

Transcrire un document médiéval en XML : Pour qui et comment ?

Alexei Lavrentev¹ & Nicolas Mazziotta²

¹ICAR, ENS-LSH, Université de Lyon

²Département de Langues et Littératures Romanes, Université de Liège

11 avril 2008, Villejuif

École doctorale

« Langage, langue : théorisation, description, transmission »

Plan (NM)

1 Introduction : l'édition de documents médiévaux (AL)

Plan (NM)

- 1 Introduction : l'édition de documents médiévaux (AL)
- 2 Présentation d'XML : techniques de base (NM)

Plan (NM)

- 1 Introduction : l'édition de documents médiévaux (AL)
- 2 Présentation d'XML : techniques de base (NM)
- 3 Transcription diffractée (AL)

Plan (NM)

- 1 Introduction : l'édition de documents médiévaux (AL)
- 2 Présentation d'XML : techniques de base (NM)
- 3 Transcription diffractée (AL)
- 4 Transcription monolithique (NM)

Plan (NM)

- 1 Introduction : l'édition de documents médiévaux (AL)
- 2 Présentation d'XML : techniques de base (NM)
- 3 Transcription diffractée (AL)
- 4 Transcription monolithique (NM)
- 5 Vers un compromis ? (AL)

Plan (NM)

- 1 Introduction : l'édition de documents médiévaux (AL)
- 2 Présentation d'XML : techniques de base (NM)
- 3 Transcription diffractée (AL)
- 4 Transcription monolithique (NM)
- 5 Vers un compromis ? (AL)
- 6 Conclusion (NM)

- 1 Introduction : l'édition de documents médiévaux (AL)
 - Pluralité des publics et des éditions
 - Exemple : les *Serments de Strasbourg* (14 février 842)
 - Limites de l'édition traditionnelle
- 2 Présentation d'XML : techniques de base (NM)
- 3 Transcription diffractée (AL)
- 4 Transcription monolithique (NM)
- 5 Vers un compromis ? (AL)
- 6 Conclusion (NM)

Pluralité des publics et des éditions

Problème

Représentation d'un document textuel historique dans une édition scientifique.

Pluralité des publics et des éditions

Problème

Représentation d'un document textuel historique dans une édition scientifique.

Solutions traditionnelles

Pluralité des publics et des éditions

Problème

Représentation d'un document textuel historique dans une édition scientifique.

Solutions traditionnelles

- Deux extrémités : **facsimilé** (photographie) vs **traduction** en langue moderne

Pluralité des publics et des éditions

Problème

Représentation d'un document textuel historique dans une édition scientifique.

Solutions traditionnelles

- Deux extrémités : **facsimilé** (photographie) vs **traduction** en langue moderne
- Intermédiaires : transcription qui s'approche plus ou moins de

Pluralité des publics et des éditions

Problème

Représentation d'un document textuel historique dans une édition scientifique.

Solutions traditionnelles

- Deux extrémités : **facsimilé** (photographie) vs **traduction** en langue moderne
- Intermédiaires : transcription qui s'approche plus ou moins de
 - ① représentation fine des **détails graphiques** de l'original pour les paléographes, linguistes étudiant les systèmes graphiques

Pluralité des publics et des éditions

Problème

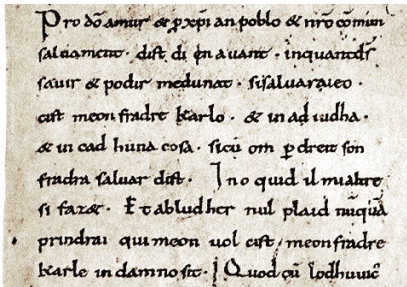
Représentation d'un document textuel historique dans une édition scientifique.

Solutions traditionnelles

- Deux extrémités : **facsimilé** (photographie) vs **traduction** en langue moderne
- Intermédiaires : transcription qui s'approche plus ou moins de
 - ① représentation fine des **détails graphiques** de l'original pour les paléographes, linguistes étudiant les systèmes graphiques
 - ② normes typographiques et usages de lecture **modernes** pour les historiens, littéraires, majorité des linguistes

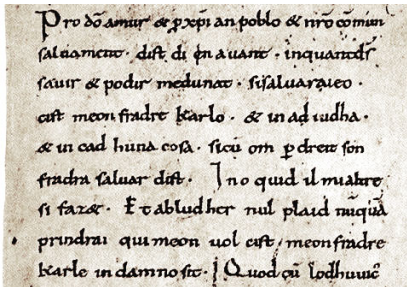
Exemple : les *Serments de Strasbourg* (14 février 842)

Manuscrit



Exemple : les *Serments de Strasbourg* (14 février 842)

Transcription « normalisée » (Lauer 1926)

