

1. Guide d'encodage de la rédaction non cyclique de l'*Histoire de Kanor et de ses freres* d'après le manuscrit BnF fr. 1446 (C)

Extensions et limitations : ce document constitue la présentation de certains des éléments ayant guidé et structuré le travail d'édition numérique des parties *rechapitulation* et *Pelyarmenus* contenues dans le manuscrit BnF fr. 1446 (C). Il présente également des pistes de réflexion pour la section *Kanor*.

Crédits : nous souhaitons dans un premier temps remercier quelques sources sans lesquelles ce travail préliminaire n'aurait pu être si rapidement construit : le projet GitHub msDesc qui propose un ensemble de modèles <msDesc> commentés et *TEI conformant*. Ce projet est porté entre autres par James Cummings (Université de Newcastle) : il propose un ensemble de patrons (en. «*template*») pour l'élément <msDesc> du *header* du fichier de l'édition en XML-TEI. C'est l'endroit où sont posées toutes les informations qui touchent à la description du manuscrit. De même, la thèse d'Ariane Pinche (CNRS), disponible également sur GitHub (EditionLiSaintConfessor), est une source d'inspiration tant pour la pertinence de ses choix éditoriaux que pour la clarté technique et structurelle de ses fichiers, dont celui-ci même s'inspire énormément. Enfin, Camille Carnaille (UniGe), ainsi que les documents préparatoires aux éditions de l'ensemble du groupe C7S, pour avoir pris le temps de m'initier aux premières nécessités d'une bonne édition numérique.

Introduction : ce petit guide procède d'une double ambition. Il a tout d'abord pour but d'expliquer de quelle manière l'édition est encodée en XML-TEI. Les principes d'encodage choisis sont déterminés pour produire des données pérennes, interopérables et faciles à transformer, cela afin de permettre non seulement une bonne intégration de ce travail à ceux de l'équipe de recherche des C7S mais aussi une souplesse pour l'édition de ce texte. L'encodage a été réalisé en suivant les recommandations de la Text Encoding Initiative (TEI), présentées dans les Guidelines P5 (<https://tei-c.org/guidelines/p5/>). L'intégralité de l'encodage proposé est *TEI conformant*. Cette édition étant également un travail de thèse, les principes dont elle découle sont l'objet d'une double articulation : présenter d'une part la progression d'une réflexion éditoriale sur ce projet, et exposer les solutions techniques à sa réalisation.

L'édition de l'*Histoire de Kanor et de ses freres* réunit les parties suivantes :

- Récapitulation du *Cycle*
- Fin du *Pelyarmenus*
- Récapitulation puis fin de *Kanor*

Le corpus numérique sera à terme constitué de trois fichiers XML-TEI distincts correspondant chacun à une partie. Ces derniers seront réunis dans un *teiCorpus* nommé *LiHistoiredeKanor.xml*. [Note: pour le moment, le nom du texte est tiré de la rubrique initiale et de l'explicit «*li histoire de Kanor et de ses freres*», mais de toute évidence le statut de héros principal du personnage «*Kanor*», a fortiori dans cette narration complexe, qui mêle plusieurs récits, sous différentes formes (résumé, chronique, etc.), et selon des modalités esthétiques variées (amplification, contraction, invention), mérite d'être interrogé.]

1.1. Structure du fichier XML

Le contenu de cette section est pour le moment encore à l'état d'ébauche. La structure du fichier XML-TEI est constituée et suffisante pour générer, déjà, un travail exploitable. Toutefois, elle est amenée à s'affiner.

La structuration des textes contenus dans les fichiers XML suit les recommandations du service CTS, *Canonical Text System* (développé pour le Homer Multitext Project grâce à Christopher Blackwell et Neel Smith). Ce système permet de mettre en place un ensemble de services basés sur l'identification des textes ou de leurs fragments grâce à une référence canonique qui prend la forme d'une URN CTS [Note: Uniform Resource Name, nom d'un standard informatique qui permet d'identifier une ressource indépendamment de sa localisation et de son accessibilité par internet, ce qui permet à cet identifiant d'être pérenne]. Ainsi, cette édition est aisément citable et archivable grâce au respect de normes internationales.

Pour différencier les différentes parties de l'œuvre qui composent *Li Histoire de Kanor et de ses freres*, nous avons décidé, à terme, d'encoder chacune d'entre elles dans un fichier XML indépendant. Cela permet, outre une taille moindre des fichiers, de distinguer plus nettement les choix d'édition et, partant, d'encodage de l'information qui ont été décidés. Les fichiers de l'édition sont nommés et identifiés de la manière suivante : «*urn:cts:froLit:jns47.jns8386.ciham-fro2*». La première partie de l'URN : «*urn:cts*» indique à quel système de référence appartient l'URN ; ici, la norme CTS. «*froLit*» signifie que le texte appartient à un corpus des textes en ancien français. Les éléments précédés de «*jns*» indiquent les identifiants de l'œuvre dans le portail Jonas développé par l'IRHT qui rassemble des répertoires de textes et manuscrits médiévaux en langue d'oc et d'oïl, «*47*» est l'identifiant générique pour un auteur anonyme, tandis que «*8386*» désigne l'œuvre, dans ce cas le roman de *Kanor*. Enfin la dernière partie de l'URN désigne l'instance éditrice : «*c7s-fro7*». Dans ce cas, nous avons mis le nom du projet auquel participe cette édition : Canoniser les *Sept Sages*, «*fro7*» faisant référence au rang «*7*» de ce texte à l'intérieur d'un répertoire fermé de textes en ancien français, celui du Cycle des *Septs sages de Rome*, qui comporte une œuvre source et six continuations, dans le cas où le groupe de recherche souhaiterait adopter cette norme pour son projet.

