



# Huma-Num

la TGIR des humanités numériques

Journées

## **PRESERVER SES DONNEES NUMERIQUES**

6-7 FEVRIER 2020 - CAMPUS CONDORCET

## **La préservation à Huma-Num**

Nicolas Larrousse



Aix-Marseille  
université

CAMPUS  
CONDORCET  
Paris-Aubervilliers

# Huma-Num

## Très Grande Infrastructure de Recherche au même titre que

### ESRF

Source Européenne de Rayonnement Synchrotron/  
European Synchrotron Radiation Facility



Type de l'infrastructure : TGIR

### FOF

Flotte Océanographique française



Type de l'infrastructure : TGIR

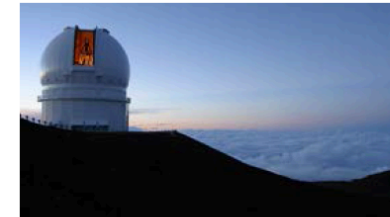
**STRATÉGIE NATIONALE  
DES INFRASTRUCTURES DE RECHERCHE  
ÉDITION 2018**

[www.enseignement-sup-recherche.gouv.fr](http://www.enseignement-sup-recherche.gouv.fr)

MINISTÈRE DE L'ÉDUCATION NATIONALE  
DE L'ENSEIGNEMENT SUPÉRIEUR  
ET DE LA RECHERCHE

### CFHT

Canada-France-Hawaii Telescope



Type de l'infrastructure : TGIR

**HN** Huma-Num  
la TGIR des humanités numériques



# Un écosystème centré sur les projets et données de la recherche



# Collaborations Nationales

- Un ensemble de correspondants dans le réseau des MSH (Maisons des Sciences de l'Homme) ;



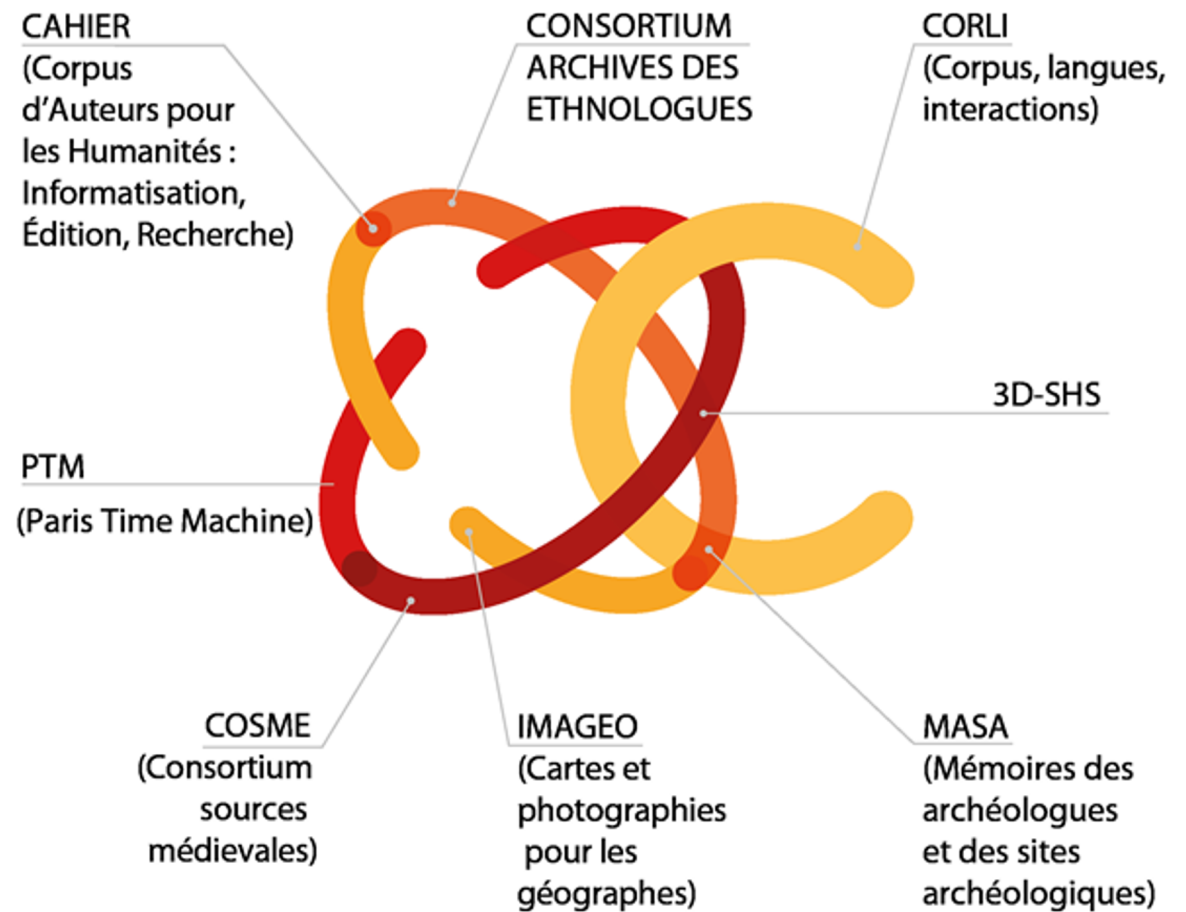
- Des collaborations avec d'autres infrastructures nationales





