



**HAL**  
open science

## Acquérir le texte à partir de l'image : les technologies OCR/HTR

Morgane Pica, Mathieu Goux

► **To cite this version:**

Morgane Pica, Mathieu Goux. Acquérir le texte à partir de l'image : les technologies OCR/HTR. École thématique. Formation à Transkribus, Université de Poitiers, Poitiers, France. 2022, 40 p. hal-04799886

**HAL Id: hal-04799886**

**<https://hal.science/hal-04799886v1>**

Submitted on 23 Nov 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

# Formation Transkribus

Acquérir le texte à partir de l'image :  
Les technologies OCR/HTR

---

5-6 mai 2022, Université de  
Poitiers

morgane.pica@ens-lyon.fr  
mathieu.goux@unicaen.fr



# Plan général de la formation

## I. Présentation générale (5 mai)

1. De l'importance de la méthode
2. Les chaînes d'acquisition des données textuelles
3. L'outil Transkribus
4. Présentation des outils du cours

## II. Import des documents et images (6 mai, matin)

1. Import et gestion des collections et utilisateurs
2. Fonctionnement des images
3. L'outil d'analyse automatique
4. Découpage et typage manuel des images
5. Entraînement d'un HTR

## III. Texte et export (6 mai, après-midi)

1. Examen des résultats de l'HTR
2. Évaluer les performances du modèle
3. Gérer les versions
4. Formats d'export

# Présentation générale :

## 1. De l'importance de la méthode

Souvent, on se lance à corps perdu dans un processus très énergivore, par ignorance des possibilités réelles de l'informatique.

On peut éviter de s'entendre dire « vous auriez pu le faire en 2 jours plutôt que 2 ans, si vous vous y étiez pris autrement », avec un minimum de méthode.

En plus de vous montrer Transkribus, nous allons vous donner des astuces pour vous permettre de gagner du temps.

Ce gain de temps ne sera peut-être pas immédiat car vos propres étapes de traitement seront un peu plus longues, mais se manifesteront dans les étapes suivantes.

# Présentation générale :

## 1. De l'importance de la méthode

Si vous réfléchissez ainsi, vos données gagneront :

- en fiabilité,
- en stabilité,
- en réutilisabilité,
- en conformité aux principes FAIR (oui!).

# Présentation générale :

## 1. De l'importance de la méthode

Transkribus est un outil très puissant... si on l'utilise correctement.

Le problème n'est pas de savoir appuyer sur les bons boutons.

Le problème est de savoir anticiper de futures utilisations, voire une (semi-)automatisation dans les étapes futures. On ne sait jamais comment un projet de transcription/édition peut évoluer.

- Il ne s'agit pas de réfléchir en ingénieur, mais d'acquérir quelques automatismes. Ces automatismes vous simplifieront aussi la vie si votre projet est interrompu et vous devez reprendre vos données plus tard.
- = « Qui peut le plus peut le moins. »

En plus de vous montrer les boutons de Transkribus, nous allons essayer de vous montrer quels automatismes peuvent vous aider, ainsi que vos éventuels collaborateurs, à vous simplifier la vie.

# Présentation générale :

## 2. Chaînes d'acquisition

### Édition numérique :

- Travail scientifique, utilisant les moyens numériques, d'adaptation pour le public d'une œuvre rendue obscure à l'œil moderne par l'état de la langue ou de l'écriture dans lesquelles elle a été initialement conçue et/ou représentée.
- Travail d'enrichissement scientifique des données brutes d'une œuvre par des moyens numériques.

L'utilisation du numérique change la forme mais non le fond de la démarche. Elle change également les possibilités de l'éditeur.

# Présentation générale :

## 2. Chaînes d'acquisition

Pourquoi une édition numérique ?

- Faciliter le partage des travaux scientifiques.
- Faciliter la recherche et l'analyse automatiques.
- Profiter des nombreuses possibilités structurelles et sémantiques des langages numériques.
- Repenser l'approche d'une œuvre en s'affranchissant des contraintes du format papier.



# Présentation générale :

## 2. Chaînes d'acquisition

« Chaîne de numérisation » :  
de quoi parle-t-on ?

L'ensemble des procédés, physiques et numériques, permettant de passer d'un support papier originel à sa représentation numérique.

# Présentation générale :

## 2. Chaînes d'acquisition

- 1) Travail initial.
- 2) Numérisation.
- 3) Acquisition du texte.
- 4) Conception de la base de données.
- 5) Encodage du texte.
- 6) Conception des outils de consultation.
- 7) Mise à disposition du public.
- 8) Maintenance.

# Présentation générale :

## 2. Chaînes d'acquisition

### 1) Travail initial :

- Travail codicologique et bibliographique.
- Définition des buts et cibles.
- Évaluation du travail à venir.

# Présentation générale :

## 2. Chaînes d'acquisition

2) Numérisation : acquisition d'images de la source.

En fonction de :

- Accès à la source.
- État de la source.
- Taille de la source.
- Matériel.
- Buts scientifiques et techniques.

# Présentation générale :

## 2. Chaînes d'acquisition



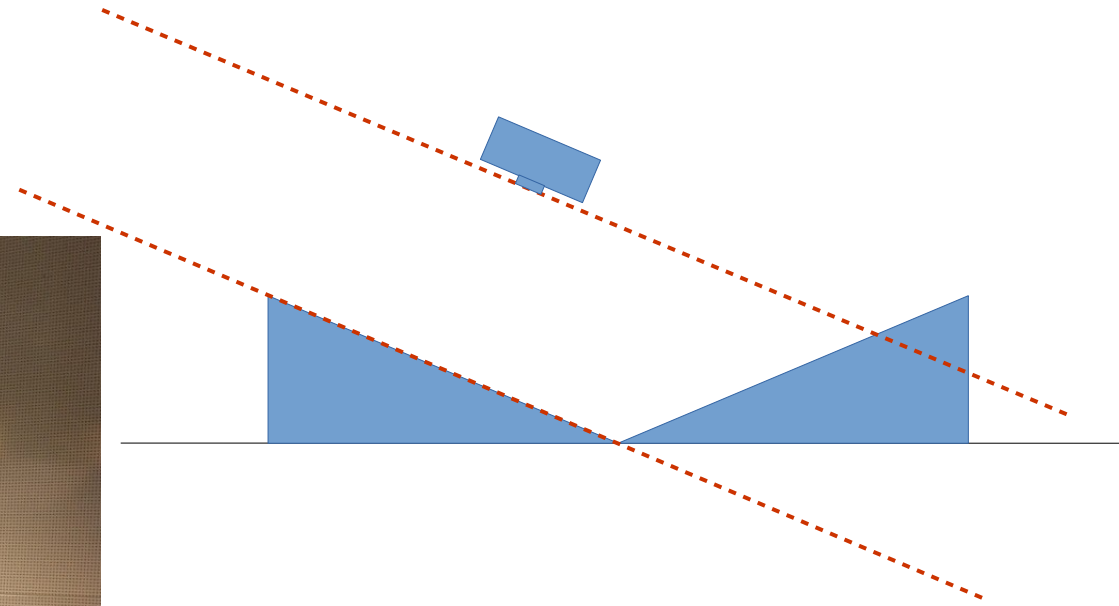
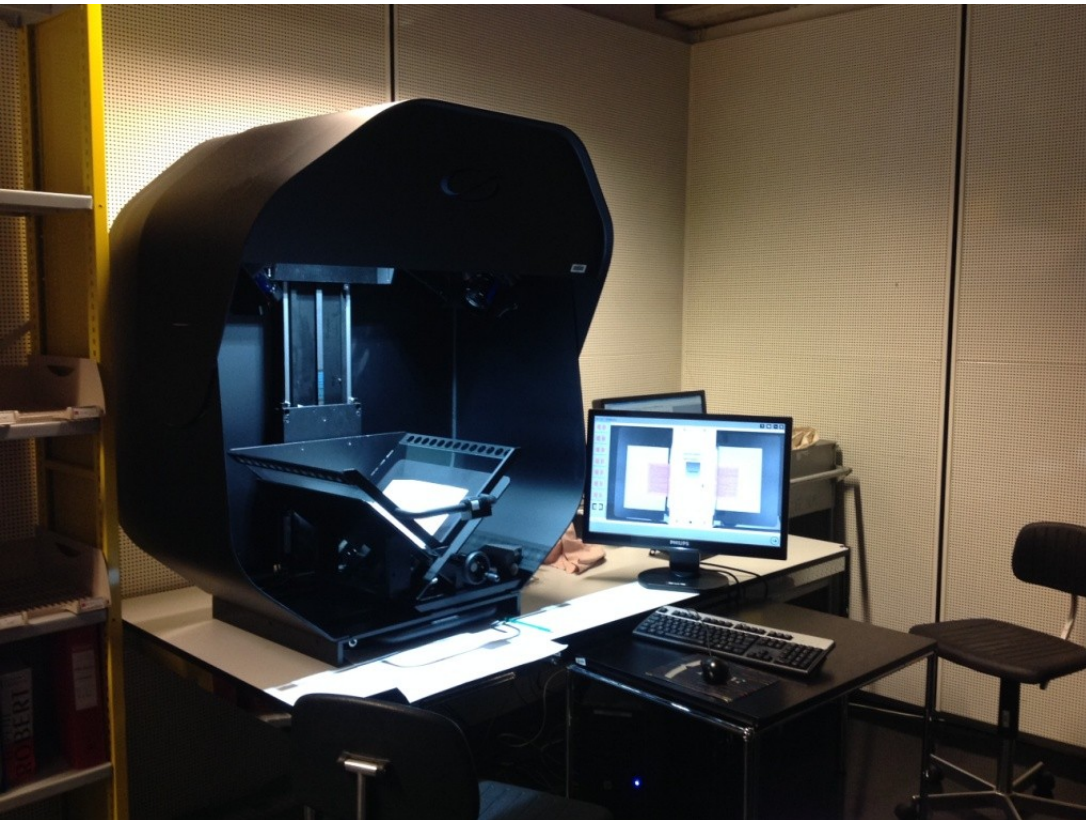
---

