

17. hiver 2025

éphéméride

Le magazine de l'École Pratique des Hautes Études – PSL

DOSSIER SPÉCIAL

Humanités : quand l'IA transcrit les textes anciens.



www.ephe.psl.eu



École Pratique
des Hautes Études



Éditorial du Président p.3

Les actualités de l'EPHE – PSL p.4

Les actualités des instituts p.6

Le Fonds de dotation de l'EPHE :
de l'ambition à l'impact concret p.10



Dossier spécial :
Humanités : quand l'IA transcrit les textes anciens. p.16

Introduction : eScriptorium p.18

eScriptorium : une aventure collective au croisement
des humanités et de l'intelligence artificielle. p.20

Portrait : Robin Tissot, le concepteur d'eScriptorium p.22

Intelligence artificielle et Sciences des données pour
les Humanités dans la stratégie de l'Université PSL. p.25

Portrait : Benjamin Kiessler, l'architecte IA de l'EPHE – PSL. . . p.26

Biblistima : observatoire des cultures écrites anciennes . . . p.28

Dunhuang : déjouer la loi de puissance p.30

MiDRASH : reconstruire la culture juive médiévale écrite. . . p.32

CATMuS : lire les langues de l'Europe au Moyen-Âge p.34

Syriac Transcribathon : un exercice collaboratif
pour décrypter des textes syriaques p.35

L'initiative Open Islamicate Texts (OpenITI) :
étudier les textes arabes p.36

ERC PRIMA : les manuscrits italiens du XVI^e au XVIII^e siècle . . p.37

DeLiCaTe : déchiffrer le patrimoine écrit du Caucase ancien . . p.38

Conclusion : « Aussi ouvert que possible, aussi fermé
que nécessaire » p.39

Lexique p.40

Nos publications p.42

Sur le terrain p.44

Éphéméride

Le magazine de l'École Pratique
des Hautes Études – PSL
Numéro 17
hiver 2025
ISSN 2679 - 7631

**Directeur
de la publication**
Michel Hochmann

Rédacteur en chef
Fabrice Auffret van der Kemp

Création graphique
Jean-Baptiste Lasserre

Remerciements :
Colin Brisson
Marc Bui
Thibault Clerice
Sébastien Couette
Serena Crespi
Jost Gippert
Benjamin Keissling
George Kiraz
Mathew Miller
Catherine Paul
Elena Pierazzo
Stéphanie Plenchette
Christine Roughan
Sarah Savant
Peter Stokes
Daniel Stökl Ben Ezra
Robin Tissot
Anne-Marie Turcan-Verkerk
Stéphane Verger

Couverture
Weißenauer Passionale,
Coligny,
Fondation Martin Bodmer,
Cod. Bodmer 127

Impression
Paragon

**École Pratique des
Hautes Études – PSL**
4 rue Ferrus 75014 Paris
www.ephe.psl.eu

L'École Pratique des Hautes Études est, depuis ses origines, un centre majeur pour ce qu'on appelle les sciences de l'érudition. On a pu considérer, à un certain moment, que ces disciplines avaient été largement remises en cause par les nouvelles approches des sciences sociales. Or, le développement spectaculaire des humanités numériques et, en particulier, de l'intelligence artificielle ont, ces dernières années, complètement bouleversé ce paysage et démontré le dynamisme et la validité des spécialités qui ont toujours été les nôtres, le déchiffrement des langues et des écritures rares ou la constitution de corpus, pour prendre quelques exemples. L'École a su développer des programmes originaux dans ce domaine et a joué un rôle pionnier. Comme l'indique Marc Bui, il faut rappeler que Denis Pelletier et Hubert Bost, deux de mes prédécesseurs, ont su fédérer les forces et lancer des initiatives fondamentales pour le développement de ces recherches. Je pense, en particulier, à la création de l'Initiative de recherches interdisciplinaires et stratégiques Scripta, développée au sein de notre Université PSL, qui a permis la naissance de la plateforme eScriptorium, dont le succès international se trouve illustré par les multiples programmes de recherche auxquels elle a contribué dans les domaines les plus divers, des manuscrits hébreux aux manuscrits médiévaux, à Princeton, à l'Université du Maryland, et, évidemment, à l'EPHE. Grâce en partie à eScriptorium, nous avons pu aussi obtenir de grands contrats de recherche européens, comme l'ERC Synergy MiDRASH.

Au cours des mêmes années et grâce à l'action d'Anne-Marie Turcan, c'est aussi au sein de l'École qu'a été créé l'Equipex (équipement d'excellence) Bibliissima, qui structure des millions de données relatives aux textes anciens et qui va sans doute connaître bientôt une ultime métamorphose en devenant une infrastructure de recherche, l'une des très rares dans le domaine des sciences humaines et sociales. La prise en compte de l'importance du développement des humanités numériques a abouti à la création d'une direction d'études spécifique dans ce domaine, qu'occupe Peter Stokes depuis sa création, et, lui-même et ses collègues Daniel Stökl Ben Ezra et Marc Bui ont joué un rôle capital dans le développement d'eScriptorium. Ils ont pu s'appuyer sur des collaborateurs d'exception, Benjamin Kiessling et Robin Tissot, dont vous découvrirez les portraits dans ce numéro d'Éphéméride. La recherche, à l'École, est toujours étroitement liée à la formation, et nous avons participé à la mise en place du master Humanités numériques porté par l'École nationale des Chartes au sein de PSL. Ce croisement entre des disciplines rares et des technologies d'avant-garde démontre la fécondité de nos approches et la capacité d'innovation de nos directeurs d'études et maîtres de conférences, et je m'en réjouis. Nous travaillons à présent pour que ces avancées trouvent pleinement leur place dans PSL et, notamment, au sein de l'institut Prairie qui regroupe les forces de notre université dans le domaine de l'intelligence artificielle.