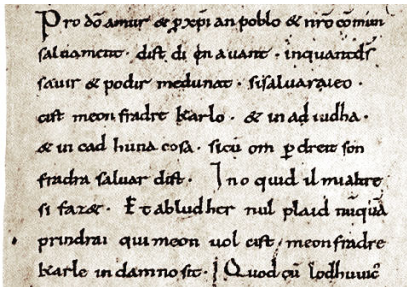


Pro deo amur & p[ro]xi an[im]o p[ro]p[ri]o & n[ost]ro comun
saluam[en]to. dist[inct]i di[ct]i q[ui] a uant. inquant[um]
saur[is] & podur medunat. si saluar[is] queo.
cist meon fradre karlo. & in ad iudha.
& in ead huna cosa. sicut om[n]i p[er] dret[ur] son
fradra saluar dist[inct]. Ino quid il mutare
si fazet. Et ab ludher nul plaid ni quid
prindrai qui meon uol cist / meon fradre
karle in damno sit. | Quod cu[m] lothhuic

« Pro Deo amur et pro christian p[ro]p[ri]o et n[ost]ro com-
mun saluam[en]to, d'ist di in^b avant, in quant Deus savir
« et podir² me dunat, si salvarai^c eo cist meon fradre
« Karlo et in aiudha^d et³ in cadhuna cosa, si cum om per
« dreit son fradra saluar dist^{e4}, in o quid il mi altres
« fazet et ab Ludher nul plaid nunquam prindrai, qui,
« meon vol, cist meon fradre Karle in damno sit. »

Exemple : les *Serments de Strasbourg* (14 février 842)

Transcription « imitative » (Koschwitz 1920)



1 Pro dō amur & p xpian poblo & nrō cōmun | salua- 6
ment. dist di en avant: inquantd | saur & podir medunat.
2 si saluarāieo. | cist meon fradre karlo. & in ad iudha. |
& in cad huna cosa. sicū om p dreit son | fradra saluar 10
dist. Ino quid il mialtre | si fazet. Et abludher nul plaid
6 nūquā | prindrai qui meon uol cist meon fradre | karle in
damno sit.

Exemple : les *Serments de Strasbourg* (14 février 842)

Édition synoptique (Wagner et Collet 1995)

Pro dō amur & px̄pian poblo & nrō cōmun
saluament · dist di ꝑn auant · inquantd's
saur & podir medunat · sisaluarai eo ·
cist meon fradre karlo · & in aḡ iudha ·
& in cad huna cosa · sicū om ꝑ dreit son
fradra saluar dift · I no quid il mialtre
si faz& · Et abludher nul plaid nūquā
prindrai qui meon uol cist meon fradre
karle in damno sit ·

Pro Deo amur et pro christian poblo et nostro
commun saluament, d'ist di in auant, in quant
Deus sauir et podir me dunat, si saluarai eo
cist meon fradre Karlo et in aiudha et in
cadhuna cosa, si cum om per dreit son fradra
saluar dift, in o quid il mi altresī fazet. Et ab
Ludher nul plaid nunqua·m prindrai qui meon
uol cist meon fradre Karle in damno sit.

Limites de l'édition traditionnelle

Limitations liées au *medium* et à la tradition

Limites de l'édition traditionnelle

Limitations liées au *medium* et à la tradition

- ❶ Choix limité : impossible de satisfaire tous les publics

Limites de l'édition traditionnelle

Limitations liées au *medium* et à la tradition

- 1 Choix limité : impossible de satisfaire tous les publics
- 2 Coût et « lourdeur » d'éditions synoptiques

Limites de l'édition traditionnelle

Limitations liées au *medium* et à la tradition

- 1 Choix limité : impossible de satisfaire tous les publics
- 2 Coût et « lourdeur » d'éditions synoptiques
- 3 Absence de stabilité dans la tradition « diplomatique »

Limites de l'édition traditionnelle

Limitations liées au *medium* et à la tradition

- 1 Choix limité : impossible de satisfaire tous les publics
 - 2 Coût et « lourdeur » d'éditions synoptiques
 - 3 Absence de stabilité dans la tradition « diplomatique »
- Quelles solutions offrent le *medium* électronique ?

- 1 Introduction : l'édition de documents médiévaux (AL)
- 2 **Présentation d'XML : techniques de base (NM)**
 - Texte et balisage
 - Structure hiérarchique
 - Conclusion : Potentiel et limitations
- 3 Transcription diffractée (AL)
- 4 Transcription monolithique (NM)
- 5 Vers un compromis ? (AL)
- 6 Conclusion (NM)

Texte et balisage

Texte et balisage

```
<balisage>texte</balisage>
```

Texte et balisage

Texte et balisage

```
<balisage>texte</balisage>
```

Balise de début (nom : *balisage*)

Texte et balisage

Texte et balisage

```
<balisage>texte</balisage>
```

Balise de fin (même nom)

Texte et balisage

Texte et balisage

```
<balisage>texte</balisage>
```

```
<balisage attribut="valeur">texte</balisage>
```

