



**HAL**  
open science

## Hagiographie bourguignonne 2.0 – corpus électronique et nouvelles recherches

Eliana Magnani

► **To cite this version:**

Eliana Magnani. Hagiographie bourguignonne 2.0 – corpus électronique et nouvelles recherches :  
Compte-rendu de la journée d'études du 6 décembre 2021. 2021. halshs-03500876

**HAL Id: halshs-03500876**

**<https://shs.hal.science/halshs-03500876>**

Submitted on 22 Dec 2021

**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 - ShareAlike 4.0 International License



## Hagiographie bourguignonne 2.0 – corpus électronique et nouvelles recherches

Compte-rendu de la journée d'études du 6 décembre 2021

Eliana Magnani

---



### Édition électronique

URL : <https://journals.openedition.org/cem/18495>

ISSN : 1954-3093

### Éditeur

Centre d'études médiévales Saint-Germain d'Auxerre

### Référence électronique

Eliana Magnani, « Hagiographie bourguignonne 2.0 – corpus électronique et nouvelles recherches », *Bulletin du centre d'études médiévales d'Auxerre | BUCEMA* [En ligne], Collection CBMA, Les journées d'études, mis en ligne le 22 décembre 2021, consulté le 22 décembre 2021. URL : <http://journals.openedition.org/cem/18495>

---

Ce document a été généré automatiquement le 22 décembre 2021.



Les contenus du *Bulletin du centre d'études médiévales d'Auxerre (BUCEMA)* sont mis à disposition selon les termes de la Licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Partage dans les Mêmes Conditions 4.0 International.

---

# Hagiographie bourguignonne 2.0 – corpus électronique et nouvelles recherches

Compte-rendu de la journée d'études du 6 décembre 2021

Eliana Magnani

---

1 La rencontre « à distance » qui s'est tenue le 6 décembre 2021 visait à reprendre les travaux collectifs autour du programme CBMA, à la suite des ateliers réunis en 2018<sup>1</sup> et des reports successifs des rencontres prévues en décembre 2019 et en mars 2020<sup>2</sup>.



2 L'accent a été mis sur le sous-corpus hagiographique, car la mise à disposition d'un ensemble interrogeable de 325 textes hagiographiques latins, renseignés avec des nombreuses métadonnées, offre une extensive matière pré-organisée pouvant servir à des recherches les plus variées, mais aussi l'occasion de mesurer les lacunes que les recherches à venir pourraient combler à l'aide du numérique<sup>3</sup>. En vue de confronter les résultats et de partager les découvertes de ces investigations nouvelles, trois ateliers sont d'ores et déjà prévus en 2022<sup>4</sup>, où seront conjuguées des différentes approches, en variant les échelles d'analyse, de près et de loin, et diverses thématiques : tradition manuscrite, érudition moderne et historiographie ; datation, attribution et spatialisation ; philologie, sémantique et intertextualité. Alors que le « corpus hagiographique bourguignon » est un ensemble composé de productions irréductibles à une composante purement régionale, l'ouverture à d'autres espaces (Italie, Normandie, Provence, Septimanie, entre autres), les comparaisons, la prise en compte de la réticularité

dont témoignent ces écrits devront aussi orienter les travaux. De même, l'articulation de ces textes avec d'autres formes de productions écrites, voire figurées, devront aussi retenir l'attention.

- 3 Après l'introduction générale, le rappel des actualités du projet et la présentation du profil du corpus « hétérogène » que nous détaillons ci-dessous, quatre présentations sont revenues sur des points divers de l'utilisation de la documentation hagiographique bourguignonne : l'emploi du sous-corpus hagiographique pour la création d'un module lexical médiéval pour l'application *Collatinus* (Yves Ouvrard, Philippe Verkerk), les débuts de l'enquête sémantique sur la « tristesse » et l'utilisation des grands corpus pour une histoire des émotions (Gabriel Castanho), le dossier hagiographique du martyr Baudile de Nîmes et la translation de ses reliques en Bourgogne (*BHL* 1047) (Fernand Peloux), et le dossier de Jean de Réôme à l'égard des analyses sémantiques et des recherches diplomatiques et archéologiques récentes (Nicolas Perreaux)<sup>5</sup>.

## Nouveautés du programme CBMA

- 4 Les principales nouveautés du programme CBMA, en 2020 et 2021, en matière de documents saisis et mis à disposition sont liées aux évolutions récentes d'un projet qui s'inscrit dans la durée et a été conçu comme une plateforme dynamique, toujours ouverte aux nouvelles incorporations. Sans refaire toute la généalogie du projet, qu'on peut lire ailleurs<sup>6</sup>, il n'est peut-être pas inutile de rappeler qu'il est né en 2004 avec le propos de réunir les documents diplomatiques en latin antérieurs à 1300 relatifs à la Bourgogne comprise d'après les juridictions ecclésiastiques médiévales – les diocèses d'Auxerre, Langres, Nevers, Autun, Chalon et Mâcon. Cet ensemble était, en principe, monotypologique, mais avec la saisie, par exemple, de la *Bibliotheca Cluniacensis* (1614), ou encore d'autres éditions qui mélangeaient chartes et d'autres types de documents comme *Les chartes et documents de l'abbaye de Cîteaux* par Jean Marilier (1961), divers types documentaires sont entrés très tôt dans la collection, même s'ils restaient marginaux<sup>7</sup>. L'idée qui avait germé depuis longtemps d'ouvrir le CBMA à d'autres types de documents a été formalisée en 2017 avec le passage des « Chartes » au « Corpus », et la décision assumée de constituer un corpus hétérogène et de le prolonger jusqu'à la fin du Moyen Âge<sup>8</sup>.
- 5 Deux sous-corpus principaux ont été mis en œuvre, celui de l'hagiographie latine et celui de l'épigraphie, mais d'autres sont en cours d'incorporation<sup>9</sup>. Avec le sous-corpus épigraphique, réalisé en collaboration avec le CESCUM de Poitiers, en plus des inscriptions éditées dans les volumes du *Corpus des Inscriptions de la France Médiévale*, nous avons récupéré aussi les inscriptions des XIV<sup>e</sup> et XV<sup>e</sup> siècles qui avaient été recensées dans les fichiers papier à Poitiers mais pas publiées<sup>10</sup>. Cela a fait entrer des nombreux textes en langue vernaculaire

dans le corpus, mais ce n'était pas non plus une nouveauté car nous avons déjà des lots de documents en ancien français dans la base, notamment depuis la saisie en 2016 des presque 10 000 pièces justificatives publiées par Ernest Petit dans les 9 tomes de son *Histoire des Ducs de Bourgogne de la race capétienne* (1885-1905), dont beaucoup sont des registres<sup>11</sup>.

