

Projet *Khartês*
Méthode d'édition de chartes originales
Philosophie d'encodage

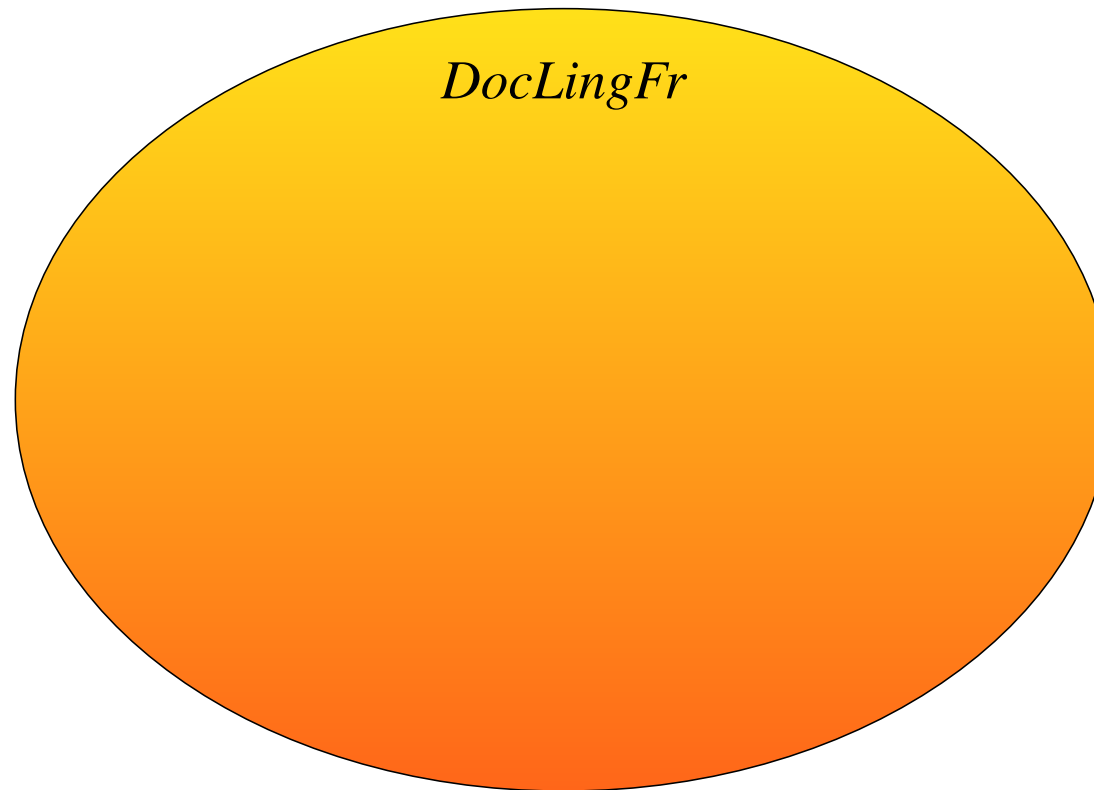
Nicolas Mazziotta (LLR)
10 novembre 2005, CPM