Texte et balisage

Texte et balisage

```
<balisage>texte</balisage>
```

```
<balisage attribut="valeur">texte</balisage>
```

Attribut (paire *clé* = "*valeur*")

Texte et balisage

Texte et balisage

```
<balisage>texte</balisage>
```

```
<balisage attribut="valeur">texte</balisage>
```

Application : le XHTML (pour les pages web)

```
<h1>Introduction</h1>
```

```
<p>Texte de l'introduction</p>
```

Texte et balisage

Texte et balisage

```
<balisage>texte</balisage>
```

```
<balisage attribut="valeur">texte</balisage>
```

Application : le XHTML (pour les pages web)

```
<h1>Introduction</h1>
```

```
<p>Texte de l'introduction</p>
```

Texte

Texte et balisage

Texte et balisage

```
<balisage>texte</balisage>
```

```
<balisage attribut="valeur">texte</balisage>
```

Application : le XHTML (pour les pages web)

```
<h1>Introduction</h1>
```

```
<p>Texte de l'introduction</p>
```

Balisage

Texte et balisage

Texte et balisage

```
<balisage>texte</balisage>
```

```
<balisage attribut="valeur">texte</balisage>
```

Application : le XHTML (pour les pages web)

```
<h1>Introduction</h1>
```

```
<p>Texte de l'introduction</p>
```

h1 indique les limites du titre

Texte et balisage

Texte et balisage

```
<balisage>texte</balisage>
```

```
<balisage attribut="valeur">texte</balisage>
```

Application : le XHTML (pour les pages web)

```
<h1>Introduction</h1>
```

```
<p>Texte de l'introduction</p>
```

p marque le paragraphe

Texte et balisage

Texte et balisage

```
<balisage>texte</balisage>
```

```
<balisage attribut="valeur">texte</balisage>
```

Application : le XHTML (pour les pages web)

```
<h1>Introduction</h1>
```

```
<p>Texte de l'introduction</p>
```

Disjonction entre **forme** et le **fond**

Structure hiérarchique

Hiérarchie du balisage

Une balise peut en **contenir** une autre :

```
<balise><balise>Texte</balise></balise>
```

Structure hiérarchique

Hiérarchie du balisage

Une balise peut en **contenir** une autre :

```
<balise><balise>Texte</balise></balise>
```

Il n'y a pas de limites à l'emboîtement. Mais le balisage ne peut se chevaucher.

Structure hiérarchique

Hiérarchie du balisage

Une balise peut en **contenir** une autre :

```
<balise><balise>Texte</balise></balise>
```

Il n'y a pas de limites à l'emboîtement. Mais le balisage ne peut se chevaucher.

Application : XHTML

```
<div>
  Txt du cadre
  <div> Txt du 1er sous-cadre </div>
  <div> Txt du 2e sous-cadre </div>
</div>
```

Structure hiérarchique

Hiérarchie du balisage

Une balise peut en **contenir** une autre :

```
<balise><balise>Texte</balise></balise>
```

Il n'y a pas de limites à l'emboîtement. Mais le balisage ne peut se chevaucher.

Application : XHTML

```
<div>
  Txt du cadre
  <div> Txt du 1er sous-cadre </div>
  <div> Txt du 2e sous-cadre </div>
</div>
```

Structure hiérarchique

Hiérarchie du balisage

Une balise peut en **contenir** une autre :

```
<balise><balise>Texte</balise></balise>
```

Il n'y a pas de limites à l'emboîtement. Mais le balisage ne peut se chevaucher.

Application : XHTML

```
<div>
  Txt du cadre
  <div> Txt du 1er sous-cadre </div>
  <div> Txt du 2e sous-cadre </div>
</div>
```

Principes utiles pour **indiquer la structure** des textes à l'ordinateur

Conclusion : Potentiel et limitations

XML nous aide...

Conclusion : Potentiel et limitations

XML nous aide...

- À exprimer notre analyse dans un langage **compréhensible par l'ordinateur** (traitement automatique)

Conclusion : Potentiel et limitations

XML nous aide...

- À exprimer notre analyse dans un langage compréhensible par l'ordinateur (traitement automatique)
- À **différencier la forme du fond** (mise en forme automatique)

Conclusion : Potentiel et limitations

XML nous aide...

- À exprimer notre analyse dans un langage compréhensible par l'ordinateur (traitement automatique)
- À différencier la forme du fond (mise en forme automatique)
- À **partager** nos données (standardisation)

Conclusion : Potentiel et limitations

XML nous aide...

- À exprimer notre analyse dans un langage compréhensible par l'ordinateur (traitement automatique)
- À différencier la forme du fond (mise en forme automatique)
- À partager nos données (standardisation)
- À **multiplier les vues** possibles

Conclusion : Potentiel et limitations

XML nous aide...

