

# Matthias GILLE LEVENSON

*ATER, Aix-Marseille Université*

Né le 22/05/1993

✉ [matthias.gille.levenson@ens-lyon.fr](mailto:matthias.gille.levenson@ens-lyon.fr)

🌐 [matgille](#)

🆔 0000-0001-9488-5986

👤 [mgillelevenson](#)

Docteur ès lettres qualifié par les sections 7 et 14 du CNU.  
Spécialités : études hispaniques, bas Moyen Âge castillan, philologie computationnelle.  
Ancien membre scientifique de la Casa de Velázquez – EHEHI.  
Agrégé d'espagnol (2017).  
Ancien élève de l'École Normale Supérieure de Lyon.

Mise à jour : octobre 2024

## Table des matières

Parcours professionnel	1
Diplômes obtenus	2
Formation, concours et qualifications	2
Séjours universitaires et de recherche à l'étranger	3
Production scientifique	3
Expérience d'enseignement	8
Projets financés en cours	12
Intégration dans des comités éditoriaux et de lecture	12
Activités d'encadrement et participation à des jurys	12
Compétences	13
Responsabilités organisationnelles	13

## Parcours professionnel

- 2024–2025 **Attaché Temporaire d'Enseignement et de Recherche**, Aix-Marseille Université, Centre Aixois d'Études Romanes, UR 854
- 2023–2024 **Post Doctorant**, *École nationale des chartes*, Centre Jean Mabillon, École Normale Supérieure de Lyon, associé au CIHAM, UMR 5648.
- 2022–2023 **Attaché Temporaire d'Enseignement et de Recherche**, *ÉNS de Lyon*
- 2020–2022 **Membre scientifique**, *Casa de Velázquez – EHEHI*
- 2018–2020 **Doctorant contractuel**, *ÉNS de Lyon*
- 2013–2018 **Élève fonctionnaire**, *ÉNS de Lyon*

## Diplômes obtenus

- 2023 **Doctorat ès Lettres, spécialité Études Hispaniques, École Normale Supérieure de Lyon**  
Titre de la thèse de doctorat : *Le Regimiento de los príncipes et sa glose : étude et édition numérique de la partie sur le gouvernement de la cité en temps de guerre (III, 3)*. Thèse réalisée sous la direction de Carlos Heusch (ENS de Lyon) et de Jesús R. Velasco (Yale University).
- 2016 **Master 2 en langues hispanophones et lusophones, École Normale Supérieure de Lyon**  
Titre du mémoire de Master 2 : *La version B du “Regimiento de los príncipes glosado”*. Note obtenue : 19.
- 2014 **Licence LLCER espagnol, Université Lumière-Lyon 2**

## Formation, concours et qualifications

- 2024 **Qualifications aux fonctions de Maître de conférences par le CNU**  
— Section 7 : Sciences du langage ; no de qualification : 24207362400.  
— Section 14 : Langues et littératures romanes ; no de qualification : 24214362400.
- 2018–2023 **Préparation du diplôme de Doctorat, soutenu le 8 septembre 2023**  
— *Visiting Assistant in Research* à l'Université Yale (septembre – décembre 2019).  
— Suivi (sur invitation) de l'atelier DTS (*Distributed Text Services*) « Texts and API » à Hambourg, du 15 au 16 juillet 2019.  
— Suivi de l'école d'été de l'IRHT « Le livre médiéval au regard des méthodes quantitatives » du 25 au 28 juin 2019.
- 2017–2018 **4<sup>e</sup> année d'ÉNS ; préparation de la thèse et mise à niveau technique et méthodologique.**  
— Participation à l'atelier 2018 du consortium CAHIER (Montpellier, 26–29 juin 2018).  
— Participation à la session 2018 de la formation EDEEN (École d'été en humanités numériques) à Grenoble, organisée par l'Université Grenoble-Alpes et la MSH Alpes (Grenoble, 28 mai–2 juin).  
— Suivi d'un atelier en édition numérique à Cambridge (Royaume-Uni) en avril 2018, dans la continuité du DEMM (*Digital Encoding of Medieval Manuscripts*).
- 2016–2017 **Agrégation externe d'Espagnol.**
- 2015–2016 **Master II d'études hispanophones et lusophones, ÉNS de Lyon**  
— Suivi des cours du Master *Patrimonio Textual y humanidades digitales* de l'Université de Salamanque dans le cadre du programme Erasmus.  
— Mémoire préparé sous la direction de Carlos HEUSCH<sup>ID</sup> et de FRANCISCO BAUTISTA<sup>ID</sup> : *La version B du Regimiento de los príncipes glosado*. Note obtenue : 19.  
— Suivi de la formation *Digital Encoding of Medieval Manuscripts* (DEMM) financée par l'Union Européenne. Trois semaines de stage à Vienne (Abbaye de Klosterneuburg), Lyon (ÉNS de Lyon), Londres (Queen Mary University), entre janvier 2015 et mai 2016.
- 2014–2015 **Master I d'études hispanophones et lusophones et hispanophones, ÉNS de Lyon**  
— Spécialisation en philologie castillane et en étude de textes du bas-Moyen Âge. Mémoire préparé sous la direction de Carlos Heusch : *Edición crítica de dos textos de Pedro de Chinchilla*. Note obtenue : 18.  
— Stage d'ecdotique des Sources chrétiennes (Lyon, février 2015).