### Sources :

- p. 316, Très Ancien Coutumier, Paris, Bibliothèque Sainte-Geneviève, in ms. fr. 1743
- Au delà des murs Créations - <https://www.audeladesmurscreations.com/product-page/ref-113h-housse-pour-113-113p>

# Présentation générale :

## 2. Chaînes d'acquisition



# Présentation générale :

## 2. Chaînes d'acquisition

### 3) Acquisition du texte :

- Acquisition manuelle (lecture et saisie humaines).
- Automatisation :
  - OCR (Optical Character Recognition) : caractères majoritairement prédéfinis, conçu pour les textes contemporains imprimés.
  - HTR (Handwritten Text Recognition) : apprentissage progressif du programme permettant de prendre en compte les variantes d'une même lettre dans une écriture manuscrite.
    - Sortie en texte brut.

# Présentation générale :

## 2. Chaînes d'acquisition

### Junicode

Projet de créer des caractères Unicode pour retranscrire les sources antiques et médiévales dans une police gratuite et standard.

Œ œ þ ʒ Ɔ ɣ ꝥ ft ā t̄ ð ʳ

Développée par Peter Baker à l'université de Virginie dans le cadre du projet MUF1.

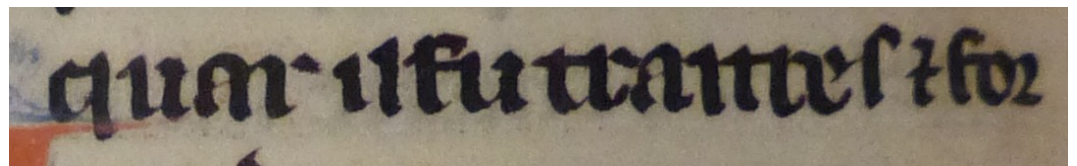


# Présentation générale :

## 2. Chaînes d'acquisition

### Entraînement d'un modèle d'HTR

- Transcription initiale manuelle d'un extrait représentatif.
- Synchronisation texte-image pour le modèle  
(caractère par caractère).



quar il fu traictes ] fo2

quar il fu traictes ] fo2

# Présentation générale :

## 2. Chaînes d'acquisition

### Entraînement d'un modèle d'HTR

- Transcription initiale manuelle d'un extrait représentatif.
- Synchronisation texte-image pour le modèle.
- Apprentissage initial suivant la synchronisation.
- Test initial sur de nouvelles pages.
- Correction du test.
- Reprise de l'apprentissage, et encore, et encore... jusqu'à satisfaction.

# Présentation générale :

## 2. Chaînes d'acquisition

4

DE JURISDICTION.

ARTICLE PREMIER.

Le Bailli, ou son Lieutenant, connoît de tous crimes en premiere instance.

Bailli signifie la même chose que Gardien ; comme Baillie signifie Garde & Protection. D'où vient que Baillifre, dans les anciennes Ordonnances & quelques Coutumes, signifie Tuteur ou Curateur. Le Bailli donc étoit comme le conservateur du Peuple & des Loix : mais ayant été remarqué, que les Baillis s'attachoient plus volontiers aux exercices des Armes qu'aux fonctions de Judicature, le Roi Louis XII ordonna que leurs Lieutenans seroient licenciés en l'Etude du Droit Romain, & leur attribua toute la Jurisdiction ; de forte qu'il ne reste aux Baillis pour marque de leur ancien pouvoir, que la préférence, la voix honoraire non délibérative, & leur nom mis dans le titre des Sentences (1) : les Ordonnances d'Orléans, de Moulins & de Blois ayant disposé, que les Baillis fussent des Officiers Militaires & de Robe-courte. Quand donc cet Article attribue au Bailli la connoissance de tous crimes en premiere instance, cela s'entend de ses Lieutenans : mais cette compétence est limitée par plusieurs exceptions ; car il y a des Crimes prévôtaux & préfidiaux, dont la connoissance en premiere instance appartient aux Prévôts des Maréchaux de France, aux Vice-Baillis & aux Préfidiaux. Voyez l'Ordonnance criminelle, au Titre de la compétence des Juges, aux Articles XI, XII & XV. Le Vicomte de l'Eau, qui a la police sur la Riviere de Seine jusqu'à

Un Arrêt du Conseil du 10 Décembre 1744, fait Règlement entre les Avocats & Procureur du Roi des Jurisdicions ordinaires, Règlement utile pour affermir la paix entre ceux qui, par état, doivent veiller à la tranquillité publique : mais on s'écrite journellement de nouvelles difficultés sur son exécution. Voyez un Règlement de la Cour de 6 Juillet 1763, & le Règlement de Justice du 18 Juin 1769.

La Stance & les fonctions des Officiers d'un même Siège dépendent ordinairement des Concordats & de la possession qui n'est point abusive.

(1) Le Bailli étoit autrefois un Officier très-important, il réunissoit en lui les fonctions militaires & civiles ; il étoit établi pour veiller, dit l'ancien Coutumier, au maintien des Etats de nos Ducs, conserver les intérêts de la Patrie, & faire régner au-dedans la paix & la tranquillité, en réprimant les attentats, soit par la force, ou par la févérité des loix. Nous avons conservé la formule de son serment, qui nous retrace, avec le tableau de ses devoirs, les prérogatives de sa dignité. Comme le commandement des armes ne lui permettoit pas toujours de s'occuper de la distribution de la justice, il se faisoit représenter par des Lieutenans qu'il créoit & destituoit à son gré, il les multiplioit suivant la distance des lieux & l'étendue de son Bailliage : ces Officiers n'auroient pu sans doute l'empêcher de juger ; aussi dès que le Bailli cessoit d'être Guerrier, il redevenoit Magistrat, & & il jugeoit lui-même ; rarement ses jugemens étoient attaqués, on s'en plaignoit quelquefois. Les grands pouvoirs font assez près de l'abus, la carrière séduisante des armes porta peu à peu le Bailli à négliger les fonctions de la justice ; dès qu'il commença de les mépriser, il en fit un honneux trafic ; il salut, pour rétablir l'ordre, multiplier les Ordonnances. Dans les derniers temps, on ne lui a enfin conservé, dans les tribunaux, que des honneurs, & les révolutions ont amené ses troupes réglées. Ainsi nos Loix anciennes nous ont appris ce que le Bailli a été, nous savons par nos usages ce qu'il est.

