



# Journée Humanités Numériques & IA

(Evénement affilié à PFIA 2025)  
vendredi 4 juillet, 2025, Dijon, France

# Journée HN-IA

Objectif : faire un point sur les usages de l'**intelligence artificielle** dans le domaine des **humanités numériques**.

Suite de la journée organisée en mai 2024 sur les défis posés par les humanités numériques aux approches développées en intelligence artificielle.

**Thématique spécifique** : le lien entre les sources de connaissances mobilisées par les chercheurs en sciences humaines et sociales et les données produites, manuellement ou automatiquement.

## Organisation

- CP composé de 14 chercheurs en sciences du numérique et en SHS
- 9 soumissions - 5 acceptés
- Double relecture

**La journée est portée par l'action SAD-2HN de MADICS et soutenue par l'AFIA**

# Action SAD-2HN

## Des Sources Aux Données Historiques en Humanités Numériques

Les **humanités numériques** constituent un domaine à l'**interface** des arts, lettres, sciences humaines et sociales et des sciences du numérique.

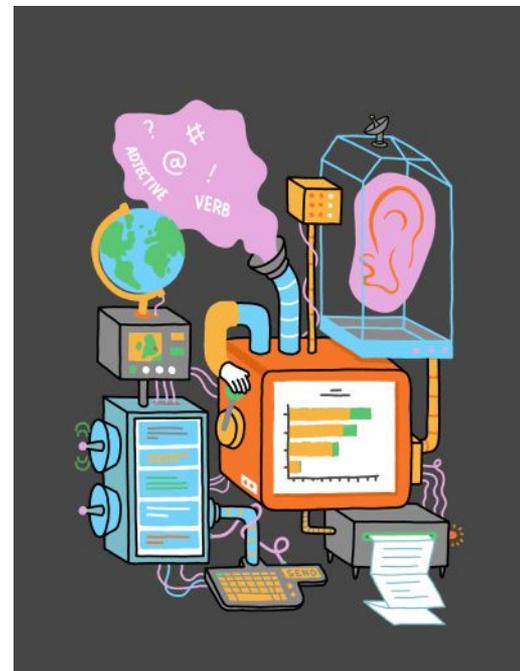
Forte interdisciplinarité :

- les sciences du numérique permettent de nouvelles perspectives et changent les pratiques en SHS
- les approches et modèles utilisés en sciences du numérique doivent être adaptés

Une **donnée historique**

- constitue le marqueur d'une **réalité passée** reconstruite (à partir de témoignages, de sources primaires...)
- est **interprétée** une fois extraite (manuellement ou automatiquement) par des chercheurs **pour créer des faits** qui servent à soutenir l'**analyse et la compréhension** de cette réalité ancienne.

L'action SaD-H2N s'intéresse à trois défis qui concernent ce type de données.



# Trois défis

- 1. Caractériser les données historiques en HN** et les domaines concernés par les dimensions
  - temporelle
  - incertaine
  - incomplète
  - évolutive
- 2. Étudier le lien entre données et sources**
  - la donnée extraite relève d'une interprétation de l'expert
  - le lien est indispensable pour que la donnée soit exploitable par les experts
- 3. Analyser les données comme support de l'étude critique des sources**
  - aller au delà de la production et du partage de données
  - proposer des approches pour analyser les comportements et pratiques dont les données sont des marqueurs

**Méthode** : animer la communauté sur ces défis, identifier les verrous scientifiques et dresser des pistes de recherche et de collaborations.

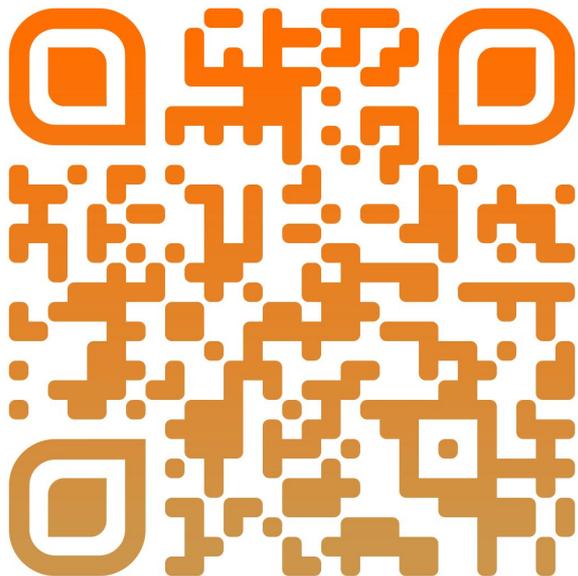
## Quelques exemples...

