

# ISIDORE 2030 : un programme de recherche et de développement pour ISIDORE

Huma-Num Lab

Stéphane Pouyllau, ingénieur de recherche

Adam Faci, **post-doctorant**

Léa Maronet, **doctorante**

Émile Provendier, **post-doctorant**

Web : <https://www.huma-num.fr/hnlab>

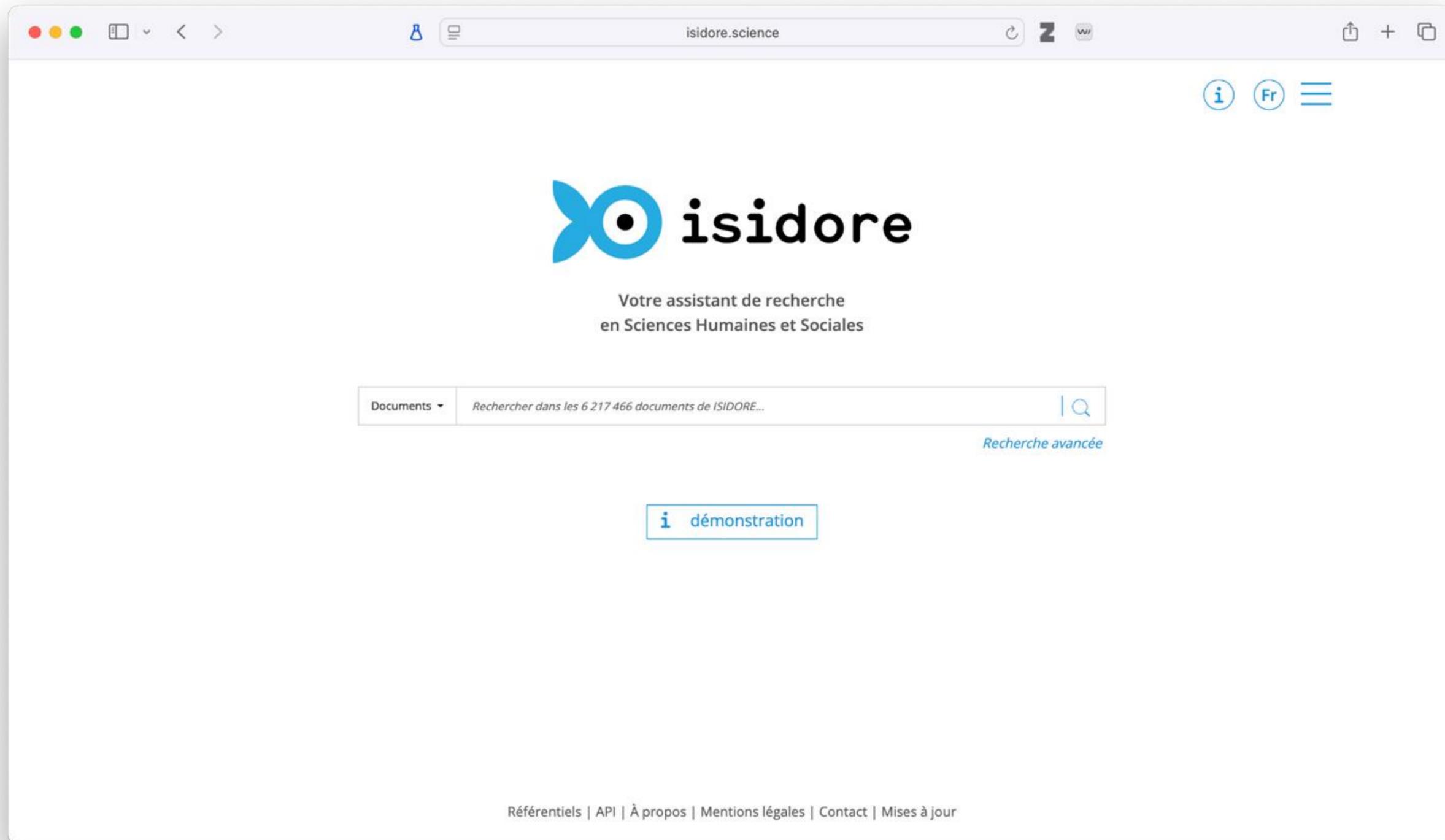
Blog : <https://hnlab.huma-num.fr/blog>

**Rencontres 20**  
**Huma-Num 25**  
3-4-5 juin  
Marseille

# Introduction

---

“Et vous allez mettre de l'IA dans Isidore ?”



“Bien sûr, c’est si facile...”

The screenshot shows a web browser window with the URL `my.fluidtopics.net`. The page features a dark navigation bar with links for Administration, Analytics, Ma Bibliothèque, and Mon Compte. The main content area displays the ISIDORE logo and the tagline "Votre assistant de recherche en sciences humaines et sociales". A search bar is present with a language dropdown set to "Français" and a placeholder text "Tapez quelque chose à chercher...". Below the search bar, there is a "NOUVEAU CHAT" section with a chat bubble containing the message "Bienvenue ! Comment puis-je vous aider aujourd'hui ?". At the bottom, there is a text input field labeled "Écrire un message".

Administration Analytics Ma Bibliothèque Mon Compte

 isidore

Votre assistant de recherche en sciences humaines et sociales

Français

NOUVEAU CHAT

Bienvenue ! Comment puis-je vous aider aujourd'hui ?

Écrire un message

**Bienvenue**

Votre assistant de recherche est disponible pour toutes demandes. Depuis cette page vous pouvez : lancer vos recherches, demander des états de l'arts, des synthèses, des traduction d'articles, etc.

A tout moment, vous pouvez changer de modèle d'IA pour affiner vos recherches.

→ EXPLORE

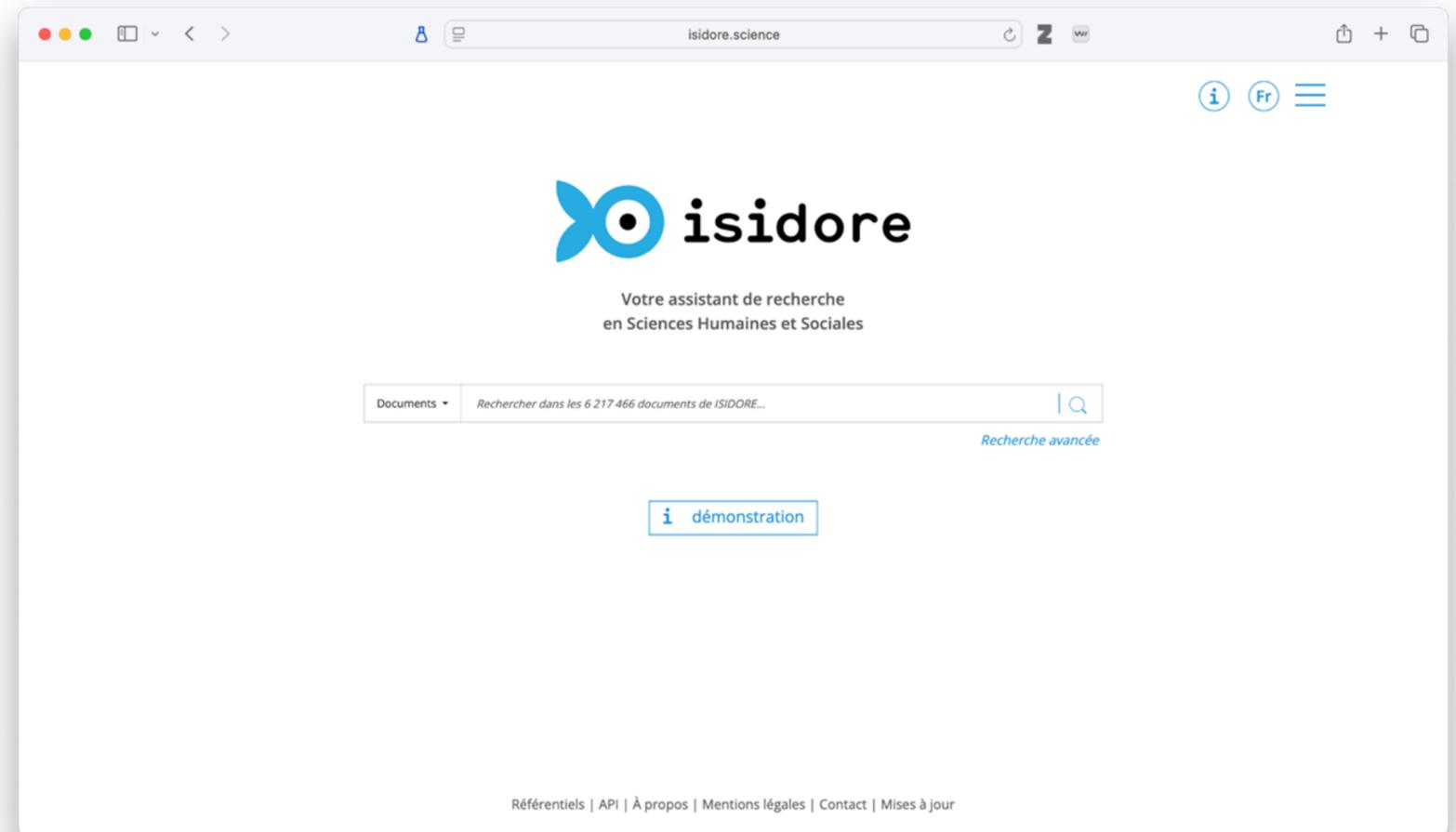
**Votre veille**

Utiliser ISIDORE pour faire suivre des thématiques de recherche en SHS.

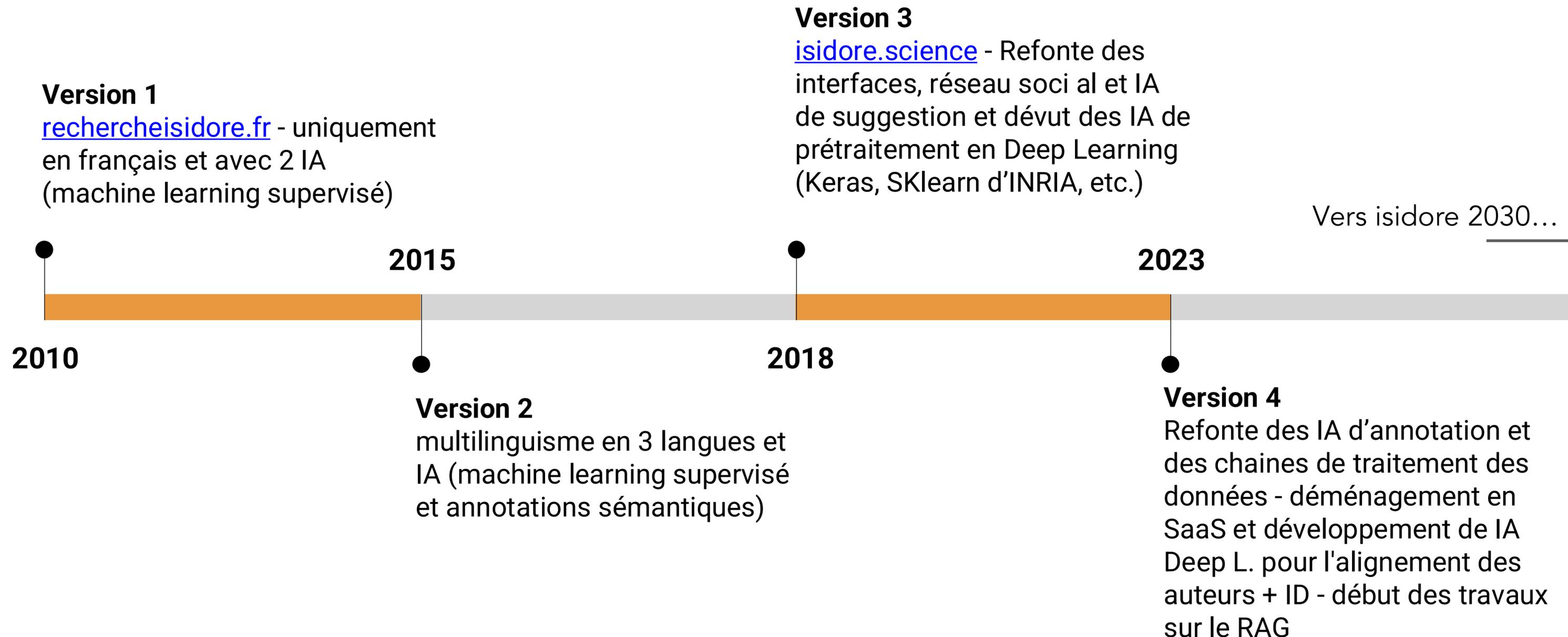
→ EXPLORE

# Isidore : un instrument de recherche depuis 2010

- 6,7 millions de documents des SHS
- 12 000 gisements de données et publications du monde entier
- 1 mise à jour quotidienne (+ ou - 100 000 traitements / jour)
- 900 000 concepts sémantique en SHS pour l'enrichissement venant de 13 bibliothèques de recherche et équipes de recherche
- + de 190 langues et dialectes indexés pour le texte intégral et 3 langues d'enrichissement
- De l'IA "classique" pour le traitement des données depuis 2010
- Une API
- 3 "produits" : isidore.science, isidore suggestions, isidore in Zotero



# Isidore : un instrument de recherche avec de l'IA depuis 2010



# Méthodologie de recherche du HN Lab depuis 2021

- Innovation par l'usage : fondée sur les retours des communautés SHS et accompagnée par des doctorants, des postdocs, des chercheurs (c'est la mission HN Lab).
- Forte dimension recherche SHS & informatique.
- Développement interne : modélisation, preuves de concept.
- Industrialisation co-construite en externe : prestation, accord-cadre CNRS, LabCom.
- Terrain d'application des travaux du HN Lab : ISIDORE



# Méthodologie de recherche du HN Lab depuis 2021

HN Lab Log

Archives du tag 'Séminaire de recherche'

Le Notebook lowtechisé, utopie ou chemin des humanités numériques ?

