

JOURNÉE D'ÉTUDE

# LA RECONNAISSANCE DES ÉCRITURES MANUSCRITES ET SES USAGES DANS LES ARCHIVES

MARDI 29 NOVEMBRE 2022

LÉGION D'HONNEUR

NUMÉRO D'ORDRE  
DES MATRICULES

26.612

Mod. N° 470

Nom

Héroult

Prénoms

Jean Auguste

Qualité

ou

grade

Membre du Conseil Municipal et  
Pressure (2 Lignes)

né le

25 Mars 1797

à

Niort (Deux Lignes)

# LA RECONNAISSANCE DES ÉCRITURES MANUSCRITES ET SES USAGES DANS LES ARCHIVES

MARDI 29 NOVEMBRE 2022  
# HTRArchives

La reconnaissance par ordinateur des écritures manuscrites (ou HTR – *Handwritten Text Recognition*) est une technologie qui s'appuie sur l'intelligence artificielle afin de transcrire automatiquement des documents : leur contenu est transformé en données prêtes à être exploitées pour des usages numériques.

Son application aux archives patrimoniales s'est considérablement développée ces dernières années. Aux Archives nationales, c'est à partir de 2015 que la technologie a commencé à être éprouvée, avec le projet HIMANIS autour des registres médiévaux de la chancellerie royale. D'autres initiatives y ont vu le jour depuis, pour traiter différentes typologies d'archives : LectAuRep (répertoires des notaires parisiens), eNDP (registres capitulaires de Notre-Dame de Paris) et SIMARA (rétroconversion des instruments de recherche manuscrits anciens). Au sein du réseau des archives, plusieurs projets sont apparus concomitamment : « Lettres en Lumières » dans les Archives de la Côte-d'Or (registres des États de Bourgogne du XVIII<sup>e</sup> siècle), POPP et SOCFACE (recensements de population conservés aux Archives départementales),

« TabelNorm » en Seine-Maritime (registres de tabellions du XV<sup>e</sup> siècle), RegistrIA puis ArchivIA au Service historique de la Défense.

Ce succès démontre que la technologie a fait ses preuves et qu'elle peut s'appliquer à des fonds allant du Moyen Âge à l'époque contemporaine. Dans ce contexte, les Archives nationales organisent une journée d'étude qui mettra en exergue des projets conduits dans le réseau des archives, afin d'illustrer les exploitations que l'on peut faire des données issues de l'HTR : moteurs de recherche dans les contenus textuels, indexation et reconnaissance d'entités nommées, récupération et restructuration des informations, pour enrichir les instruments de recherche, réutilisation des informations à des fins de recherche scientifique.

Ces retours d'expérience mettront également l'accent sur les liens nécessaires qui doivent se nouer entre les services d'archives et différents partenaires : structures de recherche en sciences humaines et sociales, en humanités numériques, en informatique, partenaires commerciaux qui livreront eux aussi leur point de vue.



JE M'INSCRIS  
<https://www.archives-nationales.culture.gouv.fr/journees-d-etude-colloques-et-conferences>

9 h 30 **Accueil et introduction à la journée d'étude**  
par Bruno Ricard, directeur des Archives nationales

9 h 40 **Qu'est-ce que l'HTR ?**  
par Christopher Kermorvant

10 h **Connaître les projets, les acteurs, les enjeux de l'HTR**  
par Jean-François Moufflet et Aurélia Rostaing

10 h 30 Échanges avec le public ; pause

**L'HTR au service de l'enrichissement des métadonnées archivistiques**  
Présidence : Florence Clavaud

11 h 00 **SIMARA : faciliter la rétroconversion des inventaires anciens (xviii<sup>e</sup>-xx<sup>e</sup> siècle)**  
par Jean-François Moufflet et Christopher Kermorvant

11 h 30 **LectAuRep : un projet de recherche et développement pour la transcription automatique de répertoires de notaires**  
par Aurélia Rostaing et Hugo Scheithauer