1.1.1. Structurer le *teiheader*

Seuls les éléments principaux et représentatifs sont exposés ici. Une travail plus fin et conscientisé sera rédigé pour la fin de la thèse.

Le <teiHeader> comporte trois grandes sections :

1.1.1.1. Le fileDesc

Le <fileDesc> comporte lui-même :

Le <titleStmt> indique le titre et l'auteur du document édité.

Le <editionStmt> est une déclaration des différents acteurs du projet d'édition. Sa lecture est transparente (de manière générale, le langage de balisage XML-TEI propose des noms, en anglais, assez compréhensibles : on ne les commentera pas ici) :

```
<editionStmt>
  <edition n="N1">Édition numérique</edition>
  <respStmt xml:id="FPZ">
    <resp when="2022">Éditeur</resp>
    <name type="person">Florian-Pierre Zanardi</name>
    <!-- ORCID 0009-0005-2868-0006 -->
  </respStmt>
  <respStmt xml:id="YFJ">
    <resp when="2022">Co-directrice</resp>
    <name type="person">Yasmina Foehr-Janssens</name>
    <!-- ORCID 0000-0002-2939-0343 -->
  </respStmt>
  <respStmt xml:id="MP">
    <resp when="2022">Co-directrice</resp>
    <name type="person">Marylène Possamai-Pérez</name>
    <!-- ORCID 0000-0002-9250-370X -->
  </respStmt>
  <respStmt xml:id="SV">
    <resp when="2022">Superviseur</resp>
    <name type="person">Simone Ventura</name>
    <!-- ORCID 0000-0003-2773-9940 -->
  </respStmt>
  <respStmt xml:id="MTM">
    <resp when="1978">Éditrice d'une précédente édition critique</resp>
    <name type="person">Meradith Tilbury McMunn</name>
  </respStmt>
</editionStmt>
```

Le <publicationStmt> offre un aperçu plus détaillé des principales informations sur l'édition : nom de l'éditeur, date, et extension précise du travail d'édition. Cette section comporte déjà un certain nombre de gestes interprétatifs sur l'objet matériel et esthétique.

```
<publicationStmt>
  <authority>UniGE (Université de Genève, Suisse)</authority>
  <authority>Lyon-II (Université Lumière-Lyon-II, France)</authority>
  <publisher>Florian-Pierre Zanardi</publisher>
  <availability status="restricted">
    <licence>usage pour les membres de l'équipe</licence>
  </availability>
  <date>2023-2028</date>
  <extent>Édition complète de la section nommée Li Histoire de <persName key="Kantor">Kantor</persName> et de ses frères contenue dans le ms. BnF fr. 1446, qui contient une récapitulation du cycle des continuations des Sept sages de <placeName key="Rome">Rome</placeName>, une version abrégée de la fin du <persName key="Pelyarmenus">Pelyarmenus</persName> part le témoin B) ainsi qu'une version remaniée, peut-être lacunaire quant à l'objet même, du roman de <persName key="Kantor">Kantor</persName>.</extent>
</publicationStmt>
```

Le <sourceDesc> est une partie importante du *TEI-header*. Il apporte toute information jugée pertinente par l'éditeur sur les manuscrits collationnés ou simplement consultés et utilisés pour l'édition. Pour le moment, il ne contient que le manuscrit de base, C [Note: *Le sigle correspond au manuscrit 1446 de la Bibliothèque Nationale de France*]. L'ensemble des données sont contenues dans un <msDesc>, dont le patron a été repris du projet *msDesc*. Cette partie correspond, toutes choses égales par ailleurs, à une notice codicologique en format XML-TEI. Nous avons choisi de mener l'enquête aussi loin que nous le permettent nos connaissances, afin d'apporter le maximum d'informations sur ce manuscrit si curieux [Note: *Sources : Derolez, Albert, The Palaeography of Gothic Manuscript Books : From the Twelfth to the Early Sixteenth Century, Cambridge University Press, 2003. ; Careri, Maria et alii, Album de manuscrits français du XIIIe siècle : mise en page et mise en texte, Viella, 2001. ; Careri, Maria, et alii, Livres et écritures en français et en occitan au XIIe siècle : catalogue illustré, Viella, 2011. ; Camps, Jean-Baptiste, La Chanson d'Otinel. Édition complète du corpus manuscrit et prolégomènes à l'édition critique, 2016.*]. Nous présentons uniquement ici les éléments saillants.

