

# L'IA générative comme mémoire artificielle

## Usages, communs et auctorialités distribuées

**Julien Schuh**

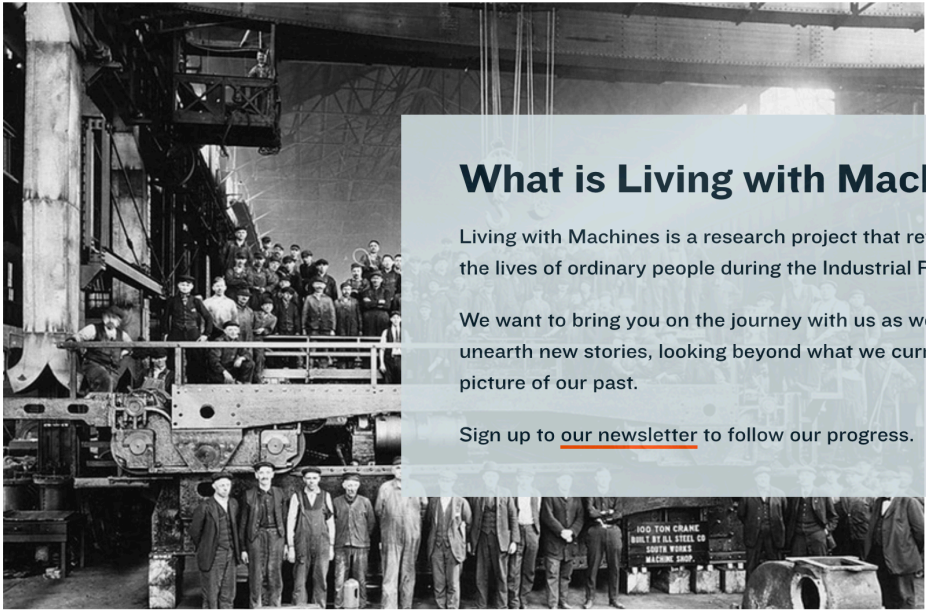
Université Paris Nanterre (CSLF / MSH Mondes)

livingwithmachines.ac.uk

**LIVING WITH MACHINES**

About Achievements Exhibition Events Blog Team Contact Us

Twitter GitHub Email

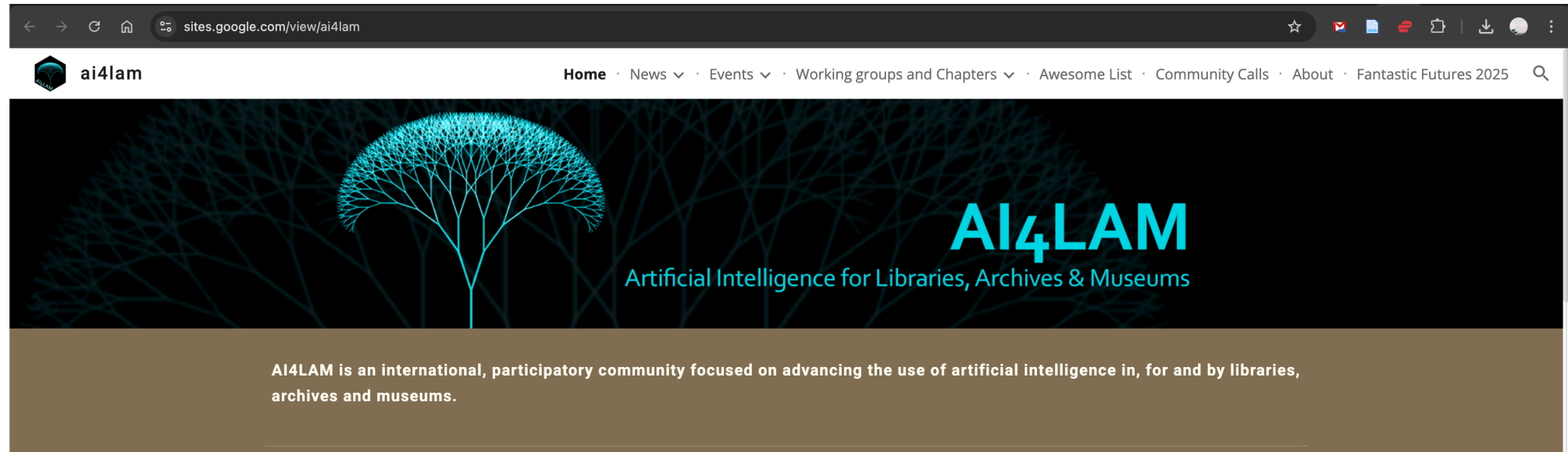


### What is Living with Machines?

Living with Machines is a research project that rethinks the impact of technology on the lives of ordinary people during the Industrial Revolution.

We want to bring you on the journey with us as we use research and technology to unearth new stories, looking beyond what we currently know to reveal a richer picture of our past.

Sign up to [our newsletter](#) to follow our progress.



ai4lam Home · News · Events · Working groups and Chapters · Awesome List · Community Calls · About · Fantastic Futures 2025

# AI4LAM

Artificial Intelligence for Libraries, Archives & Museums

AI4LAM is an international, participatory community focused on advancing the use of artificial intelligence in, for and by libraries, archives and museums.

## October 2025 Community Call

The October Community Call is set for Tuesday, October 21, 8:00 AM California | 11:00 AM Washington DC | 16:00 UK | 17:00 Oslo & Paris | 01:00 Sydney. We will focus on archives. Luis Antonio Morro González (CEO and Founder of Transkriptorium AI, SL, a Spinoff of the Universitat Politècnica de València - UPV) will discuss TRANSKRIPTORIUM, and Markus Stauffiger (4eyes GmbH) will present on Swiss Archippanion archival work support. More information to follow.



## AI4LAM Welcomes Dr. Ines Vodopivec to Lead Strategic Development

January 2, 2025 – AI4LAM is pleased to announce that Associate Professor Dr. Ines Vodopivec has joined the community to help guide its strategic development. Her appointment as Secretary General comes as AI4LAM prepares to strengthen its organizational framework and expand its role as a collaborative platform for advancing AI in the cultural heritage sector.

# Plan de la communication

1. Les usages de l'**IA générative** dans la recherche
2. L'IA comme **mémoire artificielle**
3. **Cartographier** les espaces latents
4. **Communs** vs extraction privée
5. Exemples de **projets** (autour du consortium pictorIA)

← → ↻ 🏠 🌐 revue30.org/documents/atelier\_ia\_30\_janvier/ ☆ 📧 📁 📄 📌 📂 📎

# Revue 3.0

Écrire Transmettre Découvrir

fr | en À propos Équipe Partenaires Évènements Projets Documents Blog

## Compte-rendu de l'atelier IA du 30 janvier 2025 (Frédéric Clavert)

### Introduction des Ateliers IA par Marcello Vitali-Rosati

L'idée de ces ateliers est de commencer échanges et réflexions communes sur l'implémentation d'algorithmes de ce qu'on appelle l'IA dans nos pratiques d'écriture et de publications d'articles scientifiques.

Dans la communauté DH, la réflexion sur l'implémentation d'algorithme remonte à 70 ans. Bien que l'introduction des *transformers* ait apporté des changements, on ne peut pas parler d'une révolution. Notre objectif avec cet atelier est de s'éloigner de la généralité que "IA = ChatGPT" et de la tendance actuelle à l'adoption d'applications *mainstream* sans réflexion.

Ces ateliers se veulent être un lieu de réflexion sur la théorie et les questions infrastructurelles liées à l'implémentations de technologies relevant de l'IA. C'est aussi l'opportunité de réfléchir aux besoins sur ce sujet et potentiellement de financer des projets pilotes et expérimentations dans le cadre du partenariat Revue3.o.

### Présentation par Frédéric Clavert de la fonctionnalité *explain code* du JDH

Le *Journal of Digital History*, est une revue initiée en 2020, avec une première publication en 2021. L'intégration de technologies de l'IA y a davantage été pensé pour le lecteur que pour les auteur·ice·s.

sciedirect.com/science/article/pii/S0160791X2500003X?via%3Dihub

ScienceDirect Journals & Books Help Search My account Sign in

View PDF Download full issue

Technology in Society  
Volume 81, June 2025, 102813

# Generative Artificial Intelligence (GenAI) in the research process – A survey of researchers' practices and perceptions

Jens Peter Andersen<sup>a</sup>, Lise Degn<sup>a</sup>, Rachel Fishberg<sup>a</sup>, Ebbe K. Graversen<sup>a</sup>, Serge P.J.M. Horbach<sup>b</sup>, Evanthia Kalpazidou Schmidt<sup>a</sup>, Jesper W. Schneider<sup>a</sup>, Mads P. Sørensen<sup>a</sup>

https://doi.org/10.1016/j.techsoc.2025.102813

Under a Creative Commons license

Open access

### Highlights

- Survey of Danish researchers shows three views of GenAI: "work horse," "language assistant only," & "research accelerator."
- Language editing and data analysis tasks seen as positive uses; experiment design and peer review more controversial.
- Technical fields and junior researchers report higher GenAI usage, with no gender differences.
- GenAI usage is restricted to a limited number of use cases, with more intensive use associated with more positive attitudes.
- Factor analysis shows varied comfort levels with GenAI in terms of

### Recommended articles

The politics of participatory epidemiology: Technologies, social media and influenza...  
Health Policy and Technology, Volume 6, Issue 2, 2017, ...  
Gabriel Blouin-Genest, Allison Miller

Impact of national university patenting on innovation: Researcher analysis in Japan  
Technology in Society, Volume 81, 2025, Article 102806  
Kazuyuki Motohashi, ..., Seokbeom Kwon

Who's afraid of automation? Examining determinants of fear of automation in six...  
Technology in Society, Volume 81, 2025, Article 102782  
Renata Włoch, ..., Satia Rožnyek

### Article Metrics

Citations

Citation Indexes	19
------------------	----

Captures

Mendeley Readers	126
------------------	-----

Mentions

News Mentions	3
---------------	---

PLUMX View details

FEEDBACK

Outline

- Highlights
- Abstract
- Keywords
- 1. Introduction
- 2. Methods
- 3. Results
- 4. Discussion
- 5. Conclusion and implications

CRediT authorship contribution statement

Ethics declaration

Declaration of competing interest

Acknowledgements

Appendix A. Supplementary data

Data availability

References

Show full outline

Cited by (19)

Figures (4)



## Notre mission

L'Obvia identifie les enjeux sociétaux de l'intelligence artificielle et du numérique, et contribue à des solutions qui placent les êtres vivants et la biosphère au centre de leur cycle de développement et d'utilisation.

Notre communauté de recherche produit des connaissances ouvertes qui renforcent les capacités individuelles et collectives, en collaboration avec la société civile, les acteurs publics, l'industrie et les développeurs.

L'intégration croissante des outils d'IA dans nos sociétés s'effectue dans un contexte marqué par un manque flagrant de régulation, laissant place à des usages trop souvent opaques, à des rapports de force déséquilibrés et à une gouvernance technologique largement dominée par des intérêts privés ou étatiques. Ce manque d'encadrement est souligné dans ce rapport par l'ensemble des axes de recherche de l'Obvia, qui mettent en lumière des zones grises entourant le développement et l'usage de l'IAG.

OBVIA, *État de la situation sur les impacts sociétaux de l'intelligence artificielle et du numérique*, Québec, OBVIA, 2025. URL: <https://www.obvia.ca/ressources/etat-de-la-situation-sur-les-impacts-societaux-de-lintelligence-artificielle-et-du-numerique-2025>

eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=OJ:L\_202401689

RÈGLEMENT (UE) ...

- > préambule
- > dispositif
- Formules finales
- annexe I
- annexe II
- annexe III
- annexe IV
- annexe V
- annexe VI
- annexe VII
- annexe VIII
- annexe IX
- annexe X
- annexe XI
- annexe XII
- annexe XIII

Journal officiel de l'Union européenne

FR  
Série L

2024/1689 12.7.2024

**RÈGLEMENT (UE) 2024/1689 DU PARLEMENT EUROPÉEN ET DU CONSEIL**  
**du 13 juin 2024**  
**établissant des règles harmonisées concernant l'intelligence artificielle et modifiant les règlements (CE) n° 300/2008, (UE) n° 167/2013, (UE) n° 168/2013, (UE) 2018/858, (UE) 2018/1139 et (UE) 2019/2144 et les directives 2014/90/UE, (UE) 2016/797 et (UE) 2020/1828 (règlement sur l'intelligence artificielle)**

**(Texte présentant de l'intérêt pour l'EEE)**

LE PARLEMENT EUROPÉEN ET LE CONSEIL DE L'UNION EUROPÉENNE,

vu le traité sur le fonctionnement de l'Union européenne, et notamment ses articles 16 et 114,

vu la proposition de la Commission européenne,

après transmission du projet d'acte législatif aux parlements nationaux,

vu l'avis du Comité économique et social européen <sup>(1)</sup>,

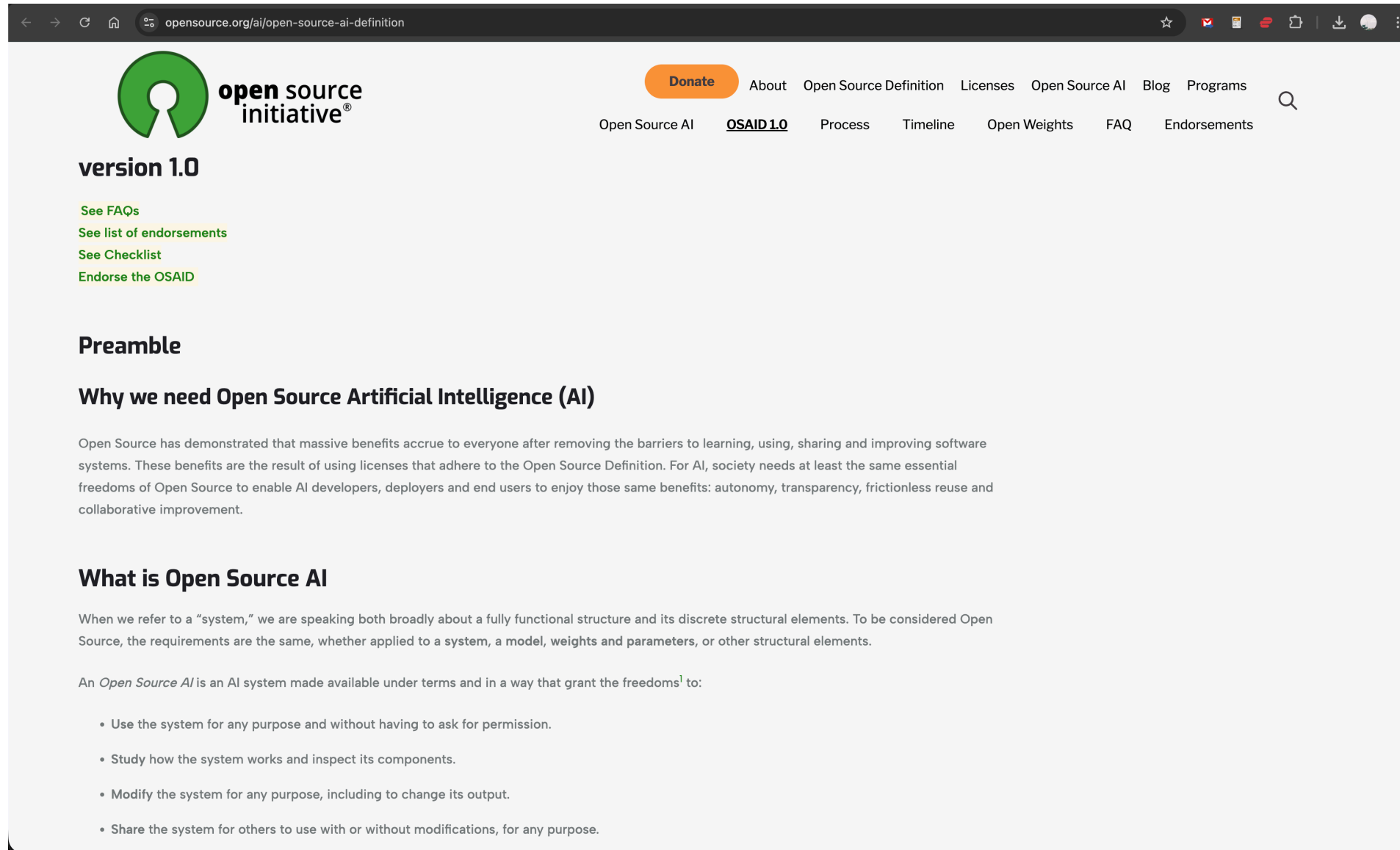
vu l'avis de la Banque centrale européenne <sup>(2)</sup>,

vu l'avis du Comité des régions <sup>(3)</sup>,

statuant conformément à la procédure législative ordinaire <sup>(4)</sup>,

considérant ce qui suit:

- (1) L'objectif du présent règlement est d'améliorer le fonctionnement du marché intérieur en établissant un cadre juridique uniforme, en particulier pour le développement, la mise sur le marché, la mise en service et l'utilisation de systèmes d'intelligence artificielle (ci-après dénommés «systèmes d'IA») dans l'Union, dans le respect des valeurs de l'Union, de promouvoir l'adoption de l'intelligence artificielle (IA) axée sur l'humain et digne de confiance tout en garantissant un niveau élevé de protection de la santé, de la sécurité et des droits fondamentaux consacrés dans la Charte des droits fondamentaux de l'Union européenne (ci-après dénommée «Charte»), y compris la démocratie, l'état de droit et la protection de l'environnement, de protéger contre les effets néfastes des systèmes d'IA dans l'Union, et de soutenir l'innovation. Le présent règlement garantit la libre circulation transfrontière des biens et services fondés sur l'IA, empêchant ainsi les États membres d'imposer des restrictions au développement, à la commercialisation et à l'utilisation de systèmes d'IA, sauf autorisation expresse du présent règlement.
- (2) Le présent règlement devrait être appliqué dans le respect des valeurs de l'Union consacrées dans la Charte, en facilitant la protection des personnes physiques, des entreprises, de la démocratie, de l'état de droit et de l'environnement, tout en stimulant l'innovation et l'emploi et en faisant de l'Union un acteur de premier plan dans l'adoption d'une IA digne de confiance.
- (3) Les systèmes d'IA peuvent être facilement déployés dans un large éventail de secteurs de l'économie et dans de nombreux pans de la société, y compris transfrontières, et peuvent facilement circuler dans toute l'Union. Certains États membres ont déjà envisagé l'adoption de règles nationales destinées à faire en sorte que l'IA soit digne de confiance et sûre et à ce qu'elle soit développée et utilisée dans le respect des obligations en matière de droits fondamentaux. Le fait que les règles nationales divergent peut entraîner une fragmentation du marché intérieur et peut réduire la sécurité juridique pour les opérateurs qui développent, importent ou utilisent des systèmes d'IA. Il convient donc de garantir un niveau de protection cohérent et élevé dans toute l'Union afin de parvenir à une IA digne de confiance, et d'éviter les divergences qui entravent la libre circulation, l'innovation, le déploiement et l'adoption des systèmes d'IA et des produits et services connexes au sein du marché intérieur, en établissant des obligations uniformes pour les opérateurs et en garantissant la protection uniforme des raisons impérieuses d'intérêt général et des droits des citoyens dans l'ensemble du marché intérieur sur la base de l'article 114 du traité sur le fonctionnement



The screenshot shows the Open Source Initiative (OSI) website page for the Open Source AI Definition. The browser address bar shows the URL: `opensource.org/ai/open-source-ai-definition`. The page features the OSI logo (a green circle with a white keyhole shape) and the text "open source initiative®". A navigation menu includes "Donate" (in an orange button), "About", "Open Source Definition", "Licenses", "Open Source AI", "Blog", and "Programs". Below the menu, there are links for "Open Source AI", "OSAID 1.0", "Process", "Timeline", "Open Weights", "FAQ", and "Endorsements". A search icon is also present.

**version 1.0**

[See FAQs](#)  
[See list of endorsements](#)  
[See Checklist](#)  
[Endorse the OSAID](#)

## Preamble

### Why we need Open Source Artificial Intelligence (AI)

Open Source has demonstrated that massive benefits accrue to everyone after removing the barriers to learning, using, sharing and improving software systems. These benefits are the result of using licenses that adhere to the Open Source Definition. For AI, society needs at least the same essential freedoms of Open Source to enable AI developers, deployers and end users to enjoy those same benefits: autonomy, transparency, frictionless reuse and collaborative improvement.

### What is Open Source AI

When we refer to a "system," we are speaking both broadly about a fully functional structure and its discrete structural elements. To be considered Open Source, the requirements are the same, whether applied to a system, a model, weights and parameters, or other structural elements.

An *Open Source AI* is an AI system made available under terms and in a way that grant the freedoms<sup>1</sup> to:

- Use the system for any purpose and without having to ask for permission.
- Study how the system works and inspect its components.
- Modify the system for any purpose, including to change its output.
- Share the system for others to use with or without modifications, for any purpose.

pleias

PRODUCT TEAM PARTNERS BLOG

# OUR FOUNDATION MODELS

**PROFICIENT AT KEY ENTREPRISE TASKS**

Our models are highly performant at key tasks, such as RAG, agentic search, data transformation, and reasoning over complex multilingual documents in administrative & legal & scientific fields.

**LEAN & FAST**

Our models are currently the most energy-efficient on the market: beating models 10x times their size (Qwen, Llama 3.2, Mistral...), they run (fast) on consumer-grade GPU and in CPU-only environments, without quantisation.

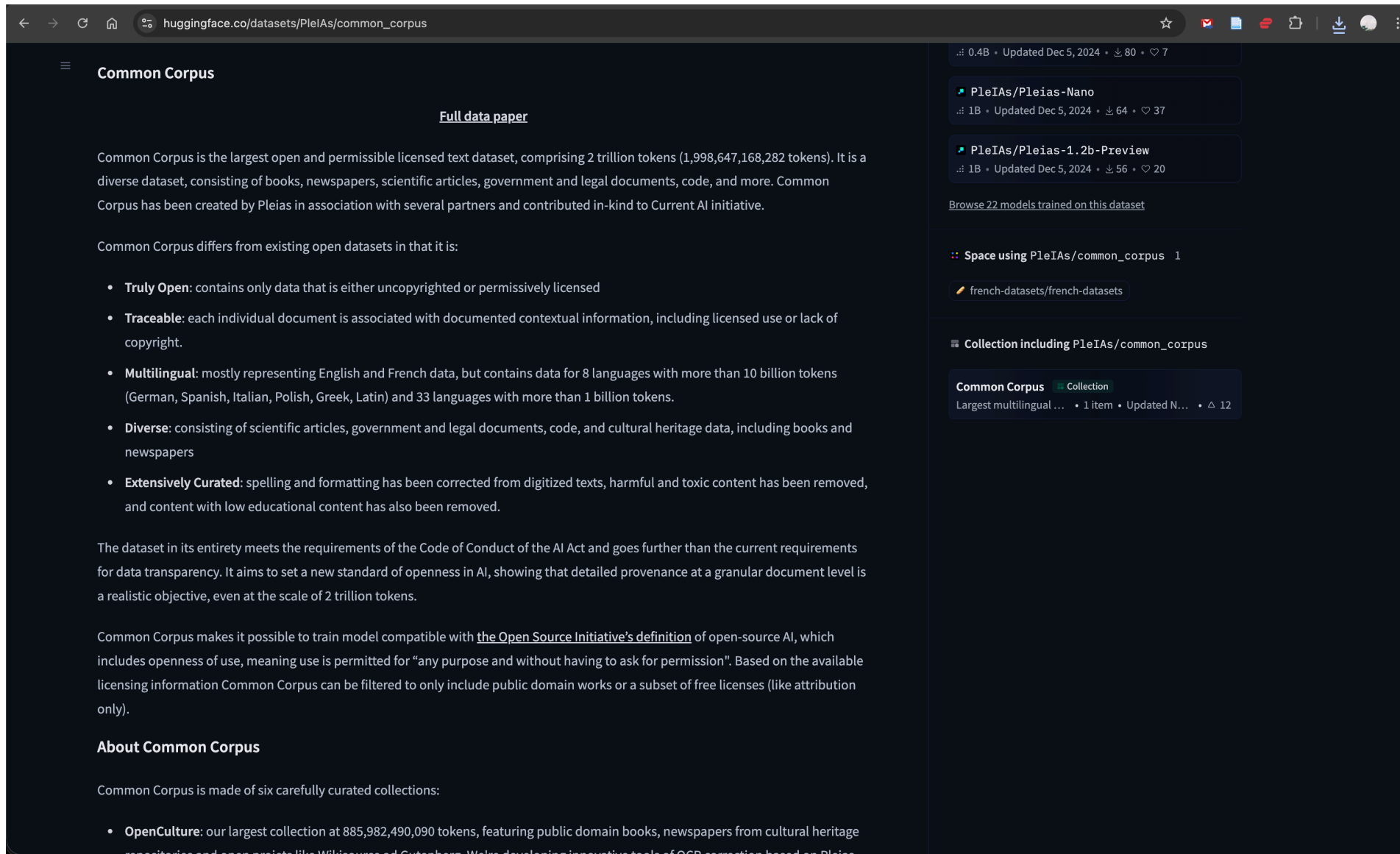
**COMPLIANT & TRANSPARENT**

Open code. Open weights. Open corpora. When we say transparent, we mean it. Designed for the most regulated industries, from banking to healthcare.

## TEST PLEIAS 1.0 TODAY

Get Started

## AI ASSISTANTS FOR REGULATED INDUSTRIES AND SENSITIVE USE CASES

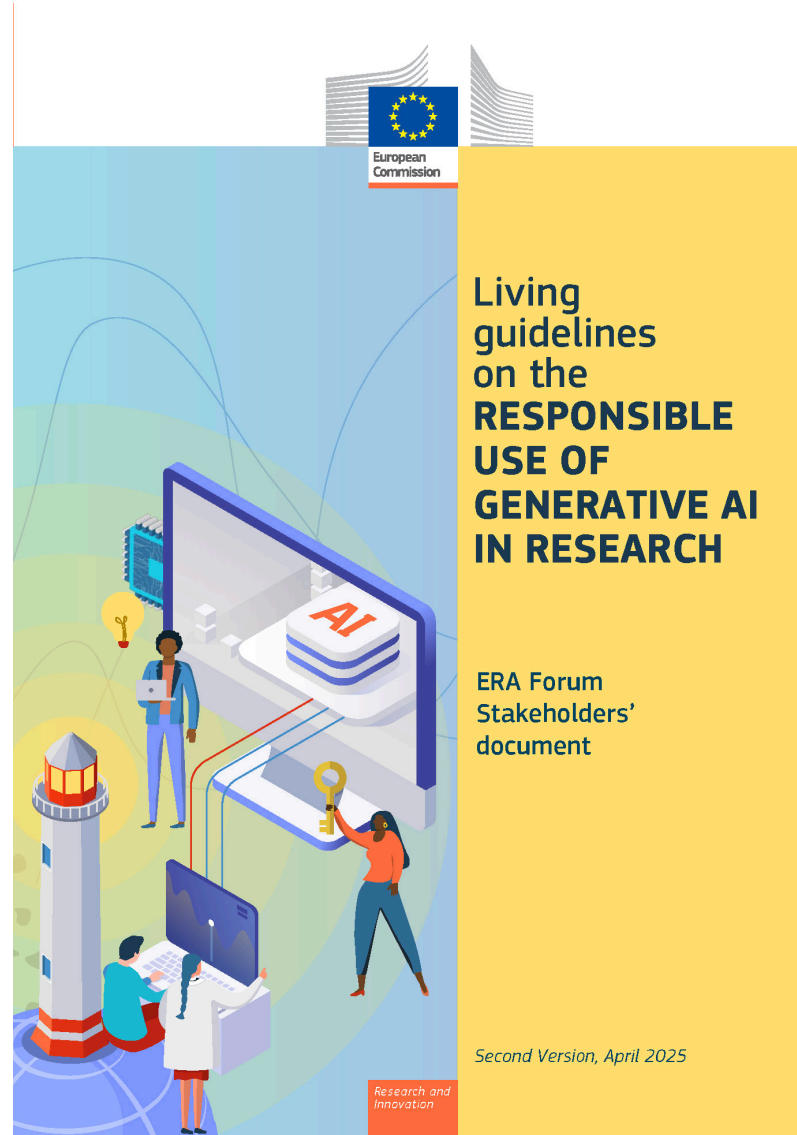


The screenshot shows the Hugging Face dataset page for 'PleIAs/common\_corpus'. The page is in dark mode and contains the following content:

- Common Corpus** (Section Header)
- [Full data paper](#) (Link)
- Common Corpus is the largest open and permissible licensed text dataset, comprising 2 trillion tokens (1,998,647,168,282 tokens). It is a diverse dataset, consisting of books, newspapers, scientific articles, government and legal documents, code, and more. Common Corpus has been created by Pleias in association with several partners and contributed in-kind to Current AI initiative.
- Common Corpus differs from existing open datasets in that it is:
- **Truly Open:** contains only data that is either uncopyrighted or permissively licensed
  - **Traceable:** each individual document is associated with documented contextual information, including licensed use or lack of copyright.
  - **Multilingual:** mostly representing English and French data, but contains data for 8 languages with more than 10 billion tokens (German, Spanish, Italian, Polish, Greek, Latin) and 33 languages with more than 1 billion tokens.
  - **Diverse:** consisting of scientific articles, government and legal documents, code, and cultural heritage data, including books and newspapers
  - **Extensively Curated:** spelling and formatting has been corrected from digitized texts, harmful and toxic content has been removed, and content with low educational content has also been removed.
- The dataset in its entirety meets the requirements of the Code of Conduct of the AI Act and goes further than the current requirements for data transparency. It aims to set a new standard of openness in AI, showing that detailed provenance at a granular document level is a realistic objective, even at the scale of 2 trillion tokens.
- Common Corpus makes it possible to train model compatible with [the Open Source Initiative's definition](#) of open-source AI, which includes openness of use, meaning use is permitted for "any purpose and without having to ask for permission". Based on the available licensing information Common Corpus can be filtered to only include public domain works or a subset of free licenses (like attribution only).
- About Common Corpus** (Section Header)
- Common Corpus is made of six carefully curated collections:
- **OpenCulture:** our largest collection at 885,982,490,090 tokens, featuring public domain books, newspapers from cultural heritage repositories and open projects like Wikisource and Gutenberg. We're developing innovative tools of OCR correction based on Pleias

On the right sidebar, there are several sections:

- Dataset statistics: 0.4B • Updated Dec 5, 2024 • ↓ 80 • ♥ 7
- Other datasets:
  - PleIAs/Pleias-Nano**: 1B • Updated Dec 5, 2024 • ↓ 64 • ♥ 37
  - PleIAs/Pleias-1.2b-Preview**: 1B • Updated Dec 5, 2024 • ↓ 56 • ♥ 20
- [Browse 22 models trained on this dataset](#)
- Space using PleIAs/common\_corpus** 1
  - [french-datasets/french-datasets](#)
- Collection including PleIAs/common\_corpus**
  - Common Corpus** (Collection): Largest multilingual... • 1 item • Updated N... • △ 12

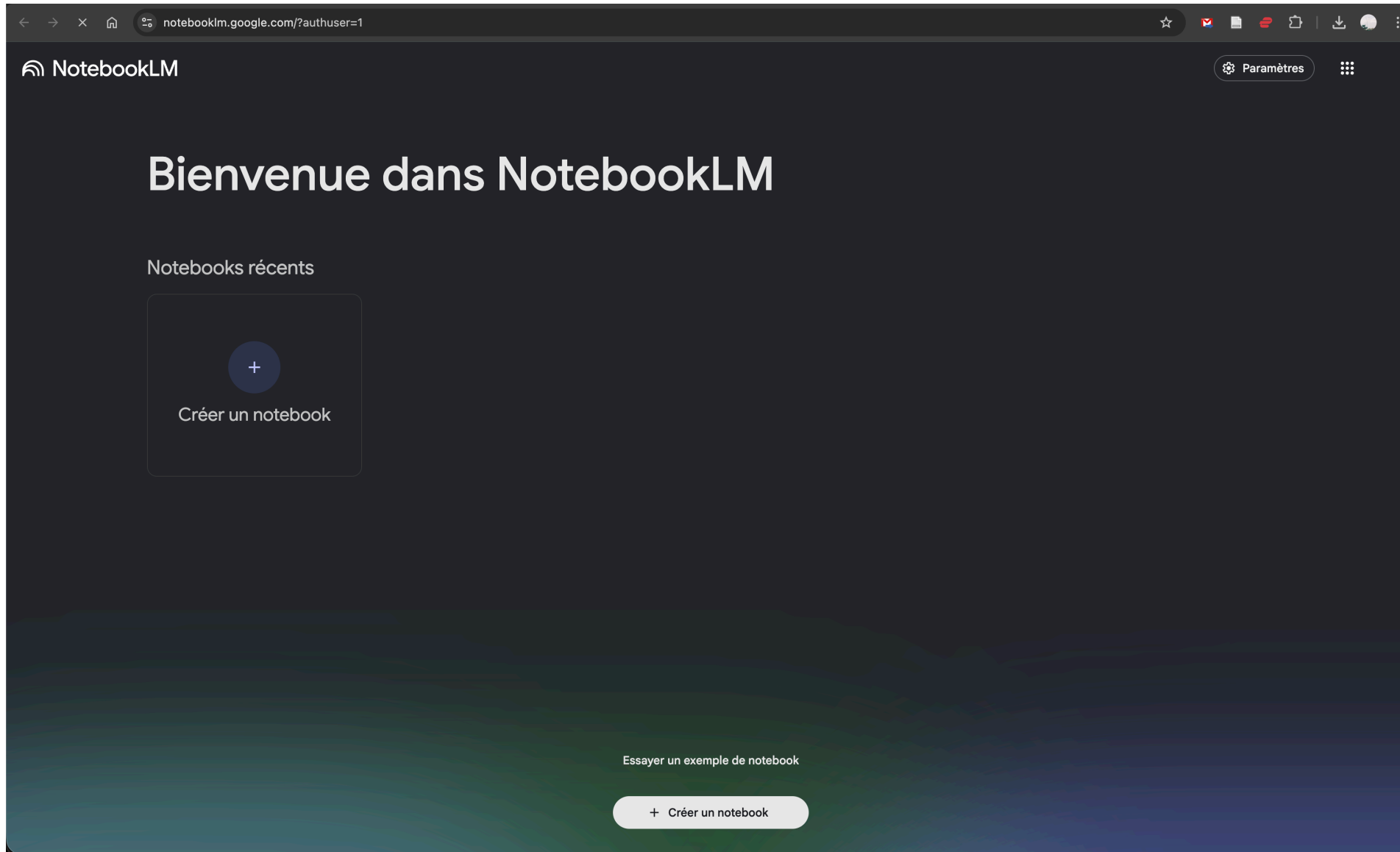


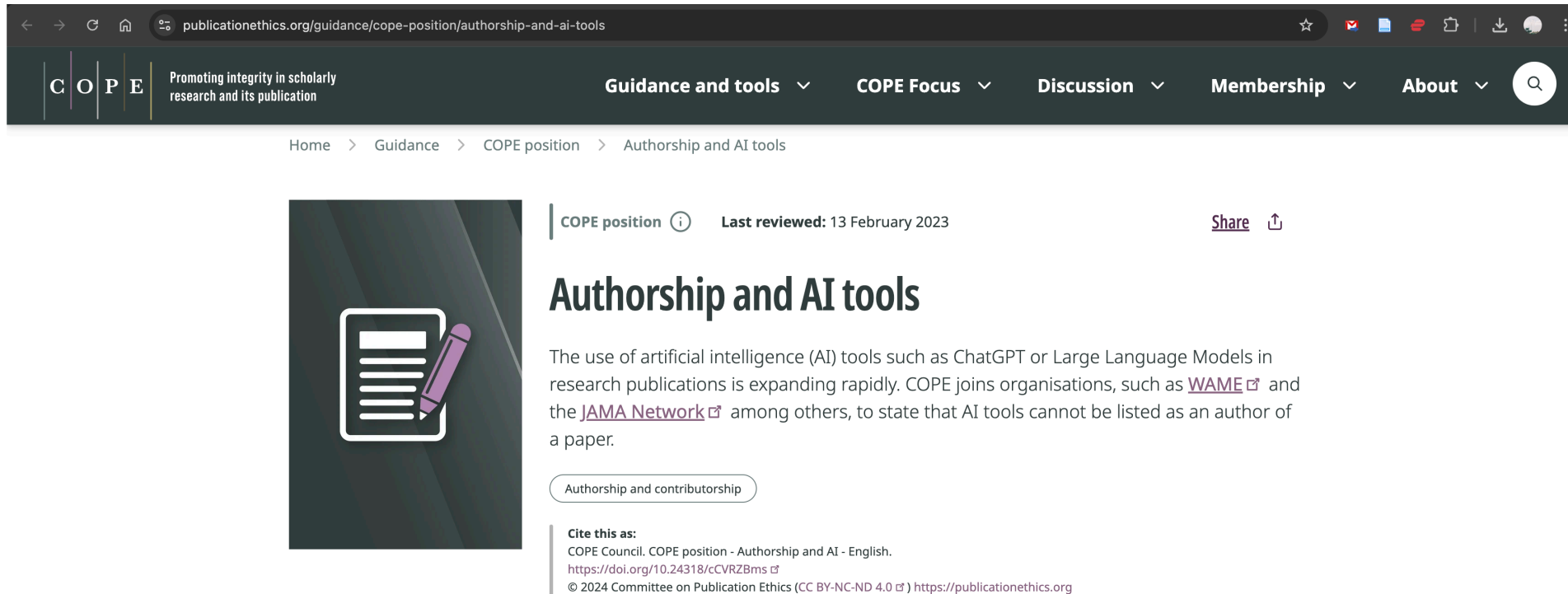
Building on the commonalities of the currently emerging guidelines from various stakeholders, the key principles behind these guidelines for the responsible use of generative AI in research are:

- **Reliability** in ensuring the quality of research, reflected in the design, methodology, analysis and use of resources. This includes aspects related to verifying and reproducing the information produced by the AI for research. It also involves being aware of possible equality and non-discrimination issues in relation to bias and inaccuracies.
- **Honesty** in developing, carrying out, reviewing, reporting and communicating on research transparently, fairly, thoroughly and impartially. This principle includes disclosing that generative AI has been used.

- **Respect** for colleagues, research participants, research subjects, society, ecosystems, cultural heritage and the environment. Responsible use of generative AI should consider the limitations of the technology, its environmental impact<sup>16</sup> and its societal effects (bias, diversity, non-discrimination, fairness and prevention of harm). This includes the proper management of information, respect for privacy, confidentiality and intellectual property rights, and proper citation.
- **Accountability** for the research from idea to publication, for its management and organisation, for training, supervision and mentoring, and for its wider societal impacts. This includes responsibility for all output a researcher produces, underpinned by the notion of human agency and oversight.

Commission européenne, DG Recherche et Innovation, *Living Guidelines on the Responsible Use of Generative AI in Research*, Bruxelles, Union européenne, 2025. URL: [https://research-and-innovation.ec.europa.eu/document/download/2b6cf7e5-36ac-41cb-aab5-0d32050143dc\\_en](https://research-and-innovation.ec.europa.eu/document/download/2b6cf7e5-36ac-41cb-aab5-0d32050143dc_en)





The screenshot shows a web browser displaying the COPE website. The URL is [publicationethics.org/guidance/cope-position/authorship-and-ai-tools](https://publicationethics.org/guidance/cope-position/authorship-and-ai-tools). The page features a dark header with the COPE logo and navigation menus for 'Guidance and tools', 'COPE Focus', 'Discussion', 'Membership', and 'About'. A breadcrumb trail reads 'Home > Guidance > COPE position > Authorship and AI tools'. The main content area includes a dark square icon with a document and pencil, a 'COPE position' label with an information icon, and a 'Last reviewed: 13 February 2023' date. The title 'Authorship and AI tools' is prominently displayed. The text explains that the use of AI tools like ChatGPT or Large Language Models in research is growing, and COPE, along with WAME and the JAMA Network, has stated that AI tools should not be listed as authors. A 'Share' button is visible. A 'Cite this as:' section provides the COPE Council's name, the page title, the DOI (<https://doi.org/10.24318/cCVRZBms>), and the copyright notice for 2024.

COPE position ⓘ Last reviewed: 13 February 2023 [Share](#) ↗

## Authorship and AI tools

The use of artificial intelligence (AI) tools such as ChatGPT or Large Language Models in research publications is expanding rapidly. COPE joins organisations, such as [WAME](#) and the [JAMA Network](#) among others, to state that AI tools cannot be listed as an author of a paper.

Authorship and contributorship

**Cite this as:**  
COPE Council. COPE position - Authorship and AI - English.  
<https://doi.org/10.24318/cCVRZBms>  
© 2024 Committee on Publication Ethics (CC BY-NC-ND 4.0) <https://publicationethics.org>

## COPE position

The use of artificial intelligence (AI) tools such as ChatGPT or Large Language Models in research publications is expanding rapidly. COPE joins organisations, such as [WAME](#) and the [JAMA Network](#) among others, to state that AI tools cannot be listed as an author of a paper.

AI tools cannot meet the requirements for authorship as they cannot take responsibility for the submitted work. As non-legal entities, they cannot assert the presence or absence of conflicts of interest nor manage copyright and license agreements.





The screenshot shows the website of the Office Français de l'Intégrité Scientifique (Ofis). The main navigation bar includes links for 'Abonnez-vous à l'infolettre', 'Espace RESINT', and 'English'. The header features the Ofis logo and the slogan 'Pour une culture partagée de l'intégrité scientifique'. The main content area is titled 'Publication de la version française du Code de conduite européen pour l'intégrité en recherche (2023)'. It includes a sub-header 'Publié le 9 janvier 2025' and 'Actualités'. The article text discusses the European Code of Conduct for Research Integrity (Code ALLEA) and its French translation, available since December 2024. It mentions that the code is a reference text for research integrity, recognized by the European Commission, and describes the responsibilities of all actors in research. The article also notes that the 2023 revised edition includes changes in data management practices, the General Data Protection Regulation (GDPR), and recent developments in artificial intelligence. A sidebar on the right, titled 'Les autres actualités', lists three other news items: 'Recommandations de l'Ofis relatives au traitement des signalements : les principales nouveautés en une fiche synthétique' (30 octobre 2025), 'L'Ofis publie l'édition 2025 de ses recommandations sur les modalités de dessaisissement d'un RIS' (15 octobre 2025), and 'Formation à destination des référents et référents à l'intégrité scientifique nouvellement nommés : Prendre en main sa fonction. Les éléments fondamentaux de la fonction de RIS' (16 septembre 2025). At the bottom of the main article, there are two buttons: 'Consulter la version française du Code de Conduite Européen 2023' and 'Consulter la version originale en anglais'.

ofis-france.fr/publication-de-la-version-francaise-du-code-de-conduite-europeen-pour-lintegrite-en-recherche-2023/

Abonnez-vous à l'infolettre Espace RESINT English

**Ofis**  
OFFICE FRANÇAIS  
DE L'INTÉGRITÉ SCIENTIFIQUE

Pour une culture partagée  
de l'intégrité scientifique

Qui sommes-nous? À propos de l'intégrité scientifique Recherche & prospective Ressources Espaces thématiques Notes et avis de l'Ofis

## Publication de la version française du Code de conduite européen pour l'intégrité en recherche (2023)

Publié le 9 janvier 2025 Actualités

Le Code de conduite européen pour l'intégrité en recherche (ou Code ALLEA), dans sa traduction française préparée en collaboration avec l'Office français de l'intégrité scientifique, est disponible depuis décembre 2024.

Il s'agit du texte de référence en matière d'intégrité de la recherche, reconnu par la commission européenne pour tous les projets de recherche financés par l'UE. Il décrit la responsabilité de tous les acteurs en matière de promotion de l'intégrité scientifique.