Plan

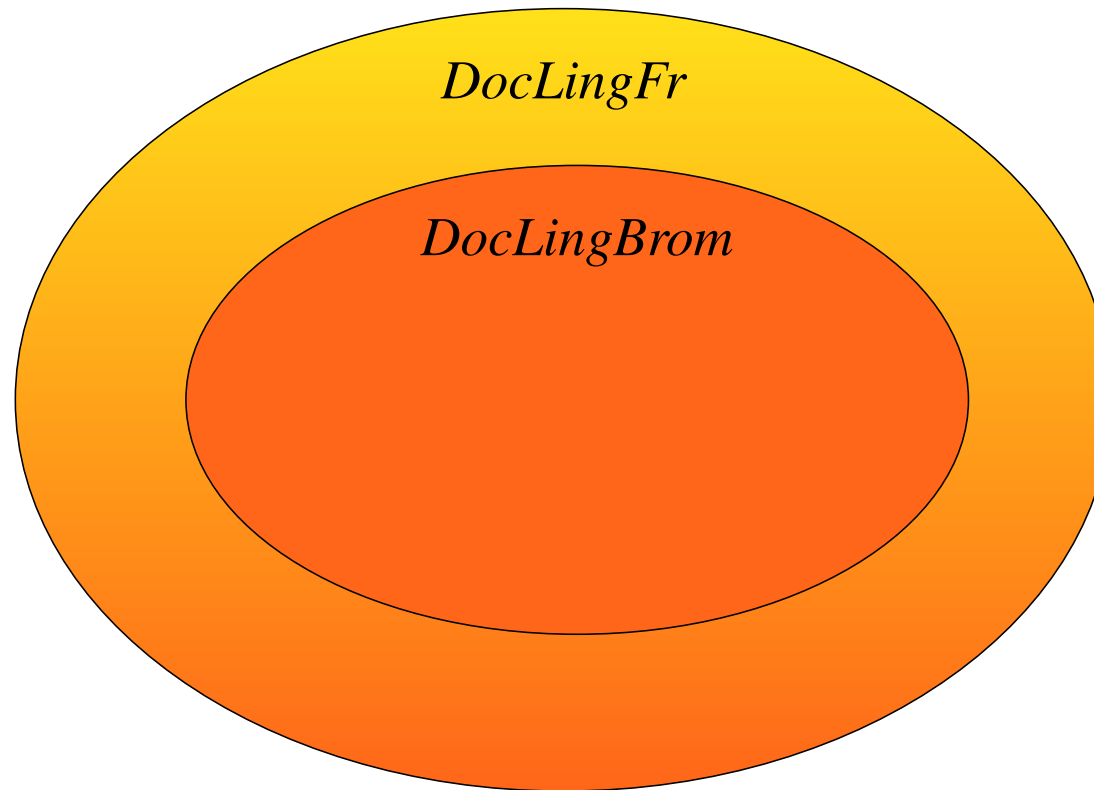
- 1 Projet *Khartês* : présentation
- 2 Problématique
- 3 Édition plurielle : approche traditionnelle
- 4 Édition plurielle : nouvelles technologies
- 5 Applications et résultats

