

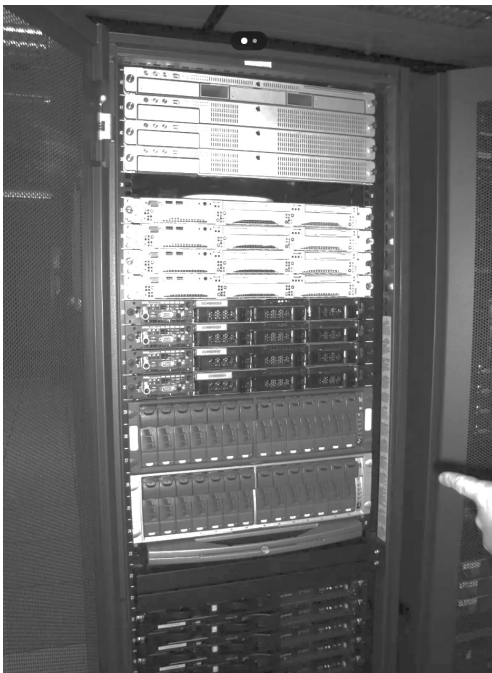
L'IR* Huma-Num au CC-IN2P3 : un hébergement fondateur

Olivier Baude
Huma-Num IR*



Petite et grande histoire d'un hébergement fondateur...

22 novembre 2006



16 juin 2026



Huma-Num^{IR*} en 2026 :



- (Très) grande **infrastructure de recherche (IR*)** de la feuille de route nationale du MESRE
- **Priorité stratégique** de l'institut CNRS SHS
- Impliquée dans les 6 **programmes structurants** en Sciences Humaines et sociales
- Co-créatrice et porteuse du nœud national de **l'ERIC DARIAH**
- IR* labellisée **Centre de Référence Thématique** de Recherche Data Gouv
- Co-porteuse de **l'EQUIPEX+ COMMONS**
- Créatrice et responsable de **Nakala l'entrepôt de confiance** pour les SHS et de deux entrepôts disciplinaires
- **Service versant** (et coordinatrice d'opérateurs SHS) pour un archivage CC-IN2P3 / CINES-Archives Nationales



IR SHS

HN Huma-Num^{IR*}



...

IR SHS et Information Scientifique



...

IR services numériques (calcul & réseaux)



Genci/Jean ZAY...

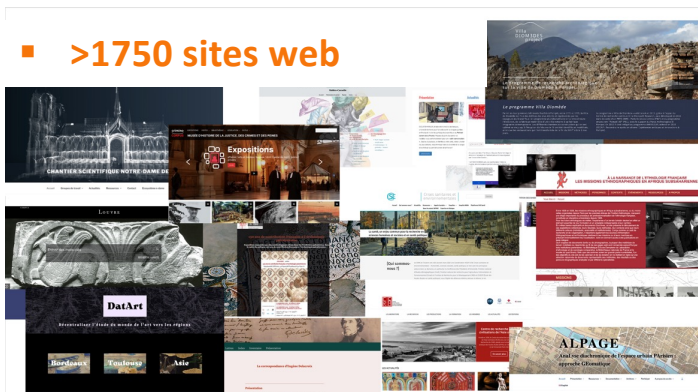
Infrastructures de recherche européennes



...