Projet : Actus  
26 Mai 2025 • Séminaire de recherche, Sciences humaines et sociales, Notebook, Humanités numériques

Des territoires de l'histoire. A la recherche des pratiques numériques discrètes

Projet : Actus  
11 Avril 2025 • Séminaire de recherche, Sciences humaines et sociales, Histoire, Humanités numériques, Numérisation, Sources historiques

Écriture et bidouillage: formats, bricoles et ateliers de production de la pensée

Projet : Actus  
4 Avril 2025 • Séminaire de recherche, Sciences humaines et sociales, écritures numériques, humanités numériques, digital humanities, édition

IA multimodale et IA générative pour l'indexation vidéo pédagogique

HN Lab Log

Archives du tag 'IA'

Expérimentations du "Retrieval-augmented generation" dans ISIDORE : des preuves de concept pour avancer

Projet : IA  
16 Mai 2025 • retrieval-augmented generation, infrastructure de recherche, sciences humaines, sciences sociales, ISIDORE, ingénierie des connaissances, intelligence artificielle, RAG, IA

IA multimodale et IA générative pour l'indexation vidéo pédagogique

Projet : Actus  
27 Février 2025 • Séminaire de recherche, Sciences humaines et sociales, IA, Vidéo, Systèmes complexes

Note sur l'expérience de l'IA au sein de l'Huma-Num Lab

Projet : Actus  
8 Janvier 2025 • Sciences humaines et sociales, Données de la recherche, Intelligence artificielle, IA, Whisper, GPT, Mistral, LLAMA, MLM

Retour d'expérience du projet Chine CoREF avec Archive-  
archiver des sites Web et leurs évolutions pour constituer  
corpus pour la recherche

HN Lab Log

Retour sur l'analyse automatique de corpus de revues SHS

Voir les logs de : Revue 2.0, IA

20 Janvier 2022 • Nicolas Sauret, Stéphane Pouyllau • preuve de concept, analyse, corpus, édition scientifique, traitement des données, machine learning, deep learning

- Résultats
- Chaîne de traitement
  - Analyse des données
    - Livrables
  - Extraction des données
    - Livrables
  - Préparation des données
    - Livrables
  - Modélisation

HN Lab Log

Hackathon I - Retrieval-Augmented Generation en SHS

Voir les logs de : IA

23 Septembre 2024 • Stéphane Pouyllau, Léa Maronet, Adam Faci, Antoine Silvestre de Sacy • traitement de données, intelligence artificielle, RAG, retrieval-augmented generation, infrastructure de recherche, sciences humaines, ingénierie des connaissances



HN Lab

<https://hnlab.huma-num.fr>

# Partie 1 – Observatoire des usages

---

# Introduction - Observatoire des usages

- Travail en sociologie des usages du numérique
- Travail sur les plateformes numériques et usage des modèles génératifs en SHS
- Construire un observatoire des usages dans le cadre de l'Equipex COMMONS (<https://humanum.hypotheses.org/6466>) : nous démarrons à partir d'Isidore dans la perspective d'Isidore 2030
- Qualifier/quantifier les profils d'utilisateurs et faciliter la découvrabilité des documents
- Produire une analyse de la production et la diffusion des sciences et des techniques
- Construire des métriques pour interpréter les interactions entre les utilisateurs et les plateformes numériques

# Horizon : Isidore 2030

1. Observer ce que Isidore nous dit de la façon dont travaillent les disciplines ;
2. Tester/uniformiser/pérenniser les chaînes d'analyse des données utilisateurs ;
3. Identifier les données qui produisent de la connaissance sur les pratiques utilisateurs ;
4. Mettre en lumière la difficulté à se saisir des outils numériques

# Développement des métriques

→ 2 sources principales pour les données de connexion à la plateforme

→ Chaque source permet un apport d'information mais beaucoup de difficultés à faire le lien entre les différentes sources

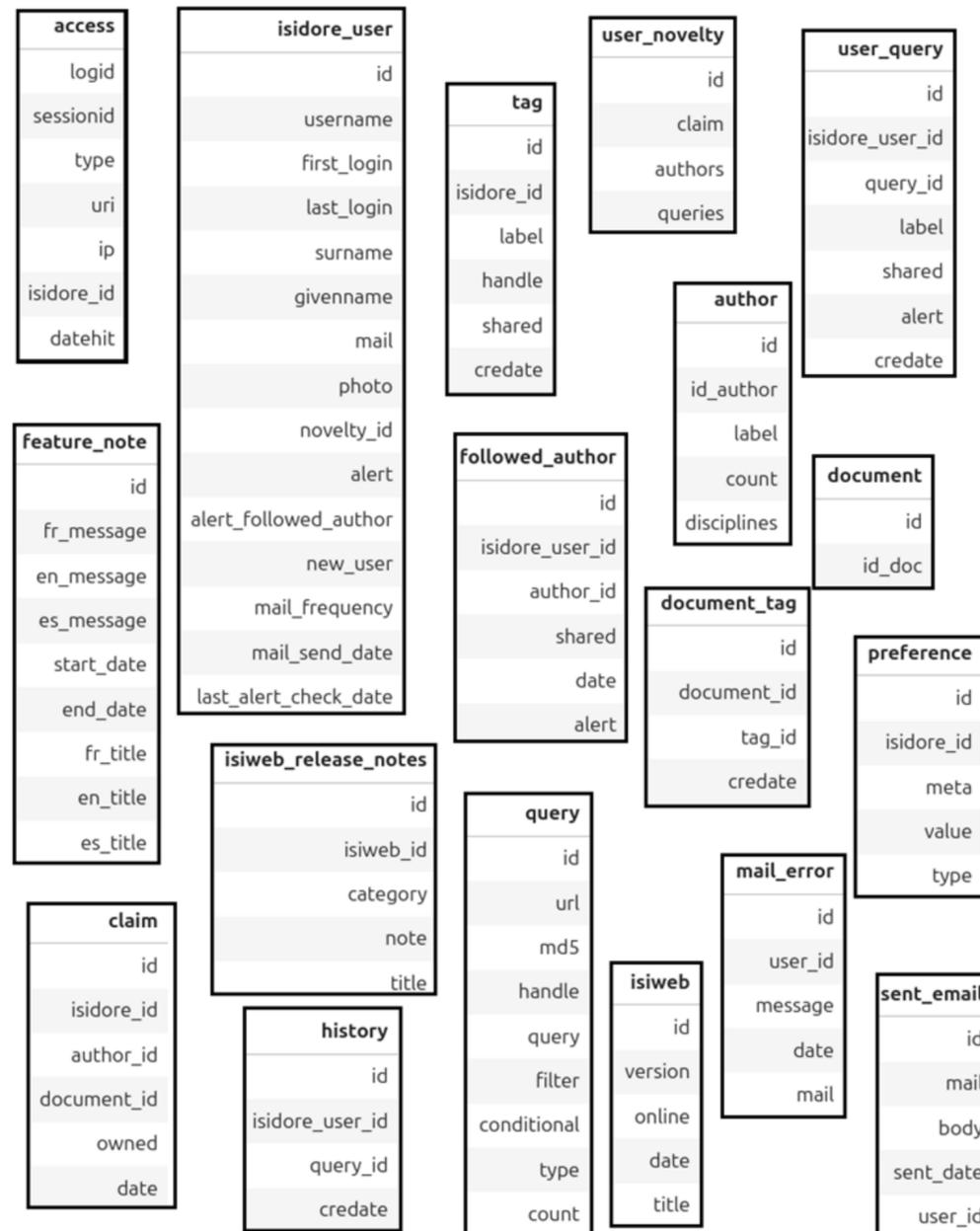
→ Développement de preuves de concept analytiques

## Sources :

- Base de données SQL application Isiweb
- Matomo

# Développement de métriques

Base SQL application Isiweb



- Données de consultation et de requêtes
- Données sur les utilisateurs avec un compte
- Données restreintes non pensées pour un observatoire des usages
- Penser les usages pour améliorer la découvrabilité des données

# Construire des profils de recherche

Des requêtes plus ou moins détaillées selon les logiques individuelles de recherche

## Documents disponibles depuis

Aujourd'hui  Une semaine  Un mois  Une année

## Par langue ou | et

- Abkhaze
- Adyguéen
- Afar
- Afrikaans
- Agni

## Par date de publication

de  à

## Par type de document ou | et

- Articles
- Autres
- Bibliographies
- Bibliothèque numérique
- Billets de blog
- Cartes et plans
- Colloques et conférences
- Cours
- Documents audiovisuels
- Données d'enquêtes

## Par discipline ou | et

- Anthropologie biologique
- Anthropologie sociale et ethnologie
- Archéologie et Préhistoire
- Architecture, aménagement de l'espace
- Art et histoire de l'art
- Démographie
- Droit
- Economies et finances
- Education
- Etudes classiques
- Etudes de l'environnement
- Etudes sur le genre
- Géographie
- Gestion et management
- Héritage culturel et muséologie
- Histoire
- Histoire, Philosophie et Sociologie des sciences
- Linguistique
- Littératures
- Méthodes et statistiques
- Musique, musicologie et arts de la scène
- Philosophie
- Psychologie
- Religions
- Science politique
- Sciences de l'information et de la communication
- Sociologie

# Développement de métriques

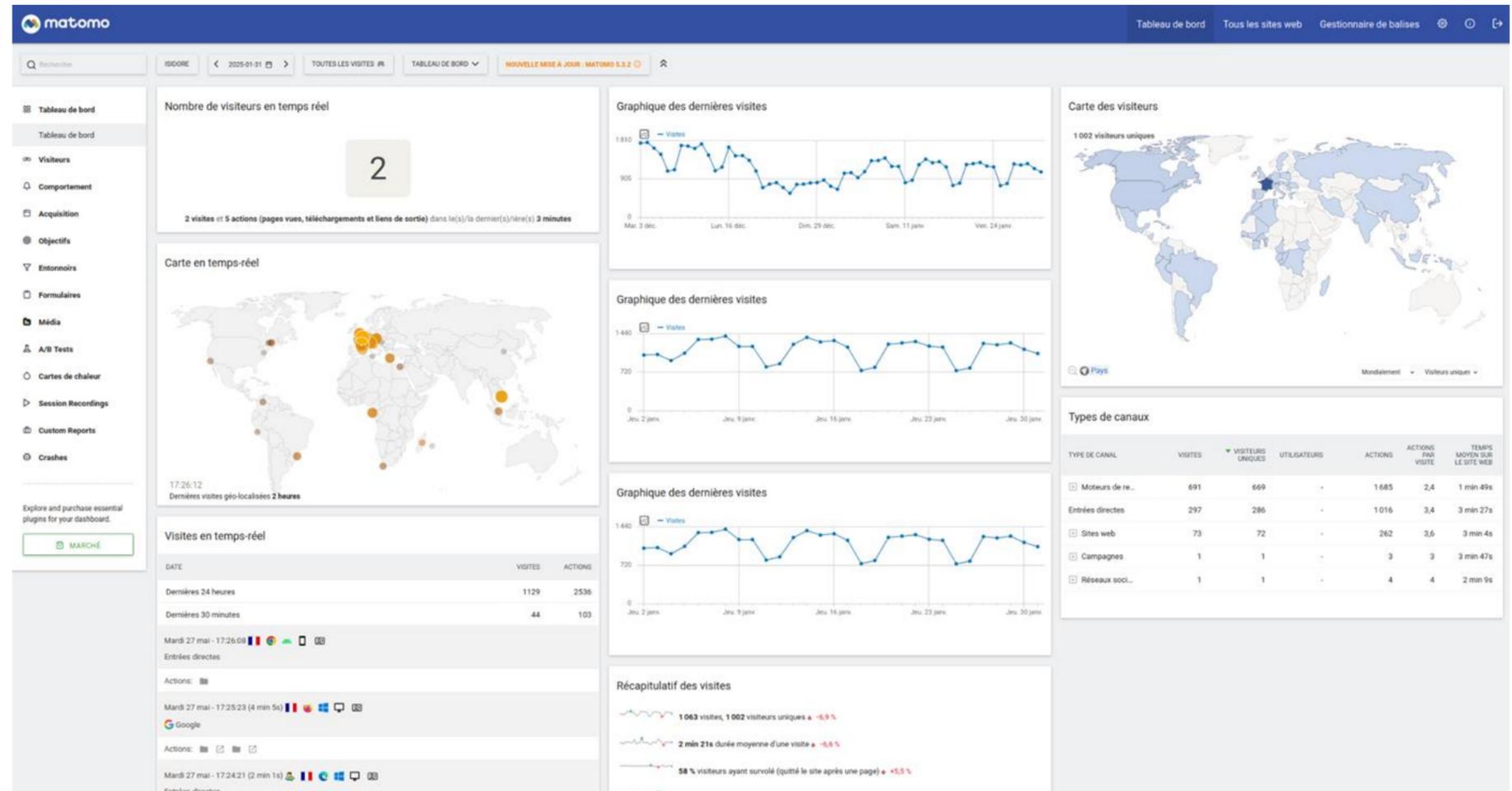
## Matomo

→ Données complètes et riches

→ Informations de navigation à travers les pages

→ Information structurée

→ pour le moment peu adapté à la mise en place d'une chaîne de traitement



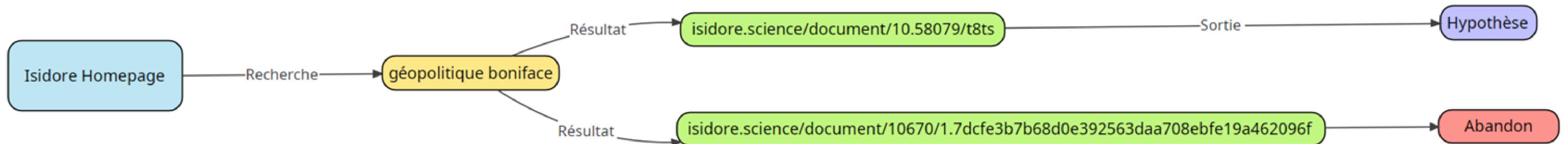
# Analyse des parcours sur la plateforme

- Analyse des parcours est présentée sur Matomo mais elle n'est pas ou peu mise en chiffre car il s'agit d'information complexe
- Permet entre autres de :
  - Déterminer accessibilité des documents
  - Estimer découvrabilité des ressources documentaires
  - Montrer la diversité des pratiques de recherche documentaire