2013-2014 **Licence LLCER espagnol**, Université Lumière-Lyon 2  
— Séminaire d'ecdotique de Gisèle BESSON, ÉNS Lyon.

2011-2013 **Préparation et obtention du concours de l'ENS de Lyon** (CPGE, Lycée Joffre, Montpellier).

2011 **Baccalauréat général scientifique, option mathématiques**, Mention TB, Lycée Joffre, Montpellier

## Séjours universitaires et de recherche à l'étranger

2020-2022 Membre scientifique de la Casa de Velázquez, à Madrid.

2019 *Visiting Assistant in Research*, Yale University, auprès de Jesús R. Velasco. Septembre-décembre 2019.

2015-2016 Échange universitaire Erasmus (M2) à l'Université de Salamanque.

## Production scientifique

Mon parcours scientifique est intimement lié au monde ecdotique et à l'histoire du texte et du livre médiéval et de la première modernité : histoire du livre, histoire du texte et des idées politiques (3, 4, 17, 18, 19, 23, 24, 25, 26), acquisition et normalisation automatisées du texte manuscrit (2, 6, 10, 33, 37, 32, 38, 40), annotation lexico-grammaticale du castillan médiéval et pré-moderne (7, 14, 34, 39), alignement et collation unilingue et multilingue de documents patrimoniaux et contemporains (9, 11, 12, 17, 36, 5), édition de textes numérique (20, 21, 26, 28, 29).

Je continue actuellement à travailler sur trois chantiers : un premier projet lancé cette année avec Lucence Ing et Jean Baptiste Camps sur l'alignement et la collation automatisée multilingue de textes médiévaux, travail essentiellement méthodologique mais qui s'appuie sur la grande tradition du *Lancelot en prose* (représentée par plusieurs textes français, une version italienne et une version castillane) ; un second chantier en cours co-porté avec Olivier Brisville-Fertin et qui inclut sur la création et la curation d'un corpus intégralement corrigé à la main d'annotations lexico-grammaticales du castillan médiéval, à des fins de production de modèles d'annotation efficaces ; un troisième chantier d'édition numérique du premier *Regimiento*, représenté par deux manuscrits conservés en Castille.

## Thèse de doctorat

**Titre** : *Le Regimiento de los príncipes et sa glose : étude et édition numérique de la partie sur le gouvernement de la cité en temps de guerre (III, 3).*

**Direction** : Carlos HEUSCH (École Normale Supérieure de Lyon) et Jesús R. VELASCO (Yale University).

**Date de soutenance** : 8 septembre 2023.

**Rapportrices** : Elena PIERAZZO (Université de Tours) et Susana ALLÉS TORRENT (University of Miami).

**Membres du jury** : Jean-Baptiste CAMPS (École nationale des chartes), Elena PIERAZZO (Université de Tours), Hélène THIEULIN-PARDO (Sorbonne Université), Marta LÓPEZ IZQUIERDO (Université Paris VIII).

**Lien vers la thèse** : tel-04337406.

**Résumé** : Ma thèse de doctorat a consisté en l'étude et en l'édition numérique de la dernière partie du *Regimiento de los príncipes* (III-3), une traduction castillane glosée du *De Regimine Principum* de Gilles de Rome. Y sont présentés tous les témoins du corpus, composé de neuf manuscrits et d'un exemplaire incunable. Elle développe une chaîne de traitement semi-automatisée qui permet l'édition du texte structuré en XML-TEI, depuis l'acquisition individuelle de chacun des témoins du corpus, par HTR, OCR ou transcription manuelle, jusqu'à la production de l'édition au format pdf. La thèse accorde une place

importante à la présentation détaillée des données produites (transcription, segmentation, lemmatisation), et s'intéresse à la question de la pérennité de la production ecdotique. La question de la collation semi-automatisée de sources livresques préalablement décrites et encodées au format XML-TEI y est essentielle, ouvrant la voie à la collation d'éléments non textuels (voir la production 36). Le corpus est étudié sous l'angle de la variation, en prenant comme axe le discours politique, afin de déterminer son homogénéité conceptuelle, pour ensuite étudier les relations entre témoins. Elle permet de mettre à jour les connaissances sur les différentes versions du texte. Enfin, le texte est étudié au prisme des idées sur la chevalerie et de la noblesse, centrale sur la partie éditée. L'édition est produite sur les 23 derniers chapitres du texte et aligne, paragraphe par paragraphe, la version supposément originale de la traduction, dite  $\omega$ , et la  $\beta$  (soit le texte édité à proprement parler), ainsi que le texte latin (en annexe), récupéré à partir du projet PALM, puis structuré en XML-TEI par mes soins.

## Articles

- accepté [1] Ariane Pinche et Matthias Gille Levenson, « Méthodes numériques et nouvelles perspectives ecdotiques : Les éditions nativement numériques des *Seint Confessor* et du *Regimiento de los Príncipes* », *Revue Mabillon*.
- [2] « Towards a general open dataset and models for late medieval Castilian text recognition (HTR/OCR). », *Journal of Data Mining and Digital Humanities (JDMDH)*, *Special Issue : Historical documents and automatic text recognition*, 2023, éd. Ariane Pinche et Peter Stokes, doi :10.5281/zenodo.7387376.
- [3] « Educación del príncipe, nobleza y caballería en el *De regimine principum* castellano (segundo y tercer libro) », *Librosdelacorte* 22, 2021, pp. 285-308. doi :10.15366/ldc2021.13.22
- [4] « L'évolution du Regimiento de los príncipes (1345-1494), conditionnée par le pouvoir politique? » in *Écritures du Pouvoir*, 2, éd. Véronique Lamazou-Duplan, Ausonius éditions, 2019, Scripta Medievalia, pp. 137-148, hal-02369116.