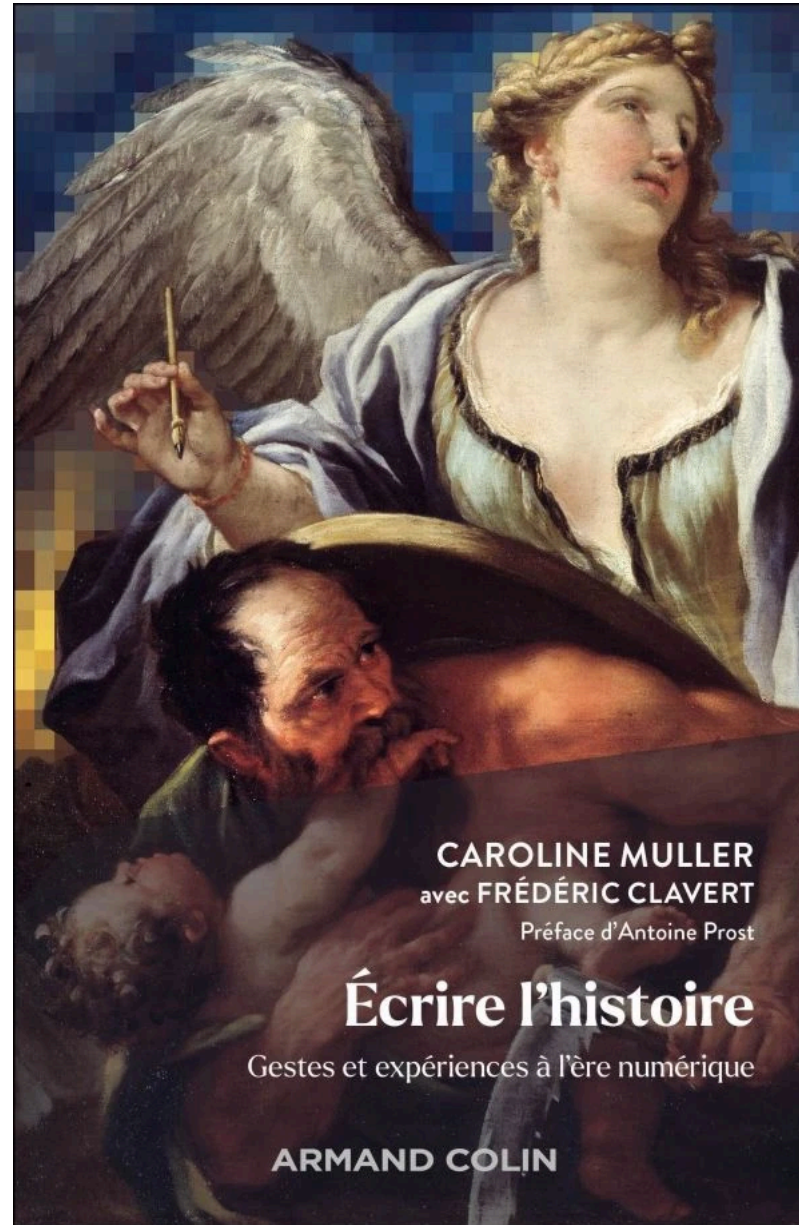
L'édition révisée en 2023 intègre notamment les changements dans les pratiques de gestion des données, le règlement général sur la protection des données (RGPD), et les développements récents en matière de science ouverte, d'évaluation de la recherche, et d'utilisation d'outils d'intelligence artificielle. Elle tient également compte de la sensibilité accrue aux mécanismes de discrimination et d'exclusion et de la nécessité de soutenir l'équité, la diversité et l'inclusion.

Consulter la version française du Code de Conduite Européen 2023

Consulter la version originale en anglais

### Les autres actualités

- Recommandations de l'Ofis relatives au traitement des signalements : les principales nouveautés en une fiche synthétique**  
30 octobre 2025  
[Lire la suite »](#)
- L'Ofis publie l'édition 2025 de ses recommandations sur les modalités de dessaisissement d'un RIS**  
15 octobre 2025  
[Lire la suite »](#)
- Formation à destination des référents et référents à l'intégrité scientifique nouvellement nommés : Prendre en main sa fonction. Les éléments fondamentaux de la fonction de RIS**  
16 septembre 2025  
[Lire la suite »](#)
- Recommandations relatives à la procédure de traitement des manquements à l'intégrité scientifique**  
21 juillet 2025  
[Lire la suite »](#)

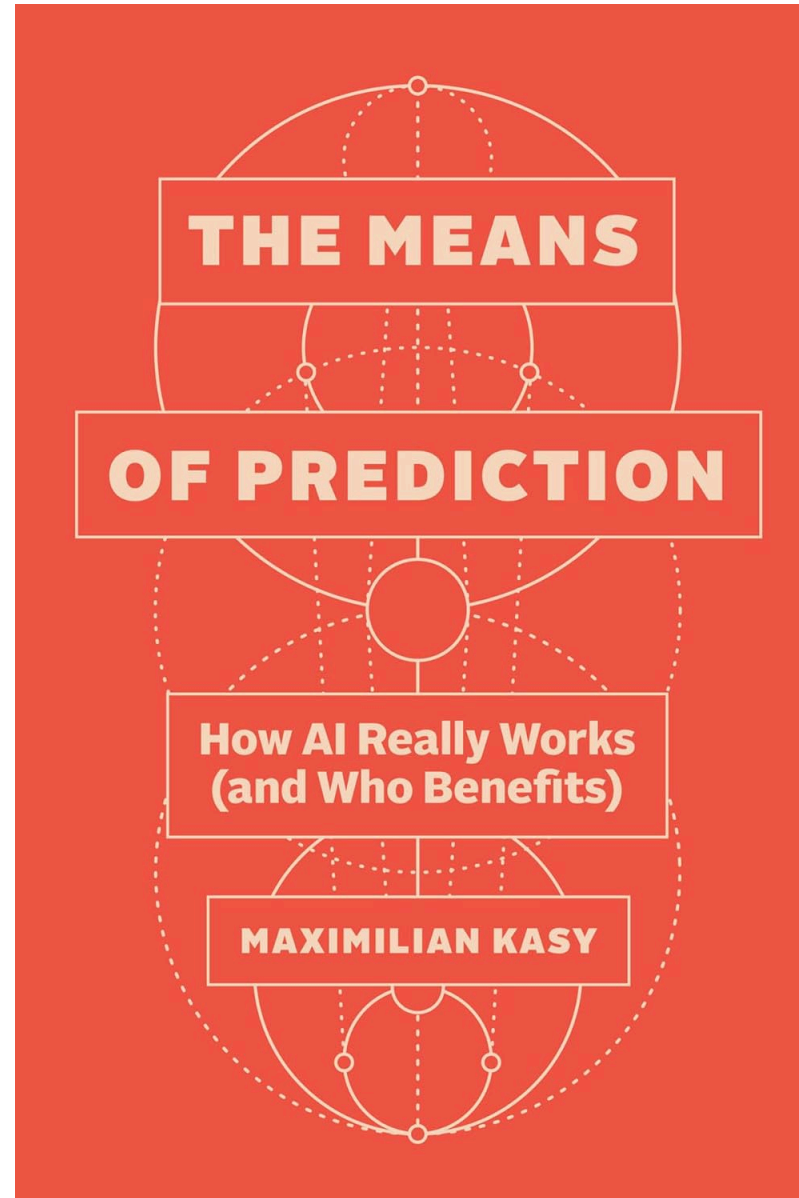


CAROLINE MULLER  
avec FRÉDÉRIC CLAVERT  
Préface d'Antoine Prost

## Écrire l'histoire

Gestes et expériences à l'ère numérique

ARMAND COLIN

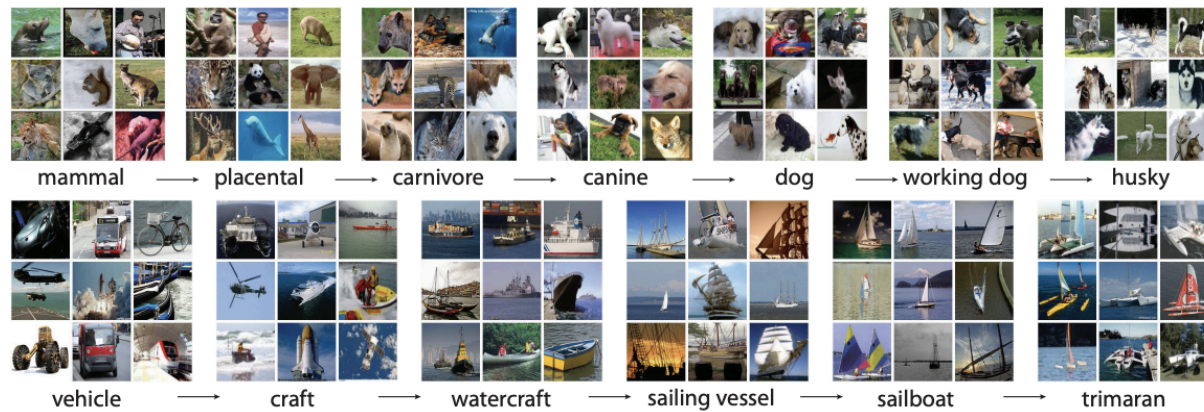
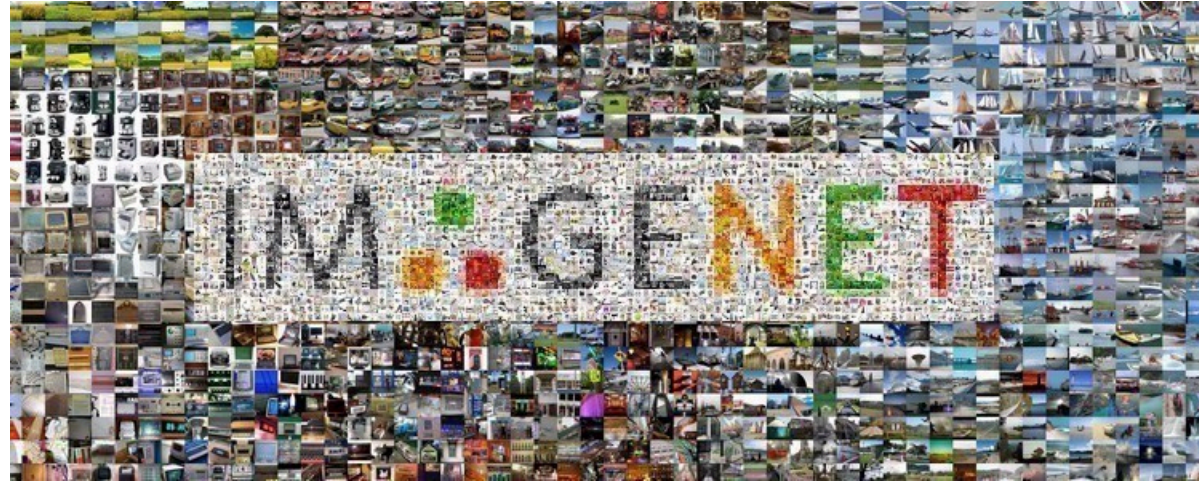


The screenshot shows a web browser window displaying a Springer Nature article. The browser's address bar shows the URL: link.springer.com/article/10.1007/s00146-024-02174-w. The page header includes the Springer Nature logo and a 'Link' label. Navigation links for 'Find a journal', 'Publish with us', 'Track your research', and 'Search' are visible. A user account icon and a shopping cart icon are also present. The main content area features a dark blue header with the article title 'The role of generative AI in academic and scientific authorship: an autopoietic perspective'. Below the title, it indicates 'Main Paper | Open access | Published: 08 January 2025' and 'Volume 40, pages 3225–3235, (2025)'. A 'Download PDF' button is available. The authors listed are Steven Watson, Erik Brezovec, and Jonathan Romic. Metrics show 7032 accesses, 11 citations, and 1 altmetric. A notification states 'This article has been updated'. The abstract begins with 'The integration of generative artificial intelligence (AI), particularly large language models like ChatGPT, presents new challenges as well as possibilities for scientific authorship. This paper draws on social systems theory to offer a nuanced understanding of the interplay between technology, individuals, society and scholarly authorial practices. This contrasts with orthodoxy, where individuals and technology are treated as essentialized'. On the right, there is a thumbnail of the journal cover 'AI & SOCIETY' and a 'Use our pre-submission checklist' link. A table of contents on the right lists sections: Abstract, Introduction, Theoretical background, Authorship, The challenges for scientific authorship and acad..., Scientific authorship as an autopoietic system, and Discussion.

The screenshot shows a web browser displaying an article on the MDPI website. The browser's address bar shows the URL `mdpi.com/2076-0752/13/5/149`. The MDPI logo is in the top left, and navigation links for Journals, Topics, Information, Author Services, Initiatives, and About are in the top center. On the right, there are buttons for 'Sign In / Sign Up' and 'Submit'. Below the navigation is a search bar with fields for 'Title / Keyword', 'Author / Affiliation / Email', 'Arts', and 'All Article Types', along with a 'Search' button and a link to 'Advanced' search.

The article's breadcrumb trail is 'Journals / Arts / Volume 13 / Issue 5 / 10.3390/arts13050149'. On the left side, there is a sidebar with the 'arts' journal logo, buttons for 'Submit to this Journal', 'Review for this Journal', and 'Propose a Special Issue'. Below this is an 'Article Menu' section with an 'Academic Editor' Francisco Tigre Moura, 'Recommended Articles', 'Related Info Link', and 'More by Author Links'. At the bottom of the sidebar is a 'Table of Contents' with a list of sections: Abstract, Introduction, The Contemporary Artwork, Distributed Authorship, Conclusions, Funding, Data Availability Statement, Acknowledgments, and Conflicts of Interest.

The main article content features a title 'The Distributed Authorship of Art in the Age of AI' by Paul Goodfellow. It includes an 'Open Access Article' badge, an 'Order Article Reprints' button, and a 'Share' icon. The author's affiliation is 'Faculty of Design, Informatics and Business, Abertay University, Dundee DD1 1HG, UK'. The article's DOI is `https://doi.org/10.3390/arts13050149`. Submission and publication dates are listed: 'Submission received: 14 July 2024 / Revised: 20 September 2024 / Accepted: 20 September 2024 / Published: 30 September 2024'. A note states '(This article belongs to the Special Issue Artificial Intelligence and the Arts)'. There are buttons for 'Download', 'Browse Figures', and 'Versions Notes'. The 'Abstract' section begins with: 'The distribution of authorship in the age of machine learning or artificial intelligence (AI) suggests a taxonomic system that places art objects along a spectrum in terms of authorship: from pure human creation, which draws directly from the interior world of affect, emotions and ideas, through to co-evolved works created with tools and collective production and finally to works that are largely devoid of human involvement. Human and machine production can be distinguished in terms of motivation, with human production being driven by consciousness and the processing of subjective experience and machinic production being driven by algorithms and the processing of data. However, the expansion of AI entangles the artist in ever more complex webs of production and dissemination, whereby the boundaries between the work of the artist and the work of the networked technologies are increasingly distributed and obscured. From this perspective, AI-generated works are not solely the products of an independent machinic agency but operate in the middle of the spectrum of authorship between human and machine, as they are the consequences of a highly distributed model of production that sit across the algorithms and the underlying information systems and data that support them and the artists who both contribute and extract value. This highly distributed state further transforms the role of the artist from the creator of objects containing aesthetic and conceptual potential to the translator and curator of such objects.' The 'Keywords' are 'AI; contemporary art; authorship; systems'. The '1. Introduction' section is partially visible at the bottom.



Source : <https://blog.roboflow.com/introduction-to-imagenet/>

The screenshot displays the DALL-E 2 web interface. At the top, a navigation bar includes the OpenAI logo, a search icon, and a 'Log in' button. Below this, a central heading reads: "DALL-E 2 can create original, realistic images and art from a text description. It can combine concepts, attributes, and styles." A horizontal menu offers options for "Image generation" (which is selected), "Outpainting", "Inpainting", and "Variations". On the left side, a vertical sidebar lists various OpenAI services: Research, Safety, ChatGPT, Sora, API Platform, For Business, Stories, Company, and News. The main content area is split into two panels. The left panel, labeled "Input", contains the text: "A photorealistic image of an astronaut riding a horse". The right panel shows the resulting image: a highly detailed, photorealistic illustration of an astronaut in a full space suit riding a white horse against a dark, star-filled cosmic background. Below the image, a color calibration bar is visible. At the bottom of the interface, a footer contains the same introductory text as the top, followed by an "Ask ChatGPT" button and an upward-pointing arrow icon.

stability.ai

Models ▾ Applications ▾ Deployment ▾ Company ▾ News Resources ▾ [Contact Us](#)

## Image Models

SD3.5 SD3.0 SDXL

### Stable Diffusion 3.5 Large

At 8 billion parameters, with superior quality and prompt adherence, this base model is the most powerful in the Stable Diffusion family. This model is ideal for professional use cases at 1 megapixel resolution.