# Analyse des parcours sur la plateforme

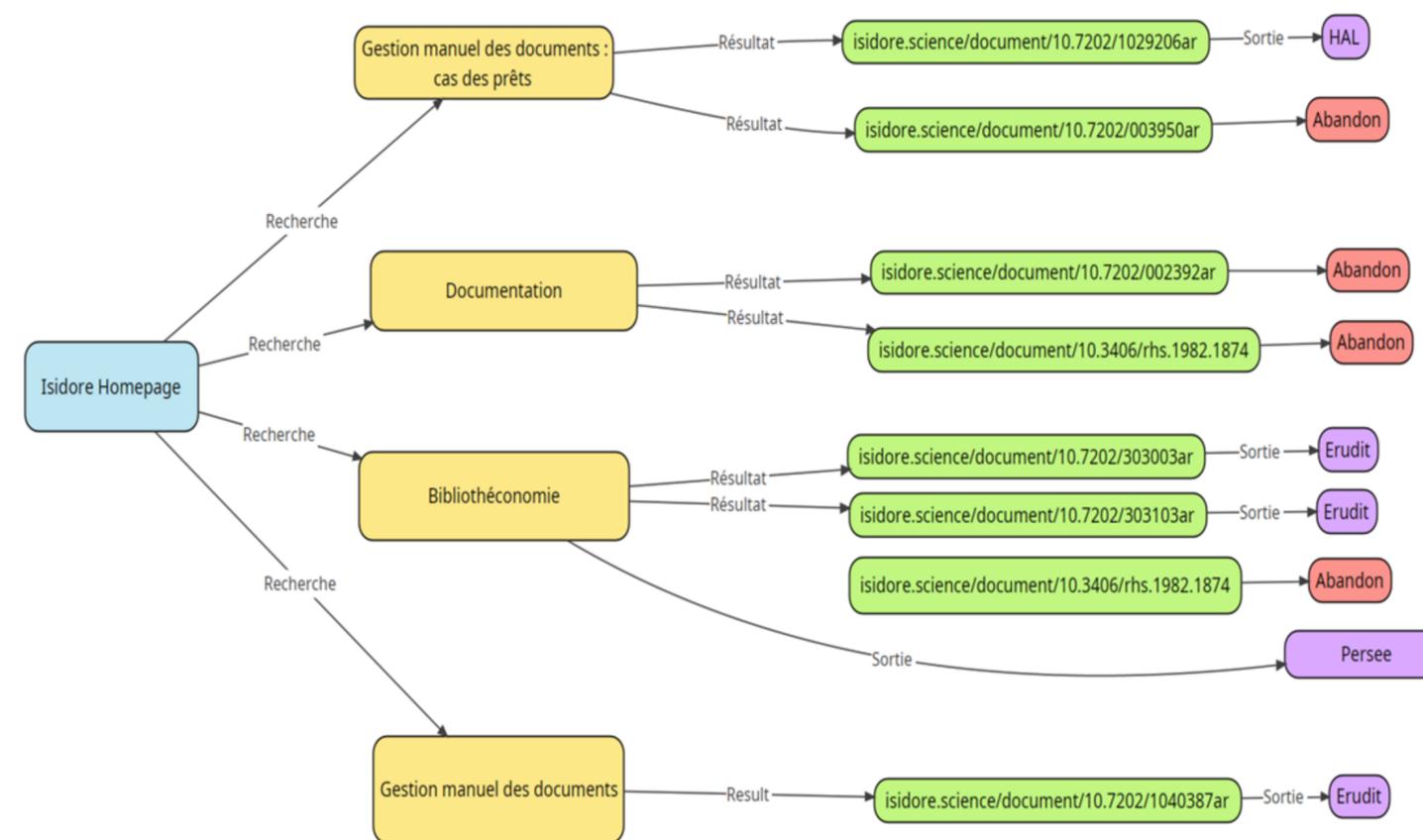
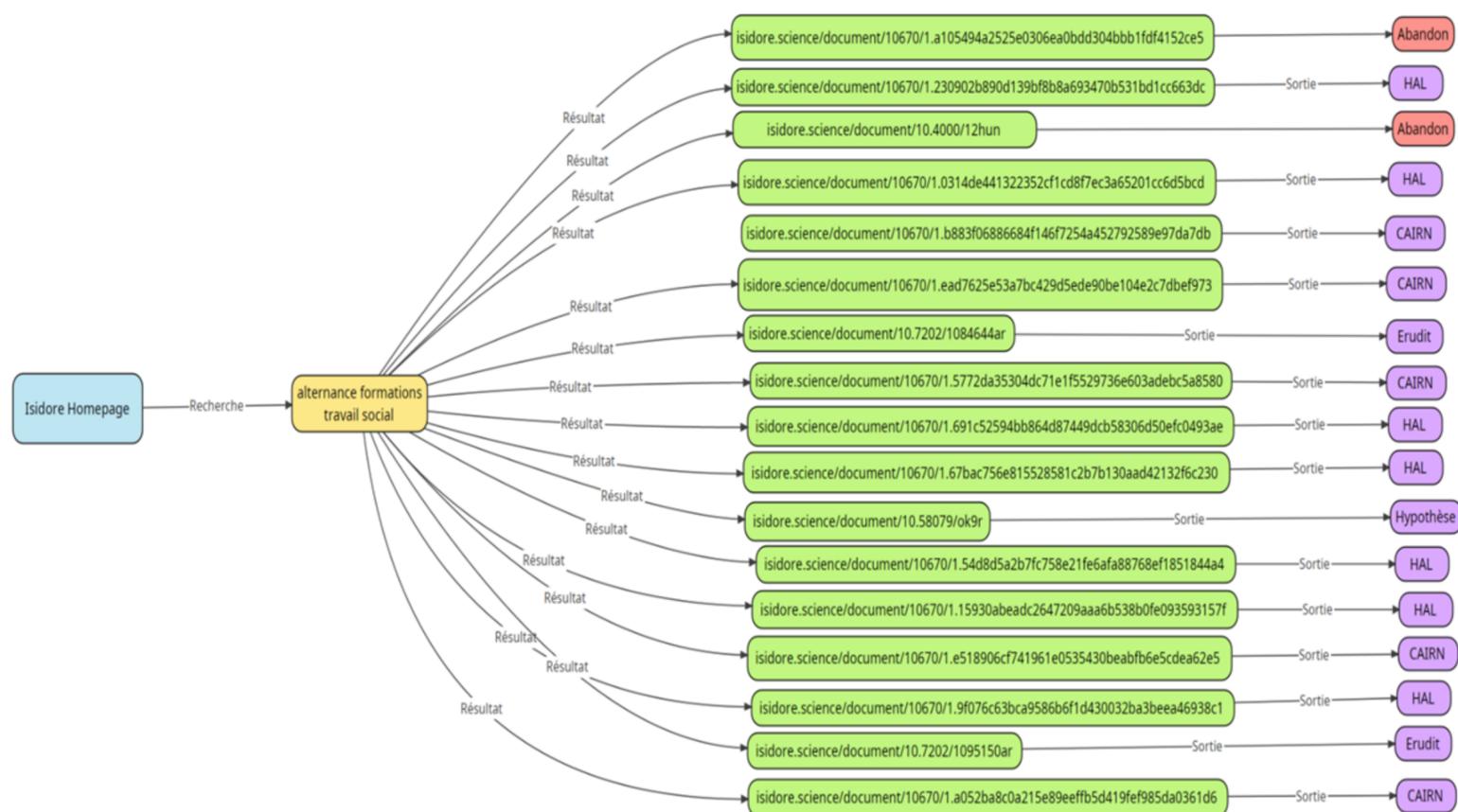
La “recherche documentaire” n’est pas une pratique unifiée

On retrouve des procédés de recherche plutôt simples :



# Analyse des parcours sur la plateforme

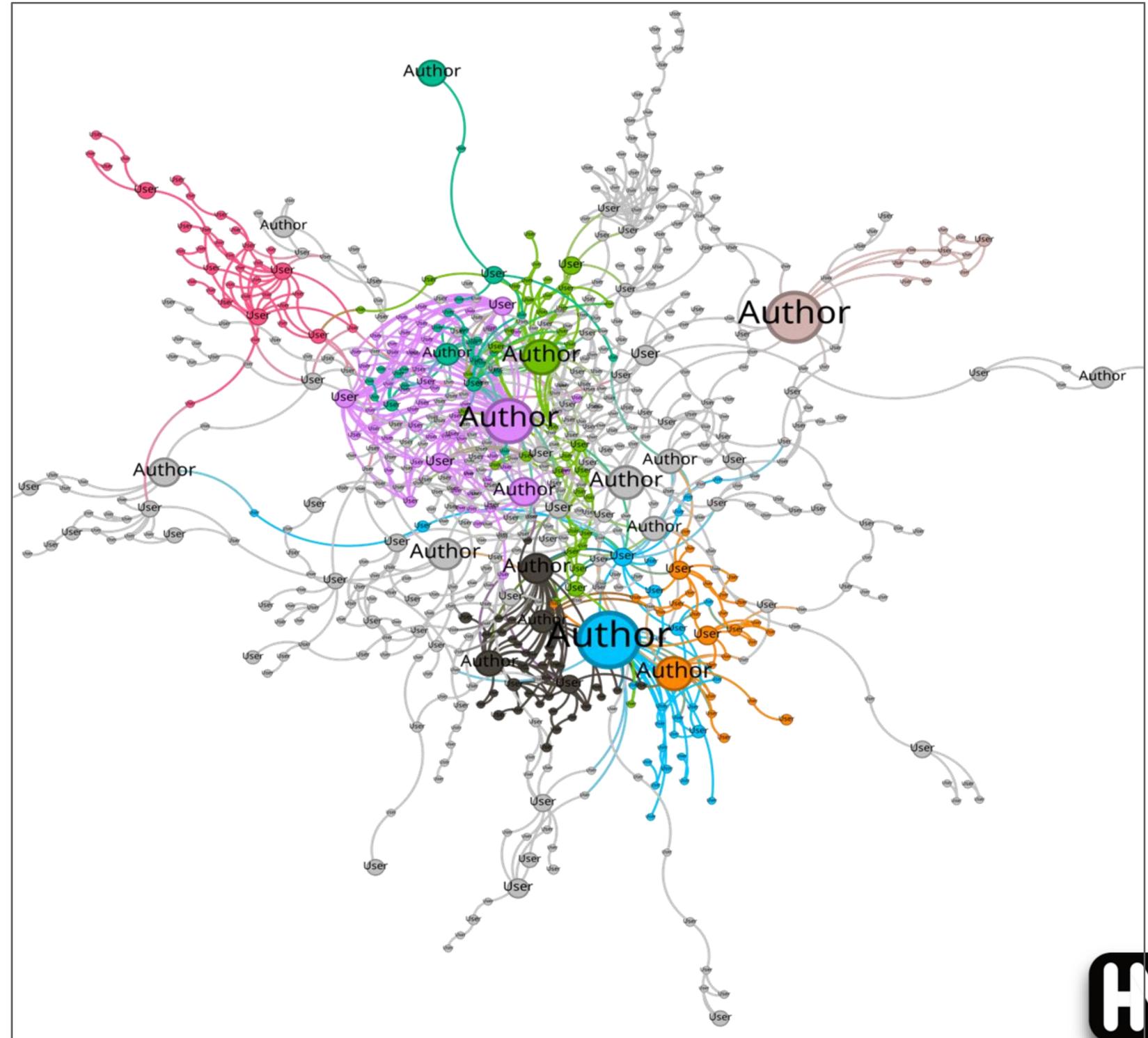
Et des formes plus complexes toutes aussi légitimes dans la recherche documentaire:



L'observatoire des usages a pour but de fournir une plateforme adaptée à toutes ces formes et de simplifier l'accès aux documents

# Rôle d'Isidore en tant que "réseau social de la recherche"

- Appliquer des méthodes d'analyse de réseau à partir des métriques d'Isidore
- Mettre en forme des communautés de connaissance
- Comprendre comment se construisent ces communautés et ce qui les relie
- Observer des rapprochements thématiques entre les usagers et les auteurs qu'ils suivent



# Partie 2 – Retrieval-augmented generation (RAG)

---

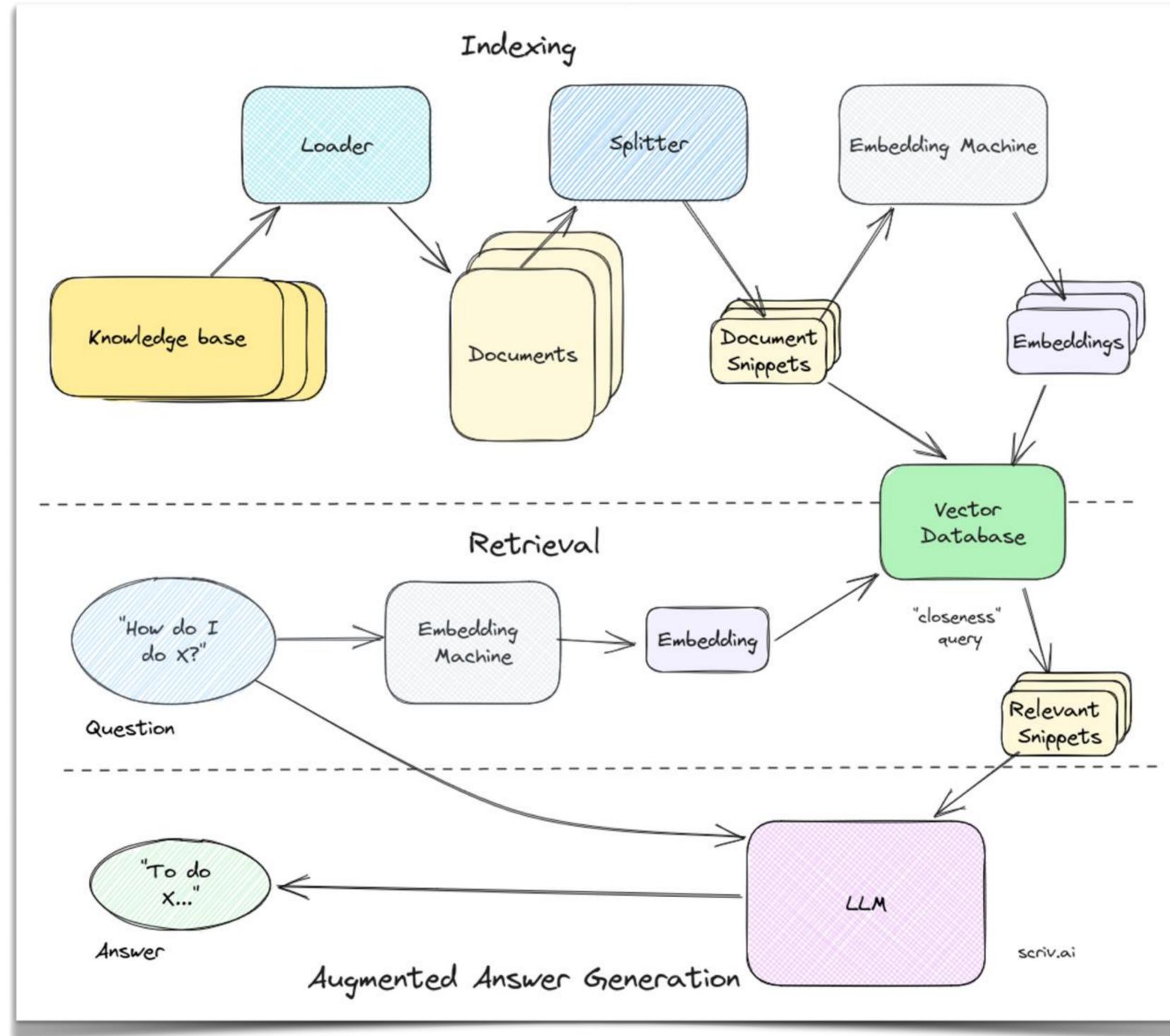
# Le RAG

- RAG (Retrieval Augmented Generation) combine deux modèles d'intelligence artificielle : un modèle de recherche d'informations et un modèle de génération de texte.
- Le modèle de recherche trouve les informations pertinentes dans une base de données, puis le modèle de génération les utilise pour créer un texte original et précis.
- Applications : réponse à des questions, chatbots, rédaction d'articles ou de résumés, traduction automatique...