4

DE JURISDICTION.

ARTICLE PREMIER.

Le Bailli, ou son Lieutenant, connoît de tous crimes en premiere instance.

Bailli signifie la même chose que Gardien ; comme Baillie signifie Garde & Protection. D'où vient que Baillifre, dans les anciennes Ordonnances & quelques Coutumes, signifie Tuteur ou Curateur. Le Bailli donc étoit comme le conservateur du Peuple & des Loix : mais ayant été remarqué, que les Baillis s'attachoient plus volontiers aux exercices des Armes qu'aux fonctions de Judicature, le Roi Louis XII ordonna que leurs Lieutenans seroient licenciés en l'Etude du Droit Romain, & leur attribua toute la Jurisdiction ; de forte qu'il ne reste aux Baillis pour marque de leur ancien pouvoir, que la préférence, la voix honoraire non délibérative, & leur nom mis dans le titre des Sentences (1) : les Ordonnances d'Orléans, de Moulins & de Blois ayant disposé, que les Baillis fussent des Officiers Militaires & de Robe-courte. Quand donc cet Article attribue au Bailli la connoissance de tous crimes en premiere instance, cela s'entend de ses Lieutenans : mais cette compétence est limitée par plusieurs exceptions ; car il y a des Crimes prévôtaux & préfidiaux, dont la connoissance en premiere instance appartient aux Prévôts des Maréchaux de France, aux Vice-Baillis & aux Préfidiaux. Voyez l'Ordonnance criminelle, au Titre de la compétence des Juges, aux Articles XI, XII & XV. Le Vicomte de l'Eau, qui a la police sur la Riviere de Seine jusqu'à

[...]

# Présentation générale :

## 2. Chaînes d'acquisition

### 4) Conception d'une base de données :

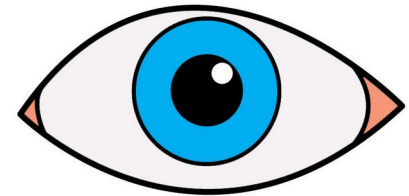
- Base de données textuelle, très souvent en XML-TEI (grandes et vastes possibilités et utilisation très large garante d'interopérabilité et pérennité).
- À nouveau, quelle finalité ?
- Quelle serait la structure logique du *corpus* ?
- Quel temps de travail a-t-on ?
- Quels outils sont prévus ?

# Présentation générale :

## 2. Chaînes d'acquisition

### 5) Encodage du texte : quelques notions.

Texte



Un **texte** est une série orale ou écrite de mots perçus comme constituant un ensemble cohérent, porteur de sens et utilisant les structures propres à une langue (conjugaisons, construction et association des phrases...). Un texte n'a pas de longueur déterminée sauf dans le cas de poèmes à forme fixe comme le sonnet ou le haïku.

L'étude formelle des textes s'appuie sur la linguistique, qui est l'approche scientifique du langage.

- Police plus grande
- Placé au-dessus
- Séparé du corps



= titre

# Présentation générale :

## 2. Chaînes d'acquisition



Texte

Chaîne de caractères,

- › Police : « Liberation Sans »
- › Taille : 22 px.
- › Couleur : noire.
- › Alignement : gauche.

# Présentation générale :

## 2. Chaînes d'acquisition



Chaîne de caractères

➤ “Texte”

Enrichissement structurel

➤ = titre

Enrichissement formel

- police
- couleur
- position...

# Présentation générale :

## 2. Chaînes d'acquisition



Enrichissement structurel

➤ = titre

XML



# Présentation générale :

## 2. Chaînes d'acquisition

- = *eXtensible Markup Language*
- Langage d'encodage ( $\neq$  programmation)
- Grammaire : langage à balises.

<titre> Texte </  
titre>

<titre type="principal"> Texte </titre>

- Finesse de l'encodage jusqu'au caractère.
- « eXtensible » = mélange de vocabulaires XML

# Présentation générale :

## 2. Chaînes d'acquisition

- = *Text Encoding Initiative*

<https://tei-c.org>

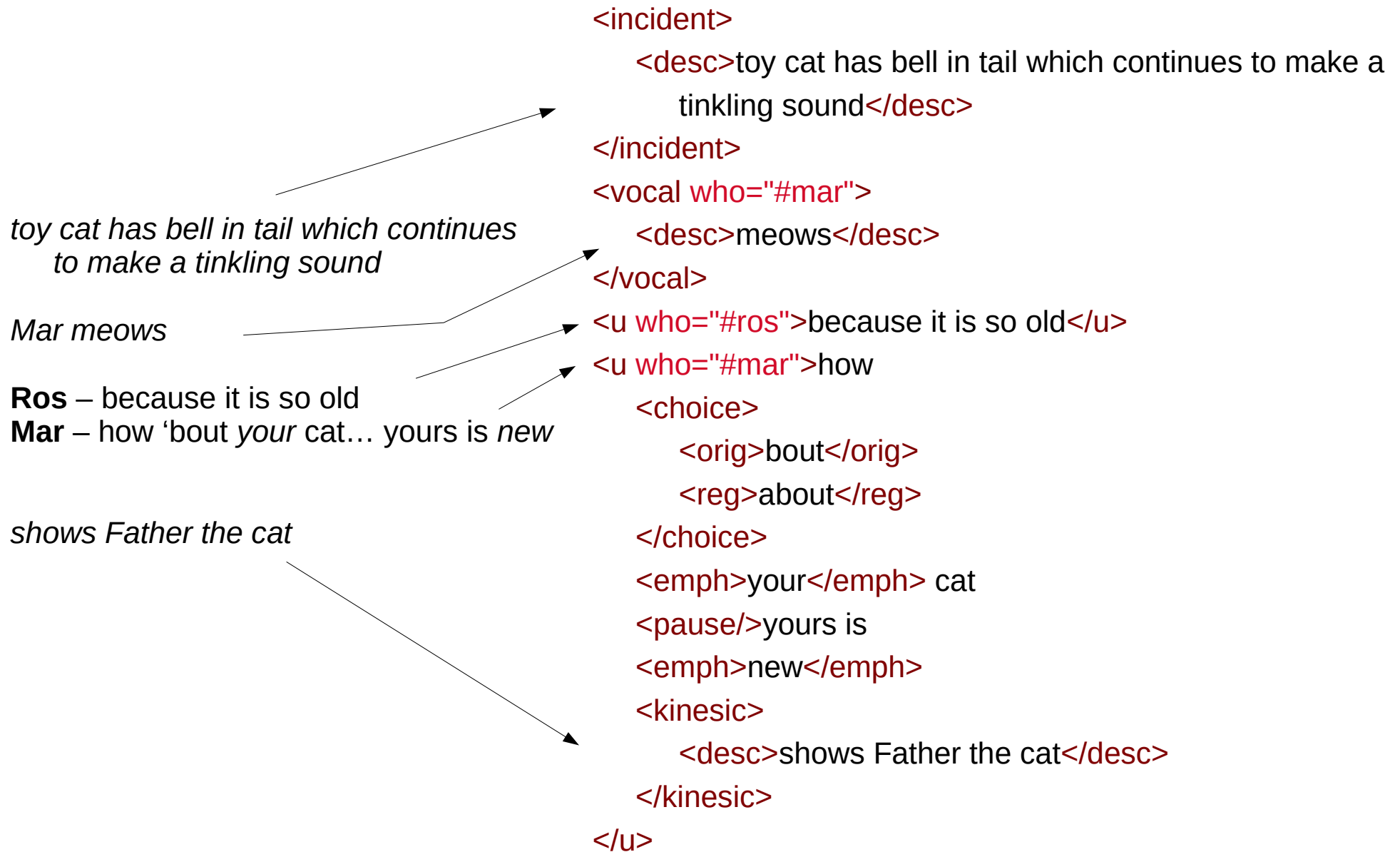
- Nom de domaine = vocabulaire spécifique

= set de balises prédéfinies

- Encodage sémantique et structurel des textes oraux et écrits, quel que soit le support originel.
- Description fine et précise des supports originels et métadonnées.