Le manuscrit BnF fr. 1446 est important à plusieurs égards. Tout d'abord, il se distingue par la présence d'un ensemble d'œuvres dont la rédaction ou la composition peuvent être qualifiées d'exceptionnelles. Outre une rédaction non cyclique et composite de la dernière des continuations du Cycle des Sept Sages de Rome, il faut signaler la présence de brouillons et ébauches attribuées à Baudouin Butor, occupant les marges inférieures des folios 70 à 109 et les folios 108 à 115, et qui sont autant de documents incontournables pour qui s'intéresse à la genèse du Roman de Perceforest. Le fr. 1446 contient aussi le manuscrit unique du Couronnement de Renart, version wallonne du Roman de Renart. Il faut signaler un témoin des Fables de Marie de France. Le manuscrit a enfin agrégé, dans sa seconde partie, trois manuscrits à l'origine séparés contenant les œuvres poétiques de Baudouin et de Jean de Condé. Le Pelyarmenus et le Kanor peuvent être lus dans quatre autres manuscrits, ici référencés V2 (BnF fr. 22549) et V3 (BnF fr. 22550), B (KBR 9245), X2 (BL Harley 4903) et G (BnF fr. 93). Le manuscrit B est celui qui se rapproche le plus du nôtre en termes de variantes pour la partie Pelyarmenus (à compléter ensuite). Notons également la source M, miniatures avec texte au dos (Musée de Reims cadre 882, III-VI).

Une recension complète des œuvres contenues dans C a été faite. Travail classique, mais nécessaire, nous avons repris et détaillé de manière plus poussée les analyses parfois parcellaires de M. McMunn, qui devait travailler sur des microfilms monochromes de qualité inférieure aux reproductions actuelles.

```
<objectDesc form="codex">
  <supportDesc material="perg">
    <support>Vélin</support>
    <extent>
      <dimensions unit="mm">
        <width>285</width>
        <height>210</height>
      </dimensions>
    </extent>
    <foliation>
      <p>iii + 210 folios (à l'encre noire en haut à droite) ; XIX-CCX  
à l'encre brune en bas à droite à partir du f. 115r (Lais du  
pellican).</p>
    </foliation>
    <condition>
      <p>Accidents en divers endroits (avec perte d'information (usure  
: f. 1rb ; mutilation : f. 2) ou sans perte (trous antérieurs  
à la rédaction (ff. 40rb, 42rb)), (déchirures (ff. 12, 21,  
22)), avec réparations (sutures : ff. 47rb, 56rb, par  
exemple). Les ff. 164 et 196 sont blancs.</p>
    </condition>
    <supportDesc>
    <layoutDesc>
      <layout columns="2" ruledLines="53">
        <p>
          <locus from="f1" to="f70"/>Document structuré en deux  
colonnes de 53 Ur. Des marques de réglure, probablement  
réalisées à la pointe sèche, sont visibles. La mise en page  
est simple, avec espacement régulier entre les colonnes. Les  
marges qui entourent la zone de texte sont laissées libres, à  
l'exception, pour la seconde main, de quelques lettres  
d'attente à côté des lettrines. Un repère est présent sur le  
f. 13rb en marge extérieure, qui renvoie vers l'ancienne cote  
La Vallière, 13 (f. 28vb), maintenant ms. BnF fr. 22550.</p>
        </layout>
      </layoutDesc>
    </objectDesc>
```

Nous reproduisons intégralement l'élément <handDesc>, car il contient un travail paléographique original. Nous y mettons au jour l'existence d'au moins trois mains ayant participé à la copie de ce roman, ce qui contredit quelques analyses faites jusqu'alors [Note: une analyse paléographique encore plus précise, avec exemples imagés, sous forme de commentaire alphabétique, est disponible dans le fichier analyses_paleographiques.pdf]