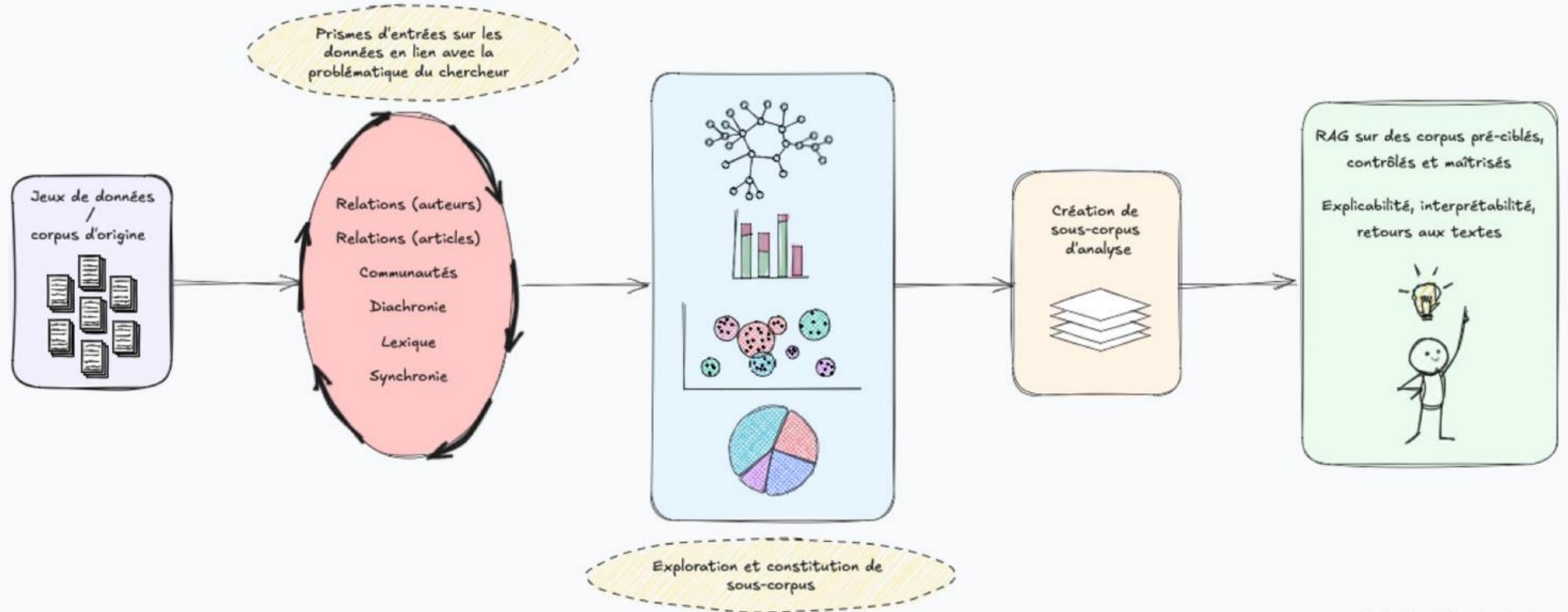
# Avantages

- Le RAG (Retrieval Augmented Generation) est un outil sur lequel nous nous concentrons car il permet de :
  - Contextualiser les requêtes sur des corpus spécifiques ;
  - Lutter contre les hallucinations en donnant le contexte ayant permis la génération ;
  - Augmenter la pertinence, la cohérence, l'explicabilité et surtout l'interprétabilité des générations ;
  - Contrôler les technologies d'IA tout en les spécialisant très fortement sur des problématiques de recherche spécifiques.

# Schéma détaillé



# Modélisation du processus de RAG multiscalaire pour des explorations et études de corpus en histoire, archéologie et histoire de l'art

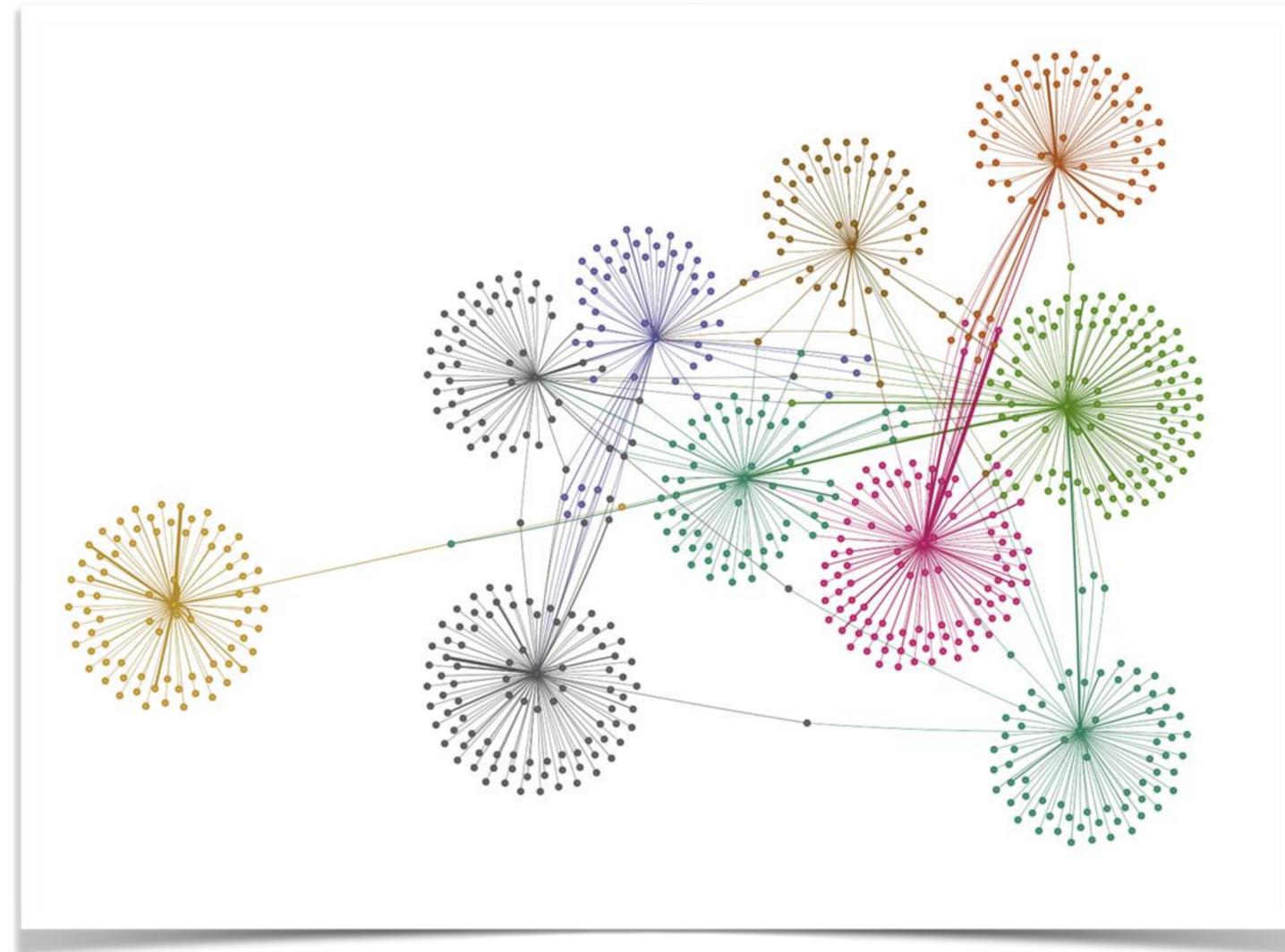


# Multiplication des prismes d'entrée sur les données

- Relations entre auteurs : Comment les auteurs sont-ils liés les uns aux autres au sein d'un domaine de recherche ?
- Relations entre les articles : Comment les articles de recherche sont-ils liés les uns aux autres ?
- Détection de communautés : Peut-on détecter des communautés de recherche ? Comment ?
- Sémantique des articles : Peut-on détecter la sémantique d'une communauté de recherche donnée ?
- Diachronie : Comment analyser un champ de recherche de façon diachronique ?
- Lexie : Comment analyser un terme spécifique (ie. un concept) au sein d'un domaine de recherche ?

# Prismes d'entrées : relations entre auteurs

## Graphs

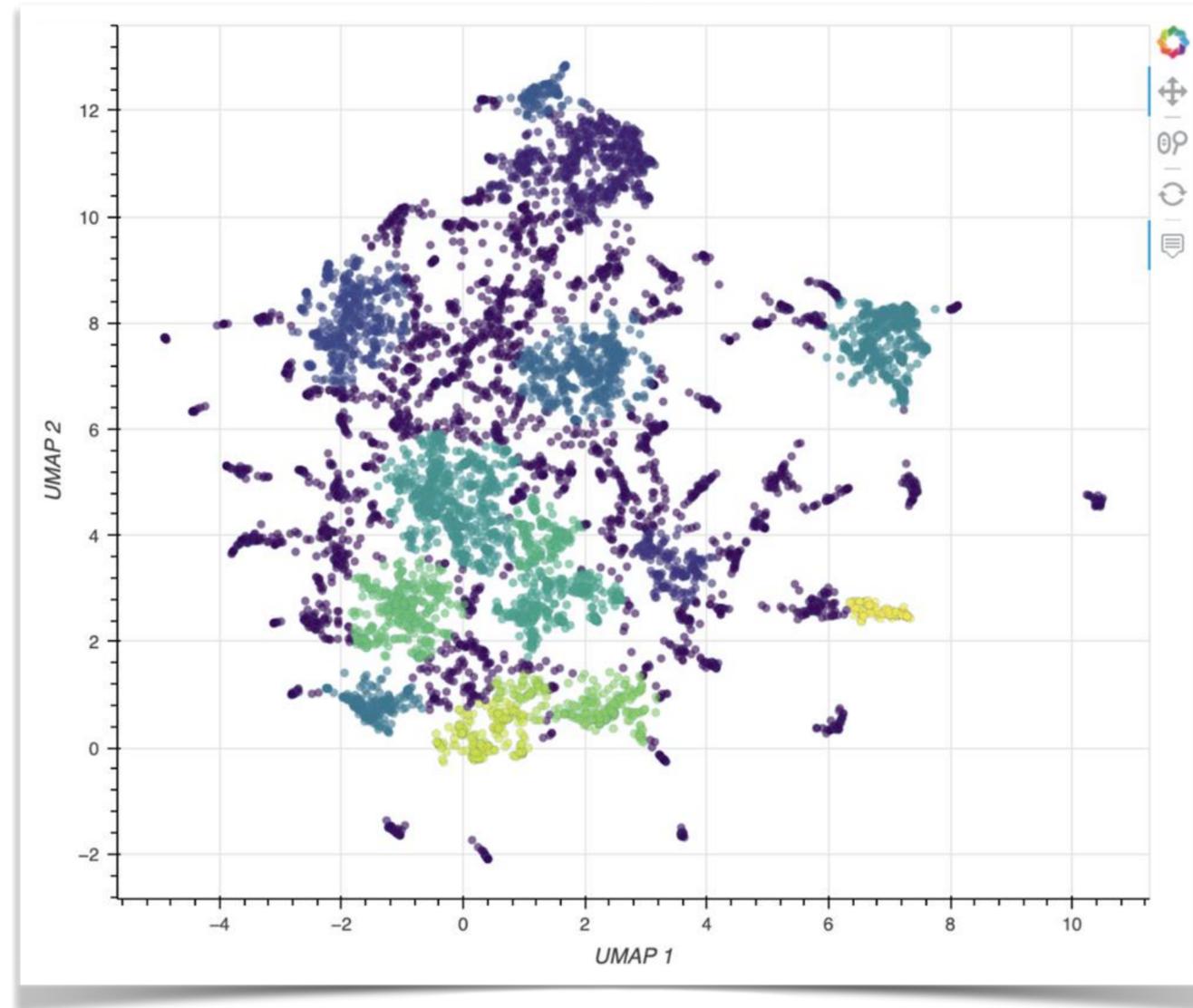


**Co-authorship dynamic network on the 10th most connected authors**

Louvain's Method

# Prismes d'entrées : par les thèmes du jeu de données

## Clustering



**Clustered communities of research papers**

TF-IDF + UMAP dimensionality reduction + DBSCAN clustering

# Intégration micro-service

**GraphToRAG** 🚀

## Présentation

GraphToRAG (GT-RAG) est une application micro-service permettant de pratiquer de la détection de communautés au sein d'un corpus de papiers de recherche. Elle est développée dans le cadre d'expérimentations sur l'intelligence artificielle menées au sein du HN Lab pour le projet de refonte du moteur de recherche académique ISIDORE à horizon 2030.