# Présentation générale :

## 2. Chaînes d'acquisition



# Présentation générale :

## 2. Chaînes d'acquisition

Document daté un 10 avril (par le scribe)  
entre 851 et 867 (par l'éditeur).

<date>

<date when="--04-10">

**un 10 avril,**

</date>

<date notBefore="0851" notAfter="0867" cert="high">

**entre 851 et 867**

</date>

</date>

# Présentation générale :

## 2. Chaînes d'acquisition

Sujet

Coutume

Commentaire français

Commentaire latin

De assise,

Entête

ne peut estre terminee  
fors en l'assise ou en les  
chiquier / ou par de  
uant le p'ce : par quoy  
la fin en puisse estre re-  
cordee.

De assise. Chap. lv.  
Assise est vne  
court en la  
quelle ce q  
est fait doit  
auoir par

Texte principal  
(gauche, droite)

Secunda additio.  
In textu ibi.

**A**ce cry zc.  
**2** Additio. Facit quod di. Bar. in l. vt vim ad si. ff. de insti. c. iure vbi dicit qd insultat? debet clamare succurrite / vel occurrere / vt non an. lit. ap. et ibi glo. ff. ad l. aquil. vt sit conuocatio hominum ad cuius auxilium debet vniuersisq; venire vt dicitur doct. in d. l. vt vim. Bar. in l. si quis in feruente in si. ff. de iur. iudeo dicit Bar. in ca. q. colla. de re iudi. q. si alium fit inuisa guerra amici etiam non vocari possunt cum iurare alleg. Bar. in l. probibitum de iure facti. libro. x. c. et possunt capere aduersarios / arg. l. ff. de malef. et d. ar. de. sed an extranei in auxilium alieius aggressi conuocari possunt occidere aggressos? dicit Bar. Bar. et communiter omnes. in d. l. vt vim q. ff. Ratio secundum Bar. in d. l. vt vim in. vj. q. est quia aggressor facit iniuriam reipublice. ideo quilibet de populo potest resistere in cap. dilecto. de sententia excommunicationis. lib. vj. et hoc maxime si sint conuocati in auxilium vt per eodem doctores. Sed an inuiti debeant iurare aggressum glo. in l. culpa. care. ff. de reg. iur. tenet q. non in contrarium videtur esse rez. in ca. quate in ti. iuncta glo. de sententia excommuni. et tex. in d. ca. dilecto. in cap. non inferenda / et ibi glo. nota. et ca. f. xxiij. q. iij. quod dicit vt possit alio. in d. l. vt vim. vj. q. colla. Sullermus le rouille alenconensis.

De assise. Chapitre. lv.

**A**ssise est vne court en la quelle ce q est fait doit auoir pardurable fermete / car se len nye ce qui a este fait es pleyz de viconte / on le peut ameder par vne simple desfrere / mais en ce qui est fait en assise / ne recoit len aucune desfrere. zc. Sur ce chapitre on doit noter par ces moys : [se] len nye ce qui a este fait es pleyz de la viconte / on le peut ameder par vne desfrere / cest a entendre que par voye de volence len est receu a dire contre ce qui a este fait en viconte / et annuler ce qui y a este fait par affermant facit contraire ou desfruct de ce qui a este fait en viconte et que len contend annuller. Item len doit noter par ce moys. Mais ce qui a este fait en assise zc. que quant on se met en record d'assise ou quel on ne se met point si no de l'accord des parties. Lombies quilz se doivent mettre de vnoit quant le cas le requiert. Le qui y est recorde et confirme a tousiours doit auoir fermete / car on ne sen peut doloir. Et la cause est pource quil est fait par record / qui est vne voye en quoy les parties se submettent de

Glose encadrante  
(gauche, droite)

Angel. in tract. malefici. in verbo / dicit? Citius se defendendo. 3af. in l. vt vim. col. ij. ff. de insti. ture. Alexan. consil. cix. incipi. vltis et vt poterit consideratis. in pumo volumi. Nec debet insultatus expectare p'nam per confusionem sed debet incipere vt debeat. Bar. et omnes. in l. c. vnde vi. l. q. etiam 3af. in d. l. vt vim. Et ibidem q. non est necesse quod imminet periculum mortis : sed factus est probare quod imminet periculum alicuius iniurie personalis / secundum Bart. in d. l. ij. col. c. vnde vi. Sed an insultatus pro defensione amici conlinguunt vel alterius occidere aggressorem? vt in glo. et per Bart. Bar. et 3af. ibidem in d. l. vt vim. Bar. in repetitio. d. l. ij. c. vnde vi. 3af. de anan. c. ij. de homicid. Et concludit q. pro defensione conlinguuntur sit licitum id quod pro suspensus : quia vt dicit 3af. in d. l. vt vim. quando fit mentio de se vel sua persona. talia verba non tantum comprehendunt propriam personam / sed etiam alios quos circa se necessarios habet per leg. ij. sed quod dignitas. ff. de pen. lega. et ibi textus notabilis. idem de contrite in via dicit Bart. et Bal. in leg. vt vim. ij. quarta questio. quia vt ibidem di. alias esset veredicta. ideo potest contumeliam suam populare. Item apud labonem q. tenetur. ff. de iniurijs. Conuenus in consil. c. iij. de defensione. col. in primo volumi. idem dicit Barto. in defensione extranei. in d. leg. vt vim. et Co. in dicto consil. Et ibidem q. licitum est alium offendere pro defensione rerum. iuno licitum est occidere : vt tenet glo. in d. l. ij. vt vim. Ratio secundum 3af. ibidem est quia res sunt secundum sanguis hominis. glo. iundo tex. in l. aduocat. c. de aduoca. Di. uer. iud. et ibi Bal. Item Bal. in l. ius nictus. c. de ope. libert. dicit vt per eundem 3af. in d. l. vt vim. equus / vt per s. frascum arretinum in consil. lxx. incipiens. sicut decet. penult. col. vbi dicit q. licet repercutere et defendere res manii non emle aut glorio in cap. dilecto. circa p. iij. q. de sententia excommuni. lib. vj. nec pro tuitione arma capere licet nec alium occidere / nisi in iudicio armis utatur / secundum Barto. et alios in d. l. ij. c. vnde vi. 3af. in 3af. andr. et 3af. in cap. olim caulam. de re iud. spo. l. i. c. iij. et 3af. in d. l. vt vim. col. post Cy. in d. l. ij. c. vnde vi. et 3af. in d. l. q. ius emem genium. in fine de iure natura. arg. et etiam an propter solas minas licitum sit alium occidere / et ibidem dicit z. concludit etiam post Bal. in l. iij. fin. c. quando lice. sine iud. se vindic. Quod si aliquis iminetur mortem et est in actu / sit licitum eum occidere alios non : quia tunc potesta est in actu que adem operatur sicut in habitu secundum s. p. politum de mart. in l. i. in p. iud. in q. ca. ff. de sicca. idem in l. penult. iij. col. ff. de quest. Bal. in l. multa. in fi. c. de lib. caus. ff. circa istam materiam defensionis. est etiam notandum q. licet debeat defendens vti moderamine vt in d. l. et d. l. vt vim. Et tamen excusatur de rāmen se defendendo non debet puniri pena mortis / sed alia minor pena arbitrio iudicis. sic dicitur eodem doct. Paul. de cast. consil. c. x. incipit in causa incertorum. col. ij. vbi di. q. licet aggressor potest fugere et insequutus ab agr. si fuit occisus / q. videtur potius ad vindictam se esse q. ad iuriam. s. amen no debet morte condemnari propter dolorem qui eum mouit ad insequendum et occidendum. l. strachus. c. de adul. idem dicit in quo libet alio actu defensionis secundum eodem doct. in d. l. vt vim. vbi inter alia dicit Bal. q. ictus non possunt dari cum mensura vt defensionis habeat iustitiam et ponderet ictus / referret se quitur c. eps. in d. consil. p. iij. incipiens. v. ff. de colla. penult. col. al. leg. ca. ff. de iur. iud. et ibi glo. et ibi glo. nota. Et quia homo intus voluit fugere / motus non sit in plenitudine intellectus / fm Bal. in cap. i. q. i. in fi. de pace iureu. fir. et quod di. idē Bal. in d. l. vt vim. iij. col. Quod insultatus non potest bene deliberare et q. sepe excedit de casto rationem. l. quicquid calore. ff. de res. iur. in cap. i. quis ius ius. ij. q. iij. et p. ius motus non sunt in potestate nostra. glo. in l. i. quis non dicam. c. de episcop. et cleric. p. ius ius ergo occidens de excessu / non de homicidio / et in iud. pena q. si non esset data occasio. ad hoc est tex. in l. si aduersum. ff. si imperatores. verii. et magis. et ibi Ang. ff. de adul. quem tex. ad hoc reputat fingulariter 3af. in d. l. vt vim. col. vbi inter alia dicit q. in duobus excessus preiudicium non dolo sed culpa processit. alleg. Sa lice. in d. l. strachus. An autem insultatus debeat fugere / dicit Bar. in d. l. vt vim. in fi. q. non nisi sit talis qui verecundiam non timet. 3af. in d. l. iud. iud. tenet q. non tenetur fugere propter periculum quod possit fugere / tamen propter casum vel altera tenent Bar. Angel. et Salice. in d. l. c. vnde vi. et Bal. in rep. per. d. l. vt vim. col. et in ca. ij. col. circa med. de consue. alleg. l. i. c. p. plagio. s. barnardus. ff. ad leg. aquil. et ibidem Cy. et no. per 3af. in cap. de ius q. magis dicit. Bal. in ca. ij. col. de iura et honeste. dicit per 3af. in d. l. c. q. qui aliter. ff. ad aquil. et quia vt dicit ibidem / fuga est refugium miserorum. Et