## Articles de conférence

- accepté [5] Matthias Gille Levenson, Lucence Ing, Jean-Baptiste Camps, « *Textual Transmission without Borders : Multiple Multilingual Alignment and Stemmataology of the "Lancelot en prose" (Medieval French, Castilian, Italian)* », Computational Humanities Research (CHR) 2024, Aarhus, Denmark.
- [6] Clérice, Thibault, Ariane Pinche, Malamatenia Vlachou-Efstathiou, Alix Chagué, Jean-Baptiste Camps, Matthias Gille Levenson, Olivier Brisville-Fertin, et al., « *CATMuS Medieval : A Multilingual Large-Scale Cross-Century Dataset in Latin Script for Handwritten Text Recognition and Beyond.* », In : Barney Smith, E.H., Liwicki, M., Peng, L. (eds) Document Analysis and Recognition - ICDAR 2024. *Lecture Notes in Computer Science*, vol 14806. Springer, Cham, 2024, doi :10.1007/978-3-031-70543-4\_11.
- [7] Simon Gabay, Thibault Clérice, Jean-Baptiste Camps, Jean-Baptiste Tanguy, Matthias Gille-Levenson. *Standardizing linguistic data : method and tools for annotating (pre-orthographic) French.*, Proceedings of the 2nd International Digital Tools and Uses Congress (DTUC '20), Oct 2020, Hammamet, Tunisia, hal-03018381. doi :10.1145/3423603.3423996

## Communications dans des congrès internationaux avec comité de sélection

- soumis [8] Lucence Ing, Matthias Gille Levenson, Jean-Baptiste Camps, « Premiers jalons de la collation multilingue du *De Regimine Principum* latin et vernaculaire », CILPR, 30 juin – 5 juillet 2024.
- [9] Matthias Gille Levenson, Célian Ringwald, Marie Flesch, Jennifer Isasi, Sofia Papastamkou, Riva Quiroga et David Valentine, « Connecter les chapitres linguistiques de Programming Historian? Vers une table conceptuelle multilingue constituée semi-automatiquement », Congrès Humanistica2024, 7-9 mai 2024.
- [10] Ariane Pinche, Thibault Clérice, Alix Chagué, Jean-Baptiste Camps, Malamatenia Vlachou-Efstathiou, Matthias Gille Levenson, Olivier Brisville-Fertin, Federico Boschetti, Franz Fischer, Michael Gervers, et al., « *CATMuS-Medieval : Consistent Approaches to Transcribing Manuscripts : A generalized set of guidelines and models for Latin scripts from Middle Ages (8th–16th century)* ». DH2024. hal-04346939
- [11] « *TEICollator : a TEI to TEI workflow* », TEI2022 : Text as data, Newcastle (Angleterre) 12-16 septembre 2022.
- [12] « *TEICollator : une chaîne de traitement ecdotique semi-automatisée* », 30e Congrès International de Linguistique et de Philologie Romane, Université de La Lagune, Canaries (Espagne), 4-9 juillet 2022. Dépôt HAL.
- [13] Matthias Gille Levenson, Gwenaëlle Patat (ordre alph.). « Programming Historian en français : Faire communauté pour le partage de ressources éducatives libres sur les méthodes numériques en sciences humaines et sociales francophones », Colloque Humanistica 2022, Mai 2022, Montréal, Canada. hal-03672420
- [14] Matthias Gille Levenson, Olivier Brisville-Fertin, María Díez Yáñez, Simon Gabay, « Construcción de un corpus de evaluación de la anotación léxico-gramatical del castellano medieval (siglos XIII-XV) », HDH2021 : scire vias, Saint Jacques de Compostelle, Espagne, oct. 2021. Lien vers le dépôt git.
- [15] Simon Gabay, Lucie Rondeau Du Noyer, Matthias Gille Levenson, Ljudmila Petkovic et Alexandre Bartz, « *Quantifying the Unknown : How many manuscripts of the marquise de Sévigné still exist ?* », DH2020, Ottawa, Canada. Abstract et diapositives.
- [16] Paul Bertrand, Ariane Pinche, Matthias Gille Levenson et Margot Ferrand, « *COSME<sup>2</sup> - Complexities : 30 Years Of Research Of Medievalists DH Concerning A Thousand Years Of Medieval Sources* » (poster). Congrès « *Digital Humanities 2019* » (DH2019), du 9 au 12 juillet 2019, Utrecht, Pays-Bas. DOI [10.5281/zenodo.3275597](https://doi.org/10.5281/zenodo.3275597)

## Communications dans des colloques ou des journées d'étude

- [17] « Vers la collation multilingue de textes romans médiévaux. Le *Regimiento de los príncipes* comme cas d'étude », Journées du Séminaire d'Études Hispaniques Médiévales : Humanités numériques et philologie romane, 5 avril 2024.
- sur invitation [18] « Primeras pistas de estudio computacional de la recepción del *Regimiento de los Príncipes* : las marcas de lectura del manuscrito K.I.5 (Biblioteca del Escorial). », Journées d'étude Cartographies du livre ibérique, CEIBA, Toulouse, 15 mars 2024.
- [19] « Chevaliers et cavaliers dans la dernière partie du *Regimiento de los príncipes* », Séminaire d'Études hispaniques médiévales 2022, co-organisation Julie Marquer et Matthias Gille Levenson, ENS de Lyon, 13 décembre 2022.