<handDesc hands="3">
 <p>Une des plus anciennes descriptions formelles du ms. BnF fr. 1446 remonte à l'édition d'Alfred Foulet du Couronnement de Renart de 1929. Dans son analyse, il indique qu'un seul copiste aurait pris en charge la partie Kanor. Des doutes sont déjà présents chez Meredith McMunn, responsable de l'édition de 1978 du Roman de Kanor ; elle indique « croil[re] voir un changement de copiste au fol. 18ra. » De manière surprenante, dans un article beaucoup plus récent et fouillé portant sur les brouillons de Baudouin Butor présents dans ce manuscrit, N. Chardonnes et B. Wahlen reconduisent l'analyse de A. Foulet et ne voient qu'une seule main pour les folios la jusqu'à 70b, soit la partie dite Kanor. L'importance d'une identification précise des mains ayant concouru à la rédaction de ce manuscrit est redoublée par les enjeux de paternité et de création des œuvres qu'il renferme : des rapports intertextuels tangibles ont en effet été établis entre le Roman de Kanor et les brouillons de Butor.</p>
 <p>Face à un flou préjudiciable à la bonne connaissance des différentes campagnes d'écriture de la section, on souhaite a minima fournir une description paléographique qui permettrait de mettre au jour les différentes mains identifiables clairement ainsi que les accidents de rédaction qu'on croit déceler.</p>
 <p>La partie <persName key="Kanor">Kanor</persName> comporte au moins trois mains.</p>
 <handNote xml:id="hand1">
 <p>f. 1ra-17vb/24va-fin</p>
 <p>Écriture : type hybrida (a à simple ove, des f et s longs filant sur la ligne, et des lettres à hastes dépourvues de boucles.) Geste d'écriture parfois assez lâche, avec à certains endroits beaucoup de corrections, d'exponctuations. Le positionnement labile des signes abrégatifs et la détérioration matérielle du manuscrit rendent parfois la lecture difficile.</p>
 <p>Formes individuelles remarquables :
 "a" : simple ove. On peut apercevoir des "a" avec crosse mais non bouclés, mais la plupart est bien à simple ove (souvent avec une petite crosse).
 g : hampe très fréquemment exagérée en retrait latéral gauche jusqu'à 2 caractères.
 s : en début de ligne, la hampe se développe dans la marge. En fin de ligne, lorsqu'elle termine le mot, elle est fréquemment en forme suscrite.
 i : marquées d'un accent oblique pour le distinguer d'autres jambages en cotexte ambigu. Non systématique et laxité positionnelle du signe.
 haut et bas de colonnes : tracés des hampes des p, j, g, s (!) pour les bas ou des hastes des l, s (!) pour les hauts parfois exagérés, jusqu'à ressembler à une fioriture décorative (folio 31v pour une vue d'ensemble).

 </p>
 <p>Abréviations : système abrégatif normal pour un texte littéraire. La laxité des emplois, pour ne pas dire leur irrégularité, rend la lecture parfois complexe (c'est le cas du tilde vertical qui agit comme un opérateur abrégatif à usage large).
Tilde avec diverses valeurs, dont nasalisation des voyelles, abréviations usuelles de p barré droit et paraphe, q barré, signe conventionnel pour -us et conjonction et. Contraction de mots fréquents jusqu'à réduction à une lettre avec tilde vertical suscrit (trés, dist, <persName key="Kanor">Kanor</persName>, <persName key="Pelyarmenus">Peliarmenus</persName>, <persName key="Garin">Garin</persName>) par une positionnement imprécis du signe abrégatif, surtout le tilde horizontal ou par un oubli du signe.</p>
 <p>Structuration : La ponctuation est assurée par deux signes différents : le point situé sur la ligne (punctus) pour distinguer les propositions apposées ou subordonnées. Suivi

```

d'une majuscule, il signale le début d'une nouvelle phrase. Le
point virgule pour marquer une ponctuation forte en fin de
paragraphe. La coupe des mots à la fin des lignes est souvent
signalée par un tiret. On retrouve comme dans le reste du
manuscrit, l'alternance typique à l'époque entre initiale rouge
filigranée de bleu et initiale bleue filigranée de rouge. Le
copiste corrige par exponctuation ses erreurs, fréquentes. La
coupe en fin de ligne est signalée par un trait d'union. Trait
fin pour la coupure de mot en fin de ligne.</p>
</handNote>
<handNote xml:id="hand2">
<p>f. 18ra-24rb</p>
<p>Écriture : type textualis ("a" à deux étages + f/s posés sur la
ligne + absence de bouclages sur les lettres à haste b/k/l, à
l'exception notable du "h") Lettres bien détachées, formalité
"libraria" avec traits réguliers.</p>
<p>Formes individuelles remarquables : "h" de type cursiva. "d"
arrondi recourbé, "y" bien représenté, pointé, avec jambage
inférieur marqué. "w" v anglais avec tilde vertical suscrit
(wide).</p>
<p>Abréviations (écarts par rapport à hand1) : "et" tironien
barré.</p>
<p>Structuration : assurée par des initiales simples colorées
rouges ou bleues alternativement, sans filigrane. Lettres
d'attente dans la marge. Pas de pied-de-mouche. Trait fin pour
la coupure de mot en fin de ligne.</p>
</handNote>
<handNote xml:id="hand3">
<p>f. 29vb(1.14)-29vb</p>
<p>Cette main, quoiqu'assez semblable à la main 1, diffère par son
trait plus appuyé, un ductus plus rond et un «et» tironien barré
(une exception). La main débute à partir du segment
prépositionnel "a ce que vos saciés". L'échantillon est court
mais suffisant pour poser l'existence d'une troisième main, qui
plus est à un lieu problématique du ms. sur le plan
codicologique. En effet, c'est ici qu'on suppose une perte de
folio(s), qui reste à établir plus précisément (pas de talon
repéré à cette endroit, ce qui semble indiquer qu'aucun folio
n'a été arraché).</p>
<p>Écriture (avec comparaison différentielle par rapport à hand1) :
hastes et hampes crochétées (p/b/l (plus que hand1). Q : hampe
plus courte voire ne dépassant pas la ligne, avec bouclage
ponctuel sur la droite (l. 21 "que"). Accents : en contexte de
jambages : traîfe, manière, auoit, u#lonie). Ponctuation :
syntactique. On retrouve le punctus elevatus après l'exclamation
"ha", qui est un marqueur d'oralité, dans un usage semblable à
hand1.</p>
</handNote>
</handDesc>

```

Les informations sur les letrines du ms. sont contenues dans l'élément <decoDesc>, sur la couverture dans <bindingDesc>. L'histoire du ms. (origines et différentes acquisitions, ainsi que l'identification et la datation de tous les ex-libris, institutionnels et individuels) est lisible dans <history>.