# Présentation générale :

## 2. Chaînes d'acquisition

<incident>

<desc>toy cat has bell in tail which continues to make a tinkling sound</desc>

</incident>

<vocal who="#mar">

<desc>meows</desc>

</vocal>

<u who="#ros">because it is so old</u>

<u who="#mar">how

<choice>

<orig>bout</orig>

<reg>about</reg>

</choice>

<emph>your</emph> cat

<pause/>yours is

<emph>new</emph>

<kinesic>

<desc>shows Father the cat</desc>

</kinesic>

</u>

*toy cat has bell in tail which continues  
to make a tinkling sound  
Mar meows*

**Ros** – because it is so old  
**Mar** – how 'bout *your* cat... yours  
is *new*

*shows Father the cat*

# Présentation générale :

## 2. Chaînes d'acquisition

### 6) Conception d'outils de consultation

Exemple : à prévoir pour un site internet.

- Serveur,
- Base de données (= fichiers XML-TEI)
- Système d'interrogation (pour naviguer dans le XML-TEI)
- Transformations (pour l'affichage du XML),
- Structure HTML/etc.

# Présentation générale :

## 2. Chaînes d'acquisition

/! TRÈS IMPORTANT /!

- Pérennité des outils et de la base de données.



# Présentation générale :

## 3. Présentation de Transkribus

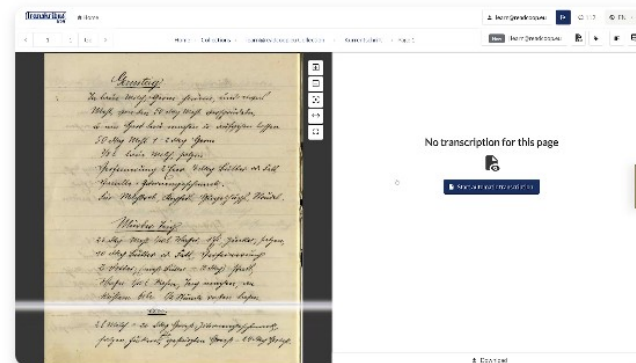


## Where AI meets historical documents

Transkribus is a comprehensive platform for the **digitisation, AI-powered text recognition, transcription and searching of historical documents** – from any place, any time, and in any language.

Try now

Sign up for free



Environnement de transcription automatique des documents, originellement développé à l'Université d'Innsbruck, sur un financement européen.

Suite à l'arrêt du financement, une coopérative a été créée pour maintenir l'outil et continuer son amélioration.

Site : <https://readcoop.eu/transkribus>

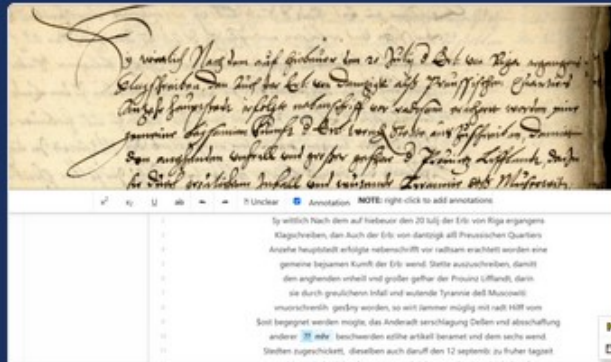
# Présentation générale :

## 3. Présentation de Transkribus

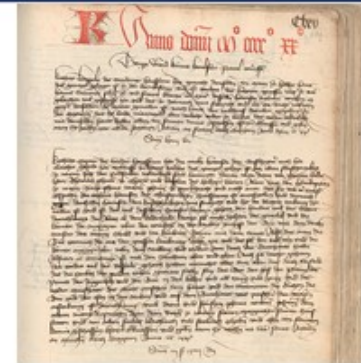
### Exceptional projects realised with Transkribus



The Pandetta Project –  
Fondazione Banco di Napoli



Recesses of Low German  
Hanseatic Days: A Hanseatic  
journey with Transkribus



Mapping Medieval Vienna: The  
digital edition of historical land  
registers supported by  
Transkribus

[See all Success Stories >>](#)

# Présentation générale :

## 3. Présentation de Transkribus

### Pricing

Training, layout analysis, using Transkribus for manually transcribing, etc. – i.e., **everything but text recognition** proper is **free of charge**. Yet, you can also use the Transkribus text recognition feature for free, as you get **500 free credits** (i.e. about 500 pages) upon signup.

Your purchases will enable us to keep Transkribus alive into the future and keep this wonderful community going. Everything you need to know about buying credits you can find [here](#).