1. Projet *Khartês* : présentation

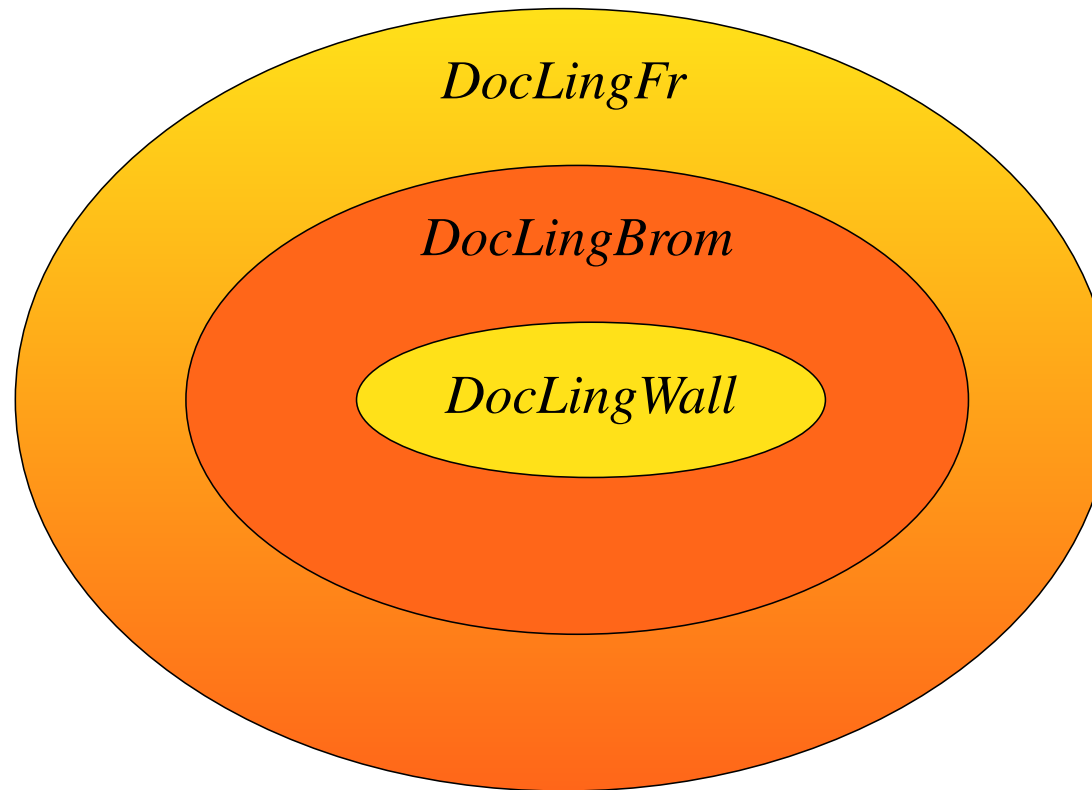
1. Projet *Khartês* : présentation



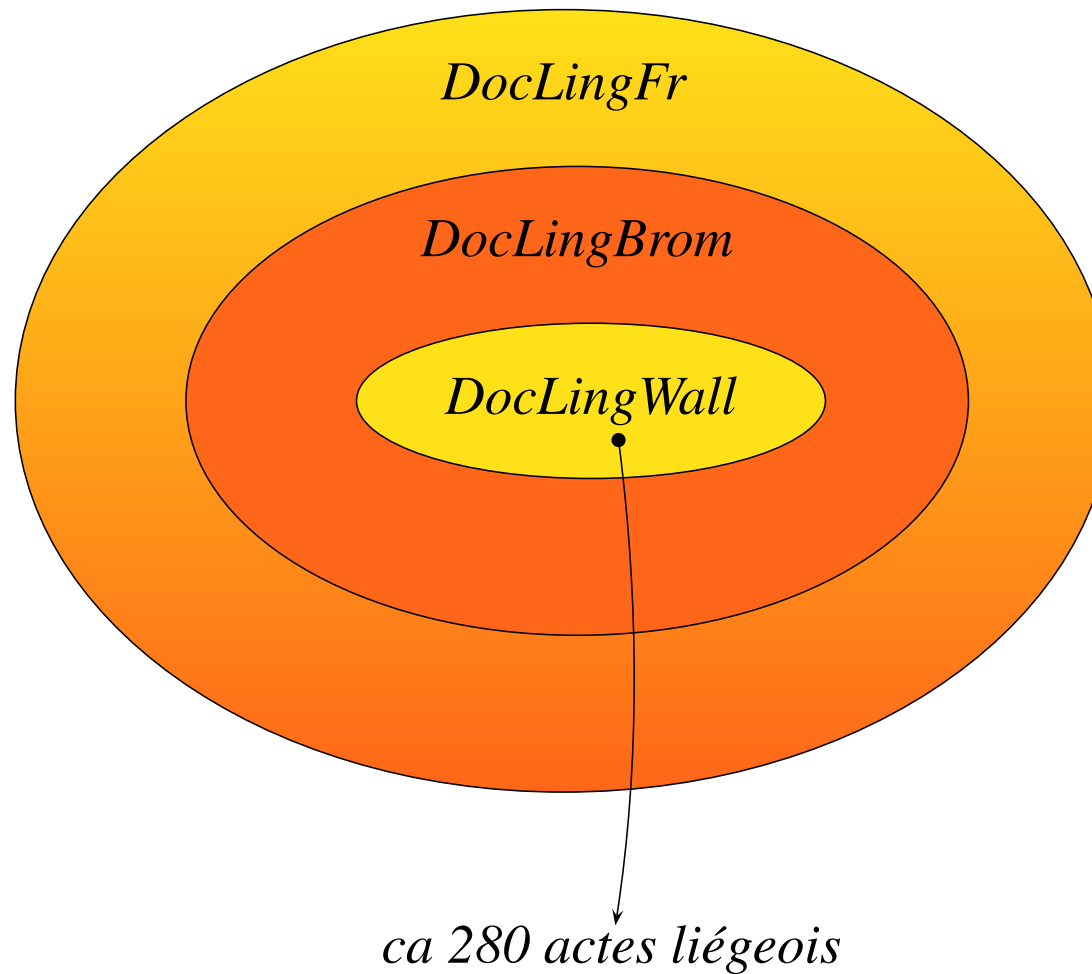
1. Projet *Khartês* : présentation



1. Projet *Khartês* : présentation



1. Projet *Khartês* : présentation



2. Problématique

2.1. Le dilemme de l'éditeur

Éditer un document = « Donner un texte à lire à un certain public »

Pratiquement, il faut :

2. Problématique

2.1. Le dilemme de l'éditeur

Éditer un document = « Donner un texte à lire à un certain public »

Pratiquement, il faut :

1. *élaborer* ;

2. Problématique

2.1. Le dilemme de l'éditeur

Éditer un document = « Donner un texte à lire à un certain public »

Pratiquement, il faut :

1. *élaborer* ;
2. *réduire*.

