pas bien formé

<auteur><nom>Dupont</auteur></nom>

Toujours pas bien formé

```
<poème type="douteux">
```

<vers>À la fin de ce vers, je dis : <locutionÉtrangère>Sic</vers>
<vers> transit gloria mundi</locutionÉtrangère></vers>

</poème>

Le modèle OHCO [Renear et al., 1993]

Introduction

Jean-Baptiste Camps & Simon Gabay

Qu'est-ce qu'une édition électronique?

Principes et fondamentaux Comment représenter

Pourquoi opter pou une édition

Outils et méthodes

> comme outil d'analyse Les outils numériq aux différentes éta de la réalisation

Une perspective d'analyse d'un texte peut être décomposée en une Ordered Hierarchy of Content Objects.

Ordered car le texte à une direction (du début vers la fin);

Hierarchy car les éléments ont entre eux une relation
hiérarchique (un chapitre contient des
paragraphes, qui contiennent des phrases...)

Content Objects d'objets de contenu quelconques, terme suffisamment vague pour englober un peu tout, selon l'analyse portée sur le texte.

Plutôt hierarchies que *hierarchy* : une par approche ou analyse d'un texte.

XML comme métalangage (eXtensible)

Introduction

Jean-Baptiste Camps & Simon Gabay

Qu'est-ce qu'une éditior électronique ?

Principes et fondamentaux

Comment représenter un texte?

Pourquoi opter p une édition

électronique?

méthodes

La modélisati

d'analyse

Les outils numérique
aux différentes étape
de la réalisation
d'une édition

XML est un métalangage, c'est-à-dire qu'il permet de définir différents langages (des applications) avec chacun leur grammaire propre (par exemple la *Text Encoding Initiative* ou TEI; l'*Encoded Archival Description* ou EAD; ou XHTML). Contrairement à d'autres langages à balise, XML ne propose pas un jeu de balise prédéfini, mais un ensemble d'éléments et de règles sur ce que doit être un document bien formé ou un document valide, et la manière de créer et mettre en œuvre des langages.

Pourquoi opter pour une édition électronique? Rappel de notre définition initiale

Introduction

Jean-Baptiste Camps & Simon Gabay

Qu'est-ce qu'une édition électronique ?

Principes et fondamentaux

Comment représent

Pourquoi opter pour une édition électronique?

Outils et méthodes

La modélisatio comme outil d'analyse

Les outils numériques aux différentes étapes de la réalisation

Édition électronique savante (Digital scholarly edition)

Une édition qui utilise les possibilités de l'outil informatique pour proposer

- une représentation critique ou enrichie sémantiquement d'une source ;
- des fonctionnalités utiles à leur compréhension et leur analyse.

L'encodage sémantique et ses avantages

Introduction

Jean-Baptiste Camps & Simon Gabay

Qu'est-ce qu'une édition électronique?

Principes et fondamentaux

Pourquoi opter pour une édition électronique?

Outils et méthodes

La modélisatio comme outil d'analyse

Les outils numériques aux différentes étapes de la réalisation

- Rend le "sémantisme implicite" (ou, plutôt, l'interprétation, l'analyse critique) d'un document compréhensible par un ordinateur et l'explicite;
- Mise en forme indépendante du document, gérée à part et (re)configurable;
- Permet une exploitation utilisant le texte encodé comme base de données (fouille de données, information retrieval, traitement automatique du langage, etc.)

Les avantages techniques d'XML

Introduction

Jean-Baptiste Camps & Simon Gabay

Qu'est-ce qu'une éditior électronique?

Principes et fondamentaux

Comment représent un texte?

Pourquoi opter pour une édition électronique?

Outils et méthodes

La modélisatio comme outil d'analyse

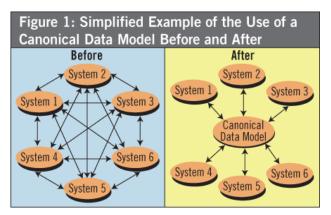
Les outils numériques aux différentes étapes le la réalisation l'une édition

- Pérennité et interopérabilité,
- multiplicité des usages à partir d'un même document
 - mise en forme;
 - exploitation (recherche avancée, analyse quantitative, etc.);
 - réutilisation;
- contenu lisible par l'humain autant que par l'ordinateur;
- environnement de travail.