## Fonctionnalités

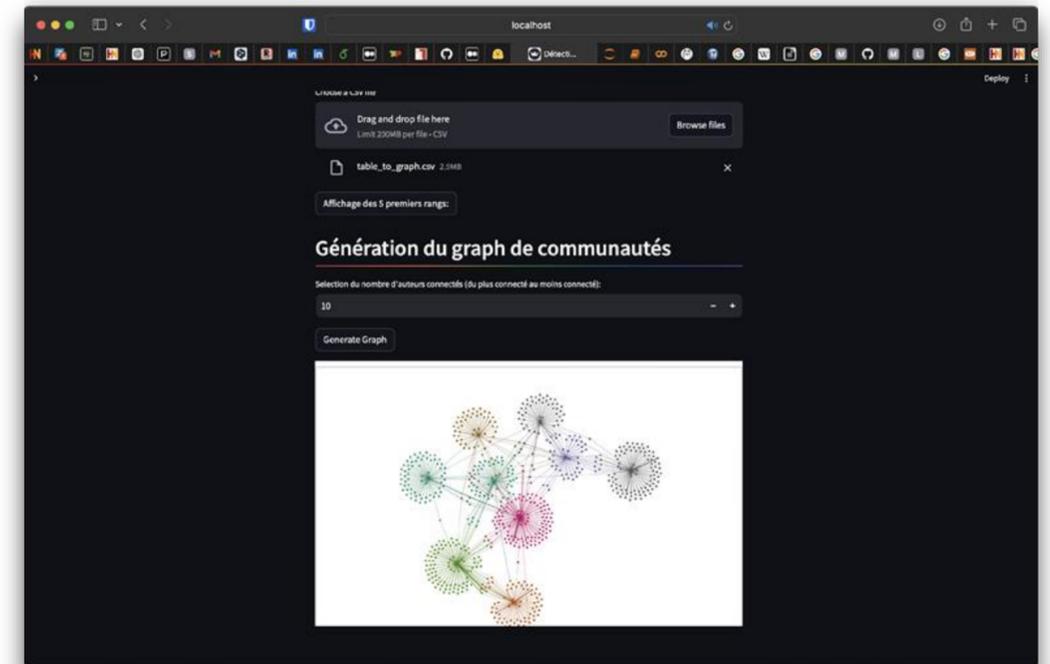
- Détection de communautés et visualisation sous forme de graphes fondés sur les co-auteurs.
- Détection de communautés et visualisation dynamique de clusters de communautés fondée sur les textes de papiers de recherche (Tf-IDF + UMAP + DBSCAN).
- Extraction des scores tf-idf d'une communauté d'intérêt et export en csv.
- *Retrieval Augmented Generation* sur les communautés cibles.

👉 Sélectionner le module sur le menu déroulant et suivez les instructions de chaque module.

## Références

- [Code source de l'application.](#)
- [Notebooks explicatifs.](#)
- [Corpus Kaggle.](#)
- [Micro-application Streamlit.](#)
- [Documentation Streamlit.](#)
- Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte et Etienne Lefebvre, « Fast unfolding of communities in large networks », *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2008, no 10, 9 octobre 2008, P10008 (DOI 10.1088/1742-5468/2008/10/P10008, Bibcode 2008JSMTE...10..008B, arXiv 0803.0476).

GraphToRAG micro-application using Streamlit framework



Community network of co-authors



Papers' clustering

# Intégration future avec Isidore 2030

- Un utilisateur peut :
  - 1°) Requête sur un champ de recherche, un auteur, un article.



Votre assistant de recherche  
en Sciences Humaines et Sociales

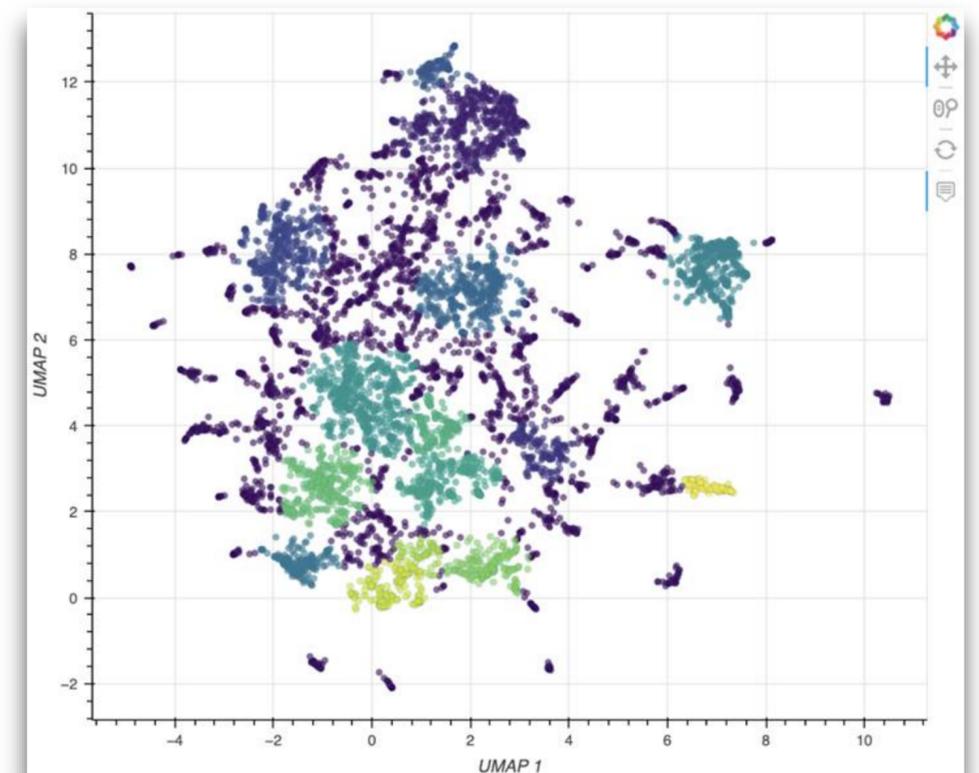
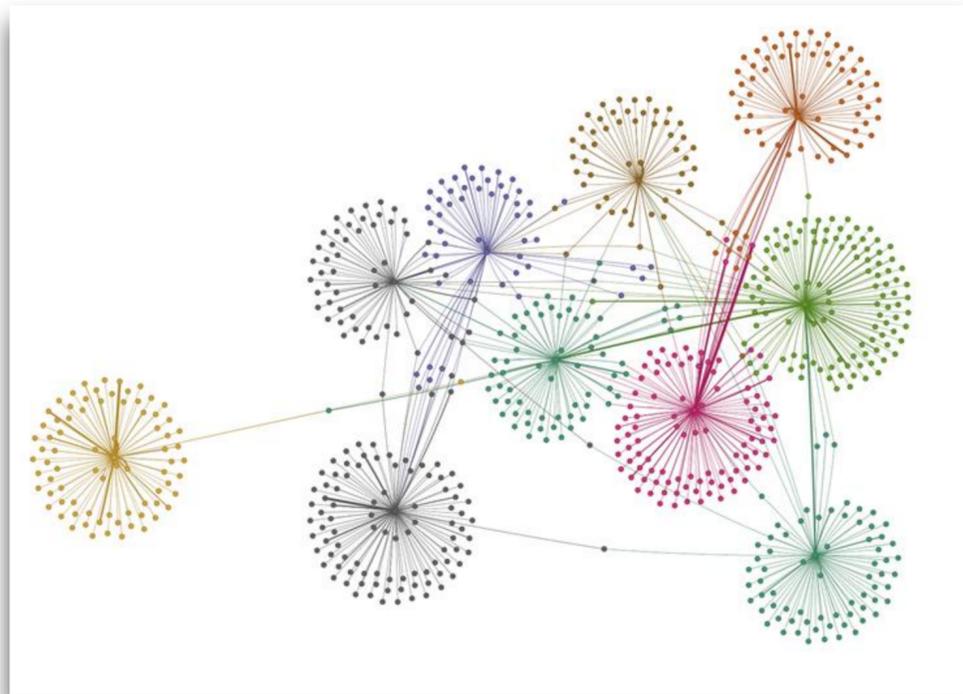
Documents ▾

Rechercher dans les 6 765 404 documents de ISIDORE...



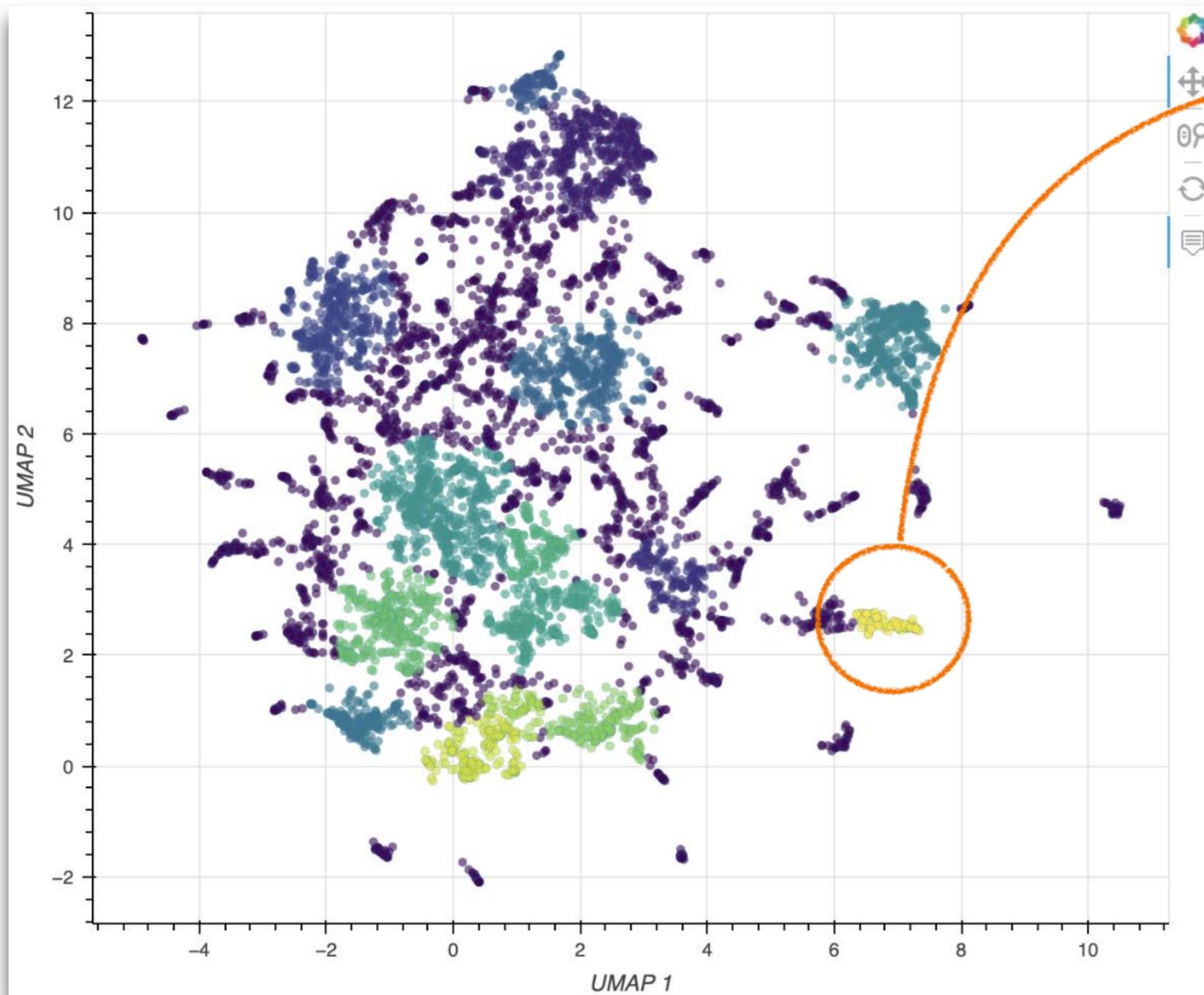
Recherche avancée

Variation des prismes d'entrées  
sur les données



# Intégration future avec ISIDORE 2030

- Un utilisateur peut :
  - 1°) Requête sur un champ de recherche, un auteur, un article.
  - 2°) Vouloir sélectionner un sous-corpus pertinent pour sa problématique de recherche.



- Cluster portant sur la question de la détection automatique de motifs littéraires, par exemple.
- Dans ce cluster, il y a 25 articles sur lesquels l'utilisateur souhaite avoir une connaissance.

**Clustered communities of research papers**  
TF-IDF + UMAP dimensionality reduction + DBSCAN clustering

# Intégration future avec ISIDORE 2030

- Un utilisateur peut :
  - 1°) Requête sur un champ de recherche, un auteur, un article.
  - 2°) Vouloir sélectionner un sous-corpus pertinent pour sa problématique de recherche.
  - 3°) Interroger ce corpus.

dragon.huma-num.fr

Home

Chargement des données et embedding

Détection de Communautés sur les co-...

Détection de Communautés sur les textes

Database

- NIPS

LLM

- open-mistral-7b
- mistral-large-latest
- llama3.2:3b
- llama3.2:1b

ID de la communauté de Louvain

- All
- Andrew Y. Ng
- Michael I. Jordan
- Eric P. Xing
- Yoshua Bengio
- Geoffrey E. Hinton
- Bernhard Schölkopf
- Klaus-Robert Müller
- Zoubin Ghahramani
- Ruslan R. Salakhutdinov

# Dragon

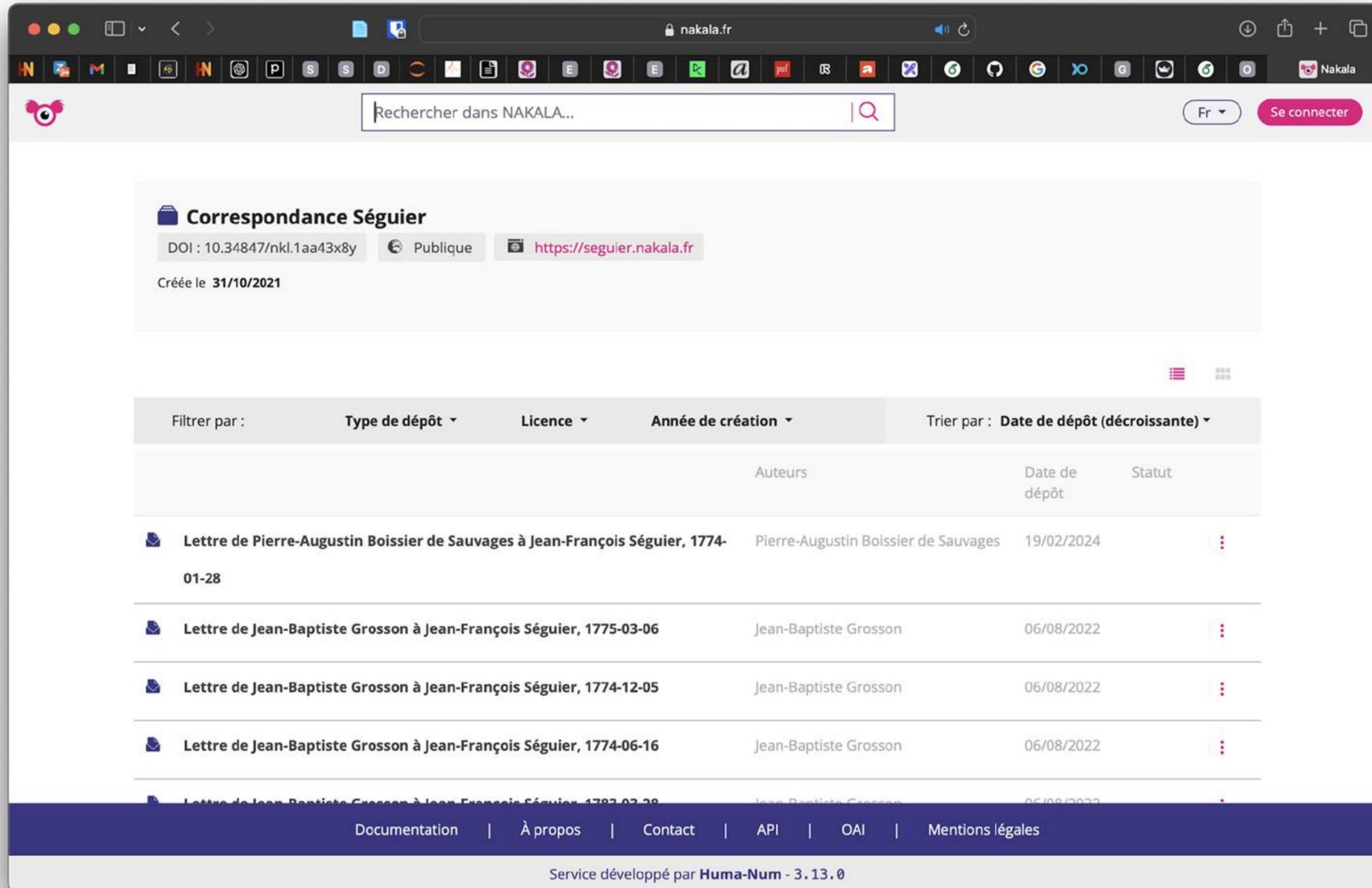
## RAG

 Comment puis-je vous aider ? Sur la colonne de gauche de cette page, veuillez choisir le LLM de base, le corpus que vous voulez interroger, ainsi que la communauté (auteur ou texte) obtenue précédemment.