# Les consortiums HN, un partage de bonnes pratiques et réseaux d'experts

- **8 consortiums**
- 80-100 équipes de recherches qui couvrent des champs variés des SHS
- Un réseau de chercheurs, ingénieurs, bibliothécaires, archivistes, etc.
- + de 50 actions de formation par an
- Environ 40% budget de la TGIR



# Quelques réalisations des consortiums

## Consortium 3D

Intégration de formats 3D au CINES  
Développement d'une application  
de création de paquets



## Consortium MASA

Modèle de DMP pour l'archéologie déposé  
sur DMP Opidor (INIST)



### CNRS: MASA - Modèle de PGD (français) - Contexte

#### Identification du projet

Identifiant du projet de recherche

*Recommandations:*

Par exemple, s'il s'agit d'un projet de recherche financé par l'Agence nationale de la recherche (ANR), indiquer ici le **numéro attributif d'aide**. Pour information, c'est aussi le numéro à indiquer dans l'archive ouverte de publications scientifiques [HAL](#) lors du dépôt d'un article ou d'une publication produite dans le cadre de ce projet ANR.

*Exemple de réponse:*

ANR-12-BSH3-0012

Titre du projet de recherche

*Recommandations:*

Reprendre l'intitulé qui a été a été déposé ou enregistré.

*Exemple de réponse:*

Néolithisation du Proche-Orient

# Réaliser une veille efficace pour les SHS à l'international



Suivre et/ou participer à des initiatives internationales

Identifier les enjeux de la structuration de la recherche aux niveaux Européens et Internationaux



**EUROPEAN OPEN  
SCIENCE CLOUD**



# SERVICES POUR LES DONNÉES NUMÉRIQUES



## STOCKER

Entreposer . Organiser

## ARCHIVER

Préservation à long terme



## TRAITER

Outils . Logiciels

## SIGNALER

Enrichissement sémantique  
Accès unifié



isidore



## DIFFUSER

Machines virtuelles  
Diffusion web



## DONNÉES DE LA RECHERCHE

Partenariat avec le CC-IN2P3  
et le CINES

## EXPOSER

Documenter . Partager



nakala

nakal(©)na

## Des services d'infrastructure classiques ...

### STOCKER

- Dès le début du projet
- Mettre à l'abri des données, des documents (Huma-Num Box)
- Travailler collectivement sur ces données (Sharedocs)

### TRAITER

- Dès le début du projet et au fil de l'eau
- Travailler les données (formats, organisation, calcul)
- Accès à des logiciels sur HN ou à distance (R Studio Pro, Server FileMaker Pro, SIG...)