HN Huma-Num^{IR*} en 2026 de l'infrastructure à l'instrument de confiance

▪ >1750 sites web

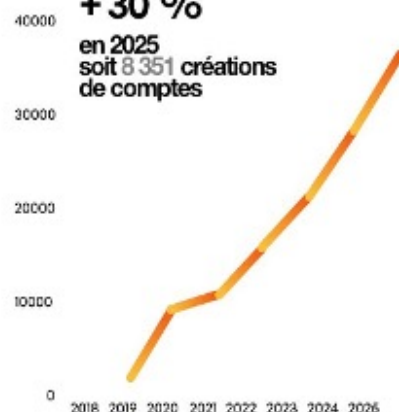


▪ 9 Consortiums-Huma-Num - 430 / 1200 chercheurs

Nombre de comptes
Utilisateurs

36 767 comptes
utilisateurs

+ 30 %
en 2025
soit 8 351 créations
de comptes

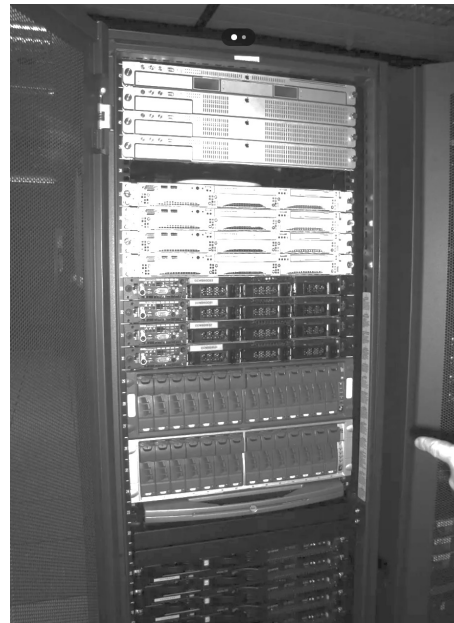
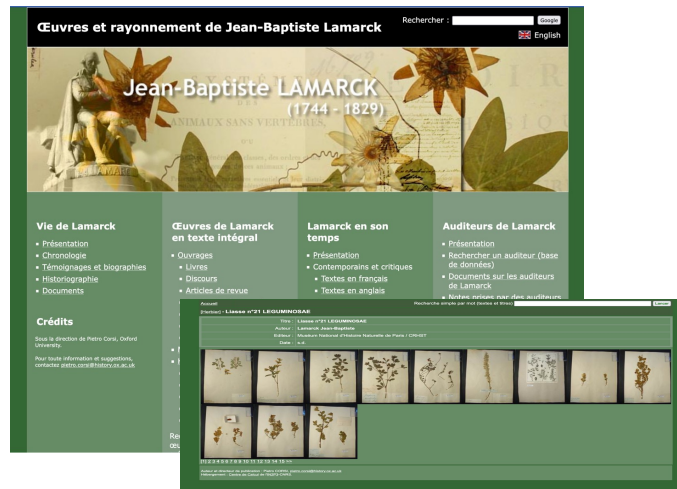


- **2054** projets de recherche hébergés
- **1450** bases de données
- **1** entrepôt certifié :
 - **1,5 million** de fichiers,
 - **850 000** dépôts,
 - **10 000** collections,
 - **2727** déposants.
- **2 Po** de données sécurisées
- **> 1000 sollicitations/mois**

1992-2004 :

Premiers besoins des SHS et accueil au CCIN2P3

- Calcul scientifique pour l'archéologie et les données 3D
- Diffusion de corpus numériques en histoire des sciences : numérisation des documents et sites Web (Lamarck et Buffon)



A cette époque, le CC-IN2P3

- Calcul scientifique, stockage, grille informatique...

Et...

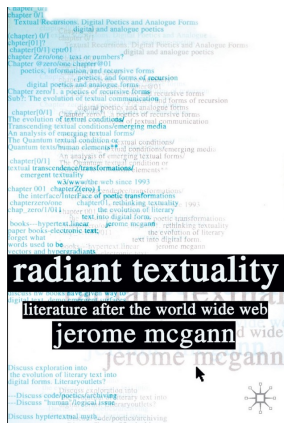
- CCSD 
- CCSD Medihal
- Premier serveur web français



Les Humanités numériques et les méthodes numériques en SHS : les années 1990 -2010

La création des centres des *Digital Humanities* :

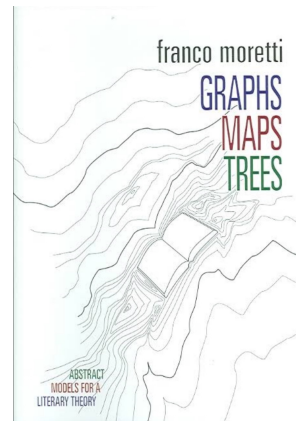
- Numérisation des sources
- Traitement informatique des archives numériques
- Une nouvelle approche critique des textes
- Une volonté de décloisonnement disciplinaire et de travail collectif



Jerome McGann : la textualité radiante

Donner accès à de grands corpus multimédias - Archives hypertextuelles, intertextuelles (rassembler des matériaux radiants autour de l'œuvre), Outiller l'exploration

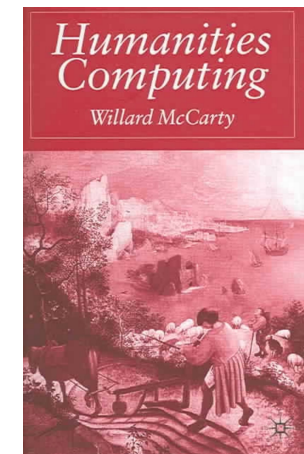
« Le caractère ouvert et non clos de l'archive numérique bouleverse les pratiques de monographie et d'éditions savantes, elles mêmes transformées par les mediums » [McGann 2008]



Franco Moretti : la lecture à distance

L'étude sur de vastes corpus, « la montée en généralité », repérage de « patterns » : Graphs, Maps, Trees

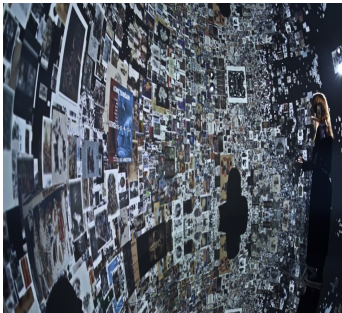
« Je dirais simplement que nous avons beaucoup à apprendre des méthodes des sciences sociales et des sciences naturelles » [Moretti, 2003]