[Download Weights](#)

### Stable Diffusion 3.5 Large Turbo


A distilled version of Stable Diffusion 3.5 Large generates high-quality images with exceptional prompt adherence in just 4 steps, making it considerably faster than Stable Diffusion 3.5 Large.

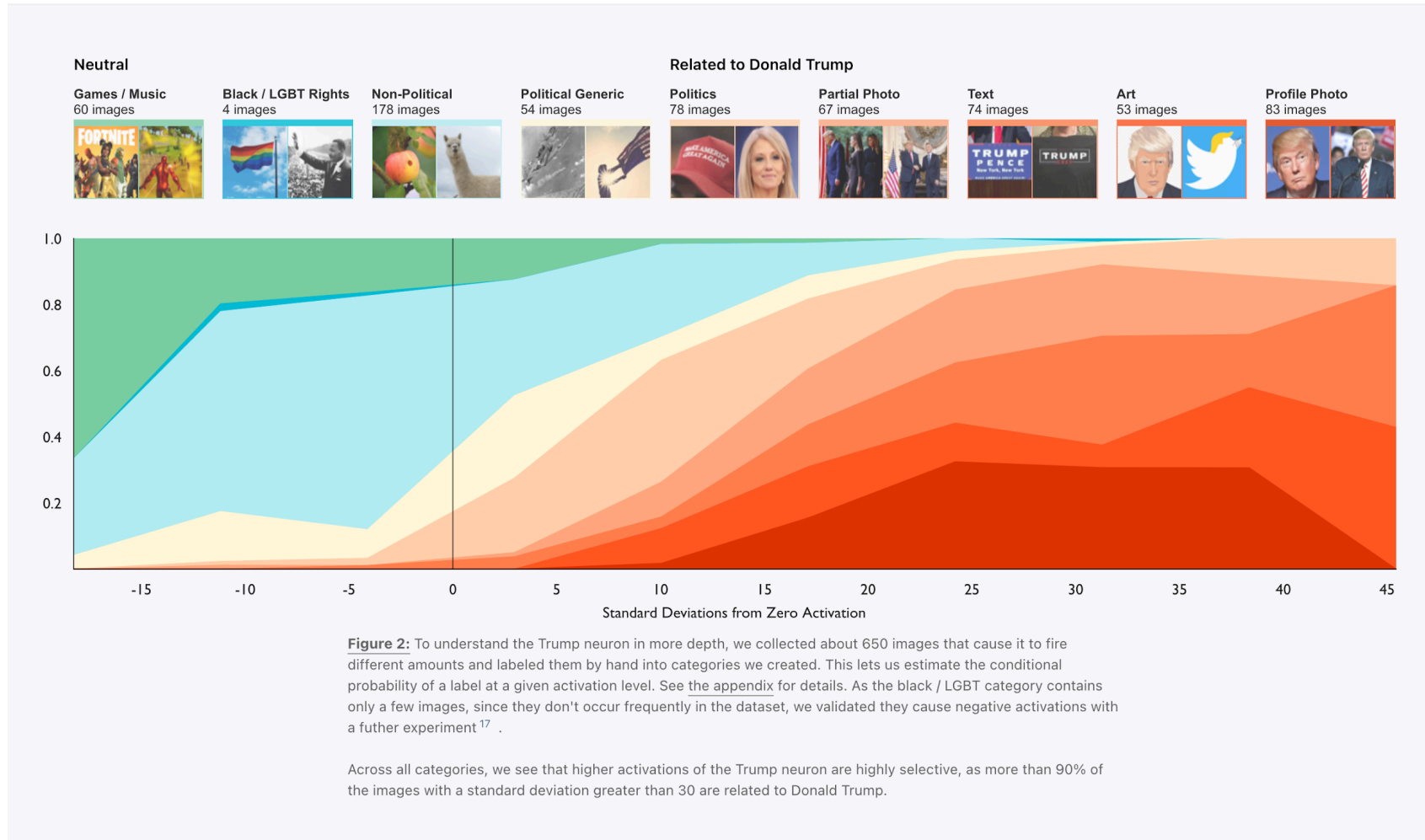
[Download Weights](#)

### Stable Diffusion 3.5 Medium

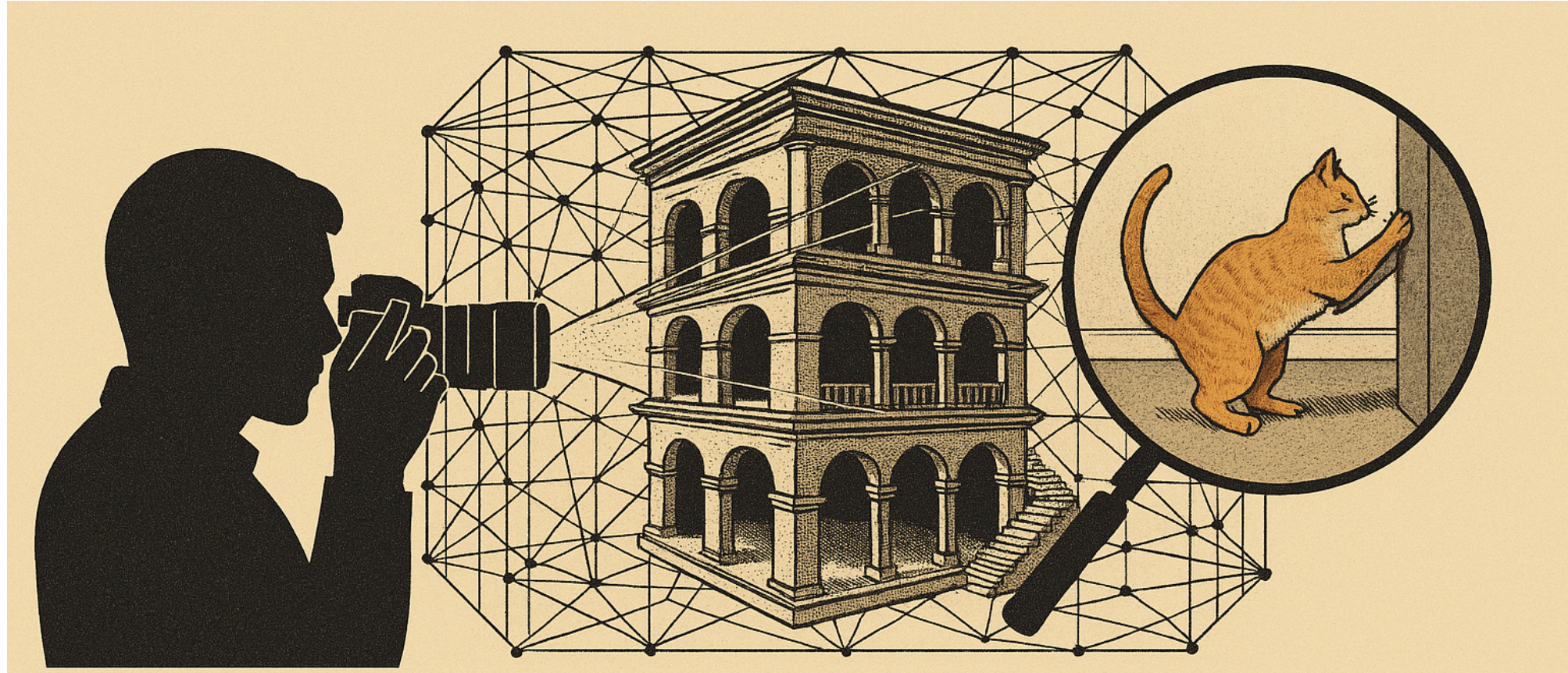
At 2.5 billion parameters, with improved MMDiT-X architecture and training methods, this model is designed to run “out of the box” on consumer hardware, striking a balance between quality and ease of customization. It is capable of generating images ranging between 0.25 and 2 megapixel resolution.

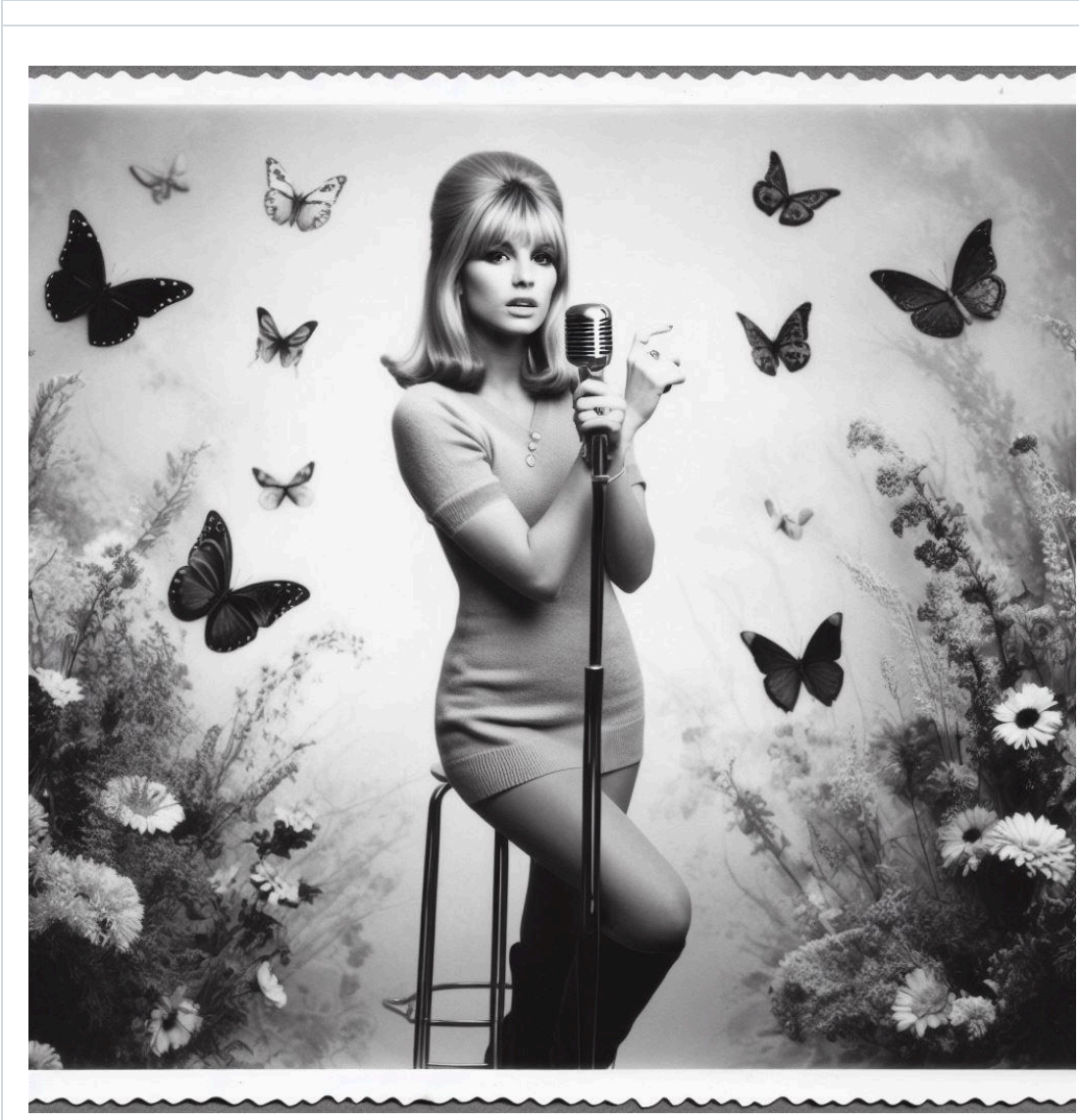
[Download Weights](#)





Source : Gabriel Goh, et al. "Multimodal Neurons in Artificial Neural Networks", *Distill*, vol. 6, n° 3, Mar. 2021, p. e30. URL: <https://doi.org/10.23915/distill.00030>.





**empreinte > data**



**engrammata > generata**



« La mémoire sociale engloutit dans les livres, en quelques décennies, toute l'Antiquité, l'histoire des grands peuples, la géographie et l'ethnographie d'un monde devenu définitivement sphérique, la philosophie, le droit, les sciences, les arts, les techniques et une littérature traduite de vingt langues différentes »

André Leroi-Gourhan, *Le geste et la parole. 2: La mémoire et les rythmes*, p. 70.

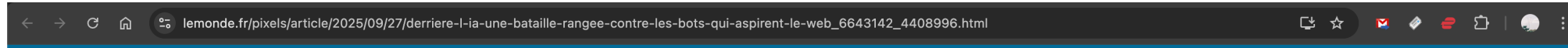
The screenshot shows a web browser at the URL `api.bnf.fr`. The page header includes the BnF logo and the text 'API ET JEUX DE DONNÉES', along with navigation links for 'ACCUEIL', 'SOURCES', 'UTILISATIONS', and 'SERVICES'. A search icon and the language 'FR' are also visible.

## Sources

Découvrez la richesse des données de la BnF

- Gallica**  
Gallica est la bibliothèque numérique de la BnF. Elle contient plusieurs millions de documents consultables et téléchargeables gratuitement.  
[EXPLORER +](#)
- Catalogue général de la BnF**  
Le catalogue général de la BnF est le catalogue en ligne qui contient la majorité des références des documents conservés sur tous les sites de la BnF, soit plus de 14 millions de documents.  
[EXPLORER +](#)
- Mandragore**  
Mandragore est la base d'indexation iconographique des enluminures de la BnF. Régulièrement enrichie depuis 1989, elle décrit aujourd'hui 205 000 enluminures issues de près de 7000 manuscrits conse.  
[EXPLORER +](#)

At the bottom of the page, a cookie consent banner reads: 'Le site bnf.fr utilise des cookies afin d'améliorer votre expérience de navigation et pour mesurer son audience. Vous pouvez choisir ceux que vous souhaitez activer'. There are buttons for 'Tout accepter', 'Tout refuser', and 'Personnaliser'.