#### FREE

##### Signing up is completely free

✔ 500 Free credits on signup

⚙ Model Training

📄 Layout Analysis

✍ Manual Transcription

Sign up

Create a free READ-COOP Account

#### PAID SERVICES

##### Handwritten Text Recognition

⚙ Process larger amounts

📄 On-demand or subscription

👥 Credits can be shared

🙌 You support the platform

See pricing

Buy credits on-demand or as subscriptions

# Présentation générale :

## 3. Présentation de Transkribus

Transkribus Expert Client v1.19.0 (01\_05\_2022\_19:11), Loaded doc: 1771 - Pesnelle, ID: 985912, Page 1, file: 1.tif [Image Meta Info: (Resolution:300.0, w\*h: 4266 \* 5791) ]

Search current document... 1 / 10 In Progress

Server Overview Layout 2

Type Text S

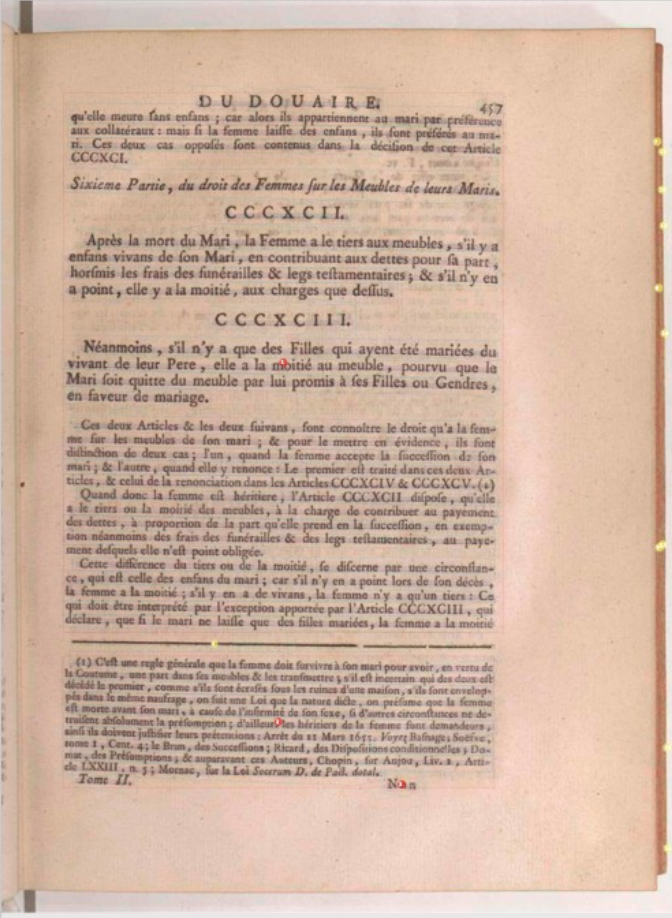
Page

TextRegion

- Line DU DOUAIRE.
- Line 457
- Line qu'elle meure sa
- Line aux collatéraux:
- Line ri. Ces deux cas
- Line CCCXCI.
- Line Sixieme Partie, c
- Line CCCXCII.
- Line Après la mort du
- Line enfans vivans de
- Line horsmis les frais
- Line a point, elle y a l
- Line CCCXCIII
- Line Néanmoins, s'il i
- Line vivant de leur Pe
- Line Mari soit quitte c
- Line en faveur de ma
- Line
- Line Ces deux Article
- Line me sur les meub
- Line distinction de de
- Line mari; & l'autre, c
- Line ticles, & celui de
- Line Quand donc la f
- Line a le tiers ou la m
- Line des dettes, à pr
- Line tion néanmoins
- Line ment desquels e
- Line Cette différence
- Line ce, qui est celle
- Line la semme a la m
- Line qui doit être inte
- Line déclare, que si l

TextRegion

- Line (1) C'est une reg
- Line la Coutume, une
- Line décédé le premi
- Line nés dans le mêm



DU DOUAIRE. 457

qu'elle meure sans enfans ; car alors ils appartiennent au mari par préférence aux collatéraux ; mais si la femme laisse des enfans, ils sont préférés au mari. Ces deux cas opposés sont contenus dans la décision de cet Article CCCXCI.

Sixieme Partie, du droit des Femmes sur les Meubles de leurs Maris.

CCCXCII.

Après la mort du Mari, la Femme a le tiers aux meubles, s'il y a enfans vivans de son Mari, en contribuant aux dettes pour sa part, horsmis les frais des funérailles & legs testamentaires; & s'il n'y en a point, elle y a la moitié, aux charges que dessus.

CCCXCIII.

Néanmoins, s'il n'y a que des Filles qui ayent été mariées du vivant de leur Pere, elle a la moitié au meuble, pourvu que le Mari soit quitte du meuble par lui promis à ses Filles ou Gendres, en faveur de mariage.

Ces deux Articles & les deux suivans, font connoître le droit qu'a la femme sur les meubles de son mari; & pour le mettre en évidence, ils font distinction de deux cas; l'un, quand la femme accepte la succession de son mari; & l'autre, quand elle y renonce: Le premier est traité dans ces deux Articles, & celui de la renonciation dans les Articles CCCXCIV & CCCXCV. (1)

Quand donc la femme est héritière, l'Article CCCXCII dispose, qu'elle a le tiers ou la moitié des meubles, à la charge de contribuer au payement des dettes, à proportion de la part qu'elle prend en la succession, en exemption néanmoins des frais des funérailles & des legs testamentaires, au payement desquels elle n'est point obligée.

Cette différence du tiers ou de la moitié, se différencie par une circonstance, qui est celle des enfans du mari; car s'il n'y en a point lors de son décès, la femme a la moitié; s'il y en a de vivans, la femme n'y a qu'un tiers: Ce qui doit être interprété par l'exception apporée par l'Article CCCXCIII, qui déclare, que si le mari ne laisse que des filles mariées, la femme a la moitié

(1) C'est une règle générale que la femme doit survivre à son mari pour avoir, en vertu de la Coutume, une part dans les meubles & les transferts; s'il est incertain qui des deux est décédé le premier, comme s'ils sont héritiers tous les deux d'une maison, s'ils sont enveloppés dans le même mariage, on fait une loi que le mortier d'ice, on présume que la femme est morte avant son mari, à cause de l'infirmité de son sexe, si d'autres circonstances ne démontrent autrement la présomption; d'ailleurs les héritiers de la femme font demander, ainsi ils doivent publier leurs prétentions: Arrêt du 21 Mars 1671. Foyez Balthaz. Socive, tome 1, Cont. 4; le Brun, des Successions; Ricard, des Dispositions conditionnelles y Doman. des Préfomptions; & auparavant ces Auteurs, Chopin, sur Anjou, Liv. 5, Article LXXII, n. 3; Morneau, sur la Loi Sacram. D. de Pail. distal. Tome II.

1-1 DU DOUAIRE.

1-2 457

1-3 qu'elle meure sans enfans ; car alors ils appartiennent au mari par préférence

1-4 aux collatéraux: mais si la femme laisse des enfans, ils sont préférés au ma-

1-5 ri. Ces deux cas opposés sont contenus dans la décision de cet Article

1-6 CCCXCI.

1-7 Sixieme Partie, du droit des Femmes sur les Meubles de leurs Maris.

1-8 CCCXCII.

1-9 Après la mort du Mari, la Femme a le tiers aux meubles, s'il y a

1-10 enfans vivans de son Mari, en contribuant aux dettes pour sa part,

1-11 horsmis les frais des funérailles & legs testamentaires; & s'il n'y en

1-12 a point, elle y a la moitié, aux charges que dessus.

1-13 CCCXCIII

1-14 Néanmoins, s'il n'y a que des Filles qui ayent été mariées du

1-15 vivant de leur Pere, elle a la moitié au meuble, pourvu que le

1-16 Mari soit quitte du meuble par lui promis à ses Filles ou Gendres,

1-17 en faveur de mariage.

1-18

1-19 Ces deux Articles & les deux suivans, font connoître le droit qu'a la fem-

1-20 me sur les meubles de son mari; & pour le mettre en évidence, ils font

1-21 distinction de deux cas; l'un, quand la femme accepte la succession de son

1-22 mari; & l'autre, quand elle y renonce: Le premier est traité dans ces deux

1-23 ticles, & celui de la renonciation dans les Articles CCCXCIV & CCCXCV. (\*)

# Présentation générale :

## 4. Présentation des outils du cours



### *Très Ancien Coutumier*

Paris, Bibliothèque Sainte-Genève  
fragment du ms. fr. 1743

Points particuliers à garder en tête :

- Manuscrit
- Graphies anciennes
- Colonnes
- Rubriques
- Initiales ornées
- Pieds-de-mouche
- Caractères anciens :
  - s long,
  - d rond,
  - abréviations,
  - ponctuation médiévale...

# Présentation générale :

## 4. Présentation des outils du cours

### Du duc, Fo. cxvij.

Qui devise a qui chascune chose etc.