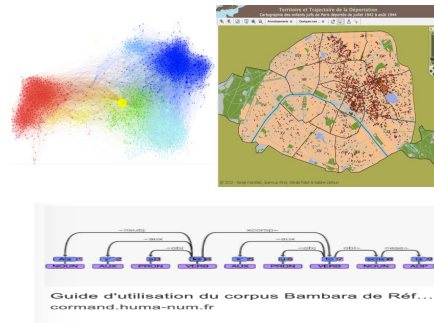
Willard McCarty : L'informatique et les humanités

« L'ordinateur est un télescope pour l'esprit. C'est l'instrument qui permet d'observer d'une autre manière et plus loin ». [McCarty, 2005]

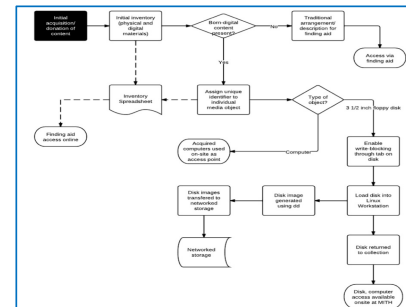
Trois besoins de stockage, réseau, calcul scientifique des SHS



1. Accéder à de grands corpus multimédias, Outiller l'exploration.



2. Traiter, visualiser les données : graphs, maps, trees...



3. Formaliser le matériau et le raisonnement pour traiter ces données

Les données comme production scientifique :

Numériser, stocker
conserver

- Traiter, explorer

Diffuser, réutiliser

Lier, intégrer à des publications

Petite et grande histoire d'un hébergement fondateur...

2005 : La DIST crée les Centres de Ressources Numériques

Les **textes** : *CRNTL*, Les **manuscrits** : *TELMA*, L'**oral** : *CRDO Paris – CRDO AIX*, L' **image** : *CN2SV*, La **3D** : *CRN3D*

Le Très Grand Équipement ADONIS au CC-IN2P3

- **2007** : serveurs pour un programme d'archivage
- **2010** : Hébergement d'ISIDORE (moteur de recherche fondé sur l'OAI-PMH et l'IA)
- **2010** : Développement d'une grille de services pour les SHS (stockage, hébergement web, mais aussi : base de données (dont XML, Filemaker, calculs,...))
- **2010** : université du TGE ADONIS « Construire une infrastructure numérique pour les SHS »

L'infrastructure de recherche Corpus-IR

2011 : Naissance de Corpus-IR et création des Consortiums disciplinaires

2013 : une infrastructure de recherche née de la fusion d'un Très Grand Equipement et d'une infrastructure pour des communautés de recherche



Très grand équipement pour les projets de recherche et leurs productions scientifiques

Corpus-IR

Infrastructure de Recherche

Communautés de recherche ayant des méthodes et pratiques numériques



Une TGIR (très grande **infrastructure de recherche**) pour :

- Faciliter le « tournant numérique » de la recherche en sciences humaines et sociales
- Développer l'appropriation par les communautés scientifiques **du cycle de vie des données numériques**,
- Proposer des **services numériques** au juste niveau et au bon moment.



Un dispositif sociotechnique

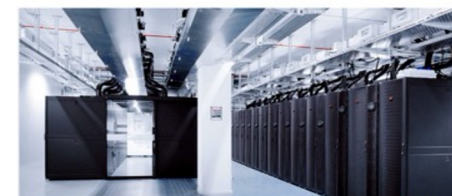
Des communautés



Des services numériques pour les données



Une infrastructure informatique



Le pilotage et l'innovation scientifiques

Les données de recherche

Les coopérations internationales

Des communautés de recherche : les Consortiums Huma-Num

Communautés de recherche

9 Consortiums-HN, autour de 150 équipes de recherche « actives » qui couvrent des champs variés des SHS

- réseaux d'experts
- thématiques et d'objets communs
- outils numériques
- principes FAIR démarche de Science Ouverte »



Mémoires des Archéologues et des sites Archéologiques plus

Porter les données archéologiques vers la Science Ouverte

Les groupes de travail

- Archives** : cartographies, guides des sources, fiches pratiques
- Documentation** : F2D machine accessible pour Data Paper
- Vocabulaire** : épigraphie et archéologie (un seul pour tous)
- Spécialisation** : inventaire et analyse des outils disponibles
- Publication** : articulation données/publication, bio-tech
- Données ouvertes** : modèles ontologiques modulaires
- Épistémologie** : ...