1.1.1.2. L'encodingDesc

L'<encodingDesc> permet de donner des informations sur l'encodage du fichier XML. Cette partie de la réflexion est encore à l'état de brouillon, quoique déjà avancée.

1.1.1.3. Le profileDesc

Le <profileDesc> contient les notices de tous les noms de personnage et de lieu apparaissant dans le roman. Chacune des notices possède un identifiant @xml : id vers lequel vient pointer chacune des apparitions du lieu ou du personnage dans le roman de *Kanor* [Note: pour le moment, l'implantation des données issues de l'édition *Pelyarmenus* dans la nôtre contient encore des erreurs. Le principe est toutefois fonctionnel, comme on peut le voir sur la sortie HTML, avec les infos-bulles.]. Cette méthode permettra à terme de générer automatiquement à partir des fichiers XML les index des noms de lieux et des noms propres. La page HTML qui reproduit l'édition en cours fait ce lien entre le balisage du texte et l'affichage d'une information biographique pour chaque personnage.

Les notices de **noms de personnages** sont regroupées dans le <particleDesc> puis <listPerson>. Ce travail a déjà été réalisé pour l'édition numérique du *Pelyarmenus*. Nous avons repris la liste, non exhaustive, des personnages de ce projet et nous l'avons complétée et adaptée, car d'une part la liste n'est pas exhaustive (pour des raisons propres à l'édition), et l'extension du texte n'étant pas la même, ses personnages ne correspondent pas entièrement. Nous souhaitons proposer une liste complète des personnages (quitte à décider de faire apparaître les personnages très minoritaires d'une autre manière dans l'*index*). Le critère numérique des apparitions nous semble peu adapté à une œuvre qui contient de nombreux passages résumptifs, où sont susceptibles d'être mentionnés une seule fois des personnages pourtant importants. Si la constitution d'un *index nominum* dans une édition critique semble aller de soi, la finalité de son emploi doit tout de même être questionnée, surtout lorsque les moyens techniques offrent une infinité d'informations stockables. Les choix actuels de l'équipe contiennent les éléments <event>, <occupation> ainsi que des informations sur les rapports, parfois complexes, relationnels entre les personnages. Ces choix sont cohérents avec l'importance du statut, du lignage et de la filiation dans le cycle. La fonction de récapitulation, si particulière à C en ce que ce témoin emprunte, parfois à nouveaux frais, de la matière narrative aux autres continuations, appelle peut-être un complément d'information, comme le moment où apparaît et disparaît un personnage. Canoniser, c'est aussi construire ou mettre au jour un réseau d'autorités. Le roman de *Kanor* construit entre autres son canon autour de l'instrumentalisation de figures d'autorité (Virgile, Aristote, Averroès, héros antiques, etc.), dont la référentialité historique, dans le roman, est parfois questionnable. Ces acteurs du roman, qui sont somme toute aussi des personnages, feront l'objet d'un traitement particulier. Par ailleurs, il pourrait être intéressant, toujours dans l'idée de construire un canon, de lier ces notices de personnages aux notices BnF des personnages, ainsi que sur le site Biblissima, afin de donner de la lisibilité aux textes.

Les notices de **noms de lieux**, regroupées dans le <settingDesc> puis <listPlace>, contiennent le nom du lieu, un identifiant grâce à l'attribut @xml:id. On souhaite renseigner à terme une courte description s'il s'agit d'un lieu propre à l'univers référentiel du cycle (le «Val Tabour» par exemple), ou bien le situer s'il s'agit d'un site historique, avec élargissement culturel si nécessaire («Thiberiadis»). Les occurrences qui apparaissent dans le texte sont signalées pas la balise <placeName> et identifiée grâce au pointeur @key.

Afin de constituer les notices des noms de lieux, nous avons utilisé les balises suivantes : [Note: Idées provenant en partie des travaux d'A. Pinche dont la pertinence me semble importante pour nos textes : la progression excentrique, au fil des cycles, de l'exploration des territoires, qui participe de l'opposition Occident/Orient (Y. Foehr-Janssens, De Jérusalem à Rome...) mais qui la dépasse, oblige à prendre en compte cette dimension dans notre étude pour mettre au jour le « symbolisme géographique » de l'œuvre (S. Seláf, Constantinople et la Hongrie dans le cycle des Sept sages de Rome). Un balisage fin des données relatives à l'espace apparaît pertinent].