What is up? 

## Récupération d'une collection NAKALA

- Analyse par graphes d'auteurs.
- Analyse de communautés sémantiques entre les documents de la collection.
- Agents de recherche via RAG.



Rechercher dans NAKALA...

Fr Se connecter

### Correspondance Séguier

DOI : 10.34847/nkl.1aa43x8y Publique <https://seguier.nakala.fr>  
Créée le 31/10/2021

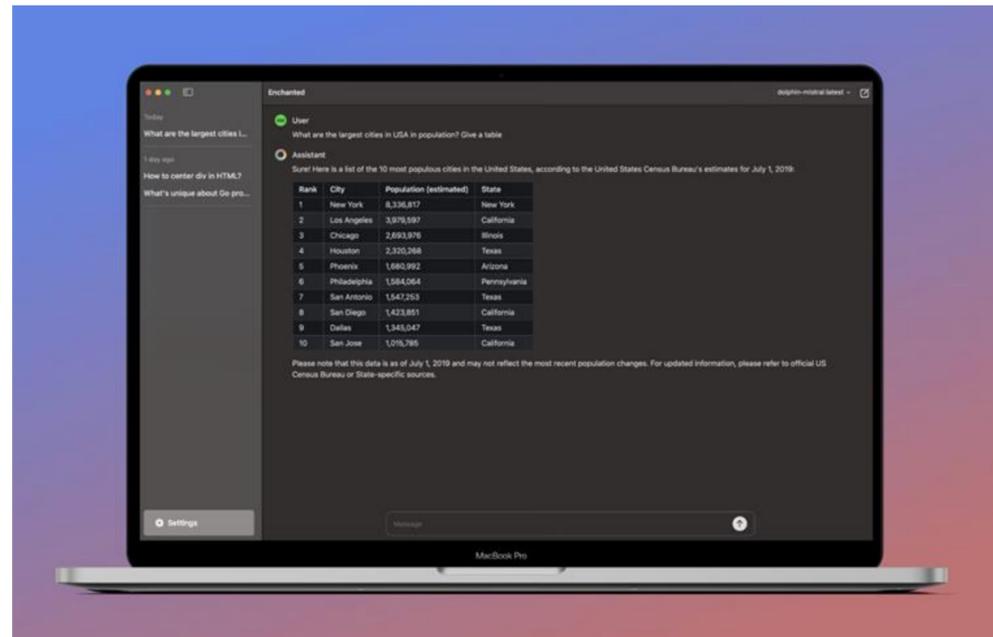
Filtrer par : Type de dépôt Licence Année de création Trier par : Date de dépôt (décroissante)

	Auteurs	Date de dépôt	Statut
Lettre de Pierre-Augustin Boissier de Sauvages à Jean-François Séguier, 1774-01-28	Pierre-Augustin Boissier de Sauvages	19/02/2024	⋮
Lettre de Jean-Baptiste Grosson à Jean-François Séguier, 1775-03-06	Jean-Baptiste Grosson	06/08/2022	⋮
Lettre de Jean-Baptiste Grosson à Jean-François Séguier, 1774-12-05	Jean-Baptiste Grosson	06/08/2022	⋮
Lettre de Jean-Baptiste Grosson à Jean-François Séguier, 1774-06-16	Jean-Baptiste Grosson	06/08/2022	⋮
Lettre de Jean-Baptiste Grosson à Jean-François Séguier, 1787-03-28	Jean-Baptiste Grosson	06/08/2022	⋮

Documentation | À propos | Contact | API | OAI | Mentions légales

Service développé par Huma-Num - 3.13.0

# Jukebox à LLM: plusieurs clients possibles



Get up and running with large language models.

Run Llama 3.3, DeepSeek-R1, Phi-4, Mistral, Gemma 2, and other models, locally.

[Download ↓](#)

Available for macOS, Linux, and Windows

Infrastructure multi-cible pour l'usage de l'IA en sciences humaines et sociales (\*)

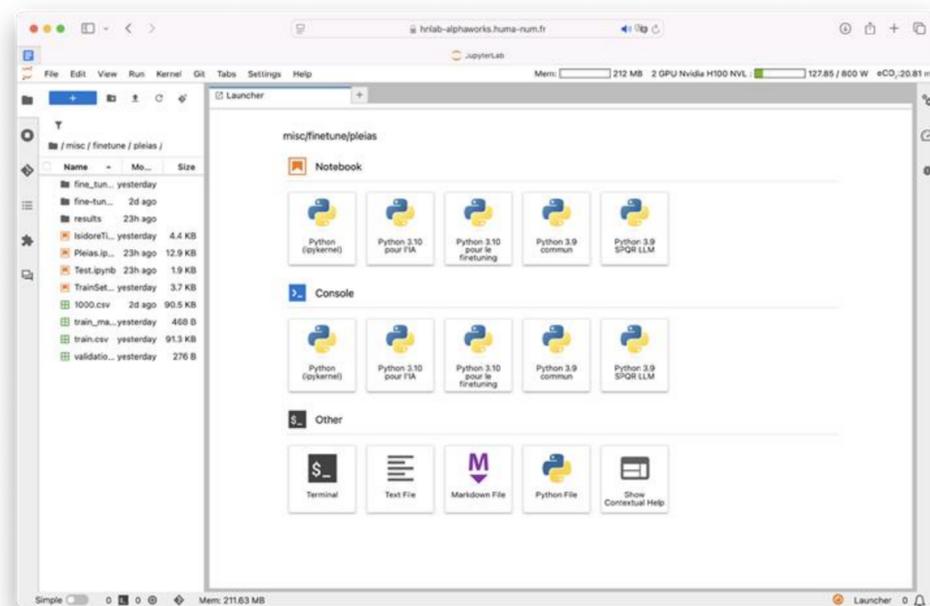
**HN** Huma-Num<sup>IR\*</sup>



**jupyterlab**  
Version 4.3.5

[CONTRIBUTOR LIST](#)

[ABOUT PROJECT JUPYTER](#)



(\*) Voir la présentation de Gérald Foliot et la documentation <https://documentation.huma-num.fr/infrastructure-IA/>

# Partie 3 – Classification et comparaison d'images

---

# Jukebox à VLM + GPU

DINOv2 pour, à partir d'une image d'exemple, retrouver d'autres images visuellement proches

## Instance Retrieval

Find art pieces similar to a given image from a large collection of art images.

DINOv2's frozen features can readily be used to retrieve images similar to a query image using a non-parametric approach: database images are simply ranked according to the similarity of their features with those of the query image.

[See results](#)   [Try another image](#)



↓ Art that looks like your image



Get up and running with large language models.

Run [Llama 3.3](#), [DeepSeek-R1](#), [Phi-4](#), [Mistral](#), [Gemma 2](#), and other models, locally.

[Download ↓](#)

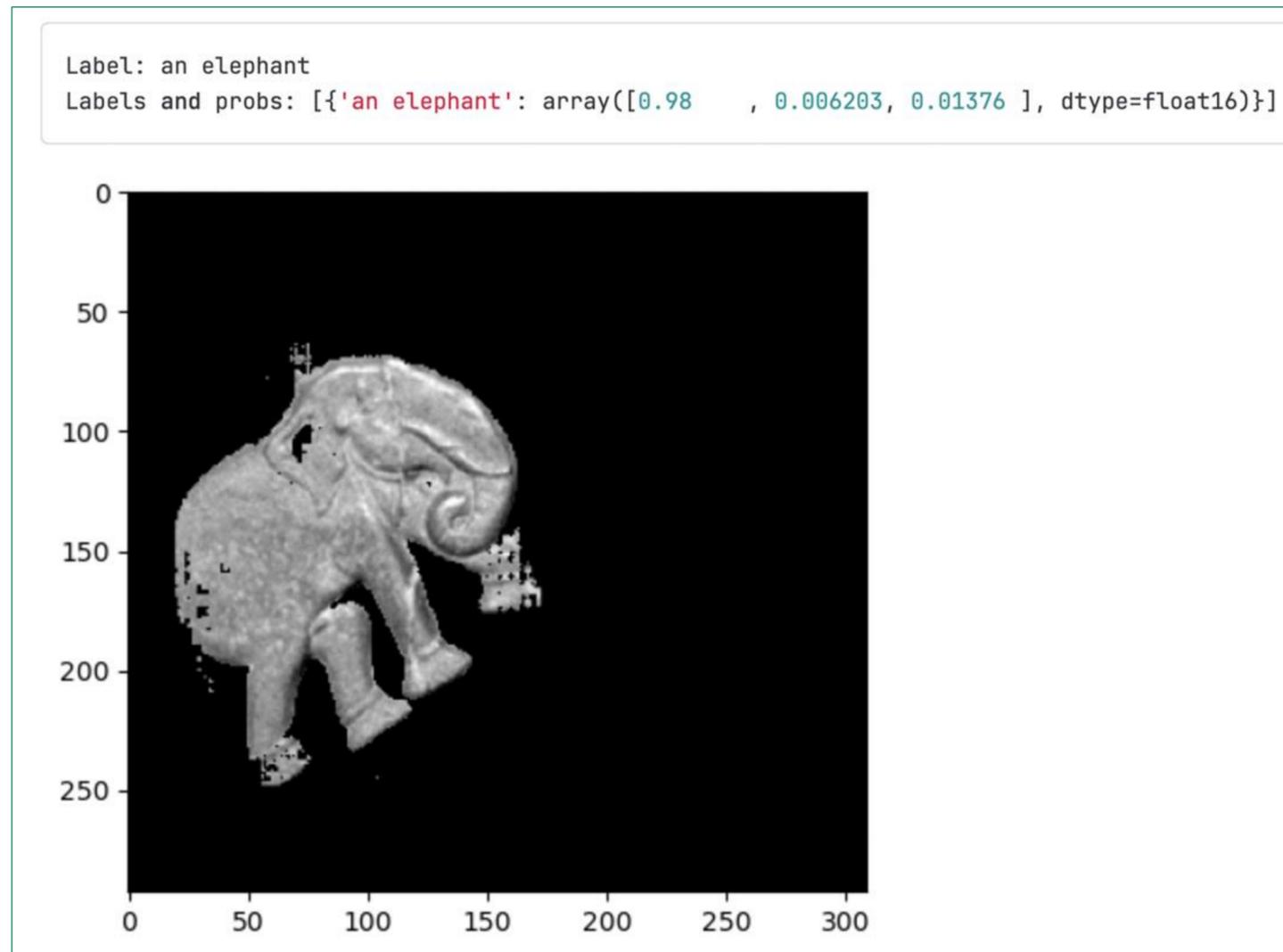
Available for macOS, Linux, and Windows

Infrastructure multi-cible pour l'usage de l'IA en sciences humaines et sociales



# Jukebox à VLM + GPU

CLIP pour la description automatique et l'association de texte à des images



Get up and running with large language models.

Run [Llama 3.3](#), [DeepSeek-R1](#), [Phi-4](#), [Mistral](#), [Gemma 2](#), and other models, locally.