Produit : Livre blanc, Modèle de plan de gestion de données, Bonnes pratiques en ligne, Plateforme web de données et exploration

Pictoria
CONSORTIUM HUMA-NUM

Programme de reconnaissance, classification, rangement d'images, Observation et Reconnaissance des formes par l'Intelligence Artificielle

Axe 1: Traitement et exploration de données
Développement et amélioration d'outils d'IA visuelle
Réflexion sur les formats et protocoles
Production de modèles et de jeux de données

Axe 2: Diffusion
Création d'une culture partagée, favorisant l'adoption des outils
Organisation d'ateliers, de séminaires et de formations

ARTICULES - RESEAUX - ARCHÉOLOGIE
ET ARTICLES - DOCUMENTS

Atelier « Intelligence artificielle et analyse de corpus modulaires »
- BREF, 2 juillet 2025
Dans le cadre du programme "Intelligence artificielle" organisé conjointement par le CNRS et le Centre National de la Recherche Scientifique (CNRS) avec le Consortium Huma-Num, l'atelier a permis d'explorer les possibilités de l'IA dans la recherche archéologique.

Estimation des membres des 1er, 2ème et 3ème cercles des Consortiums-HN

cercle 3 Communauté informée 3410	cercle 2 Communauté active 1155
	cercle 1 Noyau dur 430

- Noyau dur des Consortiums Huma-Num
- Communauté active : personnes participant aux travaux, aux formations, etc.
- Communauté informée : abonnés aux listes de diffusion, pondération réseaux sociaux et carnets de recherche

Et un pôle pour la veille de l'innovation scientifique : le HNLab

La maîtrise d'une *infrastructure informatique* et des *services numériques* de "confiance"

Des services « classiques »...

- Gérer un projet
- Stocker et partager des fichiers
- Héberger des sites web, des bases de données, du code
- Stocker des données volumineuses
- Disposer de machines virtuelles
- Disposer de puissance de calcul (CPU, GPU)
- Utiliser des outils pour le traitement des données (textes, images, vidéos, 3D, SIG,...)
- ...



...aux services pour les données, à forte valeur ajoutée, pour préserver, signaler et publier



Préservation à long terme

Infrastructure physique au sein d'un data center souverain de l'ESR



Puissance de calcul

Hébergement de sites web,
Machines virtuelles

Stockage pour la préservation
et le partage des fichiers

Serveurs pour les bases de données
et les traitements

Applications et services proposés
par Huma-Num

Réseau



1 firewall



9 switchs

Serveurs



55 serveurs
physiques



460
serveurs
virtuels

Stockage



3 baies de
stockage
(1180 To)



1 cluster de
stockage
répliqué (1.2 Po)



1 robotique
(1 Po)

GPU



64 GPU

ACCES FITS



- Infrastructure **CNRS** (serveurs, stockage, réseau) au **centre de calcul IN2P3**
- **Infrastructure souveraine** : autonomie technique, organisationnelle et financière
- Couverture complète : **calcul, diffusion web, conservation pérenne des données**
- Prise en charge du **cycle de vie complet des données** (production → valorisation)

Huma-Num^{IR*} en 2026 : de l'infrastructure à l'instrument de confiance



Description physique de l'infra en 2025

- 4 racks sur le site du CC-IN2P3
- 7 switchs réseau
- 1 FireWall
- 1 WAF
- 52 serveurs
- 24 GPU en propre
- 3 baies de stockage

Les chiffres de la puissance de calcul en 2025

- 70 GPU actives
- 34 233h heures de calcul sur GPU
- 109 comptes sur la ferme de calcul
- 216 comptes pour l'usage de R

Une infrastructure *on-premise*

- **Un socle d'exploitation** (machines, câbles...) sur 2 types de ligne de production :
 - Type 1 : système direct (dédié) Internet → Pare-feu → Serveur physique → Stockage disque interne.
 - Type 2 : système virtualisé (mutualisé) Internet → Pare-feu → Serveurs virtuels → Serveurs physiques → Système de stockage.
- **La protection du socle** : défense périmétrique (next-generation firewall)
- **Le stockage et le partage** : conserver sans qualifier (3 services de conservation
- **L'hébergement pour diffusion** : le cluster web
- **L'hébergement pour exploiter** des productions de programme (500 machines virtuelles)
- **Des environnements pour traiter** : les diverses formes du calcul scientifique
- **Des service acquis sur étagère** : exploiter des applications communes
- **Des services pour la FAIRisation des données** : l'entrepôt Nakala, l'assistant ISIDORE, archivage



Huma-Num^{IR*} en 2026 : de l'infrastructure à *l'instrument de confiance*



Projet 2026-2030 de l'IR* Huma-Num **comme instrument de confiance** :
maîtrise scientifique, sécurité informatique souveraineté, adaptativité.

Huma-Num^{IR*} en 2026 : de l'infrastructure à *l'instrument de confiance*



Confiance :

Renforcement des capacités de stockage et de calcul
et maîtrise économique et écologique

FITS

Les services proposés aux IR/IR* par le projet FITS



Fiabilité dans un écosystème national et européen
en mouvement...

Montée en puissance de la valeur des données des
SHS.