- sur invitation [20] *Impressa manent* : en defensa de la producción de ediciones digitales en formato fijo, Congrès Ecdótica y humanidades digitales : historiografía, metodología y praxis, Salamanca (Espagne) 23-24 novembre 2023.
- sur invitation [21] « La transcription automatisée, un changement de paradigme ecdotique ? », colloque Trilobe II, 29-30 septembre 2022, ENS Lyon.
- [22] Matthias Gille Levenson, Sofia Papastamkou et Célian Ringwald (ordre alph.). « Programming Historian : un lieu de collaborations et d'interactions multiples », Colloque DHNord 2022, juin 2022, en ligne.
- [23] « Réception du roman, fiction et valeur du vrai dans trois traités de littérature politique des xiv<sup>e</sup> et xv<sup>e</sup> siècles en castillan », Journée d'études « Vérité et mensonge », ENS de Lyon, 6 mai 2021.
- [24] « Describir y prescribir el liderazgo en la Castilla bajomedieval », Séminaire du Madrid Institute of Advanced Studies (MIAS), co-animé avec Edward Holt, 22 mars 2021.
- sur invitation [25] « “E mayormente los reyes”. Educación del hijo y educación del príncipe en el segundo libro del *De regimine principum* castellano », I SEMINARIO DE CORTE Y LITERATURA, Instituto Universitario La Corte en España, UAM, 26-27 novembre 2021.
- [26] « *Editing a 15th century political treatise using the computer : a back and forth between meaning and information* », Iberian Connections seminar, Yale University, 12 novembre 2019. Lien.
- [27] « Du concept au récit : théorie politique et enxiemplo dans le *Regimiento de los príncipes* (III, 3) ». Colloque « Theorica 7 : Politia », ÉNS de Lyon, Lyon, 6 juin 2019.
- [28] « La variance et ses enjeux dans l'édition numérique du *Regimiento de los príncipes* ». Journée d'étude Humanités Numériques, Agorantic-Ciham, Avignon, 27 mai 2019.
- sur invitation [29] « Construire l'édition numérique d'une tradition textuelle complexe : propositions pour le *Regimiento de los príncipes* ». Atelier IRPALL « La fabrique du texte », Toulouse, 22 février 2019.
- [30] « L'évolution du *Regimiento de los príncipes* (1345-1494), au service du pouvoir ? » Séminaire international de formation avancée : Les cultures politiques dans la Péninsule ibérique et au Maghreb, Bielle, octobre 2018.

## Rapports et documents non évalués par les pairs

- [31] Alix Chagué, Floriane Chiffolleau, Matthias Gille Levenson, Hugo Scheithauer et Ariane Pinche, « Chaînes d'acquisition, de traitement et de publication du texte », Consortium Ariane - Axe 1. 2024. hal-04734959

## Jeux de données

[32] Thibault Clérice et al. Jeu de données d'ATR diachronique et multilingue. CATMuS : reconnaissance de l'écriture manuscrite et imprimée. Mars 2024. Huggingface.

*Résumé:* Ce jeu de données de standard *gold*, produit dans le cadre des financements HTRogène (Biblissima+) et HTRomance (Datalab BNF) portés par Thibault Clérice, contient la transcriptions de fragments de plus de 200 manuscrits médiévaux du 7<sup>e</sup> au 15<sup>e</sup> siècles, dans plus d'une demi-douzaine de langues (ancien et moyen français, italien, vénitien, latin, occitan, anglais, castillan, catalan, navarrais, flamand). Outre l'apport au corpus de mon propre corpus de thèse (voir item suivant), je suis co-responsable avec Olivier Brisville-Fertin du versant ibérique du projet et de la qualité des données produites en castillan et catalan.

[33] Jeu de données et modèles évalués – reconnaissance de l'écriture manuscrite et imprimée du castillan médiéval. Novembre 2022. 10.5281/zenodo.7386489

*Résumé:* Ce jeu de données de standard *gold* contient environ 360 pages segmentées (zones et lignes) et transcrites selon des normes allographétique et graphématiques, ainsi qu'un ensemble de modèles de transcription. Les données sont celles de mon corpus de thèse. Ce jeu de données a été évalué dans l'article 2.

[34] Jeu de données non évalué – segmentation du castillan médiéval (14<sup>e</sup>–15<sup>e</sup> siècle), mai 2022. 10.5281/zenodo.7684879.

*Résumé:* Ce jeu de données de standard *gold* contient environ 45.000 lignes transcrites selon des normes allographétiques et segmentées (c'est-à-dire, les mots étant identifiés comme des tokens individuels). Les données sont celles de mon corpus de thèse.

[35] Jeu de données non évalué – lemmatisation du castillan médiéval (14<sup>e</sup>–15<sup>e</sup> siècle), juin 2023. 10.5281/zenodo.7772058.

*Résumé:* Ce jeu de données de standard *gold* contient environ 280.000 tokens annotés en lemmes, parties du discours et morphologie. Les données sont celles de mon corpus de thèse.

## Logiciels développés

L'ensemble de ma production informatique est disponible sur l'instance gitlab d'Huma-Num ou sur Github.