- À exprimer notre analyse dans un langage compréhensible par l'ordinateur (traitement automatique)
- À différencier la forme du fond (mise en forme automatique)
- À partager nos données (standardisation)
- À multiplier les vues possibles
- À annoter de manière **incrémentale**

Conclusion : Potentiel et limitations

XML nous aide...

- À exprimer notre analyse dans un langage compréhensible par l'ordinateur (traitement automatique)
 - À différencier la forme du fond (mise en forme automatique)
 - À partager nos données (standardisation)
 - À multiplier les vues possibles
 - À annoter de manière incrémentale
- XML tire parti du *medium informatique*

Conclusion : Potentiel et limitations

XML nous aide...

- À exprimer notre analyse dans un langage compréhensible par l'ordinateur (traitement automatique)
 - À différencier la forme du fond (mise en forme automatique)
 - À partager nos données (standardisation)
 - À multiplier les vues possibles
 - À annoter de manière incrémentale
- XML tire parti du *medium* informatique

XML ne nous aide pas...

Conclusion : Potentiel et limitations

XML nous aide...

- À exprimer notre analyse dans un langage compréhensible par l'ordinateur (traitement automatique)
 - À différencier la forme du fond (mise en forme automatique)
 - À partager nos données (standardisation)
 - À multiplier les vues possibles
 - À annoter de manière incrémentale
- XML tire parti du *medium* informatique

XML ne nous aide pas...

- À **comprendre** les sources

Conclusion : Potentiel et limitations

XML nous aide...

- À exprimer notre analyse dans un langage compréhensible par l'ordinateur (traitement automatique)
 - À différencier la forme du fond (mise en forme automatique)
 - À partager nos données (standardisation)
 - À multiplier les vues possibles
 - À annoter de manière incrémentale
- XML tire parti du *medium* informatique

XML ne nous aide pas...

- À comprendre les sources
- À **choisir** ce qui nous intéresse dans les sources

Conclusion : Potentiel et limitations

XML nous aide...

- À exprimer notre analyse dans un langage compréhensible par l'ordinateur (traitement automatique)
 - À différencier la forme du fond (mise en forme automatique)
 - À partager nos données (standardisation)
 - À multiplier les vues possibles
 - À annoter de manière incrémentale
- XML tire parti du *medium* informatique

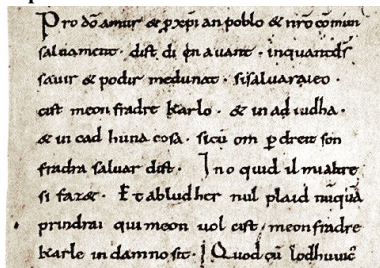
XML ne nous aide pas...

- À comprendre les sources
 - À choisir ce qui nous intéresse dans les sources
- XML est un outil qui **ne nous dispense pas de faire le travail** traditionnel

- 1 Introduction : l'édition de documents médiévaux (AL)
- 2 Présentation d'XML : techniques de base (NM)
- 3 Transcription diffractée (AL)**
 - Que voulons-nous encoder ?
 - Recommandations MENOTA pour l'édition multiniveaux
- 4 Transcription monolithique (NM)
- 5 Vers un compromis ? (AL)
- 6 Conclusion (NM)

Que voulons-nous encoder ?

Reproduction du manuscrit



Transcription diplomatique (graphématique)

Pro deo amur et pro christian poblo et nostro commun saluament, d'ist di In auant, in quant deus sauir et podir me dunat, si saluarai eo, cist meon fradre karlo, et in aiudha, et in cadhuna cosa, si cum om per dreit son fradra saluar dift. In o quid il mi altresí fazet. Et ab ludher nul plaid nunqua-m prindrai qui meon uol cist meon fradre karle in damno sit.

Transcription imitative (allographique)

Pro dō amur & p̄xpī an poblo & n̄rō cōmun saluament · dift di [c +]In auant · inquantdf faur & podir medunat · sifaluarai eo · cist meon fradre karlo · & in a[d]iudha · & in cad huna cosa · sicū om p̄dreit son fradra saluar dift · In o quid il mualtre si faz& · Et ab ludher nul plaid nūquā prindrai qui meon uol cist meon fradre karle in damno sit ·

Transcription normalisée (orthographique)

Pro Deo amur et pro christian poblo et nostro commun salvament, d'ist di in auant, in quant Deus savir et podir me dunat, si salvarai eo cist meon fradre Karlo et in aiudha et in cadhuna cosa, si cum om per dreit son fradra salvar dift, in o quid il mi altresí fazet. Et ab Ludher nul plaid nunqua-m prindrai qui, meon vol, cist meon fradre Karle in damno sit.