Huma-Num^{IR*} en 2026 : de l'infrastructure à *l'instrument de confiance*



Renforcement des capacités de stockage et de calcul et maîtrise économique et écologique : **le déménagement !**

FITS

Les services proposés aux IR/IR* par le projet FITS



Stockage

Solution de stockage de très grands volumes de données.

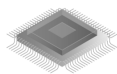
[Découvrez le stockage](#)



Hébergement

Solution d'hébergement dans un environnement fiable, sécurisé et durable.

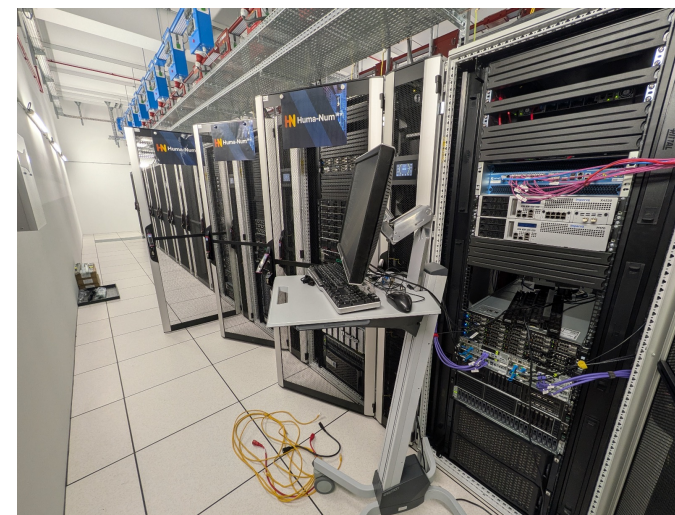
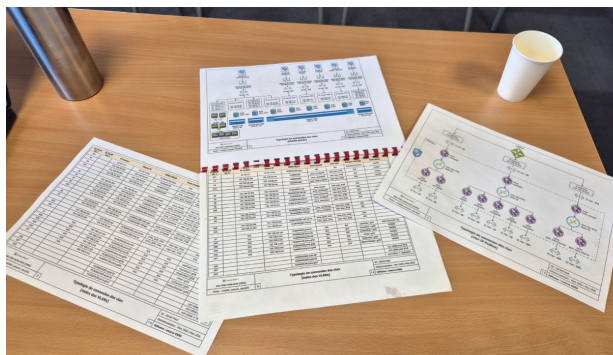
[Découvrez l'hébergement](#)



Calcul

Accès aux ressources de calcul des sites partenaires (HTC, HPC, IA).

[Prochainement](#)



Huma-Num^{IR*} en 2026 : de l'infrastructure à *l'instrument de confiance*



Confiance :

Adaptabilité à l'évolution rapide des méthodes et besoins : l'IA en SHS.

Déployer une infrastructure répondant à trois défis majeurs :

- la **sécurisation des calculs et traitements** liés à l'usage de l'IA générative ;
- la mise en œuvre d'un **écosystème d'accueil, de calcul et de développement de solutions logicielles**, permettant d'adapter l'IA et l'IA générative aux besoins spécifiques des SHS, notamment en matière de ré-entraînement de MLM et de LLM;
- le **développement de nouvelles fonctionnalités** (ShareDocs, NAKALA, ISIDORE, etc.), adaptées aux besoins des différentes communautés, avec des LLM ré-entraînés (ex. : un modèle Whisper ré-entraîné localement sur des données non publiques pour des usages internes, non public et sécurisé sans appels aux API d'OpenIA).

Petite et grande histoire d'un hébergement fondateur...

 **Huma-Num**^{IR*} en 2026 : de l'infrastructure à *l'instrument de confiance*



Confiance :

Adaptabilité à l'évolution rapide des méthodes et besoins : l'IA en SHS.

Une infrastructure pour cinq classes d'usages :

- 1. Services génériques utilisant l'IA (transcription automatique...)**
- 2. Espaces de prototypage et d'exploration (serveur interactif mutualisé)**
- 3. Traitements intensifs et adaptation de modèles (Ferme de calcul CC-In2P3 > IDRIS)**
- 4. Plateforme d'expérimentation pour programmes de recherche (Jupyter notebooks)**
- 5. Hébergement de services mutualisables et durables (serveur d'accès à des LLM, SML)**

Pour une infrastructure augmentée, souveraine et éthique

Huma-Num^{IR*} en 2026 : de l'infrastructure à *l'instrument de confiance*



Confiance :

Adaptabilité à l'évolution rapide des méthodes et besoins : l'IA en SHS.

Formation 2025 avec le CC-In2P3



Huma-Num^{IR*} en 2026 : de l'infrastructure à *l'instrument de confiance*



Confiance :

Adaptabilité à l'évolution rapide des méthodes et besoins : l'IA en SHS.