Actualités ▾

Économie ▾

Vidéos ▾

Débats ▾

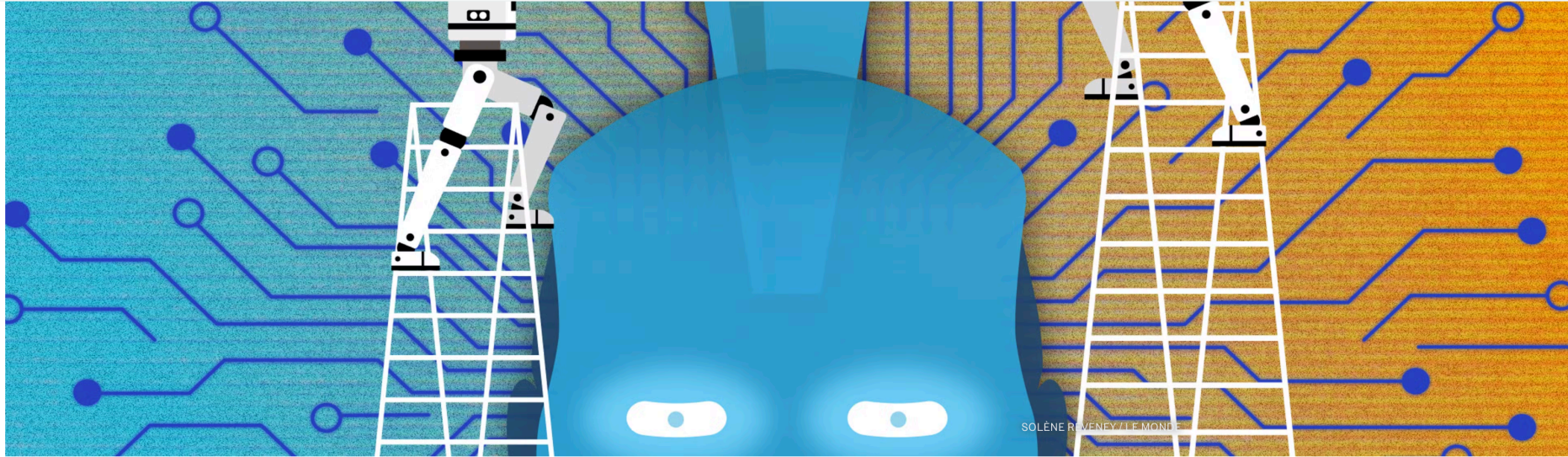
Culture ▾

Le Goût du Monde ▾

Services ▾



S'abonner



SOLÈNE RAVENY / LE MONDE

PIXELS • INTELLIGENCES ARTIFICIELLES GÉNÉRATIVES

# Des « bots » des entreprises tech aspirent le Web, et le ralentissent, pour nourrir leurs IA... mais la riposte s'organise

Par Morgane Tual

Publié le 27 septembre 2025 à 07h00, modifié le 27 septembre 2025 à 13h30

# Le consortium-HN pictoria



# Partenariats

## PARTENAIRES



AN – Archives nationales  
<https://www.archives-nationales.culture.gouv.fr/>



ArScan – Archéologies et Sciences de l'Antiquité  
<http://www.arscan.fr/>



BnF – Bibliothèque nationale de France  
<https://www.bnf.fr/fr>



CCJ – Centre Camille Jullian  
<https://cj.cnr.fr/>



AORoc – Archéologie & Philologie d'Orient et d'Occident  
<https://archo.ens.fr/>



Ariane – Analyses, Recherches, Intelligence Artificielle et Nouvelles Editions numériques  
<https://cst-ariane.huma-num.fr/>



CANAL – Cultural Analytics Lab  
<http://canal-lab.uva.nl/>



Unité de Service – Centre d'expérimentation en méthodes numériques pour les recherches en Sciences Humaines et Sociales  
<https://ceres.sorbonne-universite.fr/>



APP – Université Rennes 2, laboratoire Arts Pratiques et Poétiques  
<https://sites-recherche.univ-rennes2.fr/arts-pratiques-poetiques/>



BHVP – Bibliothèque historique de la Ville de Paris  
<https://www.paris.fr/lieux/bibliotheque-historique-de-la-ville-de-paris-bhvp-16>



CANEVAS – Consortium Huma-Num pour l'ANnotation, l'analyse Et l'archive de la Vidéo appliquées aux ctivités cientifiques  
<https://canevas.hypotheses.org/>



CRHXIX – Centre d'histoire du XIXe siècle  
<https://centrehistoire19esiecle.pantheonsorbonne.fr/bienvenue-centre-dhistoire-xixe-siecle>

# Périmètre et enjeux du consortium

## Cartographie

- Identifier les outils pertinents
- Mettre en évidence les lacunes
- Orienter le développement de nouvelles technologies

## Protocoles, référentiels et tutoriels

- Développer des protocoles partagés
- Créer une base de tutoriels
- Aborder les questions éthiques et juridiques
- Pérenniser les données et modèles coûteux

## Formations et ateliers

- Accompagner des projets existants
- Stimuler la réflexion interdisciplinaire
- Contribuer à la formation de jeunes chercheurs

## Prototypes

- Développer des prototypes réutilisables
- Tester de nouvelles techniques
- Favoriser l'intégration dans d'autres projets
- Organiser des hackathons pour explorer de nouveaux usages

## Bureau

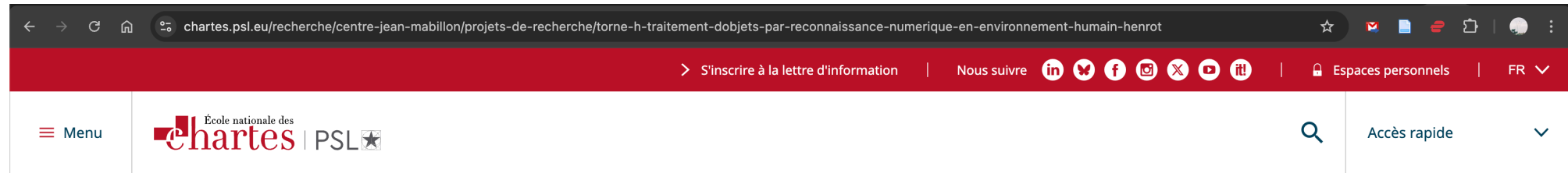
En charge de l'administration et du développement scientifique du Consortium-HN :

- Jean-Philippe MOREUX [jean-philippe.moreux@bnf.fr](mailto:jean-philippe.moreux@bnf.fr) (BnF)
- Anne-Violaine SZABADOS [anne-violaine.szabados@cnrs.fr](mailto:anne-violaine.szabados@cnrs.fr) (ArScAn)
- Julien SCHUH [julien.schuh@parisnanterre.fr](mailto:julien.schuh@parisnanterre.fr) (MSH Mondes)

## Comité de pilotage

Orientations scientifiques du consortium et de la répartition de l'enveloppe budgétaire.

- Bureau
- Direction de la MSH Mondes-UAR 3225 (accompagnement institutionnel, administratif et financier du consortium)
- Responsables des Groupes de Travail et Communautés de pratiques



Accueil / Projets de recherche / TORNE-H. Traitement d'objets par reconnaissance numérique en environnement humain / Henrot



Paul Henrot, *Escalier de l'IRSID (Institut de recherche de la sidérurgie) à Saint-Germain-en-Laye*, 1953, négatif souple. Don Marcelle Henrot, 1987 © DR © Photo : MAD Paris / Christophe Dellière

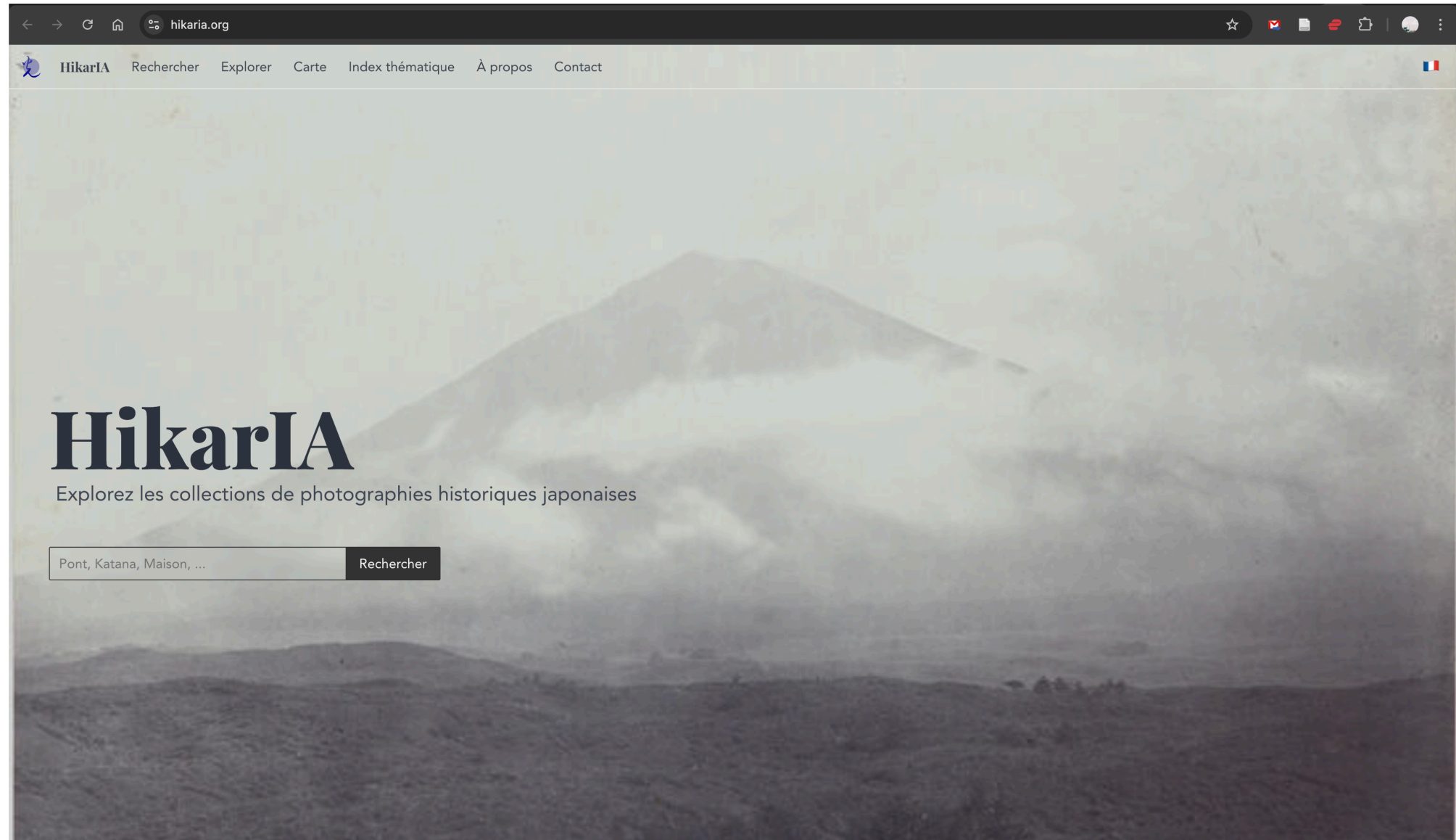
#### Projet de recherche

# TORNE-H. Traitement d'objets par reconnaissance numérique en environnement humain / Henrot

En cours

Porté par le CJM

1er octobre 2024 au 30 septembre 2025





# Projet HikarIA







Musée Guimet/TEKLIA, 2023-2026

Labélisation, détection d'objets, recherche par similarité pour une collection d'albums de photographies japonaises du XIXème siècle



Filter by type... ▾ Hide 

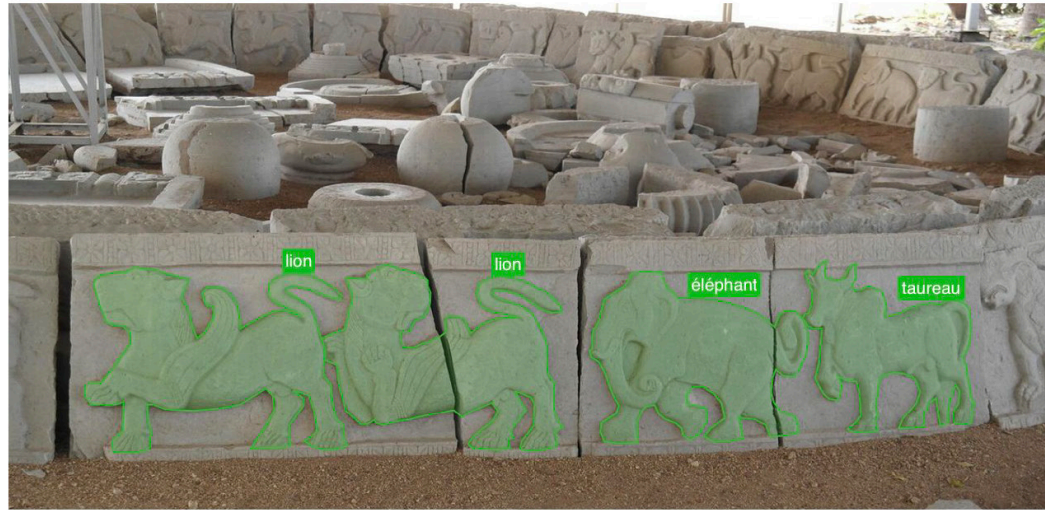
**Photograph 1** A+ 

- Sword 2 
- Sword 3 
- Sword 4 
- Sword 5 
- Sword 6 
- Map 7 

 tag	group of people
 tag	samurai
 tag	map
 tag	swords
 tag	kimono
 tag	potted plant

# Annotation d'images

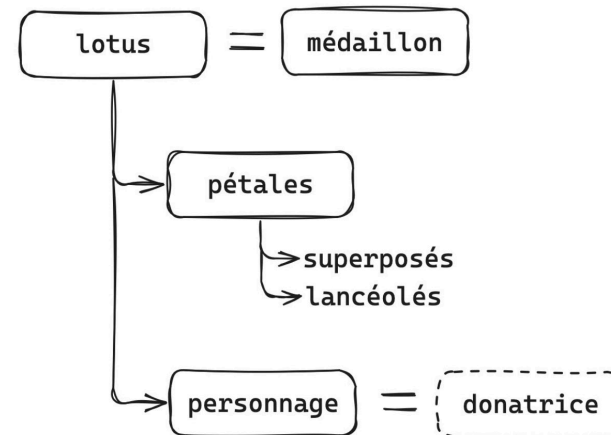
- L'annotation permet de distinguer deux types d'informations :
  - Celle structurelle : c'est-à-dire la zone d'intérêt par une forme géométrique ;
  - Celle sémantique : c'est-à-dire l'étiquette qui caractérise la forme géométrique.



# Enjeu central du groupe de travail

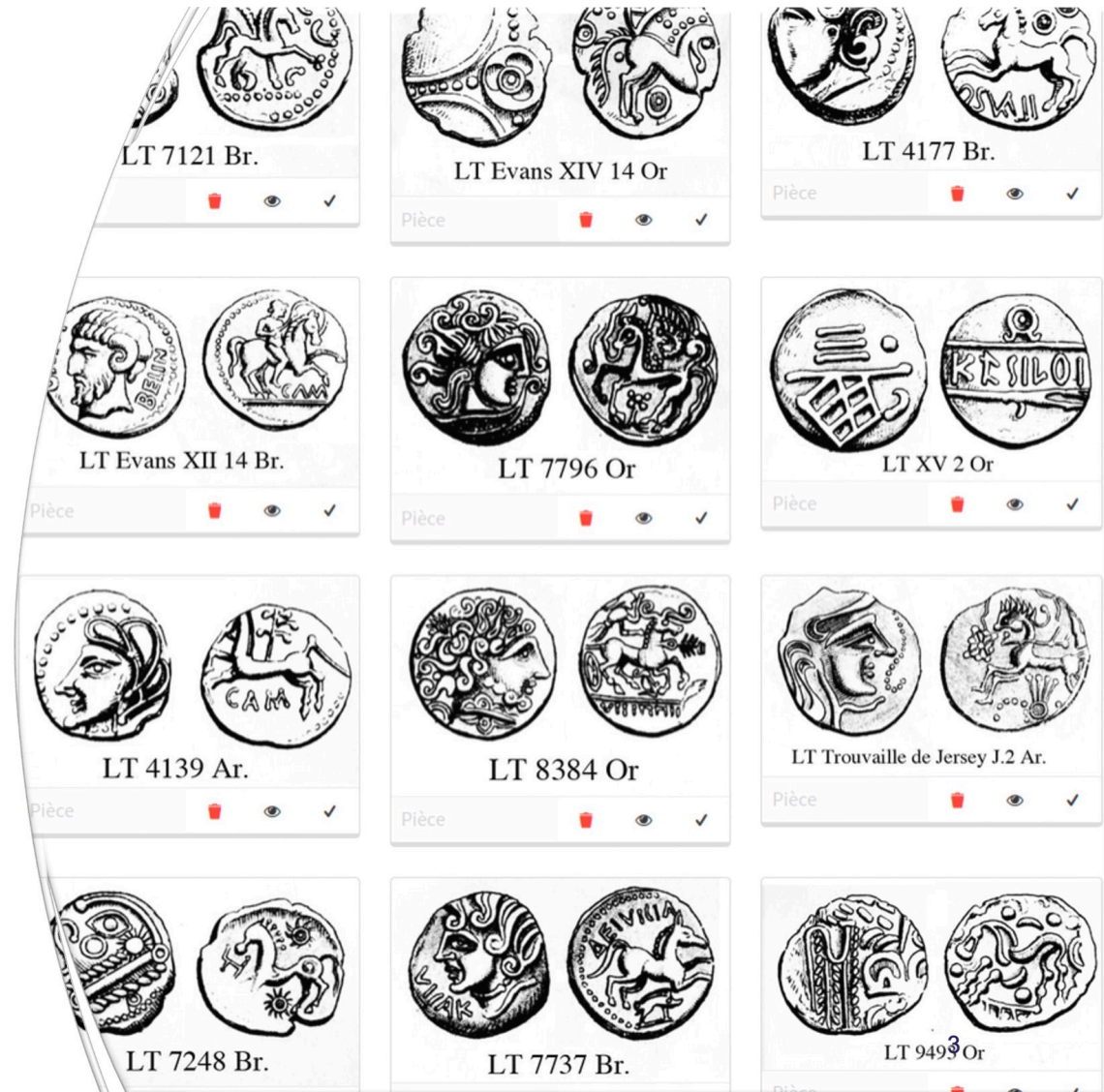
## Repenser la production des annotations

- L'enjeu est de penser les annotations depuis les besoins de recherche et non plus à partir des restrictions techniques des outils.
- Les formes géométriques doivent être enregistrées dans un format restreint et contrôlé.
- Les étiquettes produites seront ainsi multiples, à la fois parallèles et hiérarchiques.