### DIFFUSER

- A tout moment
- Diffusion rapide mais non pérenne (sites Web)

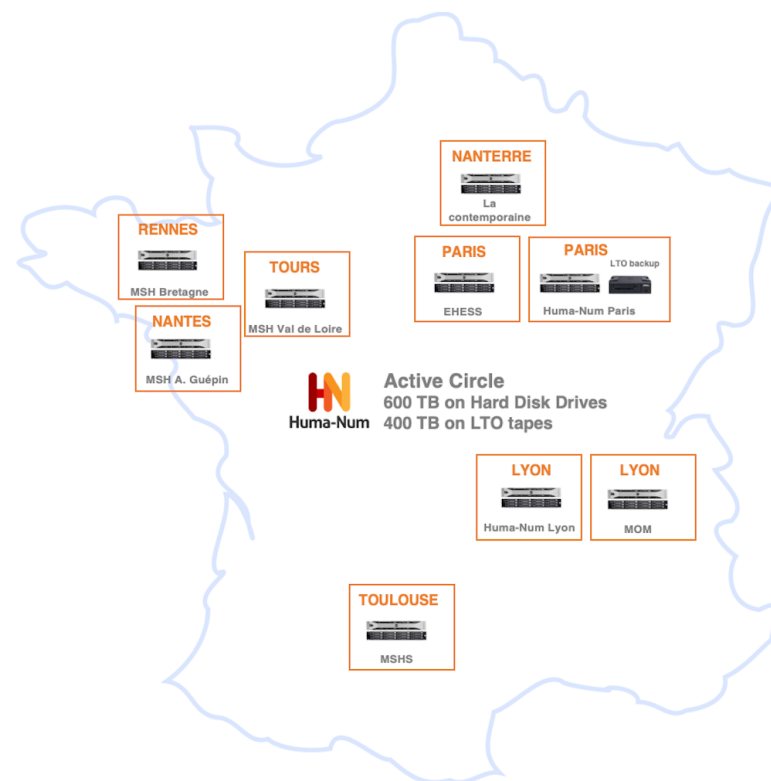
# Huma-Num Box

## Un stockage distribué pour les données « tièdes »

Maillage du territoire

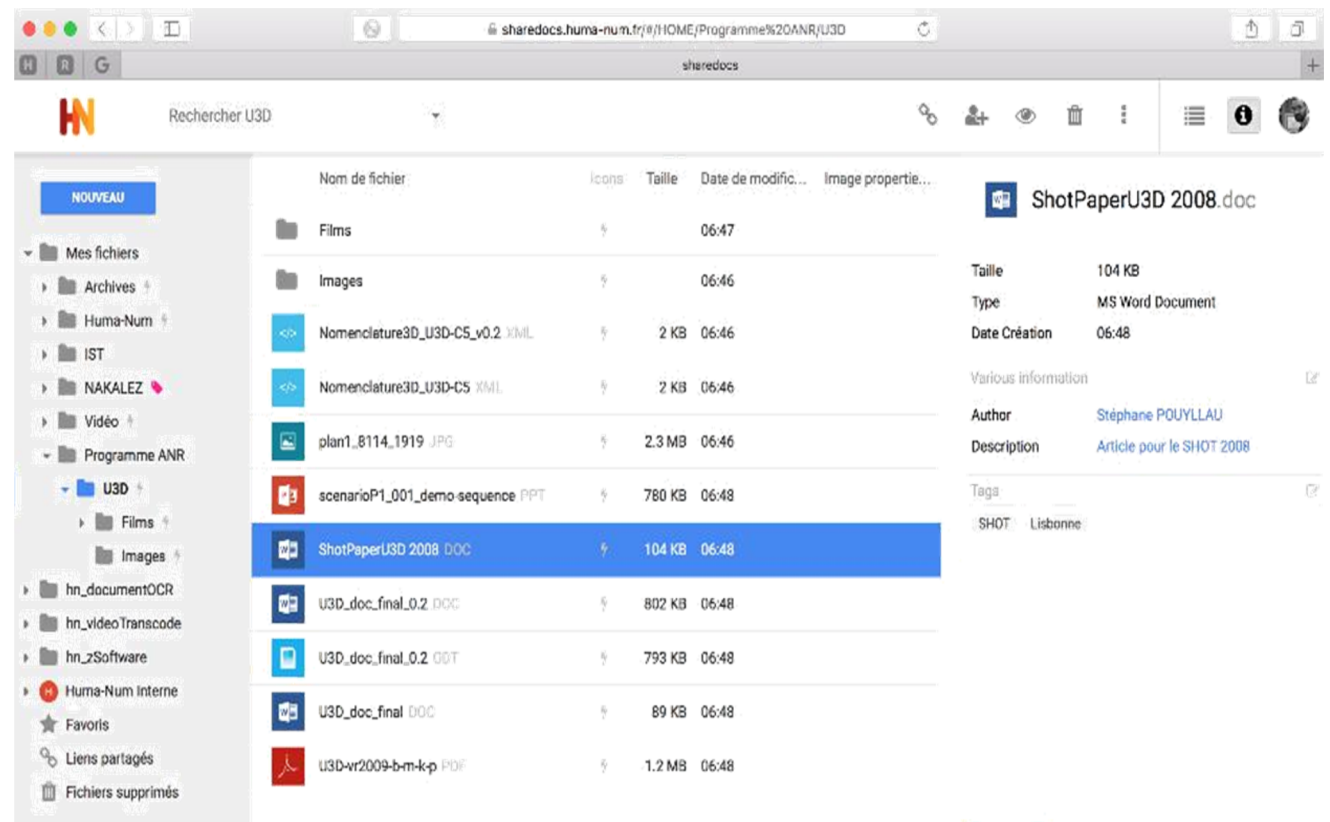
Duplication géographique

Différents supports technologiques



# ShareDocs

- **Stockage « outillé » pour les SHS**
- Stockage des données en mode « dropbox » (mais sur les serveurs d'Huma-Num, en France)
- Partage, édition, traitement des données
- Un espace de travail collaboratif, sécurisé.





# Entrepôt NAKALA



- > Browse Items
  - > Browse Items Per Country
  - > Advanced Search Items
- 

IFRA has built up a press review of East African newspapers since the late 1970's. The collection (more than 50,000 press cuttings) is