Quand le lieu est une ville identifiée, ses coordonnées sont ajoutées afin de pouvoir à terme faire une carte des lieux cités dans le corpus.

```
<place corresp="http://pleiades.stoa.org/places/423025"
xml:id="rome">
  <placeName full="yes">Rome</placeName>
  <country type="pays">Italie</country>
  <country type="civilisation">Occident</country>
  <location>
    <geo>41.891775, 12.486137</geo>
  </location>
  <note>Ville italienne, capitale de l'empire romain d'occident.</note>
</place>
```

Pour assurer la bonne identification du lien, nous avons ajouté avec @corresp un lien vers une notice externe, de préférence vers le site *Pleiades : Ancient Places* quand cela était possible, sinon vers *data.bnf*

Enfin, afin de pouvoir classer les différents lieux cités par pays, mais aussi en fonction de leur rattachement à l'Orient ou à l'Occident nous avons utilisé l'élément <country> avec @type de valeur pays pour signaler le pays et @type de valeur civilisation pour indiquer l'appartenance à l'Orient ou à l'Occident.

1.1.2. Structurer les différentes parties du texte

Le corpus est structuré à l'aide des balises suivantes :

Le texte est structuré à l'intérieur d'une balise <text>. <text> englobe l'élément <body> qui possède deux attributs. Le premier, @n, permet de spécifier l'URN CTS, tandis que le second, @xml:lang, indique la langue majoritaire du corpus, ici, l'ancien français signifié par la valeur fro. Le <body> contient des <div> avec un @type égal à « chapter », qui équivalent à une unité correspondant au **chapitre**. Ils sont numérotés à l'aide de @n. Si le texte présente un titre rubriqué, alors on le fera apparaître dans l'édition dans un élément <head>, enfant de la <div>. L'élément <p> constitue une sous-unité textuelle, le **paragraphe**. Il possède un attribut @n de valeur n qui correspond à un paragraphe dans l'édition. Chacune de ces divisions est numérotée et forme un texte ventilé en chapitres et paragraphes, chacun identifié par des numéros pour permettre une navigation plus aisée et un système de citation de l'édition efficace.

```
<div type="chapter" n="2">
  <head type="rubrique" n="2">[...]</head>
  <p n="P82"> [...]</p>
</div>
```

Les parties dialoguées sont comprises dans une balise <said> avec un attribut @who quand l'interlocuteur est un personnage nommé dans le récit [Note: le repérage de l'allocuteur n'est pas encore fait] . La valeur de l'attribut est alors un pointeur qui renvoie vers la notice du personnage dans le <teiHeader>.

```
Et dont apiela li emperere
l'escuhier et li dist : <said who="#Cassidorus" aloud="true"
direct="true">Mon ami, n'eüstes vous ore mie mout grant paour ?</said>
```

Il est précisé dans chacune des balises <said> si elles présentent un discours prononcé à haute voix ou non avec l'attribut @aloud, direct ou non avec l'attribut @direct [Note: héritage de l'édition Pelyarmenus] .

L'élément <seg> permet de signaler des éléments à l'échelle infra-paragraphe ou phrastique avec l'attribut @ana. Quatre types d'éléments sont relevés : **proverbe** : derrière ce terme se trouvent en fait des phénomènes de discours relevant à la fois d'« une extériorité et d'un figement » [Note: Paveau, Marie-Anne, Le préconstruit. Généalogie et déploiements d'une notion plastique. Florent Bréchet, Sabrina Giai-Duganera, Raphaël Luis, Agathe Mezzadri et Solène Thomas. Le préconstruit, approche pluridisciplinaire, 192, Classiques Garnier, 2017.] (ce balisage s'inscrit dans une réflexion qui lie, encore, le processus de canonisation à celui de construction de l'autorité + "portée sapientiale du cycle"). **discoursEnchassé** : pour indiquer quand un discours est intégré dans un autre discours direct (hérité de l'édition Pelyarmenus mais intéressant pour les rapports de subordination qui se construisent éventuellement lorsqu'un discours est repris). **lettre** : repérage des lettres. **chanson** : repérage des chansons (peu nombreuses dans Kanor) Ce corpus en prose contient aussi quelques vers. La question de créer un balisage spécifique pour ces rares occurrences se pose. La solution consisterait à réunir les vers dans une balise <lg> ayant un attribut @n. Chacun des vers serait englobé dans une balise <l> avec un attribut @n. Cela donnerait :

```
<seg ana="chanson">
  <lg n="1">
    <l n="1">Ne se doit desconforter, qui en vraie amors erent.</l>
    <l n="2">Rondet, ne te veil noter, d'autre chant que raconter.</l>
    <l n="3">Coment amor s'amonte, veulent chiaus de lor covent.</l>
    <l n="4">Ne se doit desconforter, qui en vraie amors erent.</l>
```

```
</lg>
</seg>
```