## Données

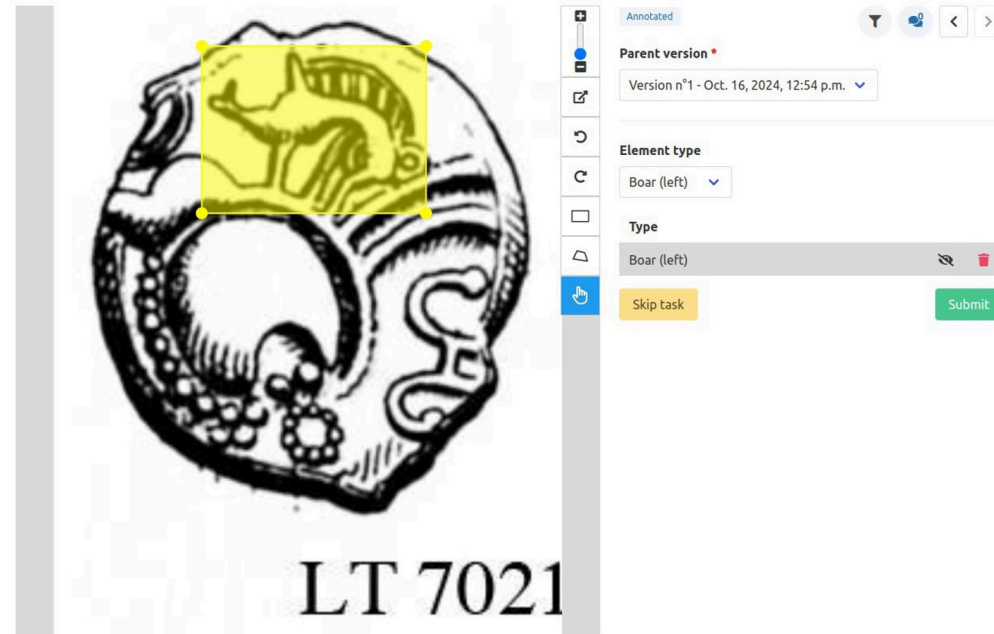
- 768 pièces : avers + revers
- Difficultés :
  - Taille des images : 274 x 442 pixels en moyenne
  - Densité de contenus
  - Variabilité de représentations d'un même motif
  - Motifs incomplets



## Annotation des données : Callico



- Annotation des 1536 faces de pièces dans Callico
  - Annotation de la localisation et de l'orientation des chevaux
- 560 chevaux annotés
- 213 orientés vers la gauche
  - 347 orientés vers la droite



[https://arkindex.teklia.com/browse/db36a6cf-7c4d-49e4-86e8-0e7a253fd9c4?worker\\_run=ef07caeb-191c-4609-ae7-72308a7201ab&type=horse\\_right](https://arkindex.teklia.com/browse/db36a6cf-7c4d-49e4-86e8-0e7a253fd9c4?worker_run=ef07caeb-191c-4609-ae7-72308a7201ab&type=horse_right)  
[https://arkindex.teklia.com/browse/db36a6cf-7c4d-49e4-86e8-0e7a253fd9c4?worker\\_run=ef07caeb-191c-4609-ae7-72308a7201ab&type=horse\\_left](https://arkindex.teklia.com/browse/db36a6cf-7c4d-49e4-86e8-0e7a253fd9c4?worker_run=ef07caeb-191c-4609-ae7-72308a7201ab&type=horse_left)

6



Séminaire-atelier pictorialIA “L'IA pour l'archéologie et les arts visuels anciens” – 17 octobre 2024, BnF DataLab

De l'intérêt de l'IA pour l'analyse des monnaies celtiques, le projet DIANet,

KATHERINE GRUEL, ELODIE PARIS, OLIVIER MASSON



DIANET



Sémiologie graphique sur les monnaies celtiques



Science des Données  
Données de la Science

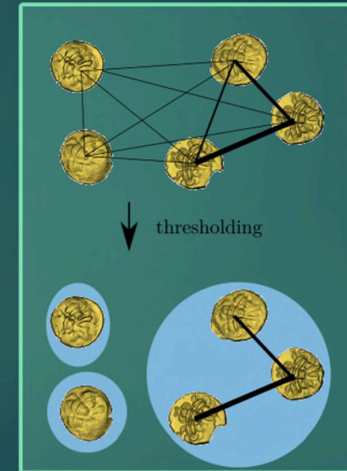
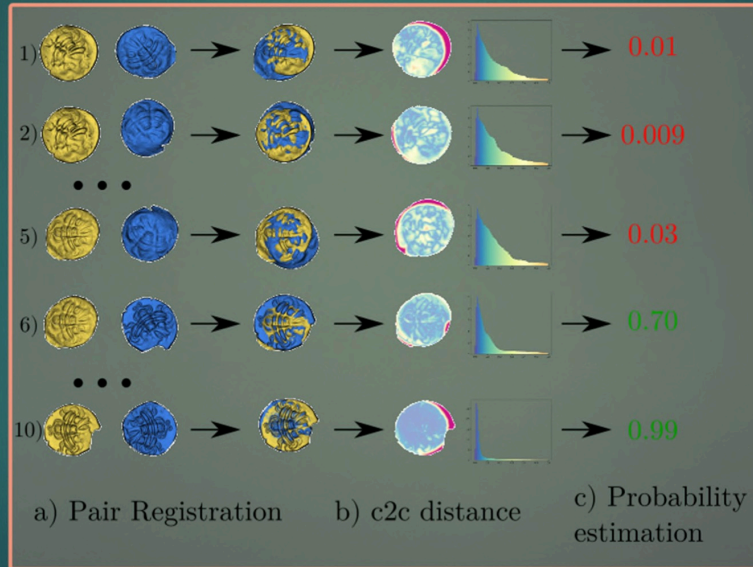


HISTOIRE ET PRATIQUES DE L'ÉCRIT



## Thèse de Sofiane Horache , PSL University, codirection Mines-Paristech et AOROC, CNRS-PSL

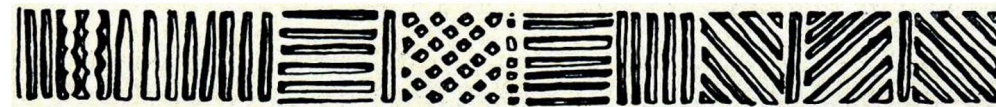
I) Acquisition II) Pairwise similarity estimation III) Clustering



## Application de l'intelligence artificielle à la reconnaissance des décors à la molette sur céramiques sigillées d'Argonne et imitations



Projet visant à faciliter l'identification et la classification automatisées des motifs décoratifs grâce à l'intelligence artificielle (IA) et aux réseaux neuronaux convolutifs (CNN).



Pablo Ciezar  
Inrap / ArScAn -

GAMMA  
Inrap

ARSCAN  
Archéologies et Sciences de l'Antiquité  
UMR 5175  
Ministère de la Culture - Université Paris 8 Vincennes Saint-Denis - Inrap

GAMMA  
Mondes antiques et médiévaux

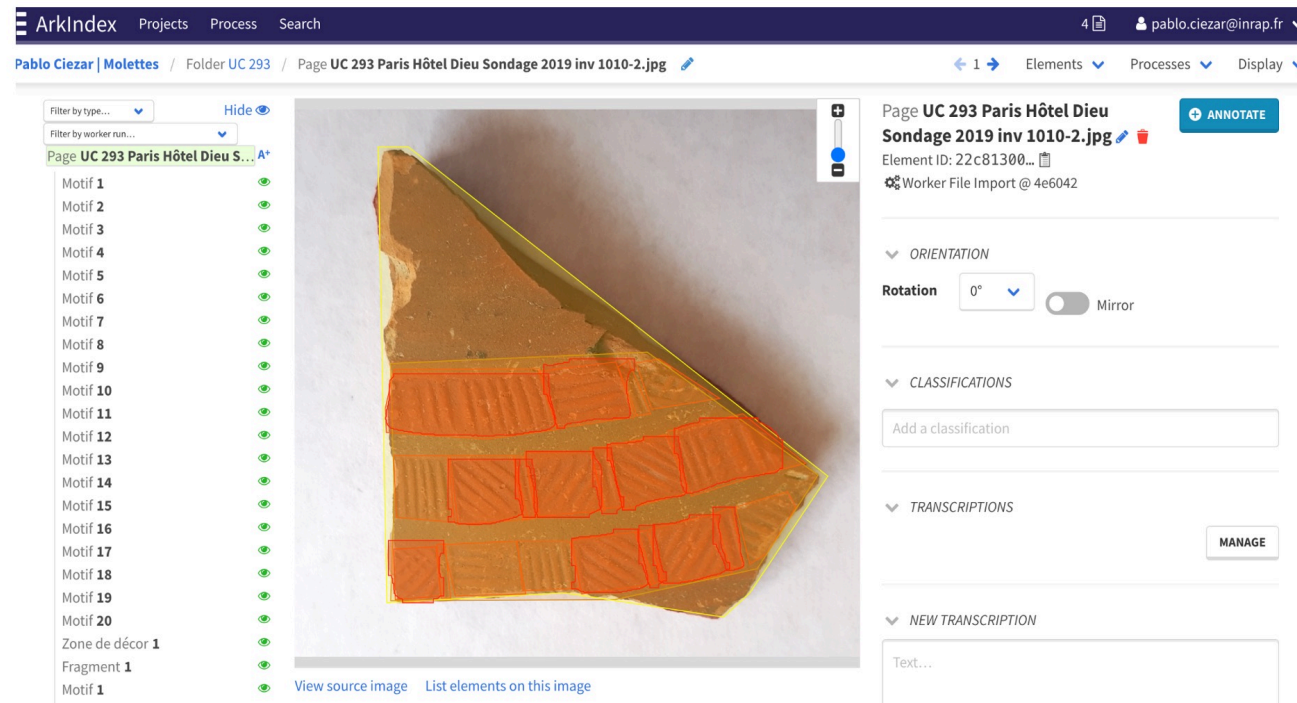
AG – pictorIA – 23 janvier  
2025

## Procéder à l'annotation précise des motifs dans chaque molette

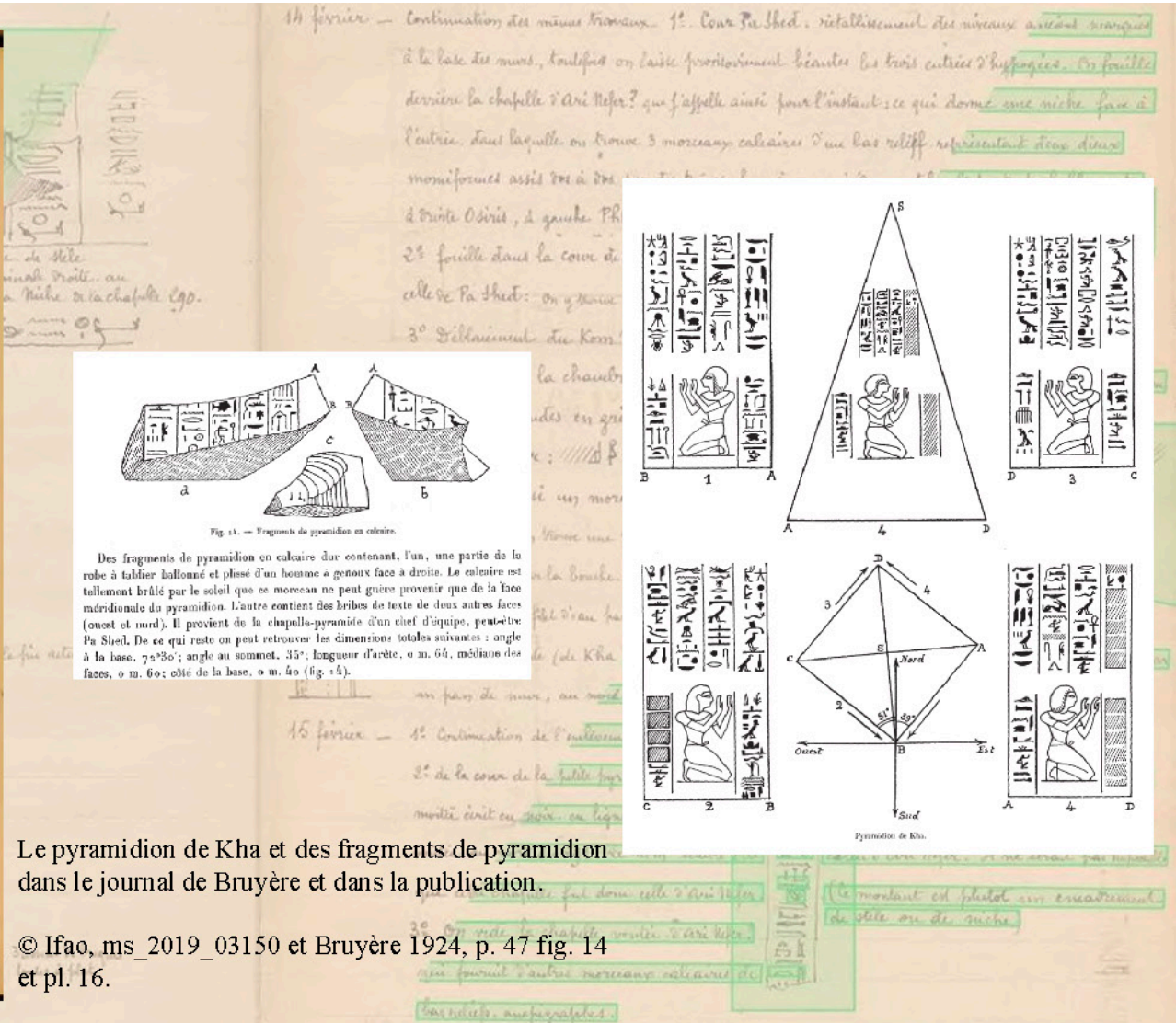
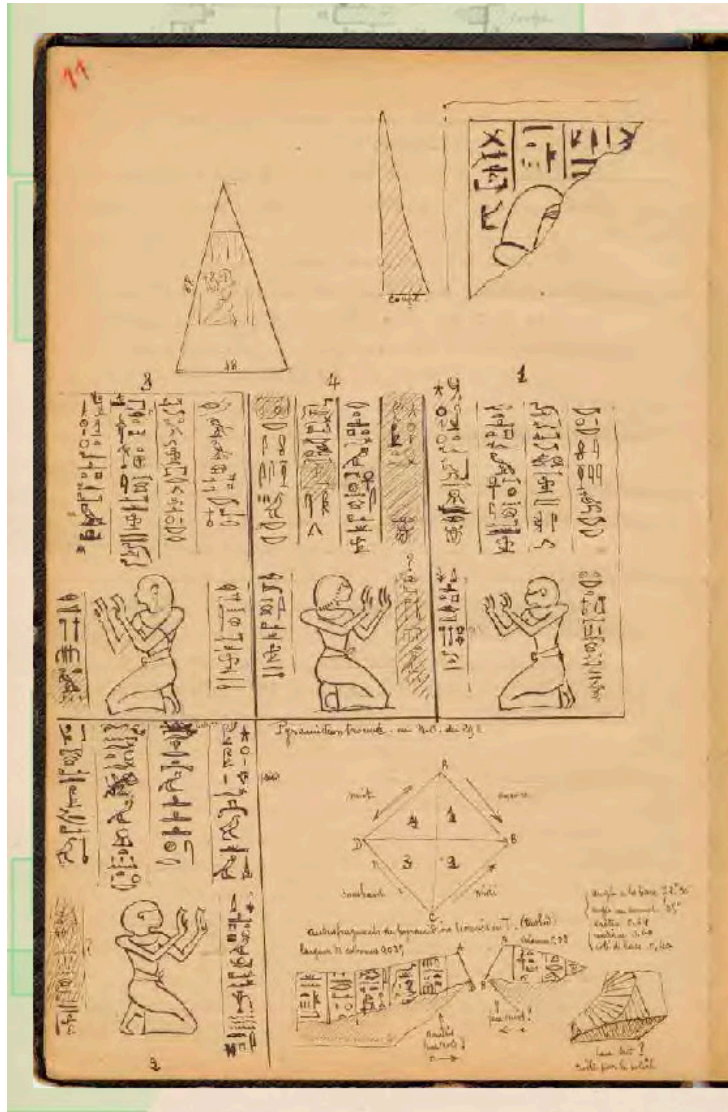
Phase 1: Constitution d'un large corpus d'apprentissage (60 jours).