Participation à la stratégie 2026-2028 de CNRS-SHS

Passage d'une posture de consommation d'outils génériques à une logique de co-construction. L'enjeu est de bâtir une infrastructure scientifique où les théories et les données des SHS ne sont pas seulement traitées par l'IA, mais participent activement à la conception de nouveaux modèles. Cette approche garantit la souveraineté de nos recherches, la maîtrise épistémologique de nos résultats et la pertinence sociale de nos analyses.

1. **Réseau métiers ingénieurs IA-SHS** : Véritable lieu de la montée en compétences, il prépare les ingénieurs à manipuler les outils de demain. C'est le préalable indispensable à toute activité scientifique pérenne.
2. **Les Trois Pôles Scientifiques** : (Langues, Textes et Cognition ; Analyse des Sources et Patrimoines ; Individus, Sociétés, Démocratie). Ils sont les lieux de la production de savoir, chargés de définir les besoins, de produire les jeux de données critiques et de valider les modèles.



Conclusion :

L'hébergement au CC-IN2P3, un élément majeur de la réussite d'Huma-Num et fondateur pour la création et le développement d'un instrument de recherche de confiance.

Bonus :

EquipeX+ FITS

CC-IN2P3 et l'IDRIS + GENCI



EquipeX+ HALiance

Lien Données - Publications



HAL SHS
sciences humaines
et sociales



EQUIPEX +

**COM
MONS**

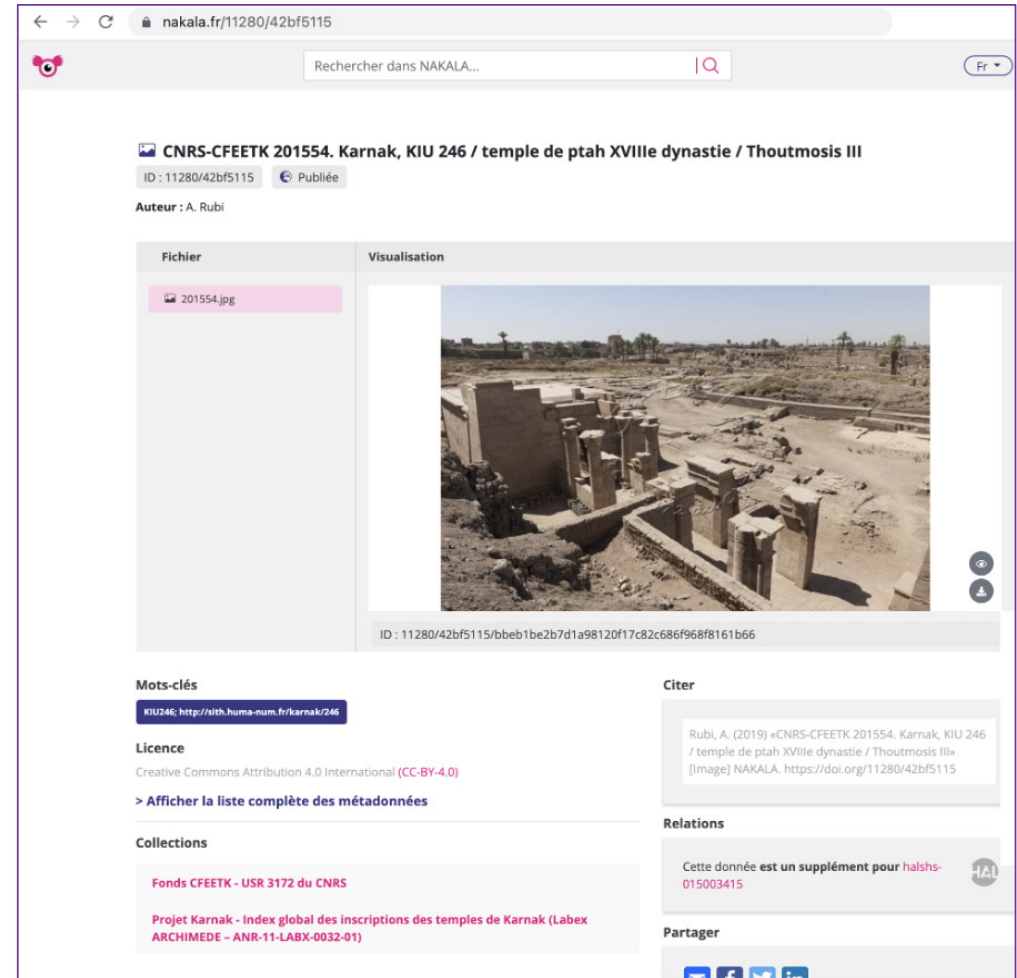
HN Huma-Num^{IR*}

Bonus :



- Pour stocker, préserver et exposer des données numériques
- Accessible en ligne, pour déposer, et pour consulter les données publiées.

Entrepôt de confiance



Bonus :

Lier données et publications :

← → ↺ nakala.fr/11280/42bf5115

Rechercher dans NAKALA...