- 6 Dans ce contexte d'extension typologique et chronologique du corpus et de prise en compte des textes vernaculaires, les dernières incorporations de textes concernent, d'une part les parties d'éditions qui avaient été écartées au tout début du projet soit parce qu'elles sortaient de la chronologie, de la typologie ou de la langue fixées, d'autre part des éditions qui manquaient encore dans le corpus, dont, parmi d'autres, le recueil de pièces pluritypologiques du tome 4 des *Mémoires concernant l'histoire d'Auxerre* de Jean Lebeuf (1855)<sup>12</sup>. Pour l'anecdote, a été saisie également la transcription du cartulaire des Dominicains de Dijon, que Nicolas Perreaux et Françoise Bornes ont (ré)découvert fortuitement en Sorbonne, dans les étagères du LaMOP, dans un mémoire de maîtrise dirigé par Michel Mollat et soutenu en 1970 par Éric Coince<sup>13</sup>. Pour l'hagiographie, le sous-corpus a été notamment complété avec le dossier de Vézelay de Marie Madeleine. Il n'est pas lieu ici d'énumérer toutes les nombreuses acquisitions récentes dont la liste détaillée a été diffusée dans *Les carnets du LaMOP*<sup>14</sup>.
- 7 Parmi les nouveautés, il faut aussi souligner, dans le cadre de la maintenant longue collaboration avec les Archives départementales de la Côte d'Or, que huit nouveaux manuscrits ont été photographiés par le photographe des archives, Frédéric Pétot, et mis en ligne sur la plateforme CBMA. Il s'agit des quatre volumes des chartes de Flavigny transcrits à la fin du XIX<sup>e</sup> siècle par l'abbé François Grignard [1 F 213-1, 1 F 213-2, 1 F 214-1 et 1 F 214-2], ainsi que le *Cartulaire du comté d'Auxerre. Actes de foi et d'hommages au comte d'Auxerre* [Cart. 6.02 (B 10434)], le *Cartulaire des fiefs et arrières-fiefs du comté d'Auxerre* [Cart. 6 (B 10435)], le *Cartulaire de la commanderie de Beaune* (1490) [Cart. 237 (H 113)], et le *Cartulaire de l'abbaye de Notre-Dame d'Oigny* (XV<sup>e</sup>-XVI<sup>e</sup> s.) [Cart. 206 (19 H)].
- 8 Depuis 2019, grâce aux procédures mises en place par Pierre Brochard, le projet CBMA a entrepris de déposer les numérisations des manuscrits conservés dans les Archives départementales de la Côte-d'Or sur l'entrepôt de données Nakala, de la TGIR Huma-Num, afin de bénéficier de leur offre de service en matière de visualisation d'images avec le point d'entrée IIIF et de les sauvegarder de manière pérenne. Chacune de ces numérisations est accompagnée d'un ensemble de métadonnées permettant son identification et sa compréhension ainsi que d'un DOI (Digital Object Identifier ou « identifiant numérique d'objet »). Grâce au logiciel Mirador et l'instance installée chez Biblissima dont nous profitons, il est désormais possible de visualiser l'ensemble de ces manuscrits avec ces nouvelles technologies<sup>15</sup>. Les manuscrits sont aussi téléchargeables entièrement en .pdf dans le site du projet<sup>16</sup>. La collection

de manuscrits numérisés CBMA contient actuellement 63 manuscrits, dont des cartulaires, des livres de comptes ou encore les tablettes en cire de Cîteaux.

- 9 Plusieurs nouveautés sont signalées dans les différentes pages du site du projet, qui a été mis à jour en septembre 2021 par Coraline Rey. À noter, en particulier, la nouvelle page dédiée au corpus épigraphique « dit provençal », avec les inscriptions relatives aux départements des Alpes-de-Haute-Provence, Hautes-Alpes, Alpes-Maritimes, Ardèche, Bouches-du-Rhône, Drôme, Gard, Lozère, Var, et Vaucluse<sup>17</sup>.

## **Le corpus « hétérogène ». Principales caractéristiques, sous-corpus et exploitations préliminaires**

- 10 Au vu des transformations dans les contenus du CBMA, il semble souhaitable de faire le point sur l'état actuel de la base de données et du corpus. Il faut bien distinguer ces deux volets du programme, ce qui relève de la base de données, où toute information, qu'elle soit ou pas associée à un texte en présence peut être utile à la recherche, et ce qui relève du corpus textuel dont nous cherchons à ce qu'il soit interrogeable avec des outils de fouille de texte, comme TXM, NoSketchEngine ou Philologic.

### **Structuration des métadonnées**

- 11 La mise au point d'une version complète et lemmatisée du corpus numérisé est en cours de réalisation, avec des tests actuellement réalisés sur la « version bêta » en vue d'ajuster au mieux la formalisation du corpus aux usages différenciés dont il peut être l'objet<sup>18</sup>. De fait, avec la compilation des sous-corpus hagiographique et épigraphique, plusieurs métadonnées spécifiques à ces types de documents ont été ajoutées à celles déjà existantes pour le corpus diplomatique. Par exemple, le n° *BHL* ou le nom du saint est propre aux textes hagiographiques, le n° du *CIFM* ou la précision du type de support est propre aux inscriptions, et ainsi de suite. Le premier point auquel nous nous sommes occupés ces derniers mois concerne le nettoyage et l'harmonisation des métadonnées (depuis juin 2019). Il fallait identifier les métadonnées communes, les fusionner le plus possible, éliminer ou fusionner des métadonnées trop peu renseignées pour être interrogées, parmi d'autres opérations. Il fallait aussi créer des nouvelles métadonnées propres, cette fois-ci, au corpus hétérogène, et indiquer notamment la typologie des documents, pour bien distinguer les sous-corpus. Nous avons retenu un jeu de neuf types : diplomatique, hagiographique, normatif, historiographique, épistolaire, administratif, théologique, liturgique, épigraphique. Tous les textes sont maintenant étiquetés avec cette typologie ce qui permet de trier le corpus sur ce critère. Une partie de ce travail a été réalisé de manière semi-automatique, mais c'est Coraline Rey qui l'a poursuivi à la main, au cas

par cas, avec toute les difficultés de choisir une catégorie pour des documents qui peuvent en fait répondre à plusieurs.