2. Problématique

2.1. Le dilemme de l'éditeur

Éditer un document = « Donner un texte à lire à un certain public »

Pratiquement, il faut :

1. *élaborer* ;
2. *réduire*.

Édition = compromis

Rendre le texte accessible = « trahir » son témoignage.

2.2. Objectif : éditer des documents pour des linguistes

Les éditions *Khartês* doivent permettre de répondre à des questions d'ordre linguistique, à plusieurs niveaux d'intégration :

2.2. Objectif : éditer des documents pour des linguistes

Les éditions *Khartês* doivent permettre de répondre à des questions d'ordre linguistique, à plusieurs niveaux d'intégration :

- textes, structures énonciatives ;
- phrases et syntagmes, structures syntaxiques ;
- mots, morphologie et lexique ;
- lettres, relation graphie/phonie, système graphique ;
- ponctuation, à tous ces niveaux ;

2.2. Objectif : éditer des documents pour des linguistes

Les éditions *Khartês* doivent permettre de répondre à des questions d'ordre linguistique, à plusieurs niveaux d'intégration :

- textes, structures énonciatives ;
- phrases et syntagmes, structures syntaxiques ;
- mots, morphologie et lexique ;
- lettres, relation graphie/phonie, système graphique ;
- ponctuation, à tous ces niveaux ;

→ Limiter la réduction des données

2.3. Responsabilité : transmettre des documents inédits

Chartes inédites →

2.3. Responsabilité : transmettre des documents inédits

Chartes inédites →

- intérêt linguistique ;
- intérêt historique.

2.3. Responsabilité : transmettre des documents inédits

Chartes inédites →

- intérêt linguistique ;
- intérêt historique.

Besoin d'enrichir les données, de rendre le texte accessible à un public relativement large...

3. Édition plurielle : approche traditionnelle

3.1. Document exemple

SLa 1236a

3.2. Principaux types d'édition

- éditions critiques ;
- éditions diplomatiques ;

3.3. Honorer l'intégrité du document : édition hybride

Le problème peut se résumer ainsi :

3.3. Honorer l'intégrité du document : édition hybride

Le problème peut se résumer ainsi :

- les niveaux d'analyse énoncés *supra* intéressent *différentes sortes de linguistes* ;
- les historiens et les philologues ne doivent être privés de l'accès à l'édition.

3.3. Honorer l'intégrité du document : édition hybride

Le problème peut se résumer ainsi :

- les niveaux d'analyse énoncés *supra* intéressent *différentes sortes de linguistes* ;
- les historiens et les philologues ne doivent être privés de l'accès à l'édition.

Chaque public a son édition → diffraction du texte en différentes éditions non liées.

3.3. Honorer l'intégrité du document : édition hybride

Le problème peut se résumer ainsi :

- les niveaux d'analyse énoncés *supra* intéressent *différentes sortes de linguistes* ;
- les historiens et les philologues ne doivent être privés de l'accès à l'édition.

Chaque public a son édition → diffraction du texte en différentes éditions non liées.

Comment ne pas rompre le lien entre les éditions ?

3.3. Honorer l'intégrité du document : édition hybride

Le problème peut se résumer ainsi :

- les niveaux d'analyse énoncés *supra* intéressent *différentes sortes de linguistes* ;
- les historiens et les philologues ne doivent être privés de l'accès à l'édition.

Chaque public a son édition → diffraction du texte en différentes éditions non liées.

Comment ne pas rompre le lien entre les éditions ?

Une édition qui intègre toutes les informations utiles permet de conserver le lien.

3.3. Honorer l'intégrité du document : édition hybride

Le problème peut se résumer ainsi :

- les niveaux d'analyse énoncés *supra* intéressent *différentes sortes de linguistes* ;
- les historiens et les philologues ne doivent être privés de l'accès à l'édition.

Chaque public a son édition → diffraction du texte en différentes éditions non liées.

Comment ne pas rompre le lien entre les éditions ?

Une édition qui intègre toutes les informations utiles permet de conserver le lien.

→ Nécessité d'établir une édition hybride.

3.4. Honorer la lisibilité : édition multiple

Mais l'édition hybride devient illisible.

3.4. Honorer la lisibilité : édition multiple

Mais l'édition hybride devient illisible.

Jakemes, li *provoft*, · **J**ohans, li **d**oinf, · li **a**rchiakene et toz li capitelez de
fain **L**amber de Liege, a toz ceaz ki *verrunt* ce letref, conoifre veriteit.