**Additio.** Sicut enim precepta iuris alteri non ledere/honorare, sic visere/interfieri sui vicium tribuere in quibus, mit de iust. c. iure. i. iustitia s. i. hoc est iura. de quibus differunt in tracta. de benef. iust. c. i. multi. lib. i. s. uillic. de rouille alieno.

In textu ibi.

Les viages fa corderent etc.

**Additio.** Sicut enim precepta iuris alteri non ledere/honorare, sic visere/interfieri sui vicium tribuere in quibus, mit de iust. c. iure. i. iustitia s. i. hoc est iura. de quibus differunt in tracta. de benef. iust. c. i. multi. lib. i. s. uillic. de rouille alieno. Faciunt no. per boici in l. de quibus. s. de legib. Sicut enim precepta iuris alteri non ledere/honorare, sic visere/interfieri sui vicium tribuere in quibus, mit de iust. c. iure. i. iustitia s. i. hoc est iura. de quibus differunt in tracta. de benef. iust. c. i. multi. lib. i. s. uillic. de rouille alieno.

qui seu. ba. poss. et fin eo talis consuetudo scripta appellat ius consuetudinarium. Zilegat. i. i. de precib. lib. vi. et no supradicta facit et appellacione iuris venit etiam consuetudo. glo. in cuncto. tex. et ibi Bal. plau. de cast. c. i. s. a. in l. prescriptio in verbo iuris. c. si contra ius vel pub. vel. et ibidem idem Bal. in quarta col. s. si videtur et si factum est compositionum in in obsequio ut pronunciant debent de iure tantum tamen in patria consuetudo iuri contraria quod debent pronunciare fin consuetudinum que preuget in patria iuri communi. alle. no. per Amos. in c. certifi. can. de sepul. ubi dicit qd petens quartam que sibi debet de iure videtur intelligere de illa que sibi debet de consuetudine. de qua queritur Bal. in l. prescriptio. s. de de. no. per Panes. in pro. clem. 2 in c. i. v. l. i. pend. ubi. in titu. de pace. c. stan. in col. in verbo nos Romanorum. Item Bal. in l. i. prescriptio. vi. col. s. i. restat videre notabiliter per Zale. no. consuetudinum. s. i. v. l. i. pend. ubi dicit qd iuris appellacione simpliciter potius inteligitur ius municipale aliqd vltra ius cōe disponere alleg. Bal. in l. lulla. c. iusticia. s. i. de hered. s. i. i. i. c. iusticia. nos. c. de doq. libe. et Zale. qd de ius. c. i. quinquagesimo. de iure iur. s. i. de rouille.

C. 39. sine la premiere distinction.

## La seconde distinction,

Du duc.

**D**uc de Normandie ou le prince est cil qui tient la seigneurie de quoy le roy etc.

Et par ce peunt on supposer qd le roy de France copie la ce texte : s. par especial ce chapitre. Et se len faitoit qstion pour quoy ce chapitre s. plusieurs autres parlent du duc. En pourroit respondre qd cest pour denoter qd laioir ce que le roy en fait seigneur, si ne soit ce pas come roy mais come duc.

C. 39. ou le texte met en la fin de ce chapit. **D**uc le plus doit faire enquerir par les baillifs et mettre en prison les larrons/les homicides/et tous les malfaicteurs etc. Il entend pas que les vicieux nen puissent enquerir; mais il appartient aux baillifs principalement et pour ce met il. par les baillifs. C. En peut sur ce faire vne telle question. Scauoir le vng homme bat vng autr. et le batu ne se plaint: si appartient au baillif enquerir et punir le malfaicteur / attendu que en ladite malefacion na mot/mebainq/ ne baro. Et aussi quil semble que le texte parle principalement des malfacions ou il ya mot ou mebainq ou autre grand cas. C. En peut respondre que ouy pour deux causes. La premiere pour ce que le texte parle generalement de tous malfaicteurs. La seconde est, que ledict texte met que le baillif doit enquerir des malfaicteurs/afin que le peuple puisse estre en pais. Laquelle cause n'auroit pas lieu se on souffroit tels delictz : mais seroit chose de scandalle cōtenner et desputer iustice et fait de mal exemple/contre le ben. de pais : et dont il pourroit enuyneur moult d'inconueniens si sur ce n'estoit pourueu par iustice.

Chapitre douzieme.

In textu ibi

Le duc de Normandie etc.

**Additio.** Ducus Romanie in iudiciis habuerit a Rellone dicitur noy s. supra in initio premissis in successib. dicitur ius fuit continuata ducatus possessio vsq ad deposita regis francie Pipilli Augusti qui conquisit et cosone francie reunitur dicitur in iudiciis et factus linea Romanie duci per. ex. an. nos vsq ex cronis facile est supputari. Sunt qui dicunt quod Ricardus fuisse Romanie ducem tempore Caroli magni et quod omnia absurdissima est cetero sic Romanie nomen erat penitus inauditum nec illius meminit Turpinus Rhenensis archiepiscopus in historia vocat ducem Romanie tunc que Romanie est dicitur et quod quia per grabellum fuit sic nominata quod ponitur Caroli simplicis RAPODIA l. i. b. v. l. 2. de viij. fo. cxxij. post Caroli magni obitum inter quos fluxerunt centum anni. Aduerte tamen quod bene reperiri in cronica Bitalanie quod fuit alias in Scythia quidam nomine Zuerberus nomine Carriagy apud Kothomagus cui ob benemeritis Pipinus concessit medietatem suam Scythie et habuit duos filios ex duabus coniugibus Robertus videlicet cognominatus Diabolus propter illius crudelitatem factus fuit vir perditissimus de quo plura fabulantur et obiit sine liberis. Alterum filium habuit Zuerberus nominatus Ricardus strenuus milite cum Carolo magno ob peccata sua concessit totam Scythiam cum titulo ducis et francie Partis sine liberis concedo nepotem suum ex sorore Zuerberum ducem nominatus Zerno beredem in Scythia habuit et regi Ludouico pio rebellio factus tandem in pretio est occisus sic Scythia redijt regi qui postea etiam cepit inuocari filio suo Carolo caluo cognominato vsq restat. Ro. gras qui. in cronica in vita Ludouici pij vocatur inquit ad se auilicos rum pimus Caroli tunc in filio ponton imperij bar. in conilio apud castum agens presente filio Ludouico Carolum militari cingulo et regis vrbibus onas data illi Scythia etc. De ducibus vide in titu. quibus dicitur dux mar. vel co. et ibi per boctus in titu. de precib. f. a. alle. per s. de s. p. terra bucatata. et quod dicitur dux eo quod ducit populum sine exercitum secundum papiam. Refert Inco. de sancto Zegio in tracta. in pun. Suis lernas le rouille Zlenconten.

In textu ibi.