CNRS-CFEETK 201554. Karnak, KIU 246 / temple de ptah XVIIIe dynastie / Thoutmosis III

ID : 11280/42bf5115

Auteur : A. Rubi

Fichier

201554.jpg

Visualisation



ID : 11280/42bf5115/bbeb1be2b7d1a98120f17c82c686f968f8161b66

Mots-clés

KIU246; <http://oith.huma-num.fr/karnak/246>

Licence

Creative Commons Attribution 4.0 International (CC-BY-4.0)

> [Afficher la liste complète des métadonnées](#)

Collections

Fonds CFEETK - USR 3172 du CNRS

Projet Karnak - Index global des inscriptions des temples de Karnak (Labex ARCHEMEDE - ANR-11-LABX-0032-01)

Citer

Rubi, A. (2019) «CNRS-CFEETK 201554. Karnak, KIU 246 / temple de ptah XVIIIe dynastie / Thoutmosis III» [Image] NAKALA. <https://doi.org/11280/42bf5115>

Relations

Cette donnée est un supplément pour [halshts-015003415](#) 

Partager





HAL
archives-ouvertes.fr

Page d'accueil Dépôt Recherche Actualités Consultation par période Mon espace Privileges

hal-01667359, version 1 Article dans une revue

À propos de deux documents d'Achmosis à Karnak

Sebastien Biston-Moulin ¹ [Détails](#)

1 CFEETK - Centre Franco-Égyptien d'Étude des Temples de Karnak

Abstract : The first part of this paper is a new examination of the carving of the date on the "year 17" block of King NebphetyRe Achmose at Karnak which led to a reconsideration of the orientation of the moon sign in his birth name during his reign as a chronological criterion. The second part deals with an unpublished fragment of the lunette of the Tempest stela stored in the Cheikh Labib magazine at Karnak which allows one of the oldest attestations of the rite of "driving the calves" to be identified.

Keywords : Neb-Phéty-Ré Achmosis Stèle de la tempête d'Achmosis Karnak

Type de document : Article dans une revue

Domaine : Sciences de l'Homme et Société / Archéologie et Préhistoire

Liste complète des métadonnées

RESSOURCES ASSOCIÉES

- HasPart**  Maucor, J. (2014) «CNRS-CFEETK 146972. Karnak, KIU 590 / bloc dit de « l'an 17 » XVIIIe dynastie / Achmosis» [Image] Nakala. <https://nakala.fr/11280/edeaa949>
- HasPart**  Bellod, A. (1975) «CNRS-CFEETK 13150. Karnak, KIU 590 / bloc dit de « l'an 17 » XVIIIe dynastie / Achmosis» [Image] Nakala. <https://nakala.fr/11280/ae65387d1>
- HasPart**  Chevrier, H. «CNRS-CFEETK 47217. Karnak, KIU 590 / bloc dit de « l'an 17 » XVIIIe dynastie / Achmosis» [Image] Nakala. <https://nakala.fr/11280/f73dbb7>

FICHIER

 3Biston-Moulin_Karnak 15.pdf

IDENTIFIANTS

- DOCID : 1667359
- HAL Id : hal-01667359, version 1

CITATION

Sebastien Biston-Moulin. À propos de deux documents d'Achmosis à Karnak. *Les cahiers de Karnak*, 2014. (hal-01667359)

EXPORTER

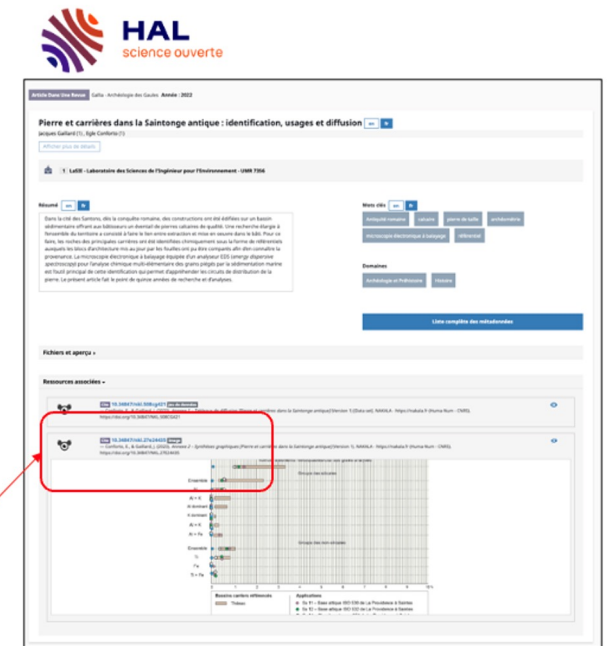
PARTAGER



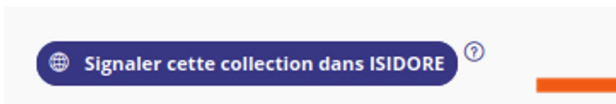
Bonus :

Mécanisme automatisé et standardisé pour lier des données dans Nakala à une publication dans HAL

En application des recommandations du CoSO



Nakala : option de signalement dans Isidore



101 collections de données déposées dans Nakala accompagnées en 2025 pour leur signalement dans Isidore



Un exemple : du document unique en ubykh aux enjeux de l'IA et des LLMs.