3.4. Honorer la lisibilité : édition multiple

Mais l'édition hybride devient illisible.

Jakemes, li *provoft*, · Johans, li *doïnf*, · li *archiakene* et toz li capitelez de fain Lamber de Liege, a toz ceaz ki verrunt ce letref, conoifre veriteit.

Édition illisible = édition qui *ne remplit pas sa mission*.
Comment produire une édition lisible ?

3.4. Honorer la lisibilité : édition multiple

Mais l'édition hybride devient illisible.

Jakemes, li *provoft*, · Johans, li *doïnf*, · li *archiakene* et toz li capitez de fain Lamber de Liege, a toz ceaz ki verrunt ce letref, conoifre veriteit.

Édition illisible = édition qui *ne remplit pas sa mission*.
Comment produire une édition lisible ?

La diversité de l'information qui nuit à la réception du message.

3.4. Honorer la lisibilité : édition multiple

Mais l'édition hybride devient illisible.

Jakemes, li *provoft*, · Johans, li *doïnf*, · li *archiakene* et toz li capitez de fain Lamber de Liege, a toz ceaz ki verrunt ce letref, conoifre veriteit.

Édition illisible = édition qui *ne remplit pas sa mission*.
Comment produire une édition lisible ?

La diversité de l'information qui nuit à la réception du message.

→ Nécessité d'établir plusieurs éditions.

3.5. Le nœud

Problèmes antagonistes → solutions antagonistes :

3.5. Le nœud

Problèmes antagonistes → solutions antagonistes :

Résoudre les problèmes de l'édition multiple ne peut se faire qu'en établissant une édition hybride.

Résoudre les problèmes de l'édition hybride ne peut se faire qu'en établissant une édition multiple.

3.5. Le nœud

Problèmes antagonistes → solutions antagonistes :

Résoudre les problèmes de l'édition multiple ne peut se faire qu'en établissant une édition hybride.



Résoudre les problèmes de l'édition hybride ne peut se faire qu'en établissant une édition multiple.

Le serpent se mord la queue !

4. Édition plurielle : nouvelles technologies

Le papier n'offre aucune autre alternative.

Il n'y a donc pas de solution traditionnelle qui permette de résoudre le dilemme.

4. Édition plurielle : nouvelles technologies

Le papier n'offre aucune autre alternative.

Il n'y a donc pas de solution traditionnelle qui permette de résoudre le dilemme.

→ Nécessité de se servir des nouvelles technologies.

4.1. Approche intuitive : traitements de texte courants

Les traitements de texte *wysiwyg* courants permettent déjà de combiner plusieurs vues d'un même texte.

4.1.1. « Remplacer par »

On peut tenter une expérience simple :

Texte hybride →→ remplacements systématiques...

4.1.1. « Remplacer par »

On peut tenter une expérience simple :

Texte hybride →→ remplacements systématiques...

Jakemes, li *provoft*, · Johans, li **doinf**, · li **archiakene** et toz
li capitelez de fain **Lamber** de Liege, a toz ceaz ki verrunt ce
letref, conoifre veriteit.

Le texte est enrichi et fidèle, donc difficilement lisible. Voyons
ce que permettent les remplacements.

4.1.1. « Remplacer par »

On peut tenter une expérience simple :

Texte hybride →→ remplacements systématiques...

Jakemes, li *provoft*, Johans, li **doïnf**, li **archiakene** et toz li capitelez de fain Lamber de Liege, a toz ceaz ki verrunt ce letref, conoifre veriteit.

Pour alléger la lecture, on peut retirer la ponctuation originale (remplacer les signes par des blancs).

4.1.1. « Remplacer par »

On peut tenter une expérience simple :

Texte hybride →→ remplacements systématiques...

Jakemes, li *provost*, **J**ohans, li **d**oïns, li **a**rchiakene et toz li capitelez de sain **L**amber de Liege, a toz ceaz ki verrunt ce letres, conoistre veriteit.

De même, celui qui n'a que faire de l'opposition ⟨s⟩ vs ⟨f⟩ peut l'évacuer (en remplaçant le dernier par le premier).

4.1.1. « Remplacer par »

On peut tenter une expérience simple :

Texte hybride →→ remplacements systématiques...