- 12 La langue des textes a aussi été renseignée complètement, pour pouvoir séparer le corpus latin du corpus français ancien.
- 13 Une autre métadonnée, la date, nous a occupé, et nous occupe toujours beaucoup. Pour un affichage chronologique des textes dans les applications de fouille de texte, il fallait aussi créer ce que nous avons appelé « date simplifiée » (en fait quatre chiffres indiquant une année) qui, quand le document n'est daté qu'approximativement, affiche l'estimation de la date au milieu de la fourchette chronologique qui lui est attribuée, ce qui permet de le placer peut-être le moins mal possible dans une liste par ordre chronologique (par exemple, un document daté du XI<sup>e</sup> siècle reçoit la « date simplifiée » de 1050). Cette opération de mise en ordre chronologique n'est pas une opération de datation, la datation à proprement parler se trouvant dans une autre métadonnée, la « date retenue », qui peut correspondre à celle donnée par l'éditeur ou à celle éventuellement corrigée ou ajoutée par le projet CBMA. La « date simplifiée » est le premier des trois éléments qui constituent l'identifiant de chaque unité documentaire, qui contient également une référence abrégée à des éléments de l'édition, ce qui rend possible la reconnaissance rapide du document, ainsi que le « numéro CBMA ». L'identifiant est compilé ainsi à partir de différentes métadonnées de la base de données. Dans l'extrait ci-dessous du résultat d'une requête faite avec TXM, quelqu'un familier de la documentation bourguignonne reconnaît rapidement que les résultats concernent les chartes de Cluny éditées par Bernard et Bruel, le catalogue de la bibliothèque de Cluny édité par Léopold Delisle, les chartes de Saint-Vincent de Mâcon éditées par Camille Ragut, celles de Molesme par Jacques Laurent, de Paray-le-Monial par Ulysse Chevalier, de Saint-Benoît-sur-Loire par Maurice Prou et Alexandre Vidier, ainsi que les coutumiers clunisiens de Bernard et d'Ulrich (Fig. 1).

Fig. 1. CBMA. Extrait du résultat d'une requête avec TXM

text_id	Contexte gauche	Pivot	Contexte droit
1079 clunv-bernardbruel 4708	eorum honore consecrato, pro redemptione anime et	corporis	secultura alterius filii mei Constantini. unam uini
1079 clunv-bernardbruel 4964	et a communiione omnium fidelium extraneus et a	corpoere	Domini nostri oriuatus per lonoum temous. id es
1079 clunv-bernardbruel 4970	eis totam suoradiclam ecclesiam. excopta sexta oarte	corporis	ecclesie. annuentibus dominis Ansoaldo scilicet
1079 clunv-delisle 20271	234. Uolumen in duo continetur Paschasius de	corpoere	et sanouine Domini. sed et oassionis ac uite oli
1079 clunv-delisle 20271	272. Uolumen in ouo continetur Paschasius de	corpoere	et sanouine Domini. habens oassionis sanctoru
1079 macon-raout 20	. Nos autem; secundum apostolum. absentes	corpoere	oresentes autem spiritu. tam iudicauimus ut orei
1080 clunv-bernardbruel 4973	leicos ab introitu omnium Ecclesiarum. et ea	corpoere	et et sanouine Domini Iesu Christi. donec penite
1080 clunv-bernardbruel 4973	si facere noluerint. ab introitu Ecclesie.	corporis	Domini et sanouinis. nost predictum terminum
1080 clunv-bernardbruel 4973	ab omnium Ecclesiarum introitu. et ea communiione	corporis	et sanouinis Domini esse seoratum. donec resi
1080 clunv-bernardbruel 4985	uxor mea. et eoo cum ea reddidi	corpus	meum et isi similiter me cum faciant. Omnes hi
1080 clunv-bernardbruel 4985	et maledictionem Dei omnioientis et celorem uindictam in	corpoere	et in anima in hoc mundo et in die iudicii incurra
1080 clunv-bernardbruel 4985	in hac uita et in alia. in	corpoere	et in anima suoer se habeat. et tota benedictio c
1080 molesme-laurent 14385	litteris disoouis commendare. Eouidem animaduertens salutem esse	corporis	et anime si ecclesias Dei studerem in alioibus :
1080 ordocluniacensis-herroott 31937	se. continuo ueniam petere debet prostrato toto	corpoere	: nullus etiam transit ante illum ubicumoue sit lox
1080 ordocluniacensis-herroott 31938	ut has candelas ad usus hominum ad sanitatem	corporum	animarum. siue in terra. siue in acuis. per inuoc
1080 paray-chevalier 7020	quia lonae discesserant a prooris. tumulatio eius	corporis	non alibi. sed totius esset in loco a se construct
1080 saintbenoit-orouidier 11902	abbas Floriacensis. ubi beatus oster monachorum Benedictus	corporis	presentia reoasare dinoscitur. humilii suocilicatii
1080 saintbenoit-orouidier 11903	suo monasterio. ubi isoe oater Benedictus oresentia	corporis	reouasare dinoscitur. auctoritate reoie celestudin
1080 ulrich 32612	que meliora uidentur. de his fiat unum	corpus	ouod de cetero teneatur. Et si hoc uolueris esse
1080 ulrich 32612	nos federauit in unitate Spiritus. in uno	corpoere	. ouod est Ecclesia. sub uno capite. ouod est iot
1080 ulrich 32612	ut ita dicam. apendiclos. toto	corpoere	prosternimur. ut duos osalmos dicamus. orimum
1080 ulrich 32612	acoiotur : et ouibus uisum fuerit ab his	corpus	Domini sumitur. propter ouod etiam tres hostile c
1080 ulrich 32612	celebranda est oro ouolbet nuor defuncto. siue	corpus	defuncti oresentis sit. siue non. dicenda est oro e
1080 ulrich 32612	explicatis faciunt procumbendo. Interea uero reconditur Dominicum	corpus	a sacerdote retro altare. Ponitur in calena aurea