Recommandations MENOTA pour l'édition multiniveaux

Principe de base : versions concurrentes

Recommandations MENOTA pour l'édition multiniveaux

Principe de base : versions concurrentes

- *Projet Medieval Nordic Text Archive* (<http://www.menota.org>)

Recommandations MENOTA pour l'édition multiniveaux

Principe de base : versions concurrentes

- Projet *Medieval Nordic Text Archive* (<http://www.menota.org>)
- Chaque mot est représenté sous trois formes :
 - 1 normalisée : transcription orthographique
 - 2 diplomatique : transcription graphématique
 - 3 « facsimilaire » : transcription allographique

Recommandations MENOTA pour l'édition multiniveaux

Principe de base : versions concurrentes

- *Projet Medieval Nordic Text Archive* (<http://www.menota.org>)
- Chaque mot est représenté sous trois formes :
 - ① normalisée : transcription orthographique
 - ② diplomatique : transcription graphématique
 - ③ « facsimilaire » : transcription allographique

```
<w>
<choice>
  <norm>sçavoyr</norm>
  <dipl>scauoyr</dipl>
  <facs>&slong;cauoyr</facs>
</choice>
</w>
```

Recommandations MENOTA pour l'édition multiniveaux

Principe de base : versions concurrentes

- Projet *Medieval Nordic Text Archive* (<http://www.menota.org>)
- Chaque mot est représenté sous trois formes :
 - ① **normalisée** : transcription orthographique
 - ② diplomatique : transcription graphématique
 - ③ « facsimilaire » : transcription allographique

```
<w>
<choice>
  <norm>sçavoyr</norm>
  <dipl>scauoyr</dipl>
  <facs>&slong;cauoyr</facs>
</choice>
</w>
```

Recommandations MENOTA pour l'édition multiniveaux

Principe de base : versions concurrentes

- Projet *Medieval Nordic Text Archive* (<http://www.menota.org>)
- Chaque mot est représenté sous trois formes :
 - ① normalisée : transcription orthographique
 - ② **diplomatique** : transcription graphématique
 - ③ « facsimilaire » : transcription allographique

```
<w>
<choice>
  <norm>scavoyr</norm>
  <dipl>scauoyr</dipl>
  <facs>&slong;cauoyr</facs>
</choice>
</w>
```

Recommandations MENOTA pour l'édition multiniveaux

Principe de base : versions concurrentes

- Projet *Medieval Nordic Text Archive* (<http://www.menota.org>)
- Chaque mot est représenté sous trois formes :
 - ① normalisée : transcription orthographique
 - ② diplomatique : transcription graphématique
 - ③ « **facsimilaire** » : transcription allographique

```
<w>
<choice>
  <norm>sçavoyr</norm>
  <dipl>scauoyr</dipl>
  <facs>&slong;cauoyr</facs>
</choice>
</w>
```

Recommandations MENOTA pour l'édition multiniveaux

Principe de base : versions concurrentes

- Projet *Medieval Nordic Text Archive* (<http://www.menota.org>)
- Chaque mot est représenté sous trois formes :
 - ① normalisée : transcription orthographique
 - ② diplomatique : transcription graphématique
 - ③ « facsimilaire » : transcription allographique

```
<w>
<choice>
  <norm>scavoyr</norm>
  <dipl>scauoyr</dipl>
  <fac>&slong;cauyr</fac>
</choice>
</w>
```

```
<w>
<choice>
  <norm>Deus</norm>
  <dipl>d<ex>eu</ex>s</dipl>
  <fac><mdvAbbr>d&slongbar ;</mdvAbbr></fac>
</choice>
</w>
```

Recommandations MENOTA pour l'édition multiniveaux

Principe de base : versions concurrentes

- Projet *Medieval Nordic Text Archive* (<http://www.menota.org>)
- Chaque mot est représenté sous trois formes :
 - 1 normalisée : transcription orthographique
 - 2 diplomatique : transcription graphématique
 - 3 « facsimilaire » : transcription allographique

```
<w>
<choice>
  <norm>ſcavoyr</norm>
  <dipl>ſcauoyr</dipl>
  <fac>&slong;cauoyr</fac>
</choice>
</w>
```

```
<w>
<choice>
  <norm>Deus</norm>
  <dipl>d<ex>eu</ex>s</dipl>
  <fac><mdvAbbr>d&slongbar ;</mdvAbbr></fac>
</choice>
</w>
```

Exploitation

Les données peuvent être affichées à la demande, par **sélection**

- 1 Introduction : l'édition de documents médiévaux (AL)
- 2 Présentation d'XML : techniques de base (NM)
- 3 Transcription diffractée (AL)
- 4 Transcription monolithique (NM)**
 - Critique de la transcription diffractée
 - **Tout-en-un : la transcription monolithique**
- 5 Vers un compromis ? (AL)
- 6 Conclusion (NM)

Critique de la transcription diffractée

Critique de la transcription diffractée

Rupture de l'unité conceptuelle du texte

Critique de la transcription diffractée

Rupture de l'unité conceptuelle du texte

- Chaque mot est exprimé **plusieurs fois** dans l'encodage

Critique de la transcription diffractée

Rupture de l'unité conceptuelle du texte

- Chaque mot est exprimé **plusieurs fois** dans l'encodage
- Les opérations d'intervention et de description sont reportées à des **endroits différents**

