



**HAL**  
open science

# Comprendre la composition des légendiers à l'aide des méthodes numériques

Ariane Pinche

► **To cite this version:**

Ariane Pinche. Comprendre la composition des légendiers à l'aide des méthodes numériques : Acquisition automatique du texte et analyses stylo-métriques. CBMAhag – Ateliers 2022, Elena Magnani, Jan 2023, Paris, France. hal-03930385

**HAL Id: hal-03930385**

**<https://hal.science/hal-03930385>**

Submitted on 11 Mar 2023

**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.

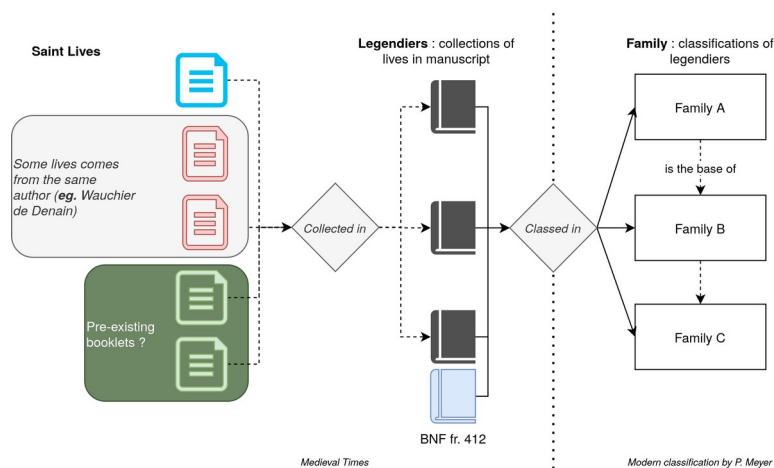


Distributed under a Creative Commons Attribution 4.0 International License

## Comprendre la composition des premiers légendiers en langue vernaculaire à l'aide des méthodes numériques : Acquisition automatique du texte (HTR) et analyses stylométriques

Ariane Pinche, CNRS, CIHAM (UMR 5648)

Les premières études sur la compilation des légendiers français remontent au début du 20<sup>e</sup> siècle avec les célèbres travaux de Paul Meyer<sup>1</sup>. Les mécanismes qui régissent leur compilation sont encore l'objet de recherches récentes<sup>2</sup>, montrant notamment que les premiers regroupements de Vies de saints en français ne proposent pas à l'intérieur de leurs principales subdivisions (apôtres, martyres, vierges et confesseurs) une organisation stable. L'ordre des Vies y est redistribué de famille en famille. Ainsi, les légendiers français sont le résultat d'un système de compilation complexe. Au niveau des manuscrits entiers, ils sont formés de grandes familles de légendiers qui se sont au cours du temps accumulées. À plus petite échelle, il semble qu'on puisse également détecter des microséries possiblement d'un même auteur, à l'instar du recueil des « *Seint Confessor* » de Wauchier de Denain (manuscrit, fr. 412 de la BnF).



L'objet de l'exposé, ici, est de proposer d'étudier les Vies de saint en langue vernaculaire avec une approche non pas par Vie ou dossier isolé, mais au niveau du légendier complet, voire des familles de légendier pour mieux comprendre les spécificités de ces compilations. Une telle approche nous permettrait :

1. d'identifier des microcompilations jusqu'ici passées inaperçues.
2. d'examiner l'histoire de leur transmission dans les légendiers postérieurs pour retracer l'évolution des récits, de leur circulation et de leurs usages.

Afin de parvenir à cette vue d'ensemble, nous appuierons notre recherche sur deux technologies numériques : la reconnaissance automatique d'écriture (HTR, *handwritten text recognition*) pour acquérir le texte des légendiers depuis la numérisation des sources, puis la stylométrie pour explorer ce vaste corpus et repérer les récits possiblement d'un même auteur.

1 Paul Meyer, « Légendes hagiographiques en français », in *Histoire littéraire de la France*, vol. 33, Imprimerie nationale, Paris, 1906, p. 328-458.

2 Anne-Françoise Labie-Leurquin, « Composition, usage et diffusion du légendier picard », in *Des saints et des livres : Christianisme flamboyant et manuscrits hagiographiques du Nord à la fin du Moyen Âge*, Turnhout, 2021 (« Hagiologia », 17), p. 79-133.

Ces deux étapes reposent sur des travaux préliminaires. Pour l'HTR, nous avons consacré l'année 2021-2022 à la confection de modèles HTR génériques capables de traiter des collections de manuscrits littéraires des 12<sup>e</sup>, 13<sup>e</sup> et 14<sup>e</sup> siècles, voir le projet CREMMALab<sup>3</sup> et le modèle Cortado<sup>4</sup>. Quant à la stylométrie, la prospection proposée ici s'appuie sur l'expérience menée pour la conférence *Digital Humanities 2019* lors de laquelle, avec nos collègues J.-B. Camps et T. Clérice<sup>5</sup>, nous avons présenté une étude stylométrique du manuscrit fr. 412 de la BnF pour en analyser la composition. C'est cette piste que nous souhaitons poursuivre et étendre à un corpus d'une trentaine de légendiers pour confronter les résultats déjà obtenus aux autres compilations et, ainsi, observer les réarrangements des Vies au fil des témoins afin de mieux comprendre les enjeux d'écriture et de réception sous-jacents.

---

3 <https://cremmalab.hypotheses.org>

4 <https://github.com/HTR-United/cremma-medieval/releases/tag/2.0.0>

5 CAMPS, Jean-Baptiste, CLÉRICE, Thibault et PINCHE, Ariane, « Noisy medieval data, from digitized manuscript to stylometric analysis: Evaluating Paul Meyer's hagiographic hypothesis », *Digital Scholarship in the Humanities*, vol. 36/Supplement\_2, octobre 2021, p. ii49 -ii71.