Jakemes, li *provost*, **Johans**, li **doïns**, li **archiakene** et toz li capitelez de sain Lamber de Liege, a toz ceaz ki verrunt ce letres, conoistre veriteit.

Mais ce type de remplacements enchaînés mène très vite à des erreurs... Il n'y a pas de moyen de contrôle.

4.1.2. Difficultés : insuffisance de ces méthodes

Mais la méthode du remplacement est insuffisante pour des raisons :

4.1.2. Difficultés : insuffisance de ces méthodes

Mais la méthode du remplacement est insuffisante pour des raisons :

1. techniques ;

4.1.2. Difficultés : insuffisance de ces méthodes

Mais la méthode du remplacement est insuffisante pour des raisons :

1. techniques ;
2. ergonomiques ;

4.1.2. Difficultés : insuffisance de ces méthodes

Mais la méthode du remplacement est insuffisante pour des raisons :

1. techniques ;
2. ergonomiques ;
3. conceptuelles ;

4.1.2. Difficultés : insuffisance de ces méthodes

Mais la méthode du remplacement est insuffisante pour des raisons :

1. techniques ;
2. ergonomiques ;
3. conceptuelles ;
4. et surtout de sécurité.

4.2. Nouvelles solutions

4.2.1. Nécessité de marquer le texte de façon explicite

L'ordinateur ne comprend pas le texte.

4.2. Nouvelles solutions

4.2.1. Nécessité de marquer le texte de façon explicite

L'ordinateur ne comprend pas le texte.

Il ne peut interpréter les ambiguïtés, comme la multiplicité des emplois de l'italique !

4.2. Nouvelles solutions

4.2.1. Nécessité de marquer le texte de façon explicite

L'ordinateur ne comprend pas le texte.

Il ne peut interpréter les ambiguïtés, comme la multiplicité des emplois de l'italique !

Pourtant, la technique du remplacement de se débarrasser de ce dont on n'a pas besoin de manière temporaire.

4.2. Nouvelles solutions

4.2.1. Nécessité de marquer le texte de façon explicite

L'ordinateur ne comprend pas le texte.

Il ne peut interpréter les ambiguïtés, comme la multiplicité des emplois de l'italique !

Pourtant, la technique du remplacement de se débarrasser de ce dont on n'a pas besoin de manière temporaire.

→ Nécessité d'expliquer à la machine ce qu'elle doit faire et comment :

1. procédures de remplacement ;
2. marquage des données.

4.2.2. Comment marquer le texte ?

Tenir compte de la spécificité du médium :

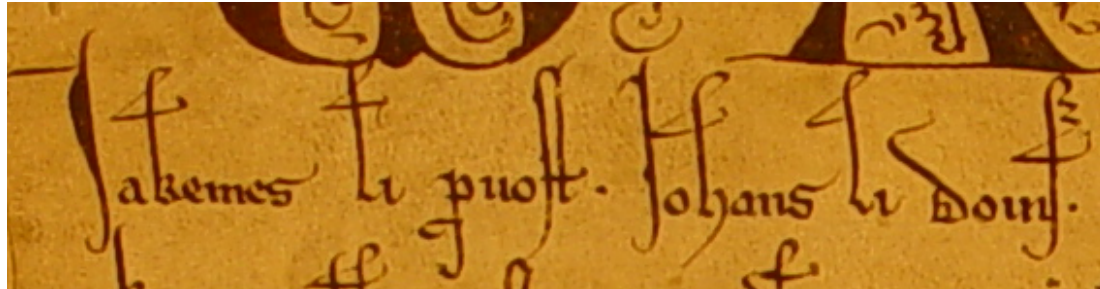
4.2.2. Comment marquer le texte ?

Tenir compte de la spécificité du médium :

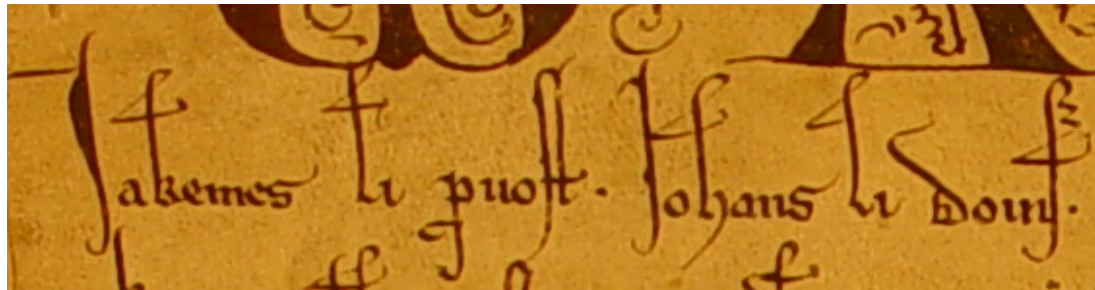
Le fichier informatique n'est pas du papier. Il peut contenir des données qui ne sont que des instructions de traitement pour un programme : l'utilisateur ne les verra jamais.