Critique de la transcription diffractée

Rupture de l'unité conceptuelle du texte

- Chaque mot est exprimé **plusieurs fois** dans l'encodage
- Les opérations d'intervention et de description sont reportées à des **endroits différents**

Conséquences :

Critique de la transcription diffractée

Rupture de l'unité conceptuelle du texte

- Chaque mot est exprimé **plusieurs fois** dans l'encodage
- Les opérations d'intervention et de description sont reportées à des **endroits différents**

Conséquences :

→ **Anomalies de mise à jour** possible ! (correction)

Critique de la transcription diffractée

Rupture de l'unité conceptuelle du texte

- Chaque mot est exprimé **plusieurs fois** dans l'encodage
- Les opérations d'intervention et de description sont reportées à des **endroits différents**

Conséquences :

- **Anomalies de mise à jour** possible ! (correction)
- **Difficultés techniques** de traiter les annotations de niveau différent (coordination)

Critique de la transcription diffractée

Rupture de l'unité conceptuelle du texte

- Chaque mot est exprimé **plusieurs fois** dans l'encodage
- Les opérations d'intervention et de description sont reportées à des **endroits différents**

Conséquences :

- **Anomalies de mise à jour** possible ! (correction)
- **Difficultés techniques** de traiter les annotations de niveau différent (coordination)

Démultiplication des représentations

Autre conséquence importante : chaque chercheur voudra une vue particulière !

Critique de la transcription diffractée

Rupture de l'unité conceptuelle du texte

- Chaque mot est exprimé **plusieurs fois** dans l'encodage
- Les opérations d'intervention et de description sont reportées à des **endroits différents**

Conséquences :

- **Anomalies de mise à jour** possible ! (correction)
- **Difficultés techniques** de traiter les annotations de niveau différent (coordination)

Démultiplication des représentations

Autre conséquence importante : chaque chercheur voudra une vue particulière ! → Il faut trouver un moyen de ne **diffracter la transcription qu'au moment de la visualisation** (à la volée)

Tout-en-un : la transcription monolithique

Principe

Tout-en-un : la transcription monolithique

Principe

- Interventions/description simultanément dans un seul document

Tout-en-un : la transcription monolithique

Principe

- Interventions/description simultanément dans un seul document
- Unité mot non diffractée : interventions au niveau des caractères

Tout-en-un : la transcription monolithique

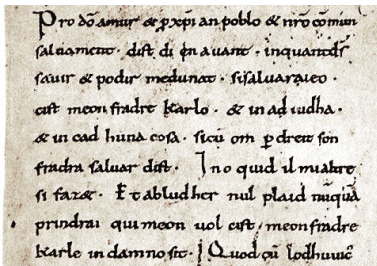
Principe

- Interventions/description simultanément dans un seul document
- Unité mot non diffractée : interventions au niveau des caractères
- Les balises explicitent la sémantique des interventions/descriptions

Tout-en-un : la transcription monolithique

Principe

- Interventions/description simultanément dans un seul document
- Unité mot non diffractée : interventions au niveau des caractères
- Les balises explicitent la sémantique des interventions/descriptions



```

pro
dō
amur
&
&pflour; X̄P̄ian
poblo
&
nrō
cōmun
&slong; aluament
.
  
```

Tout-en-un : la transcription monolithique

Principe

- Interventions/description simultanément dans un seul document
- Unité mot non diffractée : interventions au niveau des caractères
- Les balises explicitent la sémantique des interventions/descriptions

```

<w>pro</w>
<w>dō</w>
<w>amur</w>
<w>&</w>
<w>&pflour;</w>
<agglutination/>
<w>XĪian</w>
<w>poblo</w>
<w>&</w>
<w>nrō</w>
<w>cōmun</w>
<w>&slong;aluament</w>
<ponct>·</ponct>

```

Tout-en-un : la transcription monolithique

Principe

- Interventions/description simultanément dans un seul document
- Unité mot non diffractée : interventions au niveau des caractères
- Les balises explicitent la sémantique des interventions/descriptions

```
<w><lettrine hauteur='1'>p</lettrine>ro</w>  
<w>dō</w>  
<w>amur</w>  
<w>&</w>  
<w>&pflour;</w>  
<agglutination/>  
<w>XĪian</w>  
<w>poblo</w>  
<w>&</w>  
<w>nrō</w>  
<w>cōmun</w>  
<w>&slong;aluament</w>  
<ponct>·</ponct>
```