12 h **Les projets RegistrIA et ArchivIA du Service historique de la Défense : transcription, fouille de texte et enrichissement de l'indexation par l'intelligence artificielle**  
par François Touko Nicheu et Caroline Méot

12 h 30 Échanges avec le public

13 h Pause déjeuner

**L'HTR et la collecte de données nominatives**  
Présidence : Jean-Charles Bédague

14 h **La transformation des recensements de population en données démographiques : le projet SOCFACE**  
par Christopher Kermorvant, Lionel Kesztenbaum, Manonmani Restif et Mathieu Stoll

14 h 40 Échanges avec le public ; pause

**Les projets d'HTR et les liens avec les structures de recherche**  
Présidence : Dominique Stutzmann

15 h **Les registres des États de Bourgogne à la lumière de l'intelligence artificielle**  
par Édouard Bouyé et Florian Fizaine

15 h 30 **Faciliter l'accès aux registres du tabellionage normand : le projet TabelNorm**  
par Marie Grout et Pierrick Tranouez

16 h **Les registres capitulaires de Notre-Dame de Paris : une nouvelle forme d'exploitation par les historiens des archives**  
par Isabelle Bretthauer, Julie Claustre et Sergio Torres Aguilar

16 h 30 Échanges avec le public ; conclusions

## Intervenants

**Jean-Charles Bédague**, sous-directeur du pilotage, de la communication et de la valorisation des archives, Service interministériel des Archives de France

**Édouard Bouyé**, directeur des Archives départementales de la Côte-d'Or

**Isabelle Brethauer**, département du Moyen Âge et de l'Ancien Régime, Archives nationales

**Julie Claustre**, maîtresse de conférences en histoire médiévale, université de Paris 1 Panthéon-Sorbonne

**Florence Clavaud**, responsable du LAB des Archives nationales

**Florian Fizaine**, doctorant, laboratoire d'étude de l'apprentissage et du développement [LEAD], université de Bourgogne

**Marie Groult**, pôle des archives historiques, Archives départementales de Seine-Maritime

**Christopher Kermorvant**, ingénieur et docteur en informatique, expert en intelligence artificielle, société Teklia

**Lionel Kesztenbaum**, directeur de recherche à l'Institut national d'études démographique [INED]

**Caroline Méot**, chargée de médiation et de valorisation numérique, Service historique de la Défense

**Jean-François Moufflet**, département du Moyen Âge et de l'Ancien Régime, Archives nationales

**Manonmani Restif**, cheffe de projet du portail FranceArchives, Service interministériel des Archives de France

**Bruno Ricard**, directeur des Archives nationales  
**Aurélia Rostaing**, département du Minutier central des notaires de Paris, Archives nationales  
**Hugo Scheithauer**, ingénieur recherche et développement, Institut national de recherche en informatique et en automatique [INRIA]

**Mathieu Stoll**, chef du bureau de la diffusion et de la valorisation numérique, Service interministériel des Archives de France

**Dominique Stutzmann**, directeur de recherche au CNRS (Institut de recherche et d'histoire des textes)

**François Touko Nicheu**, chef du bureau de la transformation numérique, Service historique de la Défense

**Sergio Torres Aguilar**, post-doctorant, École nationale des chartes, Centre Jean-Mabillon

**Pierrick Tranouez**, ingénieur de recherche au Laboratoire d'informatique, de traitement de l'Information et des systèmes [LITIS], université de Rouen

*Avec le soutien du chapitre francophone de la communauté AI4lam (Artificial Intelligence for Libraries, Archives, Museums)*



**MINISTÈRE  
DE LA CULTURE**

*Liberté  
Égalité  
Fraternité*

**ARCHIVES  
NATIONALES**

### Archives nationales

Site de Pierrefitte-sur-Seine

59 rue Guynemer

93380 Pierrefitte-sur-Seine

Métro : Saint-Denis–Université (ligne 13)

Parking Q-Park Saint-Denis–Université

rue Toussaint-Louverture

*Entrée libre dans la limite des places disponibles*

[www.archives-nationales.culture.gouv.fr](http://www.archives-nationales.culture.gouv.fr)