Toujours autour des dates, nous nous sommes trouvés dans une impasse quand nous avons voulu appliquer aux autres types documentaires du corpus le « degré de fiabilité de la date » que nous avons retenu collectivement lors des ateliers de 2018 pour classer les textes hagiographiques. Nous avons attribué un degré de fiabilité de 1 à 3 pour la datation des textes hagiographiques : 1 = Date assurément connue, par exemple citée dans le document ; 2 = Date supposée, fourchette chronologique réduite ; 3 = Date inconnue ou fourchette chronologique large. Et cette distinction simplifiée s'avère assez problématique à être transposée aux documents diplomatiques, par exemple. Les notions floues de « fourchette chronologique réduite ou large », nous a amenés à déterminer l'ampleur de ces fourchettes et à introduire un quatrième gradient, 4, pour les documents non datés. Nous avons renommé cette métadonnée « date-précision » au lieu de « fiabilité », mais tous les documents n'ont pas encore été renseignés et nous continuons à réfléchir à ces problèmes qui sont liés à la nature même de chaque type documentaire et la façon dont ces textes s'inscrivent dans l'économie générale de la société qui les a produits.

15

Malgré les difficultés inévitables posées par la datation, toutes les unités documentaires ont été étiquetées avec une « date simplifiée » et il est possible désormais de connaître la distribution chronologique d'ensemble du corpus. La multiplication des expériences et d'analyses permettront sans doute d'affiner ces choix heuristiques qui ne sont pas sans conséquences et doivent demeurer toujours conscients lors de l'utilisation du corpus formalisé.

16

En somme, entre suppressions, fusions et créations de métadonnées, nous sommes passés d'un tableau avec 81 colonnes et 30 749 lignes en juin 2019 à un tableau de 68 colonnes et 32 220 lignes en décembre 2021. Ce travail de structuration des métadonnées a bien avancé et se poursuivra ces prochains mois.

### **Profil du corpus**

17

Grâce à ces acquis nous sommes en mesure de connaître un peu plus en détail le profil du corpus, c'est-à-dire, les distributions des unités documentaires qui sont accompagnées d'un texte.



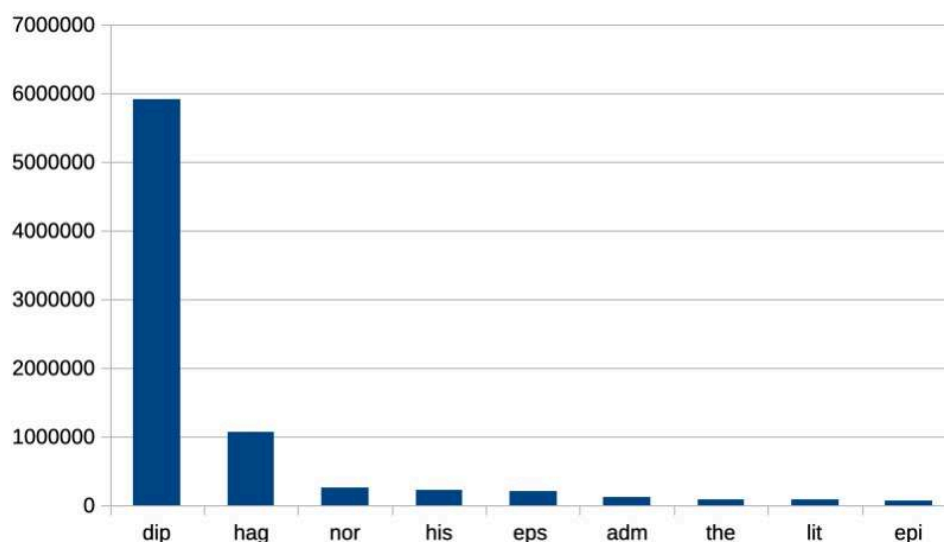
18

Le corpus contient actuellement presque 8 millions de mots (~7 988 704). Sa partition par typologie montre l'importance de la composante diplomatique qui représente plus de 73% du corpus, suivie de l'hagiographie qui représente environ 13% de l'ensemble (Tab. 1 et Graph. 1). On observe ainsi que malgré l'adjonction d'autres types documentaires, le corpus demeure foncièrement un corpus diplomatique, mais cette caractéristique pourrait évoluer avec l'intégration progressive de textes d'autres typologies.

Tab. 1. Répartition typologique du corpus CBMA par nombre de mots

Type	Étiquette	Nombre de mots
<b>Diplomatique</b>	dip	5908816
<b>Hagiographique</b>	hag	1061939
<b>Normatif</b>	nor	251398
<b>Historiographique</b>	his	221427
<b>Épistolaire</b>	eps	202239
<b>Administratif</b>	adm	118192
<b>Théologique</b>	the	82258
<b>Liturgique</b>	lit	80593
<b>Épigraphique</b>	epi	61842

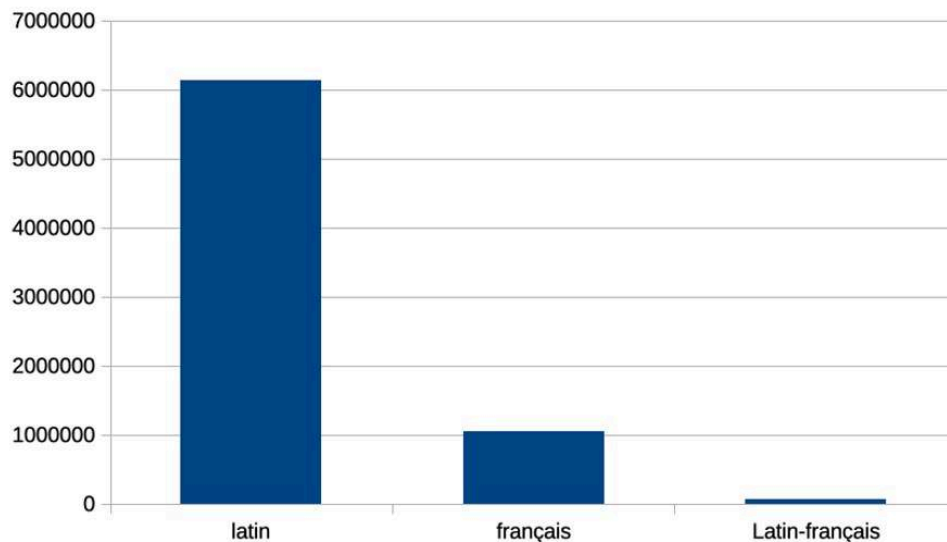
Graph. 1. Répartition typologique du corpus CBMA par nombre de mots