XML comme format pivot

Introduction

Pourquoi opter pour électronique?



extr. de Steve Hoberman. « Canonical Data Model ». Information Management Magazine, en ligne: http://www.

information-management.com/issues/2007 50/10001733-1.html (consulté le 14 nov. 2014).

Multiplicité des usages : ex. d'un chaîne éditoriale qui combine analyse et publication

Introduction

Jean-Baptiste Camps & Simon Gabay

Qu'est-ce qu'une édition électronique?

Principes et fondamentaux

fondamenta

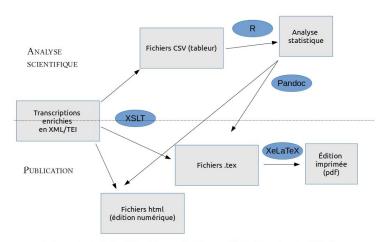
Pourquoi opter pour une édition électronique?

méthodes

La modélisatio comme outil d'analyse

es outils numériques aux différentes étapes

aux différentes étapes de la réalisation



Traitements et chaîne éditoriale pour la Chanson d'Otinel (version simplifiée)

Publier sous différents formats

Introduction

Jean-Baptiste Camps & Simon Gabay

Qu'est-ce qu'une éditior électronique ?

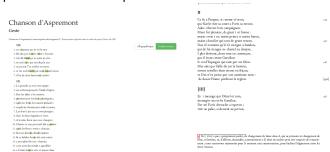
Principes et fondamentaux

Pourquoi opter pour une édition électronique?

Outils et

La modélisation comme outil d'analyse

Les outils numériques aux différentes étapes de la réalisation Convertir en xhtml pour une publication en ligne ou en La pour une édition imprimée...



Ex. de conversions XML/TEI vers d'autres formats, http://www.tei-c.org/oxgarage/.

Analyser ses données

Introduction

Jean-Baptiste Camps & Simon Gabay

Qu'est-ce qu'une édition électronique?

Principes et fondamentaux

Comment représe

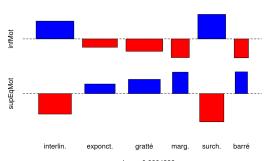
Pourquoi opter pour une édition électronique?

méthode:

comme outil d'analyse

Les outils numériques aux différentes étapes de la réalisation d'une édition

Modes de correction et niveau dans tout le corpus (modalités d'eff. >= 5)



p.value = 0.0004998

Modes de correction employés par les scribes et niveau de ces corrections (ms. de la Chanson d'Otinel)

Concevoir de nouvelles fonctionnalités

Introduction

Jean-Baptiste Camps & Simon Gabay

Qu'est-ce qu'une édition électronique?

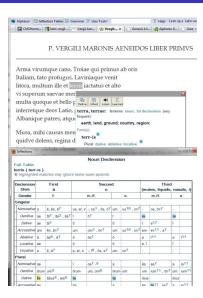
Principes et fondamentaux

Pourquoi opter pour une édition électronique?

Outils et

La modélisation comme outil

> Les outils numériques aux différentes étapes de la réalisation



Afficher les paradigmes morphologiques ou l'annotation syntaxique

Nouvelles fonctionnalités – annotation syntaxique

Introduction

Jean-Baptiste Camps & Simon Gabay

Qu'est-ce qu'une éditior électronique?

Principes et fondamentaux

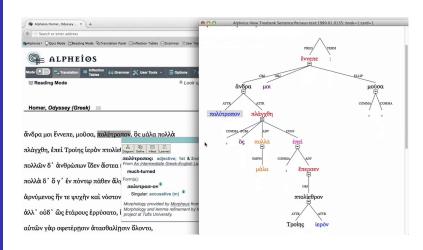
Comment représen un texte?

Pourquoi opter pour une édition électronique ?

Outils et méthode

La modélisatio comme outil d'analyse

> Les outils numériques aux différentes étapes de la réalisation d'une édition



Plan

Introduction

Jean-Baptiste Camps & Simon Gabay

Qu'est-ce qu'une éditio électronique ?