Download ↓

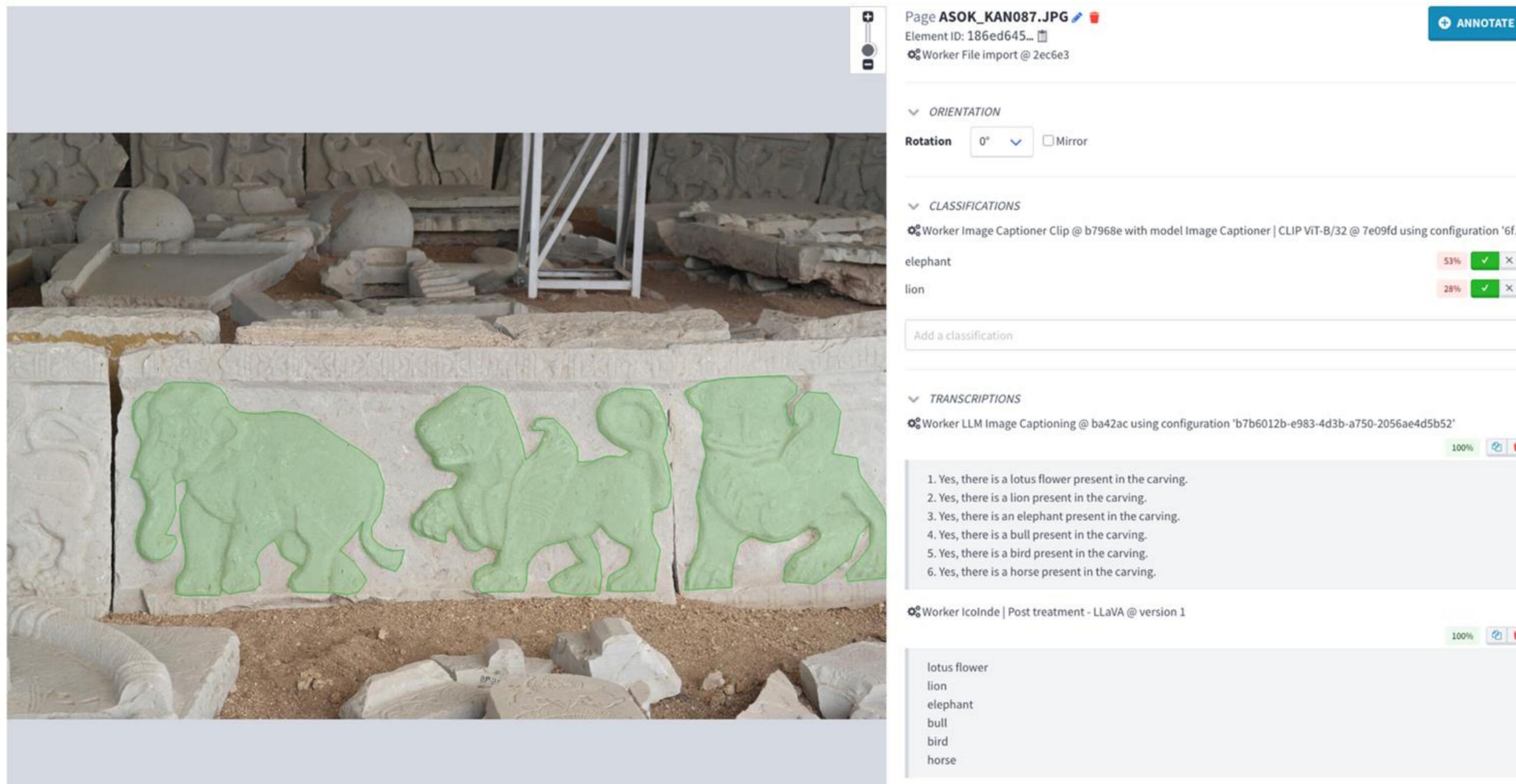
Available for macOS, Linux, and Windows

Infrastructure multi-cible pour l'usage de l'IA en sciences humaines et sociales

 Huma-Num<sup>IR\*</sup>

# Du cycle de l'innovation à l'intégration dans les services d'Huma-Num

Arkindex pour la transcription automatique de textes et la détection d'objets dans les données iconographiques



Page ASOK\_KAN087.JPG  
Element ID: 186ed645...  
Worker File import @ 2ec6e3

**ORIENTATION**  
Rotation: 0°  Mirror

**CLASSIFICATIONS**  
Worker Image Captioner Clip @ b7968e with model Image Captioner | CLIP ViT-B/32 @ 7e09fd using configuration '6f...'  
elephant 53% ✓  
lion 28% ✓

Add a classification

**TRANSCRIPTIONS**  
Worker LLM Image Captioning @ ba42ac using configuration 'b7b6012b-e983-4d3b-a750-2056ae4d5b52'  
100%

1. Yes, there is a lotus flower present in the carving.
2. Yes, there is a lion present in the carving.
3. Yes, there is an elephant present in the carving.
4. Yes, there is a bull present in the carving.
5. Yes, there is a bird present in the carving.
6. Yes, there is a horse present in the carving.

Worker Icolnde | Post treatment - LLaVA @ version 1  
100%

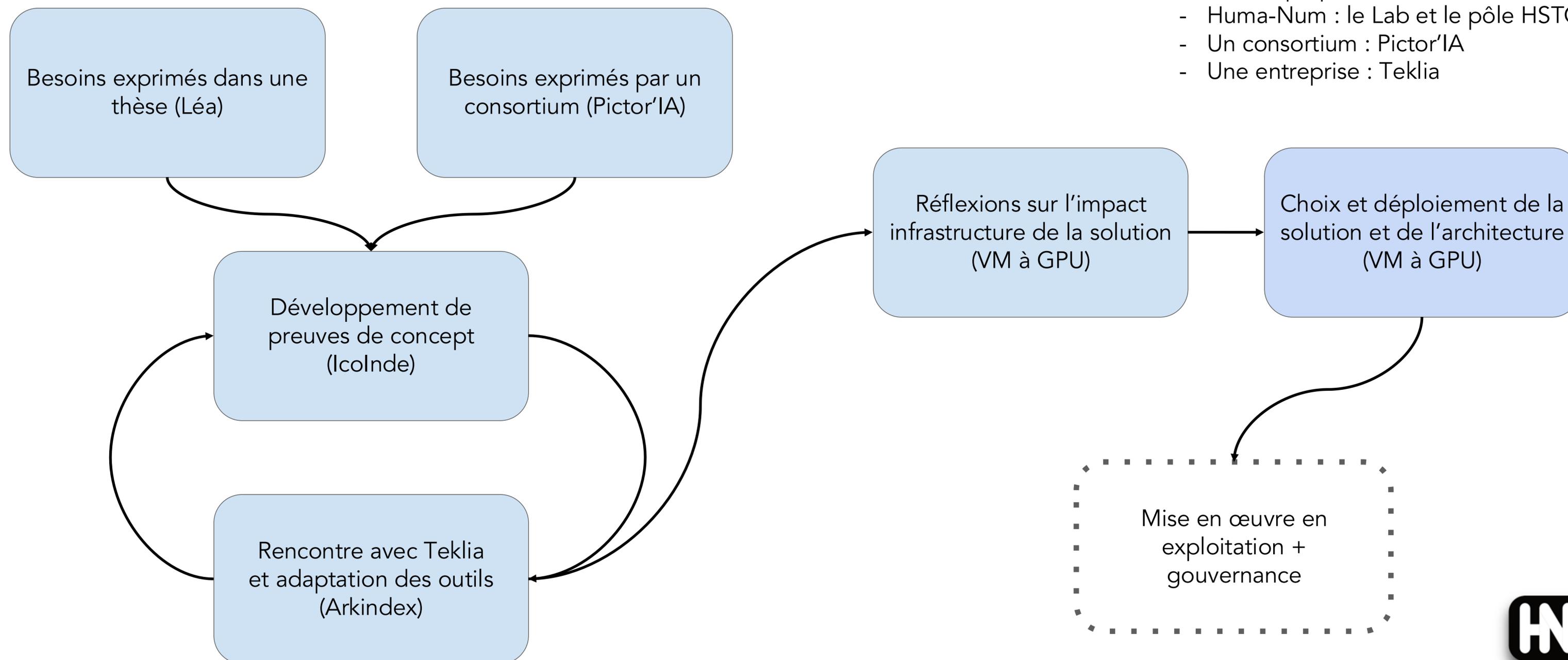
- lotus flower
- lion
- elephant
- bull
- bird
- horse

# Du cycle de l'innovation à l'intégration dans les services d'Huma-Num

Arkindex pour la transcription automatique de textes et la détection d'objets dans les données iconographiques

Acteurs impliqués :

- Huma-Num : le Lab et le pôle HSTC
- Un consortium : Pictor'IA
- Une entreprise : Teklia



# Défis actuels dans ISIDORE et NAKALA

- Amélioration des **métadonnées restreintes** pour accroître la **découvrabilité** des données dans les outils de RAG (par exemple)

**Photos\_AXIII-8** Contacter le gestionnaire

DOI : 10.34847/nkl.a300208u Publique

**Auteur :** Pierre Lacau

Archives du fonds Pierre Lacau – Centre Wladimir Golenischeff (EPHE)

**Fichier** **Visualisation**

Photos\_AXIII-8.tif



Copier l'ID Copier l'url d'intégration Copier l'url de téléchargement

**Mots-clés** Karnak **Citer**

**Photos\_XXXVIII-31b** Contacter le gestionnaire

DOI : 10.34847/nkl.436a2kos Publique

**Auteur :** Pierre Lacau

Archives du fonds Pierre Lacau – Centre Wladimir Golenischeff (EPHE)

**Fichier** **Visualisation**

Photos\_XXXVIII-31b.tif



Copier l'ID Copier l'url d'intégration Copier l'url de téléchargement

**Mots-clés** Karnak **Citer**

# Défis actuels dans ISIDORE et NAKALA

- Dans les publications, le **contenu des images** est rarement décrit en détail et la description peut-être très chronophage à produire pour les chercheurs, etc. ;
- Les moteurs comme ISIDORE ne sont pas encore assez fins pour utiliser les résultats de post-traitement des données par des IA (pas uniquement génératives).

## Le lion colossal de Bartholdi

**Fiche du document**

Auteur  
Barbant Charles ((graveur)). Fonction indéterminée

Date  
1889

Discipline  
Sciences de l'information et de la communication

Type de document  
Photos et images

Périmètre  
**Sources de la recherche**

Langue  
Français

Identifiants  
12148/btv1b10200346p  
IFN-10200346

Source  
Photographies et images fixes de Gallica

Relations  
Ce document est lié à :  
<http://catalogue.bnf.fr/ark:/12148/cb41917287h>

Collection  
Gallica

Organisation  
Bibliothèque nationale de France

Licences  
domaine public , public domain

The screenshot shows the Gallica website interface. At the top, there is a search bar with the text 'Tout Gallica' and a search button. Below the search bar, there is a navigation menu with options like 'Livres', 'Manuscrits', 'Cartes', 'Images', 'Presse et revues', 'Enregistrements sonores', and 'Partitions'. The main content area displays the title 'Le lion colossal de Bartholdi Barbant, Charles ((graveur)) 1889' and a large image of the lion sculpture. The image is a black and white photograph of the lion sculpture, which is a large, seated lion with a human-like face, known as the 'Lion Colossal' by Bartholdi. Below the image, there is a caption: 'LE LION COLOSSAL DE BARTHOLDI. Source gallica.bnf.fr / Bibliothèque nationale et universitaire de Strasbourg'.

# Défis actuels dans ISIDORE et NAKALA

- Les documents peuvent être **indexés** de manière limitée (en dessous des scores F d'ISIDORE par exemple), les images au sein des PDF ne le sont pas du tout.



## August Strindberg en héritage

**Fiche du document**

Auteur  
Anna Svenbro

Discipline  
Sciences de l'information et de la communication

Type de document  
Articles

Périmètre  
Publications

Langue  
Français

Identifiants  
ISIDORE Id: 10670/1.493b0b...  
hal: hal-02441332

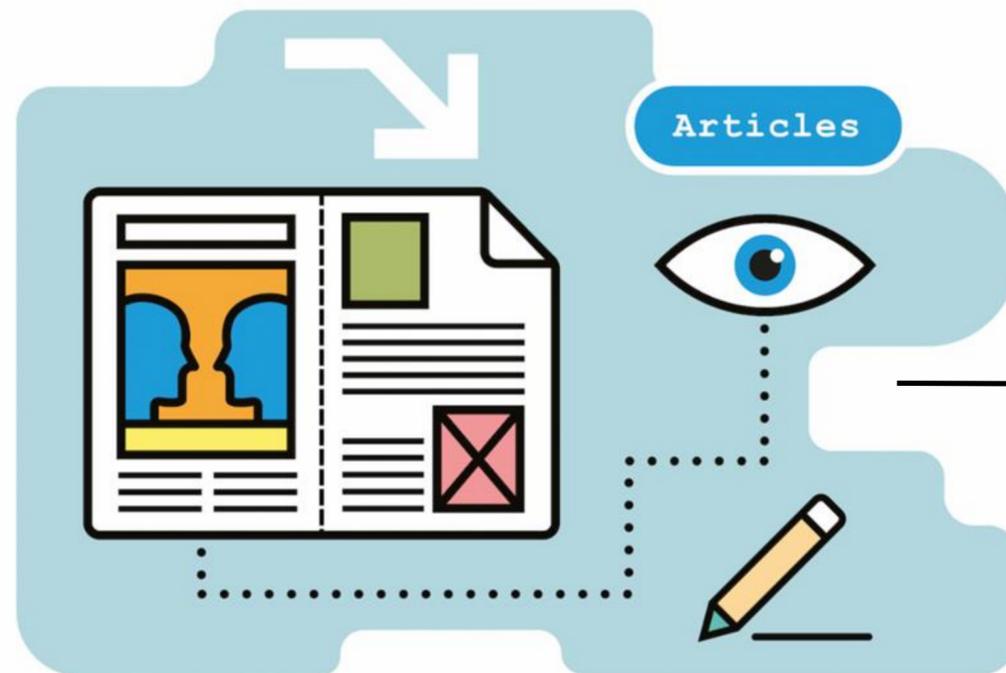
Source  
HAL SHS (Sciences de l'Homme et de la Société)

Collection  
Archives ouvertes

Organisation  
Centre pour la communication scientifique directe

### Résumé

International audience



### Fichiers et aperçu

**Fichier principal**

Svenbro\_Chroniques-063\_13.pdf (757)

2 sur 2

Zoom automatique

créateur que personnel, Strindberg aime la France, Paris, la langue et la littérature françaises, dont il a une connaissance approfondie. Parfois en butte à l'hostilité de ses compatriotes (qui vont jusqu'à le traîner devant les tribunaux suédois pour blasphème lors de la publication de littéraires et artistiques. Strindberg publie un certain nombre d'articles rédigés directement dans notre langue dans la presse française, notamment dans *L'Écho de Paris*, et écrit une partie de son œuvre directement en français, la faisant également publier dans l'Hexagone. *Le Plaidoyer* tualité, l'écologie, ou dans les débats actuels autour de la dimension culturelle du « modèle suédois ». L'empreinte qu'il a laissée en France est considérable, Strindberg ayant été, dès l'époque d'André Antoine et Lugné-Poe, ses découvreurs du Paris des années 1890, jusqu'à Antoine Vitez en passant par Antonin Artaud et Jean Vilar, une source d'inspiration et de questionnement pour les universitaires, traducteurs, écrivains, dramaturges, metteurs en scène et acteurs français. La conférence, et particulièrement la journée organisée à la BnF, permettra d'étudier le caractère testamentaire d'un certain nombre d'œuvres du dernier Strindberg, et d'examiner les diverses facettes de l'héritage strindbergien, la manière dont celui-ci a été reçu, conservé, réapproprié et transmis. Anna Svenbro

**Strindberg en héritage**  
27 septembre - 9h - 18h  
Université Paris-Sorbonne,  
Maison de la Recherche  
28 septembre  
BnF, site François-Mitterrand, petit auditorium  
29 septembre  
Fondation Grez-sur-Loing  
pour plus d'informations :  
<http://www.cric.paris-sorbonne.fr>

Chroniques de la BnF - n°62 - 13

Origine

Publication financée par une institution

# Pistes d'amélioration

- **Extraction de la connaissance** : Classification automatique des images pour améliorer les métadonnées (type de document, domaine, contenu, etc.) en se référant aux thésaurus existants (Pactols pour l'archéologie, IconoClass pour l'histoire de l'art, Théséo pour l'histoire, etc.) ;
- **Extraction des images des PDF** : Récupération des légendes (s'il y en a), classification et description automatique afin qu'elles soient elles aussi indexées ;
- **Génération d'*embeddings*** (projection dans un espace sémantique) : Produites à partir des métadonnées des images et des images elles-mêmes pour nourrir le RAG.