Tout-en-un : la transcription monolithique

Principe

- Interventions/description simultanément dans un seul document
- Unité mot non diffractée : interventions au niveau des caractères
- Les balises explicitent la sémantique des interventions/descriptions

```

<w><lettrine hauteur='1'><maj>p</maj></lettrine>ro</w>
<w><maj>d</maj><abbr résolution='eo'>ō</abbr></w>
<w>amur</w>
<w><abbr résolution='et'>&</abbr></w>
<w><abbr résolution='pro'>&pflour;</abbr></w>
<agglutination/>
<w><abbr résolution='christ'>XP</abbr>ian</w>
<w>poblo</w>
<w><abbr résolution='et'>&</abbr></w>
<w><abbr résolution='nostro'>nrō</abbr></w>
<w>c<abbr résolution='om'>ō</abbr>mun</w>
<w>
  <modern transformation='s'>&slong;</modern>
  al<modern transformation='v'>u</modern>ament
</w>
<ponct>.</ponct>

```


Tout-en-un : la transcription monolithique

Principe

- Interventions/description simultanément dans un seul document
- Unité mot non diffractée : interventions au niveau des caractères
- Les balises explicitent la sémantique des interventions/descriptions

Exploitation

Les données peuvent être mises en forme à la demande, par **construction**

Tout-en-un : la transcription monolithique

Principe

- Interventions/description simultanément dans un seul document
- Unité mot non diffractée : interventions au niveau des caractères
- Les balises explicitent la sémantique des interventions/descriptions

Exploitation

Les données peuvent être mises en forme à la demande, par **construction**

- ① Pro deo amur *et* pro_christian poblo *et* nostro commun saluament ·

Tout-en-un : la transcription monolithique

Principe

- Interventions/description simultanément dans un seul document
- Unité mot non diffractée : interventions au niveau des caractères
- Les balises explicitent la sémantique des interventions/descriptions

Exploitation

Les données peuvent être mises en forme à la demande, par **construction**

- ① Pro deo amur *et* pro_christian poblo *et* nostro commun saluament ·
- ② Pro Deo amur et pro christian poblo et nostro commun salvament

Tout-en-un : la transcription monolithique

Principe

- Interventions/description simultanément dans un seul document
- Unité mot non diffractée : interventions au niveau des caractères
- Les balises explicitent la sémantique des interventions/descriptions

Exploitation

Les données peuvent être mises en forme à la demande, par **construction**

- ① Pro deo amur *et pro_christian* poblo *et nostro commun* saluament ·
- ② Pro Deo amur et pro christian poblo et nostro commun salvament
- ③ Pro Deo amur *et pro christian* poblo *et nostro commun* salvament ·

- 1 Introduction : l'édition de documents médiévaux (AL)
- 2 Présentation d'XML : techniques de base (NM)
- 3 Transcription diffractée (AL)
- 4 Transcription monolithique (NM)
- 5 Vers un compromis ? (AL)**
 - Relativité de l'opposition *diffracté* vs *monolithique*
 - Recherche de l'ergonomie
- 6 Conclusion (NM)

Relativité de l'opposition *diffracté* vs *monolithique*

Une question de granularité ?

Relativité de l'opposition *diffracté* vs *monolithique*

Une question de granularité ?

- Il y a toujours diffraction

Relativité de l'opposition *diffracté* vs *monolithique*

Une question de granularité ?

- Il y a toujours diffraction
- Mot ou caractère ?

Relativité de l'opposition *diffracté* vs *monolithique*

Une question de granularité ?

- Il y a toujours diffraction

→ Mot ou caractère ?

```

<w><lettrine hauteur='1'><maj>p</maj></lettrine>ro</w>
<w><maj>d</maj><abbr résolution='eo'>ō</abbr></w>
<w>amur</w>
<w><abbr résolution='et'>&</abbr></w>
<w><abbr résolution='pro'>&pflour;</abbr></w>
<agglutination/>
<w><abbr résolution='christ'>XP̄</abbr>ian</w>
<w>poblo</w>
<w><abbr résolution='et'>&</abbr></w>
<w><abbr résolution='nostro'>nrō</abbr></w>
<w>c<abbr résolution='om'>ō</abbr>mun</w>
<w>
  <modern transformation='s'>&slong;</modern>
  al<modern transformation='v'>u</modern>ament
</w>
<punct>.</punct>

```

Recherche de l'ergonomie

Démarche

- Représentation **diffractée** au niveau des mots comme produit final
- Syntaxe compacte **non XML** pour l'encodage (saisie et correction)

Recherche de l'ergonomie

Quittons XML...