Principes et fondamentaux

un texte?

Pourquoi opter pou
une édition
électronique?

Outils et méthodes

La modélisation comme outil d'analyse

Les outils numérique aux différentes étape de la réalisation d'une édition

- Qu'est-ce qu'une édition électronique?
 - Principes et fondamentaux de l'édition électronique
 - Comment représenter un texte?
 - Pourquoi opter pour une édition électronique?
- 3 Outils et méthodes de l'édition électronique
 - La modélisation comme outil d'analyse : construire son propre modèle TEI
 - Les outils numériques aux différentes étapes de la réalisation d'une édition

La modélisation comme outil d'analyse : construire son propre modèle TEI

Introduction

Jean-Baptisto Camps & Simon Gabay

Qu'est-ce qu'une éditior électronique ?

Principes et fondamentaux

un texte?

Pourquoi opter po
une édition
électronique?

Outils et méthodes

La modélisation comme outil d'analyse

Les outils numérique aux différentes étape de la réalisation d'une édition

Modélisation

Opération par laquelle on établit le modèle d'un système complexe, afin d'étudier plus commodément et de mesurer les effets sur ce système des variations de tel ou tel de ses éléments composants »

(J. Giraud, P. Pamart, J. Riverain, Les nouveaux mots « dans le vent », Paris, France, 1974).

Définir un modèle adapté :

- aux documents que l'on édite;
- à la perspective scientifique dans laquelle l'édition est réalisée;
- aux moyens (temps, financement, ...) dont on dispose.

Édition d'un imprimé ancien : structure physique et logique (et niveau de granularité)

Introduction

Jean-Baptiste Camps & Simon Gabay

Qu'est-ce qu'une édition électronique ?

Principes et fondamentaux

un texte?

Pourquoi opter pour une édition

Outils et méthodes

La modélisation comme outil d'analyse

Les outils numériques aux différentes étapes de la réalisation d'une édition

- représentation de la structure physique de la source (feuillets, pages, lignes, etc.);
- représentation de la structure logique (parties, chapitres, etc.);

Granularité

degré de finesse ou précision d'un modèle, conçu comme le niveau de son plus petit composant. Plus la granularité est grande, plus on descend dans la modélisation (niveau phrase, mot, graphème, etc.) – et plus on ajoute de balises.

Le développement de modèles spécifiques

Introduction

Jean-Baptiste Camps & Simon Gabay

Qu'est-ce qu'une édition électronique ?

Principes et fondamentaux

Pourquoi opter pou une édition électronique?

Outils et méthodes

La modélisation comme outil d'analyse

Les outils numérique aux différentes étaps de la réalisation d'une édition

Les actes diplomatiques édités par l'École des chartes

Voir le modèle développé à l'École pour l'édition d'un acte en TEI : http:

//developpements.enc.sorbonne.fr/diple/schema/acte Exemple de la définition de l'élément <seg>, de son attribut @function, et de son jeu de valeur défini (localement), 'invocation', 'intitulatio', 'salutatio', etc. (qui renvoient à la version en ligne du Vocabulaire international de la diplomatique); http://developpements.enc.sorbonne.fr/diple/schema/acte#def_diplomatique

Les outils numériques aux différentes étapes de la réalisation (et de la vie) d'une édition

Introduction

Jean-Baptiste Camps & Simon Gabay

Qu'est-ce qu'une éditior électronique?

Principes et fondamentaux

un texte?

Pourquoi opter poi une édition

Outils et méthodes

La modélisatio comme outil d'analyse

Les outils numériques aux différentes étapes de la réalisation

- transcription, collation;
- établissement du texte et de l'édition;
- vérification de la cohérence et harmonisation;
- analyse des données;
- publication, diffusion, interrogation;
- indexation;
- archivage.

Alignement texte/image et OCR/HTR

Introduction

Jean-Baptiste Camps & Simon Gabay

Qu'est-ce qu'une éditior électronique ?

Principes et fondamentaux

un texte?

Pourquoi opter pour une édition électronique?