⇒ L'idée est d'améliorer la découvrabilité des tous les documents et de produire des métadonnées des images en fonction de leur contenu.

# Sur la salle dite "du couronnement" à Tell-el-Amarna



## Fiche du document

### Auteurs

Françoise Laroche-Traunecker  
Claude Traunecker

### Date

1985

### Type de document

Articles

### Périmètre

Publications

### Langue

Français

### Identifiants

handle: 10670/1.3md1rs  
halshs: halshs-00357957

### Source

HALSHS : archive ouverte en Sciences de l'Homme et de la Société

### Collection

Archives ouvertes

### Organisation

Centre pour la communication scientifique directe

### Licence

info:eu-repo/semantics/OpenAccess

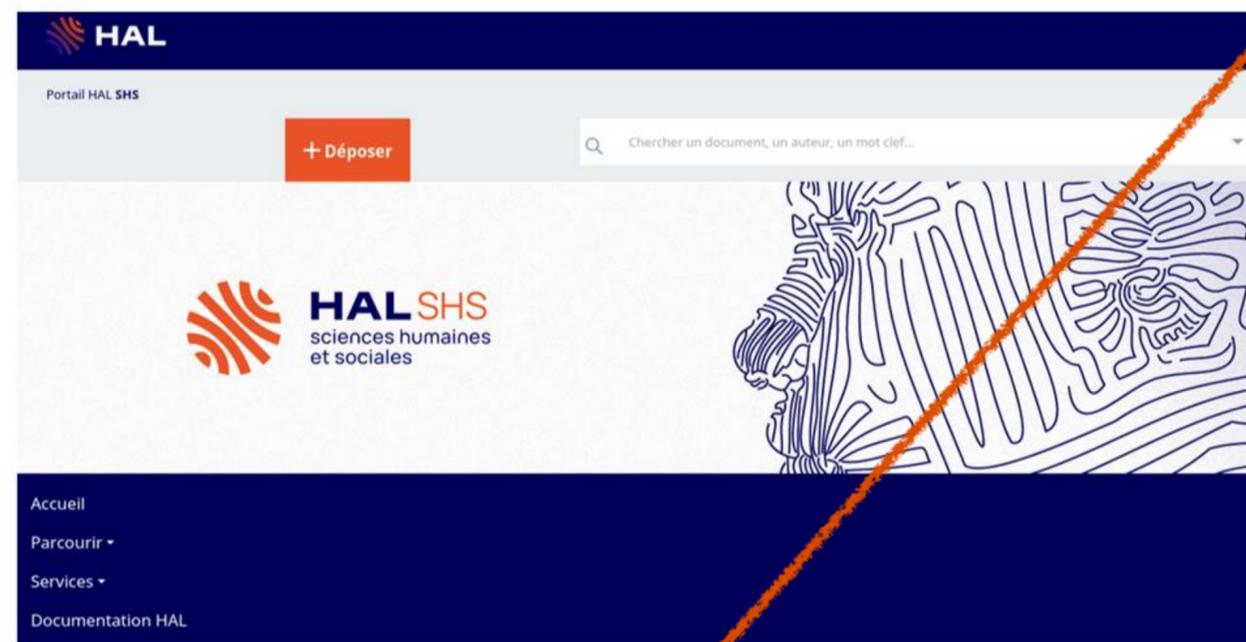
## Mots-clés

couronnement Tell-el-Amarna Akhenaton

Pendlebury fouilles

## Résumé

Le livre de J.D.S. Pendlebury consacré à ses fouilles sur le site de la ville d'Akhenaton, Tell-el-Amarna, a paru en 1936 en version française, traduit par Henri Wild. Souhaitant rendre hommage non seulement au savant mais aussi au traducteur qui permit au public francophone d'accéder en toute sécurité à de grands classiques de la littérature égyptologique anglaise et allemande, nous avons extrait de cet ouvrage le sujet de notre contribution.



Article Dans Une Revue Société d'Égyptologie, Genève Année : 1985

### Sur la salle dite "du couronnement" à Tell-el-Amarna

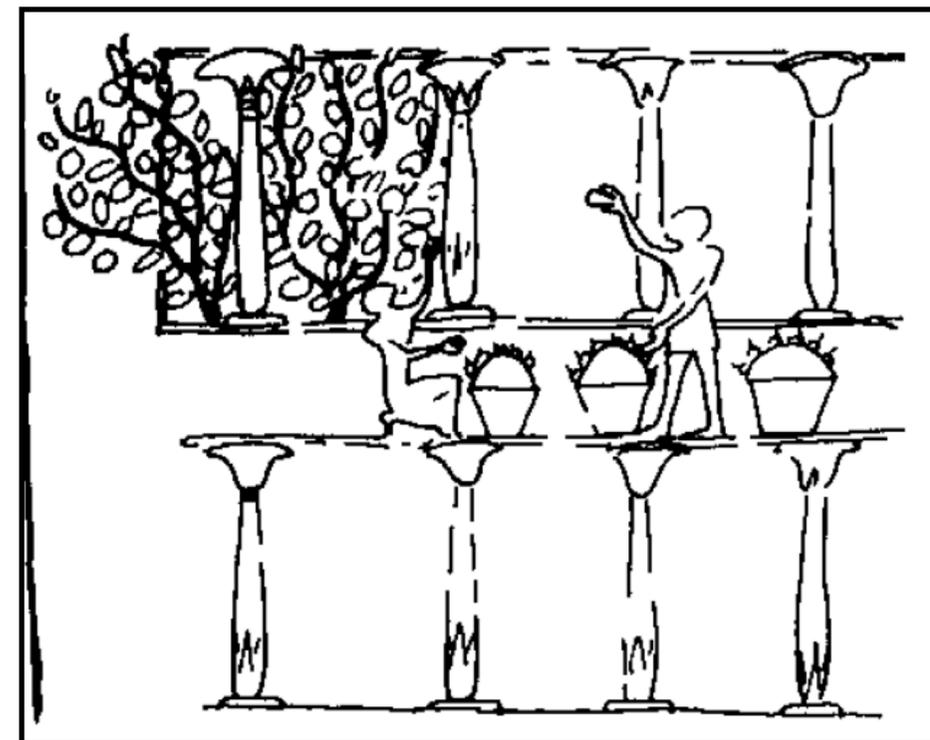
Françoise Laroche-Traunecker (1), Claude Traunecker (1)

Afficher plus de détails

1 ARCHIMEDE - Étude des Civilisations de l'Antiquité : de la Préhistoire à Byzance

Réglages

Extraction de l'image et de sa légende



7. Pergola avec plants de vigne entre les colonnes. Seuls les deux premiers plants sont reproduits. Tombe de Nebamon (TT 90).

- “
- Cette image semble être une illustration représentant une scène de jardinage ou d'agriculture dans un cadre architectural antique, avec des colonnes en style égyptien.
  - L'image est structurée en deux niveaux, chacun étant soutenu par des colonnes d'inspiration égyptienne avec des chapiteaux stylisés :
    - Au niveau supérieur, deux personnages interagissent avec des plantes placées dans des récipients ou des jardinières. L'un des personnages semble être en train de cueillir ou de manipuler une plante, tandis que l'autre pourrait également être en train de récolter ou d'effectuer une action similaire.
    - Derrière les colonnes du niveau supérieur, on observe des branches ou des vignes couvertes de feuilles, ce qui donne l'impression d'un jardin ou d'un espace verdoyant.
  - En somme, cette illustration pourrait faire partie d'un décor mural ou d'un bas-relief visant à documenter ou célébrer les pratiques agricoles et la culture des jardins dans un cadre sacré ou symbolique.
- ”

Antiquité

Art égyptien

Art méditerranéen ancien

Civilisation ancienne

Bas-relief

Fresque

Architecture antique

Colonnes égyptiennes

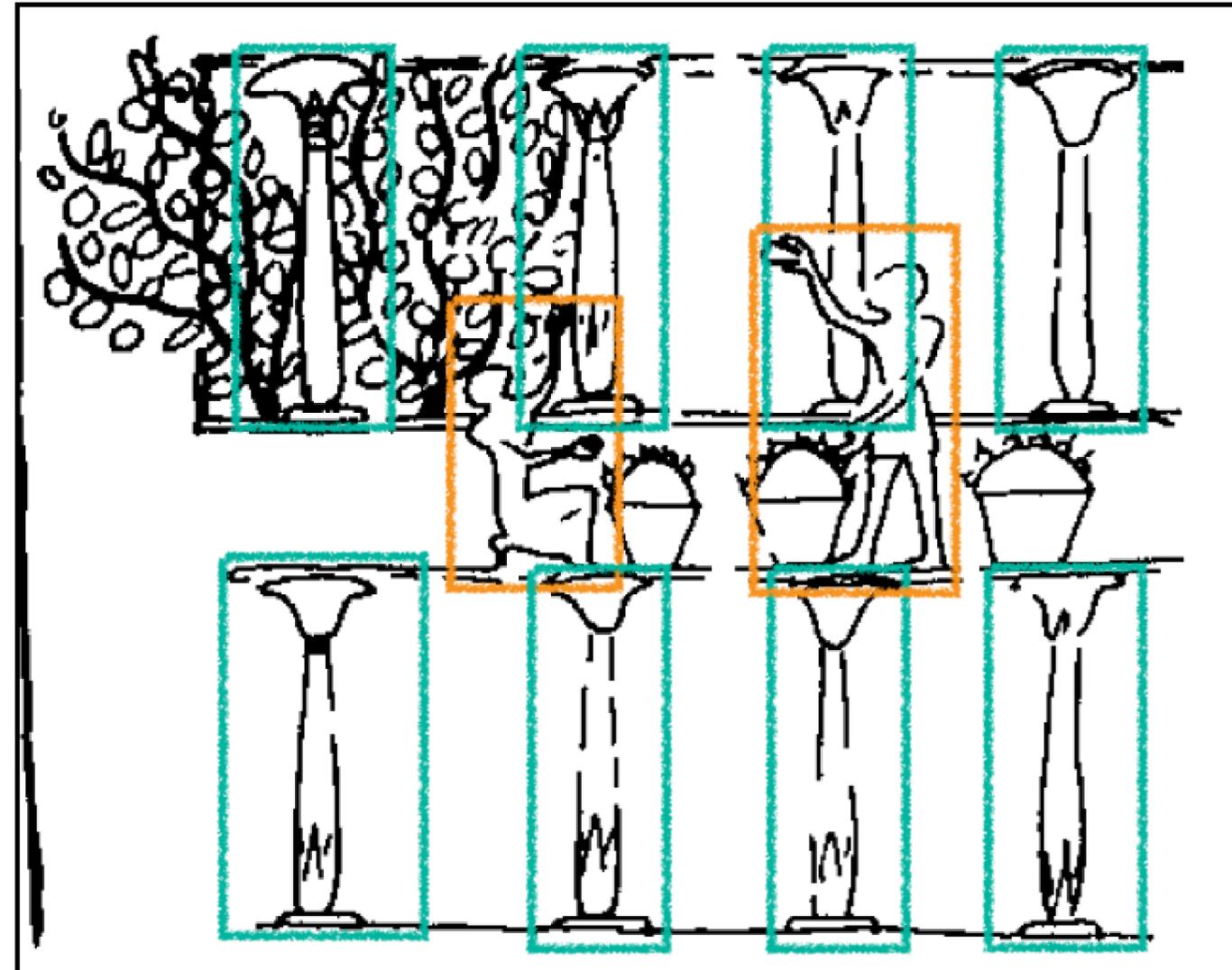
Scène quotidienne

Agriculture

Jardinage

Plantes et végétation

Identification des éléments au sein des images



colonnes  
personnages

Mise en perspective d'éléments visuellement différents mais à la sémantique commune

The image displays two overlapping browser windows from the 'isidore.science' digital archive. The left window shows the record for 'CNRS-CFEETK 73446. Karnak, KIU 38 / stèle d'enceinte de Ramsès III XXe dynastie / Ramsès III'. The right window shows 'CNRS-CFEETK 205004. Karnak, KIU 6721 / Cour sud du Ve pylône, Scène, mur nord, de la maât XXe dynastie / Ramsès III'. Orange arrows point from a rectangular box in the image of the right record to a similar box in the image of the left record, indicating a shared visual element. The interface includes search bars, document titles, and metadata sections.

**CNRS-CFEETK 73446. Karnak, KIU 38 / stèle d'enceinte de Ramsès III XXe dynastie / Ramsès III**

**CNRS-CFEETK 205004. Karnak, KIU 6721 / Cour sud du Ve pylône, Scène, mur nord, de la maât XXe dynastie / Ramsès III**

# Conclusion

---

# Pistes de réflexions

- Quels métiers sont nécessaires pour la mise en place des différentes briques mais également les exploiter et les maintenir sur le long terme) ?
  - Cela induit des cycles de formations continues et des partenariats entre les chercheurs, les savoir-faire scientifiques du traitement, l'infrastructure et des entreprises.
- Quel cycle de vie pour ces instruments ?
  - Quelles méthodes mettre en place pour stabiliser les instruments (de leur gouvernance à leur financement) tout en ayant des cycles de vie épousant l'évolution des besoins ?
- Questionnement autour du fait de ne pas s'attacher qu'aux LLM : comment adapter les outils de la GenIA aux besoins des instruments ? Comment cibler mieux les LLM, les spécialiser et utiliser de l'IA classique quand elle est meilleure ?
- Quelles conséquences sur les pratiques de recherche et quel rapport entre le chercheur, les données qu'ils exploitent et ce qu'ils publient ? En d'autres termes, comment les outils façonnent les pratiques et comment les pratiques les influencent en retour ? À l'extérieur cela pose beaucoup de question en matière de sociologie de la recherche.