19

Quant aux différentes langues en présence, le latin concerne environ 85% du corpus, le français environ 14%, avec quelques textes mixtes latin-français et des exemples tout à fait isolés de deux chartes en espagnol, ainsi que d'une inscription en grec et de trois en latin-grec qui sont entrées dans le corpus avec les inscriptions épigraphiques (Graph. 2).

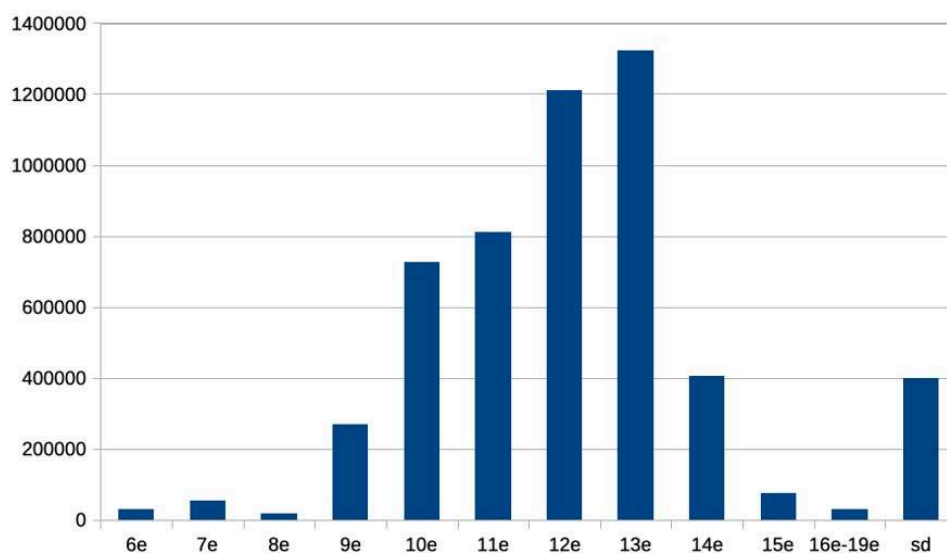
Graph. 2. Répartition linguistique du corpus CBMA par nombre de mots



20

Si on considère le corpus en latin, qui est l'objet premier du programme CBMA, la répartition chronologique par nombre de mots montre que les XII<sup>e</sup> et XIII<sup>e</sup> siècles sont les mieux représentés avec respectivement environ 23% et 25%, correspondant à presque la moitié du corpus latin. Quant aux autres siècles, le IX<sup>e</sup> représente 5%, le X<sup>e</sup> un peu plus de 13%, le XI<sup>e</sup> 15%, le XIV<sup>e</sup> un peu plus de 7%, tout comme les documents latins sans date (Graph. 3).

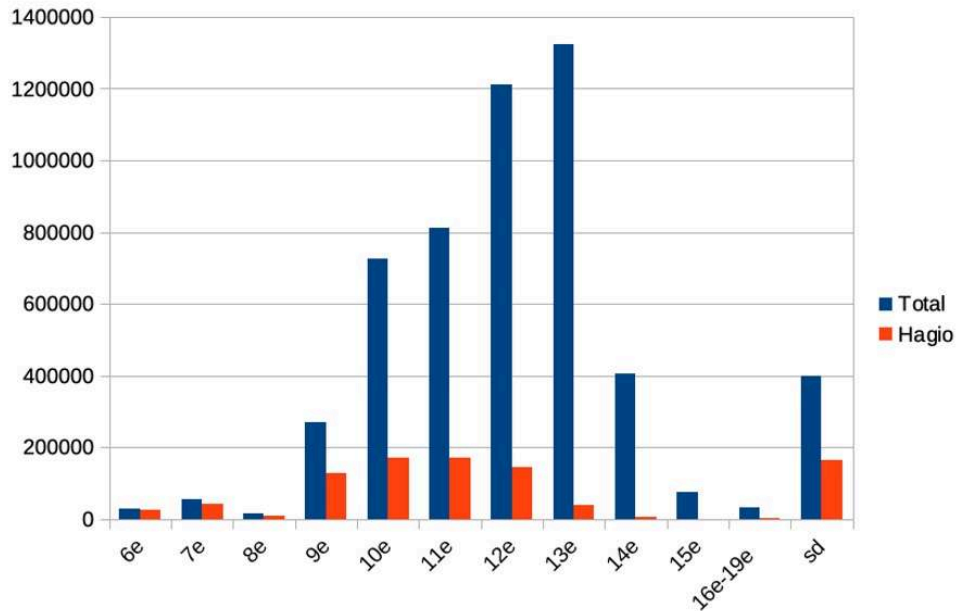
Graph. 3. Répartition chronologique du corpus CBMA latin par nombre de mots



21

Pour les siècles les moins représentés du haut Moyen Âge, la distribution chronologique du sous-corpus hagiographique est intéressante à observer car la composante hagiographique représente presque la totalité du corpus latin pour ce qui concerne les VI<sup>e</sup>, VII<sup>e</sup> et VIII<sup>e</sup> siècles, et un peu moins de la moitié environ du IX<sup>e</sup> siècle. On note aussi l'importance des textes hagiographiques parmi les documents sans date (Graph. 4). Ainsi, si d'un point de vue d'ensemble, le corpus hagiographique ne change pas le caractère éminemment diplomatique du CBMA, du point de vue de la répartition chronologique il exerce un poids considérable sur les siècles du haut Moyen Âge.