Phase 2: Développement et optimisation du prototype (120 jours).

Techniques utilisées: segmentation semi-supervisée, CNN, ResNet, EfficientNet, etc.



Défi principal : qualité des images et annotations précises.



Le pyramidon de Kha et des fragments de pyramidon dans le journal de Bruyère et dans la publication.

© Ifao, ms\_2019\_03150 et Bruyère 1924, p. 47 fig. 14 et pl. 16.

demo.arkindex.org/element/1ae9d9c9-0594-4ab5-8392-9a0151be9cfc?highlight=d3641d0d-2013-469e-8a6e-ee5ae1dd3c53

Arkindex Projects Search LOG IN REGISTER

You are currently not logged in to Arkindex. Consider [creating a new account](#) or [logging in](#) to access more resources and tools.

IFAO | Journaux de fouilles de... / Folder Journal de fouilles 1939-1955 / Page ms\_2019\_03779 ← 56 → Elements Processes Display

Filter by type...  
Filter by worker run...

Page ms\_2019\_03779

- Illustration ms\_2019\_03779\_01
- Illustration ms\_2019\_03779\_02
- Illustration ms\_2019\_03779\_03
- Illustration ms\_2019\_03779\_04
- Illustration ms\_2019\_03779\_05
- Illustration ms\_2019\_03779\_06
- Illustration ms\_2019\_03779\_07
- Illustration ms\_2019\_03779\_08
- Illustration ms\_2019\_03779\_09
- Illustration ms\_2019\_03779\_10
- Illustration ms\_2019\_03779\_11
- Illustration ms\_2019\_0377... 🔗 🔍 A+
- Illustration ms\_2019\_03779\_13
- Text line 38
- Text line 39
- Text line 42

Illustration ms\_2019\_03779\_12 🔗 🔍 🗑️

Element ID: d3641d0d...

Created by ZCochin

---

ORIENTATION

Rotation 0° ▼  Mirror

---

CLASSIFICATIONS

**Manual**

Stèle privée 👤 100% ✓ 🗑️

---

TRANSCRIPTIONS

---

METADATA

📅 date of find	1939, January 25	✎	🗑️
🌐 found_location	P_1414	✎	🗑️
🖼️ illustration_name	(122)	✎	🗑️
🏷️ other_names	Sethy Ier	✎	🗑️

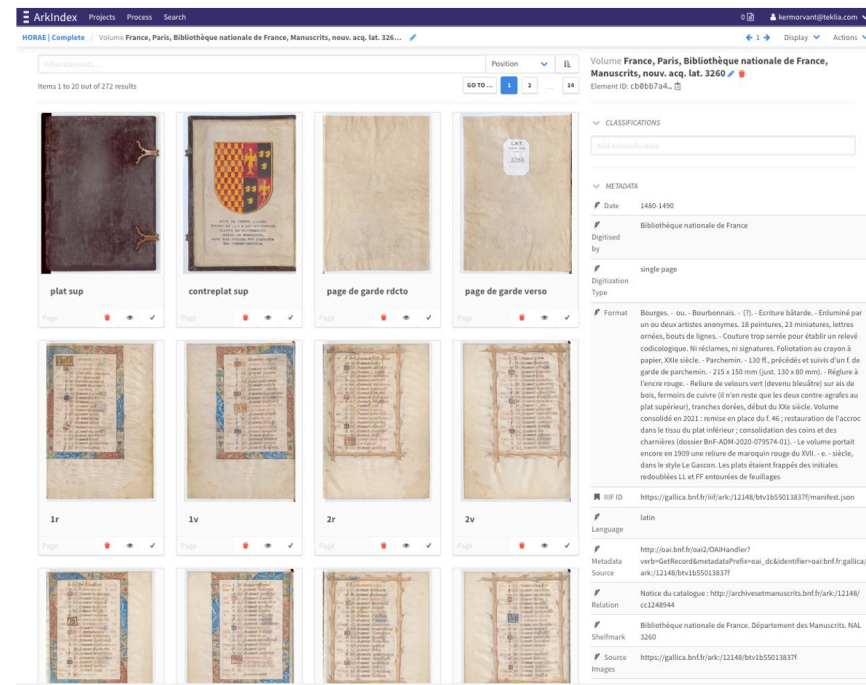
---

DATASETS

No datasets

# Arkindex : Stockage et gestion des documents

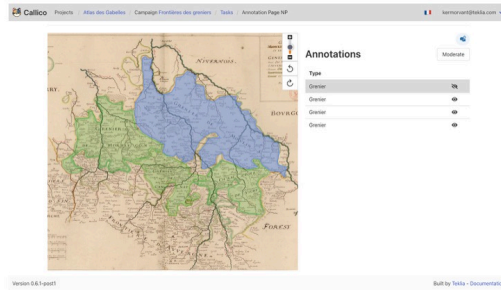
- Import web (petit corpus), S3 (gros corpus) ou par manifest IIIF
- Support des formats images et PDF
- Structuration hiérarchique des corpus complètement adaptable
- Gestion des métadonnées à tous les niveaux
- Visualisation, navigation



<https://gallica.bnf.fr/iiif/ark:/12148/btv1b55013837f/manifest.json>

# Callico : Annotation et validation

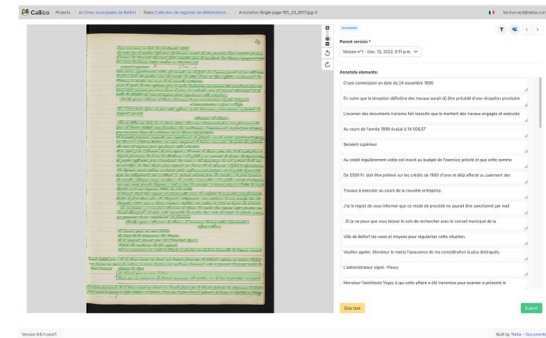
## Zonage



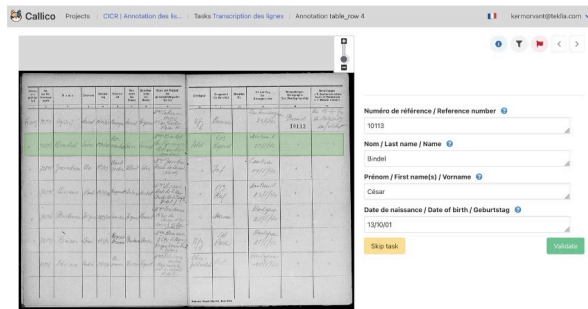
## Classification



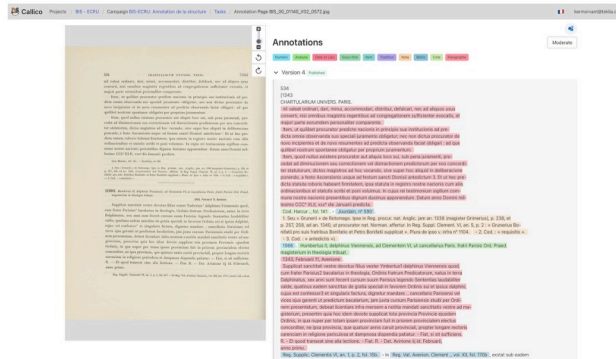
## Transcription



## Clé-valeur



## Entités nommées

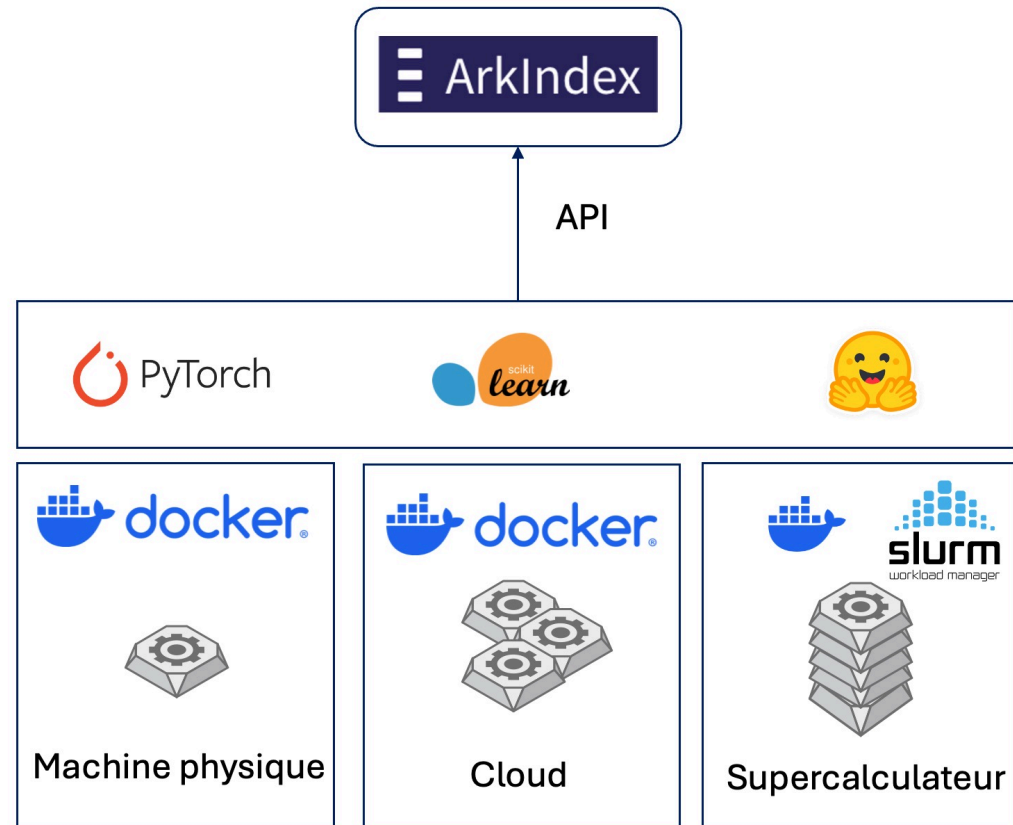


## Groupement



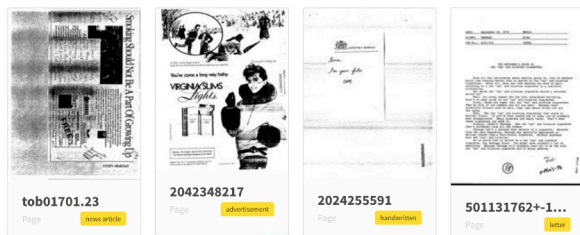
## Arkindex : Intégration de modèles/algorithmes

- Intégration de n'importe quel langage/code/modèle
- Code de base python fourni
- Intégration par API
- Déploiement par Docker
- Entraînement et inférence

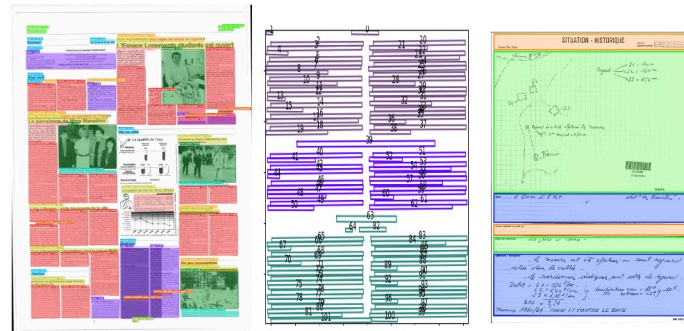


# Arkindex : applications

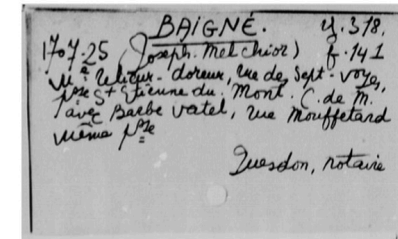
## Classification



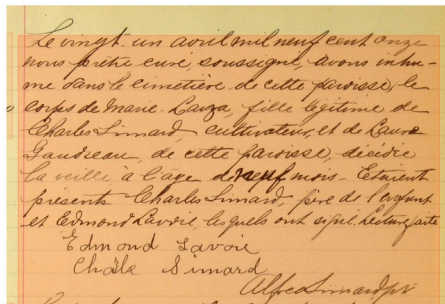
## Structuration



## OCR/HTR



## Extraction d'entités



## Analyse de photos



- tag group of people
- tag samourai
- tag map
- tag swords
- tag kimono
- tag potted plant