On peut donc pourvoir le texte d'étiquettes univoques, à quelque niveau de granularité que ce soit (texte, phrase, mot, lettre).

On peut donc pourvoir le texte d'étiquettes univoques, à quelque niveau de granularité que ce soit (texte, phrase, mot, lettre).

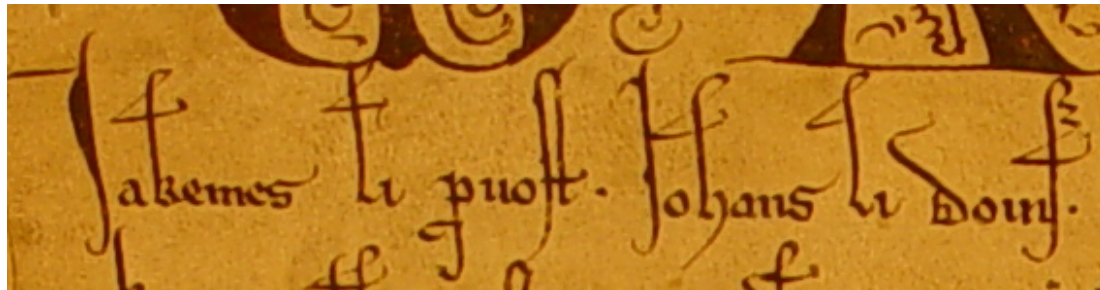


On peut donc pourvoir le texte d'étiquettes univoques, à quelque niveau de granularité que ce soit (texte, phrase, mot, lettre).



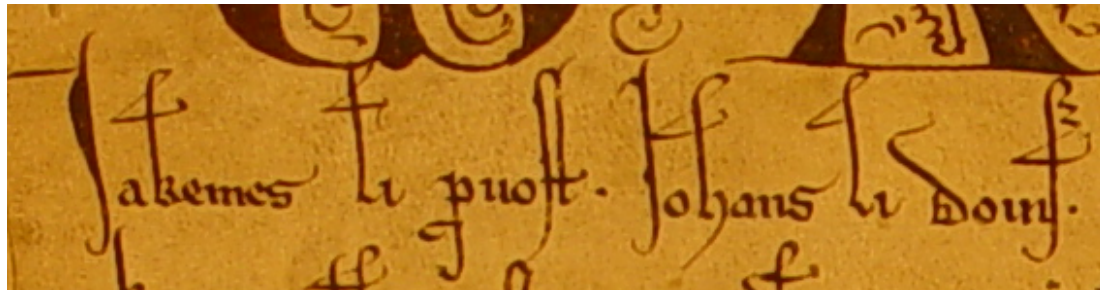
Jakemes, li *provost*, · **J**ohans, li **doinf**

On peut donc pourvoir le texte d'étiquettes univoques, à quelque niveau de granularité que ce soit (texte, phrase, mot, lettre).

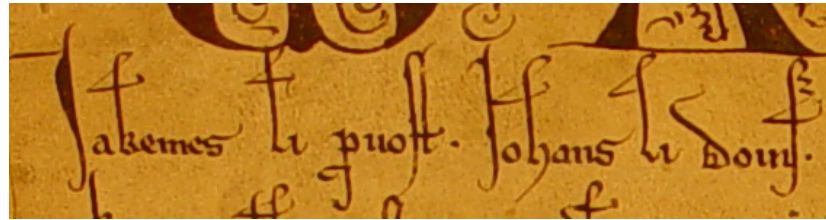


Jakemes, li *provost*, · **J**ohans, li **doinf**
jakemes, li prouost, · johans, li doins

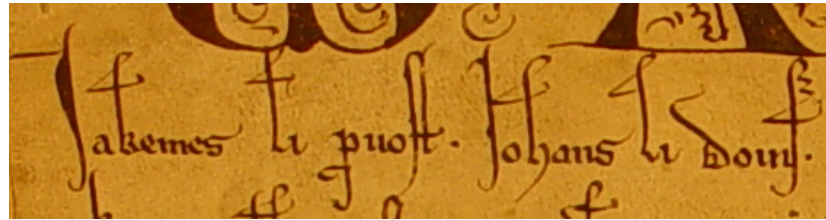
On peut donc pourvoir le texte d'étiquettes univoques, à quelque niveau de granularité que ce soit (texte, phrase, mot, lettre).