@xml:lang : les éléments du texte dans une langue étrangère (latin exclusivement) sont signalés à l'aide de l'élément `<seg>` et de son attribut `@xml:lang` [Note: un questionnement semblable à celui du balisage des vers se fait jour : il y a en fait très peu de texte latin dans le corpus.]

```
et dist en teil maniere : <seg xml:lang="lat" type="proverbe">similis similem cuerit</seg>.
```

1.2. Principes de normalisation

Si une édition imitative a bien été réalisée selon les principes établis [Note: Pinche, Ariane, Guide de transcription pour les manuscrits du Xe au XVe siècle, 2022.], il n'est pas prévu, pour le moment, d'intégrer ces résultats à l'édition numérique. Ils sont toutefois disponibles dans un document ALTO issu d'eScriptorium

1.2.1. Encodage de la mise en page du manuscrit

Afin de pouvoir produire à terme une interface proposant en regard une vue du ms. et le texte du folio édité, des informations sur la disposition du texte ont été encodées en utilisant les balises suivantes :

Chaque saut de page est indiqué avec élément `<pb>` qui contient, grâce à l'attribut `@n`, le numéro du folio et, grâce à l'attribut `@facs`, un lien vers le folio concerné du manuscrit C, numérisé sur le site Gallica. L'indication *recto* ou *verso* est contenue dans le `@n` grâce aux lettres qui suivent le numéro, *r* ou *v*.

```
<pb facs="https://gallica.bnf.fr/ark:/12148/btv1b10023851v/f7.image"
n="1r"/>
```

Seul le passage à la colonne 2, la «b», est indiqué avec l'élément `<cb>` et contient, grâce à l'attribut `@n`, la lettre de la colonne à suivre. Nous distinguons les deux faces du folio, donc il y a une colonne «a» puis «b» au verso.

```
<cb n="b"/>
```

Les lettrines, ainsi que les mots rubriqués ont été indiqués grâce à la balise `<hi>`.

```
<hi rend="decorated-initial 2 blue">Q</hi>uant li emperere
```

Ce sont les valeurs de l'attribut `@rend` qui indiquent le type de décoration présent dans le manuscrit. Voici la liste des valeurs utilisées dans le corpus :

- **rubricated** : désigne les lettres écrites à l'encre rouge.
- **decorated-initial** : désigne les lettrines décorées qui ouvrent les chapitres.

On remarque que `@rend` contient plusieurs valeurs cumulatives, bien décrites dans l'ODD : le chiffre correspond à l'espace en unité de réglure qu'occupe la lettrine, puis la couleur.

Les titres rubriqués sont contenus dans une balise `<head>` contenant un `@n` qui précise le numéro de rubrique et un `@type` qui indique qu'il s'agit d'une rubrique. Pour distinguer fond et forme, nous rebalisons le contenu de la rubrique à l'aide de la balise `<hi>`, déjà évoquée, pour faire apparaître le texte en rouge.

```
<head n="1" type="rubrique">
  <pb facs="https://gallica.bnf.fr/ark:/12148/btv1b10023851v/f7.image"
  n="1r"/>
  <hi rend="rubricated">Ichi encomence li histoire de <persName key="Kantor">Kantor</persName> et de ses freres, liqueil furent fil au noble
  <persName key="Cassidorus">Cassidorus</persName>, empereor de
  <placeName key="Constantinople">Costostinnoble</placeName> et de
  <placeName key="Rome">Rome</placeName>, liqueil furent engrené en
    l'emperis <persName key="Fastige">Fastige</persName>, ki fille fu a
    l'empereor <persName key="Fiséus">Phiseus</persName>.</hi>
</head>
```

1.2.2. Encodage des normalisations

À ce stade du travail, l'objectif est de présenter une édition normalisée de ms. 1446. Si un travail d'édition facsimilaire devait être produit, cela ne pourrait toucher qu'à une petite portion du texte, au regard du temps d'encodage que demande ce type de relevé. Toutefois, dans l'idée de présenter une vue aussi documentée que possible de ce manuscrit, nous souhaitons sélectionner un ou deux folios représentatifs du diasystème du scribe de la partie *Kantor* pour en produire une vue facsimilaire. **Les principes d'édition de notre édition (leurs esquisses, à ce stade), sont contenus dans principes_edition.pdf.**

1.3. L'apparat critique

1.3.1. Interventions éditoriales

L'état du manuscrit est variable. La lecture de certains mots est parfois incertaine, ce qui est précisé au moyen de l'élément `<unclear>`, représenté dans la sortie HTML par [élément incertain]. L'attribut `@reason` permet d'indiquer si c'est parce que les caractères sont illisibles ("illegible") ou bien si cela est dû à un effacement de l'encre ("faded") (la valeur "acertain" est une valeur de travail dans les cas qui requièrent un second regard). `@cert` précise le degré de certitude, *low* ou *high*.

```
Il mie ne s'oublia a celui jour
dont <unclear cert="high" reason="faded">l'em</unclear>perere et li
empereris s'estoient parti la nuit devant.
```

Si des caractères sont totalement illisibles voire disparus (cas des folios mutilés), l'élément `<gap>`, représenté par « [x] », est utilisé. Si une portion de texte a dû être restituée par l'éditeur, par suite d'une lacune involontaire du manuscrit (dommage matériel, par exemple), l'élément `@supplied`, représenté sur le HTML par « texte chevron portion restituée chevron » a été utilisé. `@reason` précise la raison de cette supplétion (quasiment toujours "omit-

ted"). L'élément `<note>` a été utilisé pour ajouter des informations concernant l'édition. Il contient un `@type` qui précise la nature de la note.