- données **structurées** produites par les historiens, les géographes, les responsables d'inventaires d'archives, de musées, etc.
- **documents** anciens, textuels ou iconographiques,
- **référentiels** interopérables : thésaurus, graphes de connaissances, etc.

# Pour suivre & participer à l'action

 [sadhnmadics.github.io](https://sadhnmadics.github.io)

 [sadhnmadics@pub.univ-eiffel.fr](mailto:sadhnmadics@pub.univ-eiffel.fr)



Pour s'inscrire :

**À** [sympa@pub.univ-eiffel.fr](mailto:sympa@pub.univ-eiffel.fr)

**Objet** subscribe sadhnmadics {Prénom} {Nom}

Ou : <https://sadhnmadics.github.io/about>

# Programme

## Session IC-HNIA – Ingénierie des connaissances pour les humanités numériques

10h30-11h00 : **PeGazUs : une méthode de reconstitution de l'évolution des entités géographiques à partir de données hétérogènes et fragmentaires.** Charly Bernard, Nathalie Abadie, Bertrand Duménieu et Julien Perret.

11h00-11h20 : **Développement d'une ontologie pour l'analyse de textes de décisions administratives d'Ancien Régime par des Grands Modèles de Langage.** Gilles Falquet, Christophe Chazalon, Marco Sorbi, Stéphane Marchand-Maillet et Laurent Moccozet.

11h20-11h50 : **Extraction et interprétation sémantique de tables anciennes : défis et perspectives.** Solenn Tual, Nathalie Abadie, Joseph Chazalon, Bertrand Duménieu et Julien Perret.

# Programme

## Session HNIA - APIA – Des sources aux données: enjeux de traçabilité des données pour les humanités numériques

14h00-14h20 : **Repenser les collections patrimoniales par le prisme de l'IA**. Marion Charpier et Emmanuelle Bermès.

14h20-14h40 : **Amélioration de la segmentation de scènes par l'exploitation de relations spatiales : application aux sceaux byzantins**. Ege Şendoğan, Victoria Eyharabide et Isabelle Bloch.

14h40-15h00 : **Harold: un système de requête itératif et interactif pour l'exploration de corpus de patrimoine culturel**. Prunelle Daudré-Treuil, Olivier Bruneau, Jean Lieber, Emmanuel Nauer et Laurent Rollet.

15h00-15h20 : **Quand les humanités numériques rencontrent l'IA : une enquête sur la propriété à Paris**. Aaron Parmentelat.

15h20-15h40 : **Évaluation automatique du retour à la source dans un contexte historique long et bruité : les débats parlementaires de la Troisième République française**. Aurélien Pellet, Julien Perez et Marie Puren.

15h40-16h00 : Synthèse des présentations autour des enjeux de traçabilité des données pour les Humanités numériques.

# Synthèse

Grande diversité des approches pour représenter le lien entre sources et données :

## **Dans les données elles-mêmes:**

- à l'aide de graphes de connaissances de factoïdes [pres. Charly Bernard]: représentation proche de la notion de *claims* décrite dans la présentation invitée;
- à l'aide d'ontologies contextualisées et d'annotations des triplets en RDF\* [pres. Marco Sorbi];
- en utilisant l'analyse formelle de concepts pour regrouper et annoter itérativement les sources [pres. Emmanuel Nauer].

## **Dans les métadonnées des images sources:**

- à l'aide d'annotations sémantiques des portions d'images transcrites, directement dans les manifestes IIIF [pres. Solenn Tual].

## **Sous forme implicite:**

- pas de représentation explicite formelle mais un mode d'accès au contenu des sources avec un agent intermédiaire, eg. RAG [pres. Aurélien Pellet]

# A suivre en 2025-2026

- 1) Retrouvez les résumés de la journée thématique **HNIA 2025 sur HAL**.
- 2) Une **journée thématique à l'automne 2025** : “Repenser les données historiques à l'ère des grands modèles de langues” - avec le GdR MAGIS.
  - (1)  *Public visé : acteurs des humanités numériques (GLAM, chercheur, entreprises, acteurs VGI), échange direct entre les IR (Human-Num et Progedo)*
- 3) **À partir de l'automne 2025, un webinaire bimestriel** organisé conjointement avec le GdR MAGIS et éventuellement le GDR Internet IA et Société sur des thématiques identifiées au fil de l'eau
- 4) Un **recensement des ressources de formation** à la collecte, au traitement, à l'analyse et à la diffusion des données historiques **à destination des jeunes chercheurs en humanités numériques**
- 5) Une **journée thématique HNIA** organisée dans le cadre de la conférence **humanisti.ca 2026** (Paris) ?