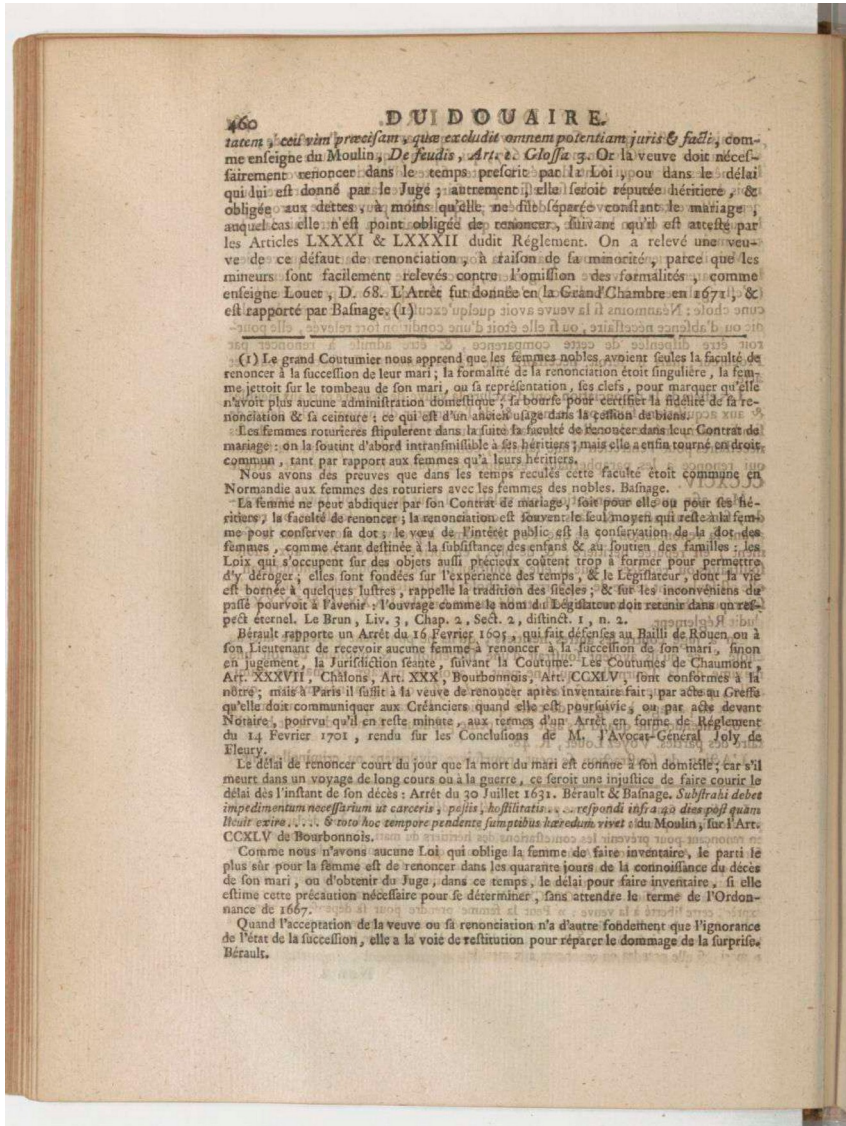
## Le grand coutumier du pays et duché de Normandie... 1539, Guillaume Le Rouillé, Rouen.

Points particuliers à garder en tête :

- Caractères « gothiques »
- Graphies anciennes
- Abréviations
- Glose encadrante
- Initiales ornées
- Pieds-de-mouche
- Appels de note en marge
- Notes additionnelles en latin
- Apparition du paratexte
- Caractères anciens :
  - s long,
  - d rond,
  - abréviations,
  - ponctuation médiévale...

# Présentation générale :

## 4. Présentation des outils du cours



### *Coutume de Normandie*

1771, Pesnelle (avocat au Parlement), en deux volumes, avec les observations de M. Roupnel de Chenilly, Rouen.

Points particuliers à garder en tête :

- Caractères « latins »
- Initiales ornées occasionnelles
- Appels de note dans le corps du texte
- Paratexte
- Caractères anciens :
  - s long.



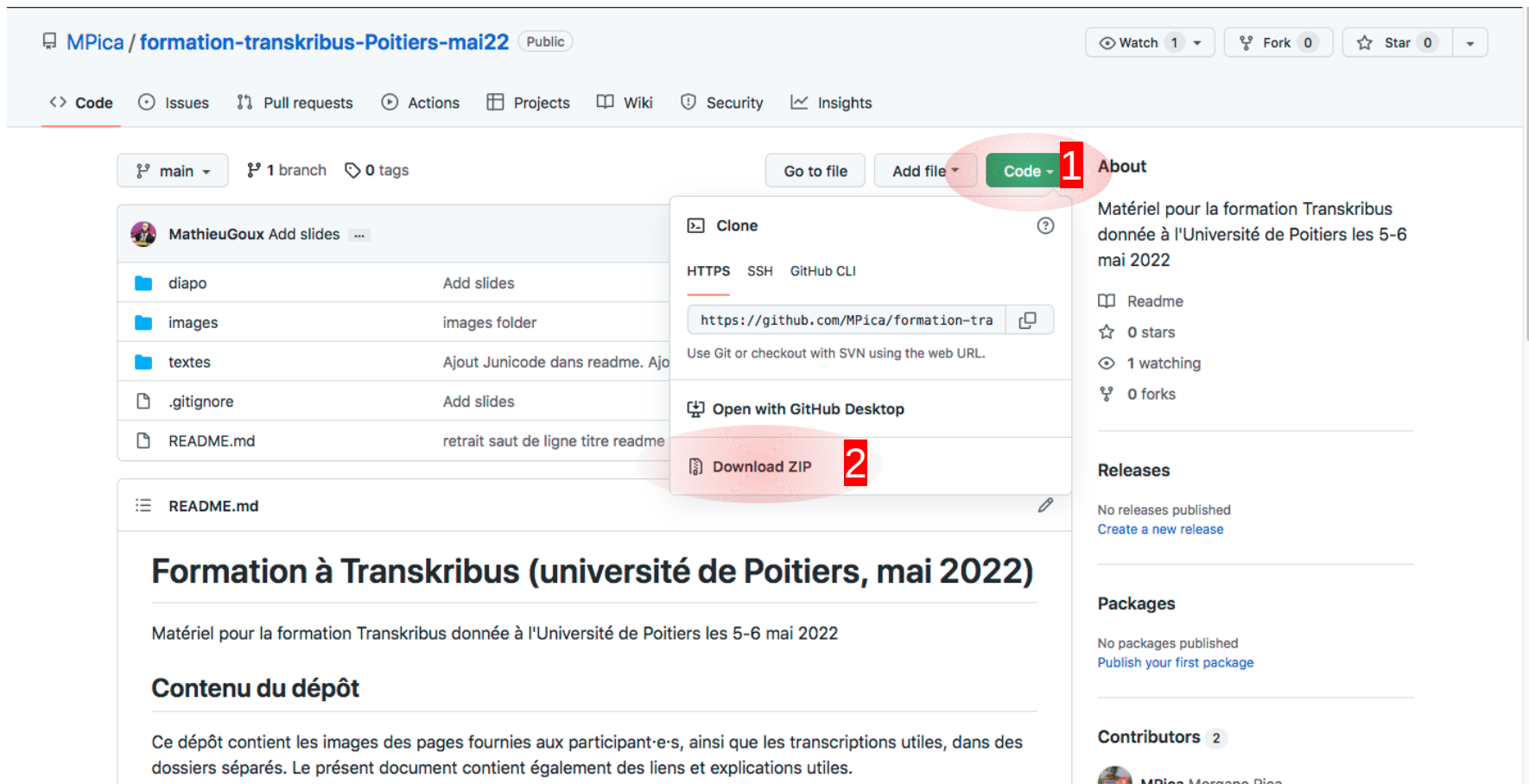


# Présentation générale :

## 4. Présentation des outils du cours

### *Dépôt GitHub*

<https://github.com/MPica/formation-transkribus-Poitiers-mai22>



The screenshot displays the GitHub interface for the repository `MPica / formation-transkribus-Poitiers-mai22`. The repository is public and has 1 watch, 0 forks, and 0 stars. The main navigation bar includes links for Code, Issues, Pull requests, Actions, Projects, Wiki, Security, and Insights. The repository structure shows a `main` branch with 1 branch and 0 tags. The file list includes `diapo`, `images`, `textes`, `.gitignore`, and `README.md`. The `Code` dropdown menu is open, showing options for cloning (HTTPS, SSH, GitHub CLI) and downloading the ZIP file. The `Download ZIP` option is highlighted with a red circle and the number 2. The `Code` button is highlighted with a red circle and the number 1. The repository description reads: "Matériel pour la formation Transkribus donnée à l'Université de Poitiers les 5-6 mai 2022". The `README.md` content includes the title "Formation à Transkribus (université de Poitiers, mai 2022)" and the text: "Ce dépôt contient les images des pages fournies aux participant-e-s, ainsi que les transcriptions utiles, dans des dossiers séparés. Le présent document contient également des liens et explications utiles."