[36] **TeiCollator** : <https://github.com/matgille/tei-collator>

TeiCollator est un outil développé en python et un ensemble de méthodes de collation semi-automatisée TEI to TEI qui s'appuie sur CollateX pour l'alignement des témoins. Il permet en particulier la collation et la comparaison d'éléments non textuels, et offre la possibilité d'obtenir en sortie autant de témoins-base que de témoins disponibles, dans une optique essentiellement bédieriste.

[37] **ClaViTranscr** : [https://github.com/matgille/transcription\\_virtual\\_keyboard](https://github.com/matgille/transcription_virtual_keyboard)

ClaViTranscr est un clavier virtuel développé en python et qui permet d'aider au travail d'établissement de vérités terrains dans le cadre de l'utilisation de logiciels de transcription automatisée comme eScriptorium. Il autorise en particulier la recherche de caractères et un ajout facilité de nouveaux caractères à la base de données.

[38] **Boudams Forks** : [https://github.com/matgille/boudams\\_like\\_tokenizer](https://github.com/matgille/boudams_like_tokenizer)

Boudams Fork est une adaptation du travail et du logiciel Boudams de Thibault Clérice. Il s'agit d'un tokéniseur fondé sur des techniques d'apprentissage supervisé (réseaux convolutifs). Cette adaptation permet l'utilisation de l'outil avec des sources XML-TEI (en particulier l'identification des mots coupés à la ligne), ainsi que la conservation de l'information de segmentation originale à l'aide d'entités XML, par ajout de classes de sortie à l'architecture originale.

[39] **Pyrrha** : <https://github.com/hipster-philology/pyrrha>

Pyrrha est une interface de post-correction d'annotations lexico-grammaticales imaginée et développée principalement par Thibault Clérice, à laquelle je contribue par des corrections de bugs, ajouts et suggestions de nouvelles fonctionnalités.

## Développement web

en cours [40] **CATMuS Guidelines** : <https://catmus-guidelines.github.io/>

Le site *CATMuS Guidelines* vise à faciliter l'utilisation et la consultation des normes édictées par le groupe CATMuS. Il consiste en un site web statique hébergé sur Github et développé à partir de fiches au format markdown, qui comprenant une fonctionnalité de d'indexation et de recherche.

## Expérience d'enseignement

Mon parcours d'enseignement m'a amené à proposer des cours de la L2 au doctorat ; les cours sont recensés ci-dessous. J'estime à environ 537 h. équivalent TD le total d'enseignement dispensé jusqu'en septembre 2024 (384 h. de cours réguliers, et 153 h. de formation ponctuelle d'une durée de 3 à 25 h).

J'ai depuis 2020 orienté une partie de mon profil pédagogique vers la reconnaissance automatique d'écriture imprimée et manuscrite ou HTR (unités d'enseignement 2, 4, 10, 11, 18, 19, 21), et me suis aussi intéressé à l'enseignement de l'annotation lexico-grammaticale (unités 18, 23, 24, 26) et à l'édition numérique (2, 9, 20, 22, 27), domaine que je maîtrise parfaitement et que je pourrais enseigner à partir de la L3. En ce qui concerne mon expérience d'enseignement en Humanités Numériques appliquées au monde hispanophone, je me suis centré sur le traitement automatisé du texte médiéval et moderne, de l'image au texte annoté (voir l'unité 18 en particulier, dont l'objet d'études était Bartolomé de las Casas).

Les cours que j'ai eu l'occasion de proposer à l'ENS de Lyon, à l'Université de Saint Quentin et à Aix-Marseille Université sont plus disciplinaires et spécifiques à mon domaine de rattachement, la langue et la civilisation hispanique : j'ai notamment une grande expérience dans les cours de traduction (version classique et moderne, tous niveaux de la L3 à l'Agrégation : 7, 15, 16, 17, pour un total de 140h). Concernant les cours de civilisation médiévale que j'ai eu l'occasion de proposer, ils ont essentiellement abordé des questions d'histoire politique et de légitimation par l'écrit, qu'il soit chronistique ou théorique, mais aussi des questions d'histoire du droit en Castille, notamment en rapport avec le programme d'Agrégation de 2023 (unités 12, 14, 25).

### *PRODUCTION COLLABORATIVE D'UN CORPUS D'ATR*

[1]

**Institution:** ÉNS de Lyon

**Niveau:** Doctorat

**Années:** 2024–2025

**Nombre total d'heures:** 15 h.

Animation d'un groupe (cinq personnes) dans le cadre du transcribathon Transcriboquest 2024 (Lyon, 11-13 septembre 2024) organisé par Thibault Clérice et Ariane Pinche. Production d'un corpus d'ATR de manuscrits littéraires.

### *HTR ET ENCODAGE XML-TEI*

[2]

**Institution:** Univ. de Poitiers / ÉNS de Lyon

**Niveau:** Doctorat

**Années:** 2023–2024 ; 2022–2023

**Nombre total d'heures:** 55 h.

Animation d'un groupe d'une demi-douzaine de personnes de la formation EnExDi 2024 (Lyon, 8-12 avril 2024) et en janvier 2023 ; séances co-animées avec Ariane Pinche sur le HTR et la transcription de sources au format XML-TEI.