Outils et méthodes

La modélisation comme outil d'analyse

Les outils numériques aux différentes étapes de la réalisation d'une édition L'alignement texte et image que permet (notamment) la TEI peut tant servir à l'établissement d'édition à visée paléographique ou linguistique, que permettre la fourniture de données utiles pour perfectionner la reconnaissance des écritures manuscrites ou l'OCR d'imprimés anciens

- Projet ORIFLAMMS, http://www. agence-nationale-recherche.fr/projet-anr/?tx_ lwmsuivibilan_pi2[CODE]=ANR-12-CORP-0010;
- Uwe Springmann and David Kaumanns, Ocrocis: a high accuracy OCR method to convert early printings into digital text, 2015,

http://cistern.cis.lmu.de/ocrocis/tutorial.pdf.

La collation des témoins

Introduction

Jean-Baptiste Camps & Simon Gabay

Qu'est-ce qu'une éditior électronique ?

Principes et fondamentaux

un texte?

Pourquoi opter pou une édition

Outils et méthodes

La modélisatio comme outil d'analyse

Les outils numériques aux différentes étapes de la réalisation d'une édition Différents outils permettent de collationner plusieurs témoins d'un même texte, d'enregistrer les variantes, voire de visualiser les différentes rédactions à partir de l'encodage de l'apparat critique.

- CollateX, http://collatex.net/
- Juxta (version en ligne :
 http://www.juxtacommons.org/
- Versioning Machine, http://www.v-machine.org/

Cohérence et harmonisation

Introduction

Jean-Baptisto Camps & Simon Gabay

Qu'est-ce qu'une éditior électronique ?

Principes et fondamentaux

un texte?
Pourquoi opter pou

Outils et

La modélisatio comme outil

Les outils numériques aux différentes étapes de la réalisation d'une édition Toutes sortes d'outils existent pour vérifier l'encodage du document et sa cohérence. À titre d'exemple, on peut citer celui développé par Marjorie Burghart pour l'apparat critique (travail en cours), http:

//ciham-digital.huma-num.fr/teitoolbox/index.php

Analyse stemmatologique

Introduction

Jean-Baptiste Camps & Simon Gabay

Qu'est-ce qu'une éditior électronique ?

Principes et fondamentaux

Pourquoi opter pour une édition

Outils et méthodes

La modélisation comme outil d'analyse

Les outils numériques aux différentes étapes de la réalisation d'une édition Des outils permettent aussi de visualiser ou analyser une tradition textuelle, et de réaliser des analyses stemmatologiques.

- Stemmaweb (Tara Andrews), en ligne : http://stemmaweb.net/
- Stemweb, sources en ligne: https://github.com/Stemweb/Stemweb
- Stemmatology (module pour R, dév. en cours), sources en ligne : https:

//github.com/Jean-Baptiste-Camps/stemmatology

Indexation, recherche, ...

Introduction

Jean-Baptiste Camps & Simon Gabay

Qu'est-ce qu'une édition électronique ?

Principes et fondamentaux

Pourquoi opter pour une édition électronique?

Outils et méthodes

La modélisation comme outil d'analyse

Les outils numériques aux différentes étapes de la réalisation d'une édition L'édition électronique offre des facilités d'indexation et de recherche dans les matériaux édités, grâce à l'encodage des éléments concernés. Certains moteurs de recherche peuvent tirer profit de cet encdoage, comme Philologic,

http://philologic.uchicago.edu/.

La TEI permet aussi de lier les données indexées, par exemple à des référentiels d'autorité comme VIAF (Virtual International Authority File).

```
<author xml:lang="lat" corresp="Gregorius I (pape ; 0540?-0604)"
    ref="http://viaf.org/viaf/100184667">
    <persName>Gregorius I papa</persName>
</author>
<title type="uniform" xml:lang="lat" corresp="Moralia in Job"
    ref="http://viaf.org/viaf/183885855">Moralium sancti
    Gregorii in Job</title>
```

Exemple tiré du mémoire de L. Lebarbey (dir. D. Stutzmann).

Transformations, imports, exports,...

Introduction

Jean-Baptisto Camps & Simon Gabay

Qu'est-ce qu'une édition électronique ?