Michel Hochmann,
Président
de l'EPHE – PSL

Je suis en tout cas très reconnaissant à tous ceux qui ont contribué à l'élaboration de ce dossier thématique qui témoigne, encore une fois, du dynamisme de notre École. Ce numéro d'Éphéméride donne d'ailleurs bien d'autres témoignages de ce dynamisme, avec le développement du mécénat, les actualités de la recherche (la découverte opérée dans le cadre de la mission archéologique de Tanis, les succès remportés à l'ERC et à l'ANR), les publications. Je regarde donc l'avenir avec confiance (malgré les difficultés budgétaires que nous traversons comme la plupart des établissements d'enseignement supérieur), et j'adresse mes vœux les plus chaleureux à notre École et à tous ceux qui la font vivre, enseignants-chercheurs, personnels administratifs, étudiants et auditeurs.

Michel Hochmann



École Pratique
des Hautes Études

PSL

édito

DeLiCaTe : déchiffrer le patrimoine écrit du Caucase ancien

Jost GIPPERT (Université de Hambourg)

Les littératures arménienne et géorgienne ont connu un développement continu depuis l'apparition de leurs alphabets au V^e siècle ap. J-C. Pendant longtemps, ces littératures ont été transmises par des manuscrits, dont beaucoup sont conservés.

Cet immense patrimoine culturel n'a été exploré que de manière partielle et beaucoup d'œuvres sont encore inédites. Afin d'accélérer ce nécessaire travail d'édition, le projet DeLiCaTe (à Hambourg) a commencé en février 2024 à développer

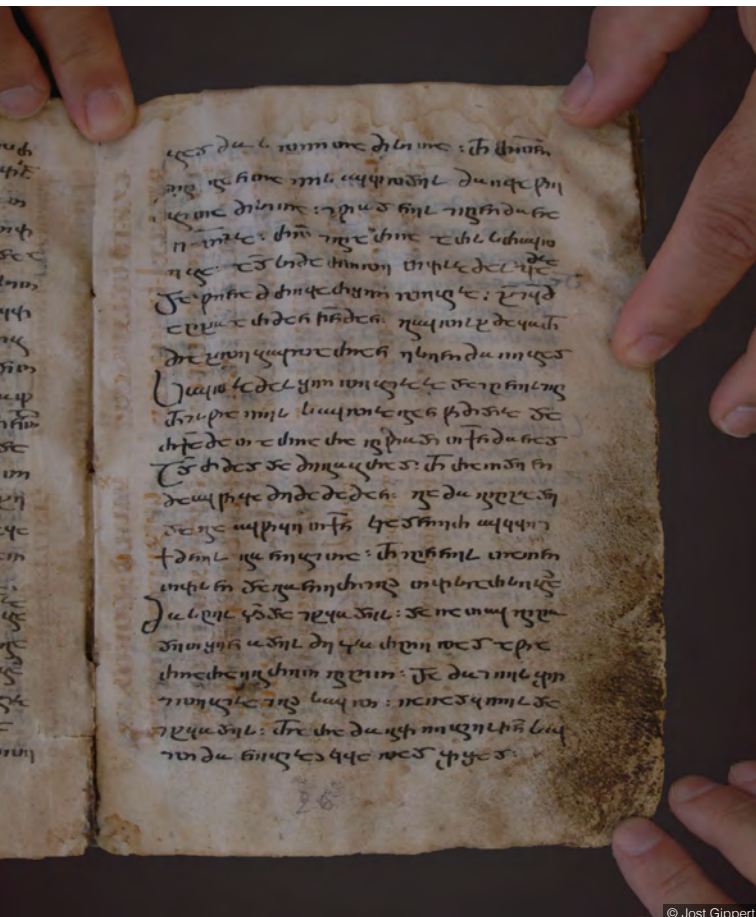
des modèles eScriptorium pour la lecture automatique des manuscrits. Sur la base des modèles préexistants pour le latin et l'hébreu, trois adaptations différentes ont été réalisées : en géorgien, pour les majuscules (taux de précision actuel : 93,8 %

et pour les manuscrits mixtes (88,7%) et en arménien (90,2 %). Ces développements ont été grandement facilités par le caractère intuitif du processus et de l'interface d'eScriptorium et par le soutien continu de Daniel Stökl Ben Ezra.

La base de développement comprenait les images de 296 pages de manuscrits qui avaient été transcrites manuellement et 118 transcriptions automatiques alignées sur des éditions existantes pour le géorgien, et 36 images transcrites manuellement pour l'arménien.

Malgré des problèmes particuliers, notamment des défauts de lisibilité dus à des dommages sur les manuscrits originaux et la qualité inégale des images disponibles, 18 554 pages géorgiennes et 6 189 pages arméniennes ont été lues avec succès au cours des quinze derniers mois.

Depuis avril 2025, nous avons installé une instance locale au CHAI de Hambourg, ce qui attire de plus en plus d'utilisateurs.



Ms. Sin. georg.
NF 13, fol. 4v:
Géorgien avec
sous-titres
albanais

© Jost Gippert