*INITIATION À HTR – SOURCES CASTILLANES MÉDIÉVALES*

[3]

**Institution:** ÉNS de Lyon

**Niveau:** Master et Doctorat

**Années:** 2023–2024

**Nombre total d’heures:** 5h.

Animation d’une formation HTR dans le cadre du Séminaire Hispanique d’Études Médiévales (Lyon, 5 avril 2024); séances co-animées avec Ariane Pinche, Olivier Brisville et Julie Bordier. Une vingtaine de participants est prévue.

*INITIATION À L’ECDOTIQUE NUMÉRIQUE*

[4]

**Institution:** Universités Grenoble-Alpes

**Niveau:** Master-Doctorat

**Années:** 2023–2024

**Nombre total d’heures:** 10 h.

Animation du colloque-école d’hiver « Du calame au clavier » (Grenoble, 17–19 janvier 2024). Communication et ateliers en HTR et sur la collation semi-automatisée à l’aide de Notebooks Python. Lien vers le matériel : github.

*SOCIÉTÉ NUMÉRIQUE ET HUMANITÉS*

[5]

**Institution:** UVSQ

**Niveau:** L2 LLCER espagnol

**Années:** 2023–2024

**Nombre total d’heures:** 18 h.

Présentation et utilisation des outils de visualisation de données (StoryMap, StoryLine, TimeLine, Datawrapper), appliqué à des thèmes de civilisation hispanophone. Présentation et application de méthodes de lecture rapide d’articles journalistiques et scientifiques sur les thèmes étudiés en cours.

*SOCIÉTÉ NUMÉRIQUE ET HUMANITÉS*

[6]

**Institution:** UVSQ

**Niveau:** L2 LLCER anglais

**Années:** 2023–2024

**Nombre total d’heures:** 18 h.

Présentation et utilisation des outils de visualisation de données (StoryMap, StoryLine, TimeLine, Datawrapper), appliqué à des thèmes de civilisation anglophone. Présentation et application de méthodes de lecture rapide d’articles journalistiques et scientifiques sur les thèmes étudiés en cours.

*VERSION*

[7]

**Institution:** ENS de Lyon

**Niveau:** Agrégation

**Années:** 2023–2024

**Nombre total d’heures:** 21 h.

Cours de version classique et moderne dans le cadre de la préparation à l’Agrégation externe d’espagnol (décembre 2023–janvier 2024).

*INITIATION À LA COLLATION AUTOMATISÉE*

[8]

**Institution:** École nationale des chartes

**Niveau:** Doctorat

**Années:** 2023–2024

**Nombre total d’heures:** 3 h.

Co-Animation d’un atelier avec Lucence Ing consacré à l’alignement et à la collation automatisée, dans le cadre des journées du cluster 7 de Biblissima+ en décembre 2023. Travail à l’aide de notebooks python. Lien vers le matériel : Alignement, Collation.

*INTRODUCTION L’ÉDITION ÉLECTRONIQUE*

[9]

**Institution:** Universidad Complutense de Madrid

**Niveau:** M2

**Années:** 2023–2024

**Nombre total d’heures:** 2 h.

Animation en novembre 2023 d’une séance de séminaire Master (Universidad Complutense de Madrid), consacré à une présentation générale de l’édition électronique, de l’image au texte édité.

*LA TRANSCRIPTION AUTOMATISÉE : INTRODUCTION AVANCÉE*

[10]

**Institution:** UNIL

**Niveau:** Master

**Années:** 2023–2024

**Nombre total d’heures:** 3 h.

Animation en octobre 2023 d’un séminaire de Master (Université de Lausanne), consacré à la transcription automatisée.

**INTRODUCTION À L'HTR** [11]  
**Institution:** ENSSIB **Niveau:** Master  
**Années:** 2023–2024; 2022–2023 **Nombre total d'heures:** 6 h.  
Participation au séminaire du master HN de l'ENSSIB/Lyon 3/ENS de Marianne Reboul en octobre 2022 et 2023; séances en collaboration avec Ariane Pinche sur la transcription automatisée.

**CIVILISATION MÉDIÉVALE HISPANIQUE** [12]  
**Institution:** ENS de Lyon **Niveau:** Agrégation  
**Années:** 2022–2023 **Nombre total d'heures:** 9 h.  
Cours d'introduction historique au siècle d'Alphonse X, dans le cadre de la préparation à l'agrégation externe d'espagnol. 3 séances au début du premier semestre 2022–2023.

**OUTILS DE LA RECHERCHE** [13]  
**Institution:** ENS de Lyon **Niveau:** M1  
**Années:** 2022–2023; 2019–2020; 2018–2019 **Nombre total d'heures:** 90 h.  
Cours de méthodologie de la recherche. Intitulé du cours : La publication scientifique : enjeux économiques, juridiques, éditoriaux et techniques. Le cours a alterné entre présentation de la publication scientifique sous ses différents prismes (juridique, scientifique, économique) et initiation à la production de documents scientifiques avec le logiciel/langage de composition  $\text{\LaTeX}$  : structuration de document, mise en page, bibliographie, diapositives.

**CIVILISATION MÉDIÉVALE HISPANIQUE** [14]  
**Institution:** ENS de Lyon **Niveau:** M1  
**Années:** 2022–2023 **Nombre total d'heures:** 30 h.  
Le sous-titre du cours est : *Droit, pouvoir, légitimité politique, d'Alphonse X aux Trastamares.*

**SÉMINAIRE DE TRADUCTION** [15]  
**Institution:** ENS de Lyon **Niveau:** M2  
**Années:** 2022–2023 **Nombre total d'heures:** 30 h.  
Le thème du cours est la réécriture. Il s'est agi de travailler à la traduction de textes ayant été repris et réécrit au cours des siècles, du Moyen Âge au 20<sup>e</sup>.