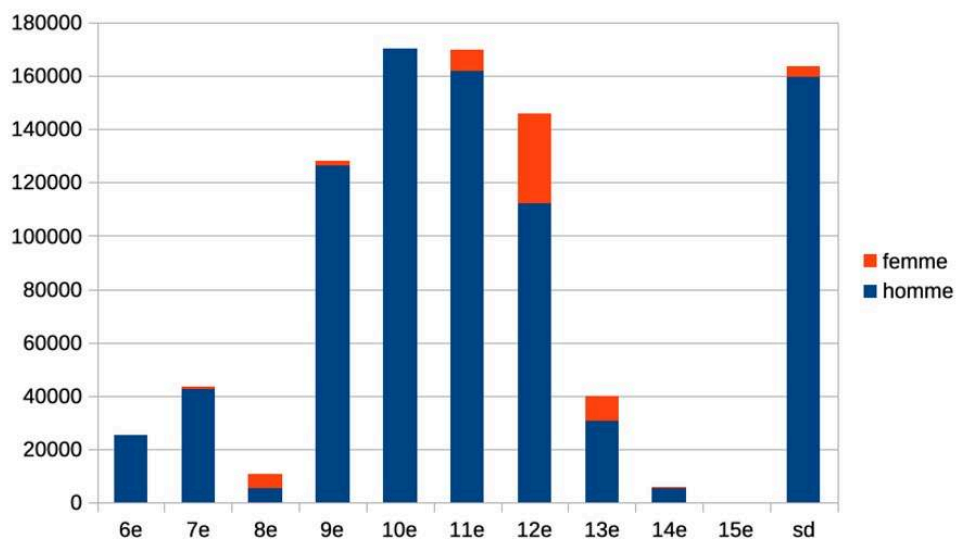
Graph. 4. Le sous-corpus hagiographique au regard de la répartition chronologique du corpus CBMA latin par nombre de mots



22

Si on resserre encore la focale au seul sous-corpus hagiographique latin, d'autres métadonnées propres à ce corpus peuvent être explorées, comme le genre des saints, leur fonction ou encore l'attribution des textes. La répartition par le genre montre l'écrasante prédominance des textes concernant des saints masculins qui couvrent presque 93% du sous-corpus. Quant aux écrits relatifs aux saintes, ils se regroupent pour plus de la moitié au XII<sup>e</sup> siècle (53%), alors que les XI<sup>e</sup> et XIII<sup>e</sup> siècles, représentent respectivement environ 12% et 14% (Graph. 5).

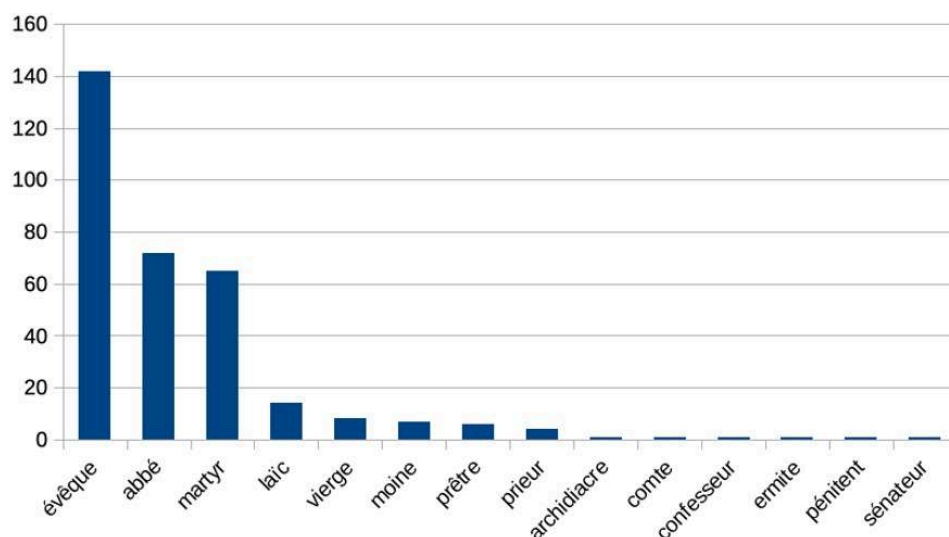
Graph. 5. Répartition par le genre du sous-corpus hagiographique CBMA par nombre de mots



23

Si au lieu du nombre de mots on considère les 325 unités textuelles du sous-corpus hagiographique, par rapport à la fonction principale des saints, la répartition montre, sans surprise, la prédominance des ecclésiastiques, dont les évêques avec 44% des textes et les abbés avec 22%. Les vierges, ici indiquant les femmes, ne représentent que 2,5% (Graph. 6).

Graph. 6. Répartition par fonction du saint du sous-corpus hagiographique CBMA par unité textuelle



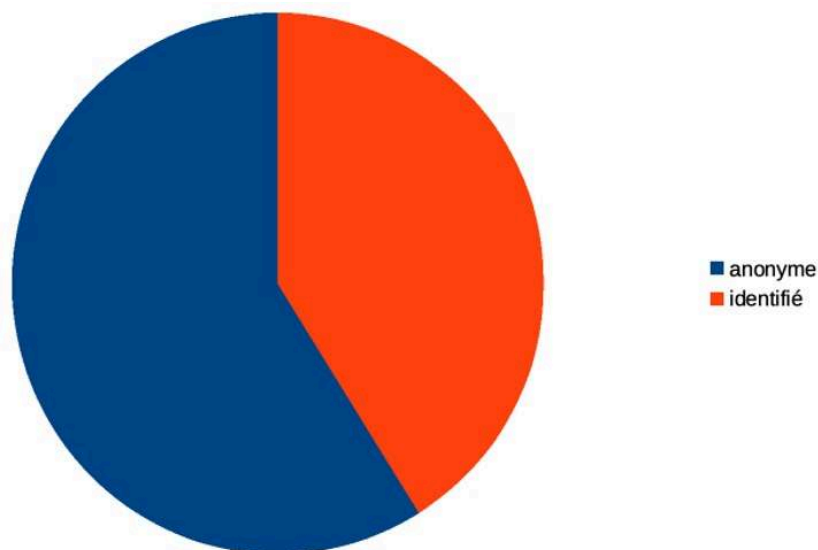
24

Le sous-corpus hagiographique a été renseigné également avec des métadonnées relatives à l'attribution et à localisation des textes, avec la géolocalisation du lieu de rédaction des textes et de l'« institution promotrice », une catégorie introduite à la suite des ateliers de 2018 pour pouvoir associer de façon plus homogène les textes à un lieu parmi souvent les plusieurs possibles. On pourrait rappeler l'exemple emblématique de la Vie d'Odon de Cluny par Jean de Salerne, rédigée en Italie vers 954, par un moine romain, disciple d'Odon, pour laquelle l'introduction de la catégorie d'« institution promotrice », en l'occurrence Cluny, permet de la situer en Bourgogne. En même temps, peu de textes sont aussi clairement datés et attribués que celui-ci.