Jakemes, li *pr*ouost, · **J**ohans, li **do**inf
jakemes, li prouost, · johans, li doins
jakemes, li **pr**ouost, · **j**ohans, li **do**ins

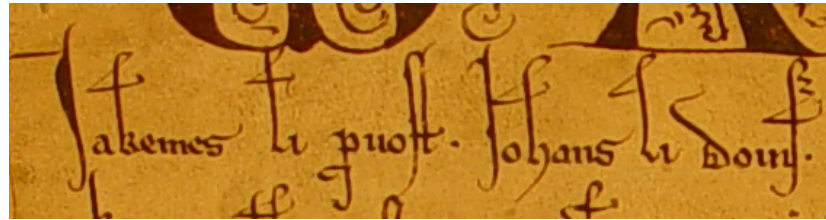


Deux types de codages :



Deux types de codages :

Inhérent : J akemes

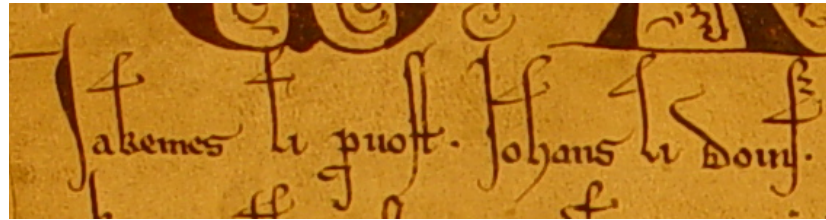


Deux types de codages :

Inhérent : J akemes

Ajouté :

XML. — `<c>J</c>akemes`



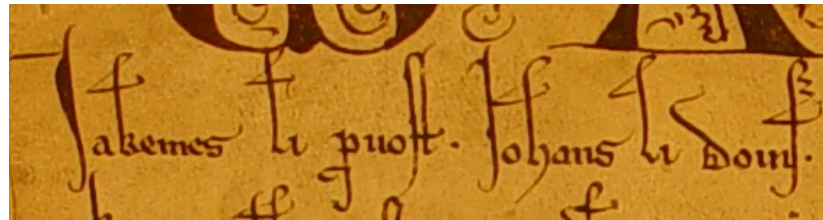
Deux types de codages :

Inhérent : J akemes

Ajouté :

XML. — <c>J</c>akemes

privé. — p[ro]°uo\$t



Deux types de codages :

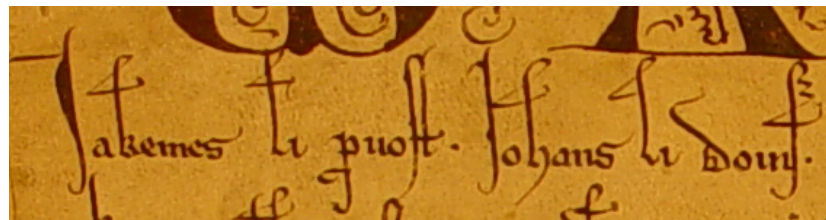
Inhérent : J akemes

Ajouté :

XML. — <c>J</c>akemes

privé. — p[ro]°uo\$t

<c>J</c>akemes, li p[ro]°uo\$t, * <c>J</c>ohans, li <c>d</c>oi"n\$



Deux types de codages :

Inhérent : J akemes

Ajouté :

XML. — <c>J</c>akemes

privé. — p[ro]°uo\$t

<c>J</c>akemes, li p[ro]°uo\$t, * <c>J</c>ohans, li <c>d</c>oi"n\$

<w><c>J</c>akemes</w> <mponct>,</mponct> <w>li</w>

<w>p[ro]°uo\$t</w> <mponct>,</mponct> <ponct>*</ponct>

<w><c>J</c>ohans</w> <mponct>,</mponct>

<w>li</w> <w><c>d</c>oi"n\$</w>

5. Applications et résultats

5.1. Vues multiples : démonstration

5.2. Partage : TEI

5.3. Prolongements

Merci