**VÉRSION CLASSIQUE** [16]  
**Institution:** ENS de Lyon **Niveau:** L3–M1  
**Années:** 2022–2023; 2018–2019 **Nombre total d'heures:** 60 h.  
Cours de version classique.

**VÉRSION MODERNE** [17]  
**Institution:** ENS de Lyon **Niveau:** L3–M1  
**Années:** 2022–2023 **Nombre total d'heures:** 30 h.  
Cours de version moderne.

**HUMANITÉS NUMÉRIQUES APPLIQUÉES AU MONDE HISPANIQUE** [18]  
**Institution:** ENS de Lyon **Niveau:** L3–M1  
**Années:** 2022–2023 **Nombre total d'heures:** 30 h.  
Le sous-titre du cours est *La Brevíssima relación de la destruyción de las Indias* : de l'image au texte annoté. Le cours a consisté en une introduction aux différentes méthodes de traitement et d'étude des sources textuelles anciennes, en utilisant une même source littéraire comme fil conducteur, la *Brevíssima relación de la destruyción de las Indias* (1552) de Bartolomé de las Casas : courte introduction à la bibliographie matérielle, OCR et travail sur l'apprentissage supervisé, segmentation, annotation lexico-grammaticale du castillan moderne.

- INITIATION À L'HTR ET À ESCRIPTORIUM* [19]  
**Institution:** UCM **Niveau:** Master et doctorat  
**Années:** 2021-2022 **Nombre total d'heures:** 4 h.  
Animation en mai 2022 d'un atelier doctoral à l'Universidad Complutense de Madrid (dans le cadre du séminaire « *Herramientas digitales* ») sur le HTR et eScriptorium.
- DIGITAL SCHOLARLY EDITIONS : MANUSCRIPTS, TEXTS AND TEI ENCODING* [20]  
**Institution:** Université Lyon II **Niveau:** Master et plus  
**Années:** 2021-2022; 2020-2021; 2019-2020 **Nombre total d'heures:** 18 heures  
Animateur secondaire (2019-2021) et principal (2022) du MOOC *Digital Scholarly Editions : Manuscripts, Texts and TEI Encoding* organisé par Marjorie Burghart dans le cadre du Master Mondes Médiévaux de l'Université Lyon II, aux printemps 2020, 2021, 2022.
- TRANSCRIPTION AUTOMATISÉE DE SOURCES MANUSCRITS AVEC ESCRIPTORIUM* [21]  
**Institution:** Casa de Velázquez **Niveau:** Master et Doctorat  
**Années:** 2021-2022 **Nombre total d'heures:** 16 h.  
Co-organisation et co-animation de la formation à la transcription automatisée de sources manuscrites avec le logiciel eScriptorium, avec María Díez Yáñez et Irene Salvo García, organisée à la Casa de Velázquez en septembre 2021 (vingt participant.es présent.es, trente en distanciel), avec la collaboration et la participation de Peter Stokes, Benjamin Kiessling, Leonor Zozaya-Montes et de Belén Almeida. Le support de la formation était la *General Estoria*. Lien vers la page de l'événement.
- DOCUMENTATION D'ÉDITIONS TEI AVEC ODD* [22]  
**Institution:** CIHAM **Niveau:**  
**Années:** 2020-2021 **Nombre total d'heures:** 12 h.  
Coorganisation et animation de la formation en ligne à la documentation TEI-ODD, avec Emmanuelle Morlock, Sarah Orsini et Ariane Pinche (mai-juin 2021, trois demi-journées). Lien vers la page de l'événement
- LEMMATISER DES TEXTES ET CORRIGER L'ANNOTATION GRÂCE A L'APPRENTISSAGE PROFOND AVEC PYRRHA* [23]  
**Institution:** **Niveau:** Mixte  
**Années:** 2020-2021 **Nombre total d'heures:** 4 h.  
Atelier coanimé avec Thibault Clérice, Lucence Ing, Ariane Pinche, Simon Gabay et Jean-Baptiste Camps en mai 2021 lors du congrès Humanistica. Lien vers la page de l'atelier et Abstract.
- INTRODUCTION À L'ANNOTATION LEXICO-GRAMMATICALE* [24]  
**Institution:** ENSSIB **Niveau:** Master  
**Années:** 2020-2021 **Nombre total d'heures:** 4 h.  
Animation d'une séance d'introduction à l'analyse grammaticale et à l'utilisation de l'outil Pyrrha, dans le cadre du master HN de l'ENSSIB, en octobre 2020.
- CIVILISATION MÉDIÉVALE HISPANIQUE* [25]  
**Institution:** ENS de Lyon **Niveau:** L3-M1  
**Années:** 2019-2020 **Nombre total d'heures:** 15 h.  
Cours d'introduction à l'histoire culturelle de la Castille sous les Trastamare. Thème : le pouvoir et sa légitimation.
- HUMANITÉS NUMÉRIQUES* [26]  
**Institution:** ENS de Lyon **Niveau:** L3-M1  
**Années:** 2019-2020 **Nombre total d'heures:** 15 h.

Cours d'introduction aux humanités numériques : l'annotation lexico-grammaticale du castillan médiéval.