Principes et fondamentaux

Comment représe

Pourquoi opter pour une édition électronique?

Outils et méthodes

La modélisation comme outil d'analyse

Les outils numériques aux différentes étapes de la réalisation d'une édition

- Feuilles de style du Consortium TEI (dév. S. Rahtz)
- oXgarage (http://oxgarage.oucs.ox.ac.uk: 8080/ege-webclient/)
- Diple (http://developpements.enc.sorbonne.fr/ diple/demo/tei_html)
- Algone (http://algone.net/)

Publication et diffusion en ligne

Introduction

Jean-Baptisto Camps & Simon Gabay

Qu'est-ce qu'une édition électronique ?

Principes et fondamentaux

un texte?

Pourquoi opter pour une édition

Outils et

La modélisatio comme outil d'analyse

Les outils numériques aux différentes étapes de la réalisation d'une édition Des Content management systems (CMS) permettent de publier (relativement) facilement de la TEI sur Internet

- CMS spécifiques : TEI Boilerplate, Kiln, Lodel, Nemo, ...
- CMS généralistes avec un module TEI (Drupal et TEI-Chi, Oméka et TeiDisplay, ...)

Analyse de corpus, statistiques

Introduction

Jean-Baptiste Camps & Simon Gabay

Qu'est-ce qu'une édition électronique?

Principes et fondamentaux

Pourquoi opter pou une édition

Outils et méthodes

La modélisation comme outil d'analyse

Les outils numériques aux différentes étapes de la réalisation d'une édition Tirer profit de l'encodage pour traiter le texte édité comme une base de données, et en extraire des informations

- base de données de variantes;
- cooccurrences de noms de personnes, de lieu;
- tables de fréquence lexicale;
- ...

que l'on pourra ensuite traiter statistiquement via un outil adapté

- TXM;
- R..

Archiver et pérenniser

Introduction

Jean-Baptiste Camps & Simon Gabay

Qu'est-ce qu'une édition électronique?

Principes et fondamentaux

un texte?

Pourquoi opter pou une édition électronique?

Outils et méthodes

La modélisation comme outil d'analyse

Les outils numériques aux différentes étapes de la réalisation d'une édition

Disponibilité,

- sur le long terme,
- via un standard,
- sous une licence adaptée.

Pour une liste des entrepôts de recherche : re3data.org, https: //www.re3data.org/

Archive d'une institution

 University of Oxford Text Archive, http://ota.ox.ac.uk/.
 Ex., UOTA 002, originellement encodé en 1976.

Entrepôt de données de la recherche

- Zenodo (CERN et OpenAIRE), https://zenodo.org/;
- Nakala (HumaNum),
- ...

Bibliographie

Introduction

Jean-Baptiste Camps & Simon Gabay

- ANDREWS (Tara), « The third way : philology and critical edition for a digital age », *Variants*, 10 (2012), http://boris.unibe.ch/43071/.
- CAMPS (Jean-Baptiste), « Où va la philologie numérique? », Fabula-LHT, 20 (2018), http://www.fabula.org/lht/20/camps.html.
- RENEAR (Allen H.), MYLONAS (Elli), et DURAND (David), «Refining our notion of what text really is: The problem of overlapping hierarchies», 1993, en ligne, http://www.ideals.illinois.edu/handle/2142/9407
- SAHLE (Patrick), Catalog of: Digital Scholarly Editions, Cologne, 2008, http://www.digitale-edition.de/.

Approfondissements: TEI et XML

Introduction

Jean-Baptiste Camps & Simon Gabay

- Ron Van den Branden, Melissa Terras, Edward Vanhoutte, TEI by Example, http://teibyexample.org/.
- TEI Consortium, TEI P5: Guidelines for Electronic Text Encoding and Interchange, http://www.tei-c.org/ release/doc/tei-p5-doc/en/html/.
- Lou Burnard, What is the Text Encoding Initiative?: How to add intelligent markup to digital resources, Marseille: OpenEdition Press, 2014, http://books.openedition.org/oep/426.
- Elliotte Rusty Harold, W. Scott Means, *XML en concentré : manuel de référence*, 3^e édition, Paris, 2005.