La chair de poisson rend intelligent

DOI: 10.34847/cocoon.20110515-308b-3c21-8f6a-4ad2995d6d13

auteur titre 1 anr

Eating fish makes you clever

Michaud, Alexis (compilateur, depositant) • Dumézil, Georges (researcher) • Sami, Tevfik (opérateur)

(prestation: 1968, dernière modification: 2010-03-08T15:12:40+01:00; mise à disposition: 2010-03-08; archivage: 2010-03-08T15:12:40+01:00; dernière modification de la notice: 2025-03-08)

Position dans le plan de classement

Collection Fungoros

La chair de poisson rend intelligente

Editeurs Laboratoire de langues et civilisations à tradition orale

Type(s) Types linguistiques: primary_text [?] Type(s) de Ressource: narrations [?]

Enregistrement sonore

Sujets Langue(s) objet d'étude: Ubykh code ISO-639-1: uby Méta-obj: Ubykh

Langue(s) Ubykh code ISO-639-1: uby

Droits: Freely accessible Copyright (c) Dumézil, Georges

Identifiant(s) doi:10.24807/FUNGOROS-0004021 acroftid:exp.knucoosn.20110515-308b-3c21-8f6a-4ad2995d6d13-0028

[?] Accéder à cette notice sur: phoscoo_000 [?] 10.34847/cocoon.20110515-308b-3c21-8f6a-4ad2995d6d13

Citation: APA Harvard MLA Vancouver

Michaud, A., & Dumézil, G. (2010, mars 1) [Chercher] COCOON. <https://cocoon.h.fr/eng/10.34847/cocoon.20110515-308b-3c21-8f6a-4ad2995d6d13>

Consulter le document

Une frégatation audio

Consulter les annotations

Eating fish makes you clever

Eating fish makes you clever

Eating fish makes you clever

Eating fish makes you clever

Partager

Code d'intégration dans une page web

dfhans:https://1007.pw/?https://cocoon.h.fr/eng/10.34847/cocoon.20110515-308b-3c21-8f6a-4ad2995d6d13

Télécharger

Format de conservation

• 100% (version: 2) taille: 8.04 Mo

51

fāx'a t'q "á-k' ábž'a k'áđ'a.n aza.x'a.š'a.na.n á-my'a.n g'a.k'a.q'á.n.

fāx'a	t'q "á-k' ábž'a	k'áđ'a.n	aza.x'a.š'a.na.n	á-my'a.n	g'a.k'a.q'á.n
un jour	deux hommes	en compagnons	étant devenu l'un pour l'autre	en chemin	entrèrent sur

Two men set out together on the road.
Un jour, deux hommes prennent la route ensemble.

52

a.fawtš.na my'áwaf a.x' ada.wtan a.k'á.na.n á-za.n fač'ála s' šb'ála x' ada.q'á.

a.fawtš.na	my'áwaf	a.x' ada.wtan	a.k'á.na.n	á-za.n	fač'ála	s' šb'ála	x' ada.q'á
pour eux manger	provision de route	eux pour acheter	étant allés	l'un	et fromage	et pain	

x' ada.q'á
il acheta

They bought some provisions for the journey. One bought cheese and bread,
Ils vont s'acheter des provisions pour la route. L'un achète du fromage et du pain,

53

áyda.xa.n-g'a s' šb'ála psá.la x' áda.n á.y.na.w.q'a.

áyda.xa.n-g'a	s' šb'ála	psá.la	x' áda.n	áy.na.w.q'a
l'autre-aussi	et pain	et poisson	ayant acheté	il apporta

the other fish and bread.
l'autre du poisson et du pain.

54



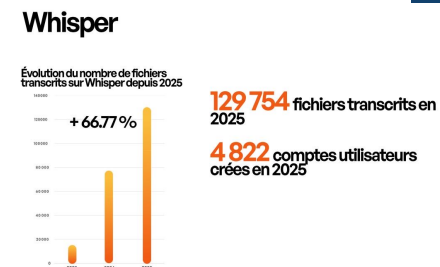
Apprendre les langues aux machines (8) - Benoît Sagot (2023-2024)
 Mathématiques et informatique - Collège de France



**COLLÈGE
DE FRANCE**

SÉMINAIRE

**Apprendre les langues
aux machines**

COLaF: Corpus et Outils pour les Langues de France et
variétés de français



Merci au CC-In2P3, à son directeur et à son équipe