#### FORMATION À XSLT

[27]

**Institution:** CIHAM

**Niveau:** Master et doctorat

**Années:** 2018–2019

**Nombre total d'heures:** 16 h.

Organisation et animation de la formation XSLT COSME 2019 en collaboration avec trois autres doctorant.es/docteurs : Simon Gabay, Université de Lausanne, Ariane Pinche, École nationale des Chartes, Jean-Paul Rehr, Université Lyon II. Lien vers le dépôt de la formation.

#### COLLES DE PRESSE

[28]

**Institution:** Lycées La Martinière–Duchère et Juliette

**Niveau:** Première année de CPGE

Récamier

**Années:** 2017–2018

**Nombre total d'heures:** 40 h (non décomptées)

Colleur dans deux lycées lyonnais. Sujets de presse, environ 40 h.

## Projets financés en cours

### Co-Porteur du projet e-Cam : (2024–)

Co-porteur avec Olivier Brisville-Fertin du projet « e-Cam » (Étiquetage lexico-grammatical du Castillan Médiéval), qui vise à la constitution de normes standard d'annotation du castillan médiéval (lemmes, parties du discours, morphologie), ainsi qu'à la création d'un corpus représentatif du castillan médiéval, annoté et intégralement corrigé à la main, dans l'optique de l'entraînement d'un modèle d'annotation plus efficace que les outils existants. Organismes financeurs : Agorantic (8000 euros obtenus), Biblissima+ (candidature en cours).

### Responsable de la section ibérique des projets HTRomance et HTRogène : (2023–2025)

Membre des projets « HTRomance » (financement DataLab BNF) et « HTRogène » (financement Biblissima+) portés par Thibault Clérice et Alix Chagué, consacrés à la production de données d'entraînement d'acquisition automatisée d'écriture manuscrite. Co-responsable avec Olivier Brisville-Fertin du versant ibérique du projet (constitution de jeux de données en castillan et catalan par supervision de vacations de transcription).

## Intégration dans des comités éditoriaux et de lecture

2023– Membre du comité éditorial des *Cahiers d'Études Hispaniques Médiévales*.

2021– Membre du comité de rédaction de *Programming historian en français* (rédacteur-en-chef adjoint).

## Activités d'encadrement et participation à des jurys

### Participation à des jurys de mémoire

2023 Jury de soutenance du mémoire de master 1 de Kélian DESNOUES. Titre du mémoire : *Un pasado continuo : estudio lingüístico del idiolecto de Cristóbal Colón*, mémoire réalisé sous la direction d'Olivier BRISVILLE-FERTIN.

### Encadrement de stages

2024 Co-encadrant du stage de M1 d'Amandine FRÉCHOU avec Olivier BRISVILLE-FERTIN. Thèmes de travail : HTR, annotations grammaticales, études lexicales.

## Compétences

### Langues pratiquées

- **Français, Espagnol** : bilingue.
- **Anglais** : compétence professionnelle et langue de travail fréquente.

### Compétences informatiques

- **Maîtrise des méthodes et outils d'édition électronique** selon le standard XML TEI/P5 : XML, ODD, XSLT, librairie python-lxml. Compréhension et utilisation du XQuery ; expérience ponctuelle dans la construction de bases de données XML. Utilisation courante des expressions régulières.
- **Utilisation courante du langage Python et maîtrise des fondamentaux de la Programmation Orientée Objet.**
- **Utilisation courante des méthodes de publication** de sources documentaires et/ou littéraires au format fixe (L<sup>A</sup>T<sub>E</sub>X, reledmac, reledpar).
- **Compétence dans l'utilisation des processus et outils de développement web** (HTML, CSS, Javascript). Expérience ponctuelle avec le framework de développement web flask.
- **Compétence dans l'utilisation des processus de publication statique** (Jekyll, jinja) adossés à Github (Github Actions).
- **Maîtrise et utilisation du logiciel de contrôle de version git** et des plateformes github et gitlab.
- **Utilisation courante des logiciels permettant la création de modèles d'apprentissage supervisé** pour le traitement de l'image et du texte ancien (reconnaissance de caractères et d'écriture manuscrite : Ocropy, Kraken, eScriptorium ; analyse grammaticale : freeling, Pie ; segmentation/tokénisation : Boudams). Maîtrise des fondamentaux de l'apprentissage supervisé. Expérience ponctuelle dans la création d'une architecture neuronale (Fork de Boudams, voir plus haut).
- **Maîtrise des systèmes d'exploitation basés sur linux** (Ubuntu, Debian). Utilisation courante de bash.

## Responsabilités organisationnelles

- Participation à l'organisation d'une journée d'études du SEHM (Séminaire d'Études Hispaniques Médiévales) du CIHAM, suivie d'une formation à l'HTR : « Humanités numériques et philologie romane », les 4 et 5 avril 2024.
- Organisation d'une séance du SEHM (Séminaire d'Études Hispaniques Médiévales) du CIHAM (coord. Julie Marquer et Carlos Heusch) : « Actualités de la recherche », le 13 décembre 2022.
- Participation à l'organisation des Journées d'Étude Doctorales 2021 du CIHAM.
- Organisation des Journées des Jeunes Hispanistes 2019 (ENS de Lyon).
- Participation à l'organisation d'un nombre important de formations ponctuelles : formations EnExDi (2024) ; Du Calame au Clavier (2024), eScriptorium (2021) ; ODD (2021) ; XSLT (2019).