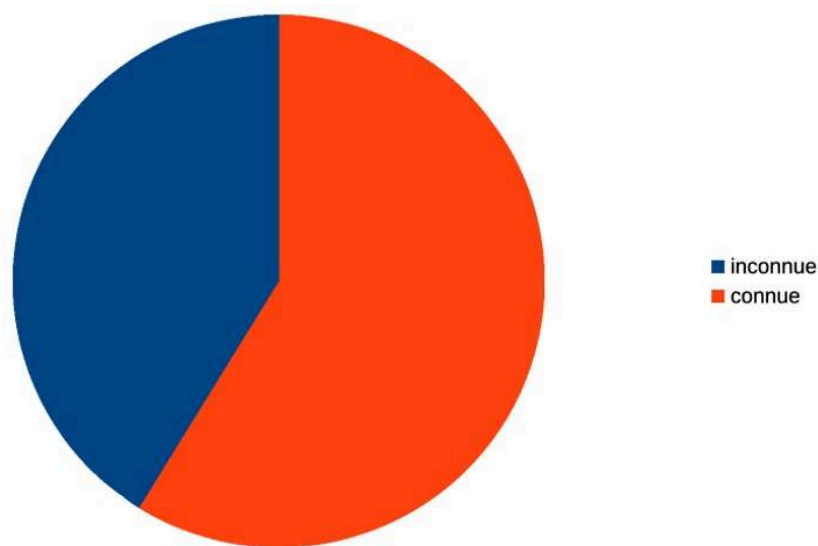
25

De fait, par rapport à leur attribution, la répartition entre les textes anonymes et les textes qui peuvent être rattachée à un « auteur » identifié est respectivement d'environ 60% et 40% (Graph. 7). Cette proportion est inversée quand on considère l'attribution non à un auteur, mais par rapport aux « institutions promotrices » (Graph. 8). Les institutions connues représentent 60% environ des textes et les inconnues 40%, ce qui montre que nous sommes mieux armés pour lier un texte à une institution, à un centre producteur, qu'à un auteur, et donc que cette forme de « localisation et d'attribution » est aussi une manière de relever des milieux d'écriture.

Graph. 7. Répartition des unités textuelles par « auteur » dans le sous-corpus hagiographique CBMA



Graph. 8. Répartition des unités textuelles par « institution promotrice » dans le sous-corpus hagiographique CBMA



26

D'autres métadonnées du corpus attendent encore d'être exploitées, comme la géolocalisation des lieux et les possibilités d'analyse spatiale qu'elle autorise. Cet aperçu général montre toutefois, du moins nous l'espérons, ce que l'utilisateur peut trouver dans CBMA et dans le sous-corpus hagiographique. C'est seulement à l'usage qu'il est possible de se rendre compte des potentialités de recherche offertes par le corpus mais aussi des améliorations nécessaires à apporter. Les impasses et difficultés inhérentes à la manipulation d'un gros corpus textuel nous ramènent toujours à des problèmes proprement historiques, ceux qui agissent sur la fabrication, la transmission et la compréhension des

documents écrits, problèmes face auxquels l'historien est appelé à faire des choix épistémologiques. La manipulation d'un corpus extensif comme le CBMA est ainsi à la fois une démarche à fort potentiel heuristique et une incitation à réfléchir sans cesse à nos pratiques.

## Rappel des communications

**Le corpus hagiographique bourguignon lemmatisé avec Collatinus 12 – Yves Ouvrard (Education nationale)**

27 Voir <https://journals.openedition.org/cem/18482>

**La « tristesse » dans les hagiographies bourguignonnes (CBMA) - Gabriel Castanho (UFRJ - LATHIMM)**

28

« Vous riez aujourd'hui. Pendant mille ans, on n'a pas ri. On a amèrement pleuré », a écrit Michelet en 1862. Sept cents ans plus tôt, Gratien écrivait dans son *Decretum – Sedeat igitur tristis, et officio suo vacet* –, en attribuant la tristesse à un groupe social, les moines. La tristesse est donc un objet historique et historiographique des médiévistes. Parmi les nombreux émotifs (REDDY) associés à la tristesse, il en a un qui mérite l'attention des chercheurs en raison de son caractère social le plus évident : l'empathique *contristare*. Lorsque l'on recherche les usages du terme dans certains *corpus* numériques, les résultats nous invitent à des réflexions plus approfondies. En effet, soit par rapport aux usages du terme, soit par rapport aux effets de source produits par le genre rhétorique des documents et même par les « effets de corpus », l'étude comparée s'impose. Le *Corpus Corporum* révèle l'existence d'une topique discursive centrée sur 5 passages bibliques souvent associés au terme. Si le CBMA diplomatique emploie 4 fois la topique, le CBMA hagiographique présente une seule occurrence. Dans les deux *corpora*, c'est le genre rhétorique du dialogue adopté en milieu monastique (clunisien) qui semble être le fonds commun. Cette communication est une esquisse d'étude à propos de cette absence, ouvrant un nouveau dossier de recherche sur la place de la tristesse dans les hagiographies bourguignonnes.

## Le dossier de Baudile de Nîmes et la Bourgogne – Fernand Peloux (CNRS - FRAMESPA)

### Retour vers le dossier de Jean de Réôme et sa datation à l'égard des analyses sémantiques – Nicolas Perreaux (Université Paris 1 Panthéon-Sorbonne - LaMOP)

---

## NOTES

1. E. MAGNANI, « Les CBMA en corpus structuré. Atelier 2. Le corpus hagiographique bourguignon. Débats et recherches », *Bulletin du centre d'études médiévales d'Auxerre | BUCEMA*, Collection CBMA, Les journées d'études, mis en ligne le 28 juillet 2018, consulté le 07 décembre 2021, <http://journals.openedition.org/cem/15493> ; E. MAGNANI, E. INGRAND-VARENNE, « Le corpus épigraphique bourguignon (VIII<sup>e</sup>-XV<sup>e</sup> siècle). Des catalogues aux applications numériques », *Bulletin du centre d'études médiévales d'Auxerre | BUCEMA*, Collection CBMA, Les journées d'études, mis en ligne le 15 novembre 2018, consulté le 07 décembre 2021, <http://journals.openedition.org/cem/15591>.