Pro
Deo
amur
et
pro
christian
poblo
et
nostro
commun
salvament

Recherche de l'ergonomie

Modernisation : majuscules, opposition *u/v*

Pro
#deo
amur
et
pro
christian
poblo
et
nostro
commun sal*uament

Recherche de l'ergonomie

Marquage des abréviations

```
Pro
#d[e]o
amur
[et]
p[ro]
[christ]ian
poblo
[et]
n[ost]ro
co[m]mun
sal*uament
```

Recherche de l'ergonomie

Traitement des espaces

```
Pro
#d[e]o
amur
[et]
p[ro]+?
[christ]i_an
poblo
[et]
n[ost]ro
co[m]mun
sal*uament
```

Recherche de l'ergonomie

Description de l'initiale

```
{{P:1:black}}ro  
#d[e]o  
amur  
[et]  
p[ro]+?  
[christ]i_an  
poblo  
[et]  
n[ost]ro  
co[m]mun  
sal*uament
```

Recherche de l'ergonomie

Traitement de &

```
{{P:1:black}}ro  
#d[e]o  
amur  
&amp;  
p[ro]+?  
[christ]i_an  
poblo  
&amp;  
n[ost]ro  
co[m]mun  
sal*uament
```


Recherche de l'ergonomie

Traitement du tilde

```

{{P:1:black}}ro
#d[e]o
amur
&amp;
p[ro]+?
[christ]i_an
poblo
&amp;
n[ost]ro
c((o&bar;))mun
sal*uament

```

Recherche de l'ergonomie

Traitement des autres abréviations

```

{{P:1:black}}ro
#((do&bar;_d[e]o))
amur
&amp;
((&pflour;_p[ro]))+?
((xp&bar;_[christ]))i_an
poblo
&amp;
((nro&bar;_n[ost]ro))
c((o&bar;))mun
sal*uament

```

Recherche de l'ergonomie

Traitement de *s* « longue » et **conversion en XML**

```
{(P:1:black)}ro
#((do&bar;_d[e]o))
amur
&amp;
((&pflour;_p[ro]))+?
((xp&bar;_[christ]))i_an
poblo
&amp;
((nro&bar;_n[ost]ro))
c((o&bar;))mun
&slong;al*uament
```

```
<w xml:id="w1030_001">
  <choice>
    <me:norm>Pro</me:norm>
    <me:dipl>Pro</me:dipl>
    <me:facs>
      <bfm:dropcap size="1" color="black">P</bfm:dropcap>
      ro
    </me:facs>
  </choice>
</w>
<w bfm:aggl="simple"
  bfm:agglCert="low"
  xml:id="w1030_002">
  <choice>
    <me:norm>Deo</me:norm>
    <me:dipl>d<ex>e</ex>o</me:dipl>
    <me:facs>
      <bfm:mdvAbbr>do&bar;</bfm:mdvAbbr>
    </me:facs>
  </choice>
</w>
```

- 1 Introduction : l'édition de documents médiévaux (AL)
- 2 Présentation d'XML : techniques de base (NM)
- 3 Transcription diffractée (AL)
- 4 Transcription monolithique (NM)
- 5 Vers un compromis ? (AL)
- 6 Conclusion (NM)**

Conclusion (NM)

Importance de la réflexion théorique

Conclusion (NM)

Importance de la réflexion théorique

- Le potentiel des outils est vaste et doit être cadré

Conclusion (NM)

Importance de la réflexion théorique

- Le potentiel des outils est vaste et doit être cadré
- Il n'y a pas de solution miracle : il faut penser le problème pour éviter les mauvaises surprises

Conclusion (NM)

Importance de la réflexion théorique

- Le potentiel des outils est vaste et doit être cadré
- Il n'y a pas de solution miracle : il faut penser le problème pour éviter les mauvaises surprises
- L'**objectif du chercheur** conditionne les moyens mis en œuvre...

Conclusion (NM)

Importance de la réflexion théorique

- Le potentiel des outils est vaste et doit être cadré
- Il n'y a pas de solution miracle : il faut penser le problème pour éviter les mauvaises surprises
- L'**objectif du chercheur** conditionne les moyens mis en œuvre...

→ Théories/outils/pratique

Conclusion (NM)

Importance de la réflexion théorique

- Le potentiel des outils est vaste et doit être cadré
- Il n'y a pas de solution miracle : il faut penser le problème pour éviter les mauvaises surprises
- L'**objectif du chercheur** conditionne les moyens mis en œuvre...

→ Théories/outils/pratique

- Les formats et les *media* sont à mi-chemin entre la théorie et la pratique

Conclusion (NM)

Importance de la réflexion théorique

- Le potentiel des outils est vaste et doit être cadré
- Il n'y a pas de solution miracle : il faut penser le problème pour éviter les mauvaises surprises
- L'**objectif du chercheur** conditionne les moyens mis en œuvre...

→ Théories/outils/pratique

- Les formats et les *media* sont à mi-chemin entre la théorie et la pratique
- Leur place est centrale (décisive),

Conclusion (NM)

Importance de la réflexion théorique

- Le potentiel des outils est vaste et doit être cadré
- Il n'y a pas de solution miracle : il faut penser le problème pour éviter les mauvaises surprises
- L'**objectif du chercheur** conditionne les moyens mis en œuvre...

→ Théories/outils/pratique

- Les formats et les *media* sont à mi-chemin entre la théorie et la pratique
- Leur place est centrale (décisive),
- mais ils ne sont pas une fin *per se*.