1.3.2. Transcription critique des représentations physiques

A été conservé un certain nombre d'informations appartenant à un état «semi-diplomatique» du texte : *[Note: chacune de ces informations est représentée d'une manière détaillée sur le fichier HTML]*

L'élément `` est utilisé pour indiquer qu'une portion de texte a fait l'objet d'une opération de suppression à un moment donné de la rédaction. Le moyen de réalisation de cette suppression est conservé par l'attribut `@rend` ("exponctuation") et le responsable par `@hand`.

```
Quant li esc<del hand="#hand1" type="exponctuation">h</del>uhiers oï çou,
```

L'élément `<hi>`, avec l'attribut `@rend` est utilisé pour décrire des portions de texte qui se distinguent par leur rendu formel ou chromatique : *lettrines* ou *rubriques*, cas déjà évoqués ailleurs.

L'élément `<add>` est utilisé pour indiquer qu'une portion de texte a été ajoutée à un moment donné de la rédaction. L'emplacement de cet ajout sur le folio est décrit par l'attribut `@place`, qui peut avoir les valeurs suivantes : "above" (la portion de texte se situe dans l'interligne), "margin" (la portion de texte se situe dans la marge, qui est par défaut la marge de gauche), "top" (la portion de texte se situe dans la marge supérieure de la page), "bottom" (la portion de texte se situe dans la marge inférieure de la page), "inline" (quand l'ajout a été fait directement sur la ligne). Le responsable, rarement identifiable, est noté dans `@hand`.

```
Dame, <add hand="#hand1" place="above">une</add> puciele les aporte de hors vile
```

1.3.3. Les corrections

Selon la main qui le copie, le texte offre différents degrés de fiabilités. Nous l'avons présenté dans l'analyse paléographique, la main *hand2* est beaucoup plus net et fiable que *hand1*. Le travail d'édition étant encore en cours, nous ne proposons ici qu'un aperçu des outils que nous comptons à terme employer. De manière générale, toutes les balises `<corr>` contiennent un attribut `@cert` (degré de certitude du geste d'édition) et `@resp` (responsabilité du geste).

Lorsqu'un manque dans le texte empêche la compréhension, nous corrigeons à l'aide de la balise `<choice>` avec l'alternative `<sic>` fermant et `<corr>` contenant l'élément restitué.

```
En ceste pensee avint que li prince<choice>
<sic/>
<corr cert="low" resp="#FPZ">s</corr>
</choice>
```

Quand le texte du manuscrit est difficilement compréhensible en raison d'un ajout de texte, la plupart du temps lié à une faute par dittographie dans le texte, les lettres en trop sont englobées dans une balise `<surplus>`. Un attribut `@reason` indique la raison de cette suppression.

```
qui ne chaçoient fors hustin et
laidire tous ciaus <surplus reason="doublon">et laidire tous ciaus</surplus> de
cui
```

Enfin, quand le texte présente une erreur ponctuelle dans le texte qui demande le recours à un autre témoin (lorsque cela est possible) pour remplacer la graphie "fautive" de C, nous avons utilisé le doublet `<sic>` et `<corr>` englobé dans une balise `<choice>`.

```
que mout <app>
<lem wit="#V3">i</lem>
<rdg wit="#C">
<supplied reason="omitted">e</supplied>n </rdg>
</app> aroit fait d'armes
```

Ce cas de figure se présente dans les parties qui reprennent de manière suivie un témoin du *Pelyarmenus*, et il se posera de manière fréquente dans *Kanor*.

1.3.4. Collations des variantes

La réflexion concernant le travail de collation est toujours en cours. Quoi qu'il en soit, l'état des lieux sur l'intérêt de ce geste d'édition est tel que : 1) la partie *rechapitulation* est propre à C (témoin unique) : il n'y a rien à collationner. 2) *Pelyarmenus* fait déjà l'objet d'une édition critique avec collation, donc s'il devait y avoir un apport pertinent, ce serait de C vers cette édition et non l'inverse (intérêt scientifique limité). 3) la partie *Kanor*, enfin, pourrait profiter d'une collation. Mais à nouveau, C peut difficilement être considéré comme un témoin d'autorité. Un travail de réédition numérique du *Kanor* devrait évidemment prévoir d'intégrer C pour en mesurer tout l'écart et toute la fécondité. Toutefois, il est bien prévu de faire des sondages signifiants, en prenant appui sur B et V notamment. Ces points de "références" nous permettront sans nul doute de mieux établir la spécificité de cette narration.