2. Programmée le 19 décembre 2019, la dernière rencontre CBMA a été reportée en raison des grèves des transports et de l'impossibilité des intervenants de se rendre à Paris. Reprogrammée pour le 19 mars 2020, elle est tombée dans les premiers jours sidérants du « premier confinement », et les différentes présentations ont été mises à disposition en ligne : *Des applications aux manuscrits. Expériences de transcriptions automatiques de manuscrits et développements du Corpus Burgundiae Medii Aevi* (<http://www.cbma-project.eu/publications/les-rencontres-cbma.html>).

3. Du point de vue historiographique, nous disposons également des panoramas importants jusqu'aux années 1030, parus dans la collection *Hagiographies : Histoire internationale de la littérature hagiographique latine et vernaculaire en Occident des origines à 1550*, vol. 1, 1994 (contribution de Dominique Iogna-Prat) et 8, 2020 (contributions de Michèle Gaillard, Alain Rauwel, Anne-Marie Verleysen).

4. Les 14 mars, 13 juin et 24 octobre 2022.

5. Le résumé ou le lien vers une notice individuelle de ces présentations seront ajoutés dans ce compte-rendu au fur et à mesure de leur réception. On trouvera le programme de la journée ici : <https://lamop.hypotheses.org/7261>

6. E. MAGNANI, « Des chartae au corpus : la plateforme des CBMA - *Chartae/Corpus Burgundiae Medii Aevi* », in *Digitizing Medieval Sources. Challenges and Methodologies / L'édition en ligne de documents d'archives médiévaux. Enjeux, méthodologie et défis*, éd. Christelle Balouzat-Loubet, Turnhout, 2019 (Atelier de Recherches sur les Textes Médiévaux, 27), p. 57-67 - <https://doi.org/10.1484/M.ARTEM-EB.5.117328> ;

7. M. MARRIER, A. Duchesne, *Bibliotheca Cluniacensis*, Paris, 1614 ; J. MARILIER, *Chartes et documents concernant l'abbaye de Cîteaux (1098-1182)*, Rome, 1961.

8. E. MAGNANI, « Un corpus structuré et hétérogène de textes latins médiévaux (Bourgogne, V<sup>e</sup>-XV<sup>e</sup> siècle) », *Bulletin du CERCOR - Centre Européen de recherches sur les congrégations et ordres religieux*, 41, 2017, p. 59-65.

9. Ces projets sont portés par le LaMOP, en collaboration avec le CESC de Poitiers pour le volet épigraphique, et ont été soutenus par le Labex Hastec (en collaboration avec l'IRHT, l'EnC - Centre Jean Mabillon, le CERCOR-LEM) et par le Consortium Sources Médiévales - Cosme2 de la



TGIR Huma-Num (CNRS). Ont participé à ces projets, entre 2017 et 2019 : Mathieu Beaud (Univ. Paris 1 – LaMOP), Pierre Brochard (CNRS – LaMOP), Hélène Caillaud (Univ. Limoges), Davide Gherdevich (UVSQ - DYPAC), Estelle Ingrand-Varenne (CNRS – CESC), Eliana Magnani (CNRS – LaMOP), Aurore Menudier (CESC), Nicolas Perreaux (Univ. Paris 1 - LaMOP), Coraline Rey (Univ. Bourgogne - Univ. Angers).

10. E. MAGNANI, E. INGRAND-VARENNE, « Le corpus épigraphique bourguignon (VIII<sup>e</sup>-XV<sup>e</sup> siècle) », *op. cit.* ; E. MAGNANI, N. PERREAUX, « A Medieval Epigraphic Corpus and its Retro-Developments (CIFM-CBMA). The Exploratory Research of the COSME2 Consortium », *Digital Scholarship in the Humanities*, 36, Suppl. 2 (2021), p. ii189–ii197, <https://doi.org/10.1093/llc/fqaa069>.

11. E. PETIT, *Histoire des Ducs de Bourgogne de la race capétienne, avec des documents inédits et des pièces justificatives*, Paris, 9 tomes, 1885-1905.

12. J. LEBEUF, *Mémoires concernant l'histoire civile et ecclésiastique d'Auxerre et de son ancien diocèse*, tome 4, Auxerre, Paris, 1855.

13. N. PERREAUX, « La découverte d'une édition de cartulaire oubliée : le fonds documentaire du LaMOP et le cas des dominicains de Dijon », *Les carnets du LaMOP*, 2020, <https://lamop.hypotheses.org/6828>.

14. « CBMA – *Corpus Burgundiae Medii Aevi*. Nouveautés mai-octobre 2021 », *Les carnets du LaMOP*, 2021, <https://lamop.hypotheses.org/7267>.

15. <https://manuscrits.cbma-project.eu/>

16. <http://www.cbma-project.eu/manuscrits.html>

17. [http://www.cbma-project.eu/editions/textes\\_epigraphiques/corpus-epigraphique-provençal.html](http://www.cbma-project.eu/editions/textes_epigraphiques/corpus-epigraphique-provençal.html). Voir A. MENUDIER, « Le corpus épigraphique provençal : premier bilan et comparaison avec le corpus bourguignon », *Bulletin du centre d'études médiévales d'Auxerre | BUCEMA*, Collection CBMA, Les journées d'études, mis en ligne le 19 mai 2020, consulté le 08 octobre 2021, <http://journals.openedition.org/cem/17076>

18. Ce travail est réalisé par l'équipe du projet : Pierre Brochard, Gabriel Castanho, Eliana Magnani, Nicolas Perreaux, Coraline Rey, ainsi que Alain Guerreau.

---

AUTEUR

**ELIANA MAGNANI**

CNRS, LaMOP