made up of Kenyan Newspapers (Daily Nation, the Standard, the East African, Kenya Times), Ugandan Newspapers (New Vision, The Monitor) and Tanzanian Newspaper (The Guardian). With the support of the Institut Français, more than 15,000 pages of press articles have been digitised. These archives are classified by thematic National and international politics (elections, political parties, personalities, public politics, etc.), Society (morality, civil society, religion, deviance, union, etc.), Economy (commercial policies, national and regional economy, urban planning, land use, etc.), Agriculture (farmers associations, union, crops, etc.).

## Recently Added Items

### Yeltsin's op due



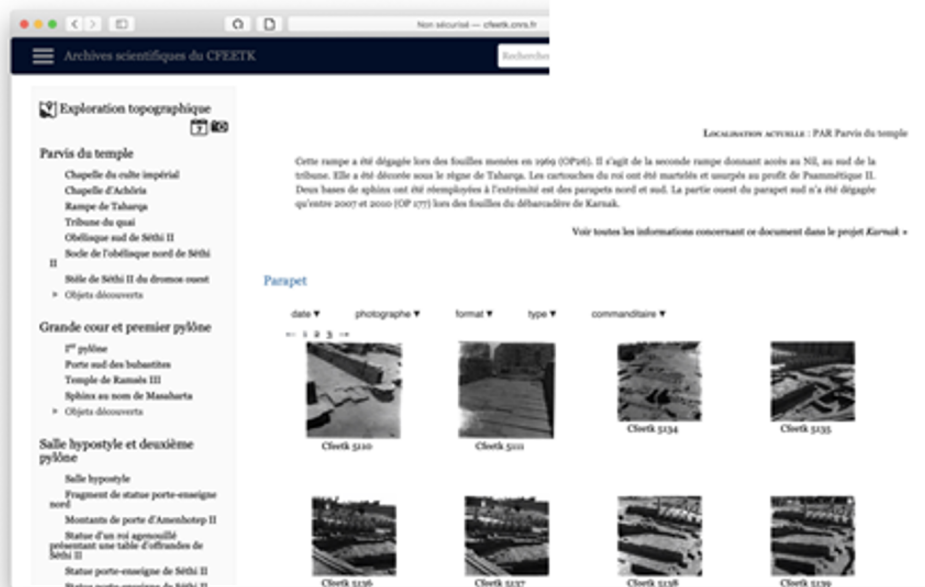
Health

### More defence groups urged



Security

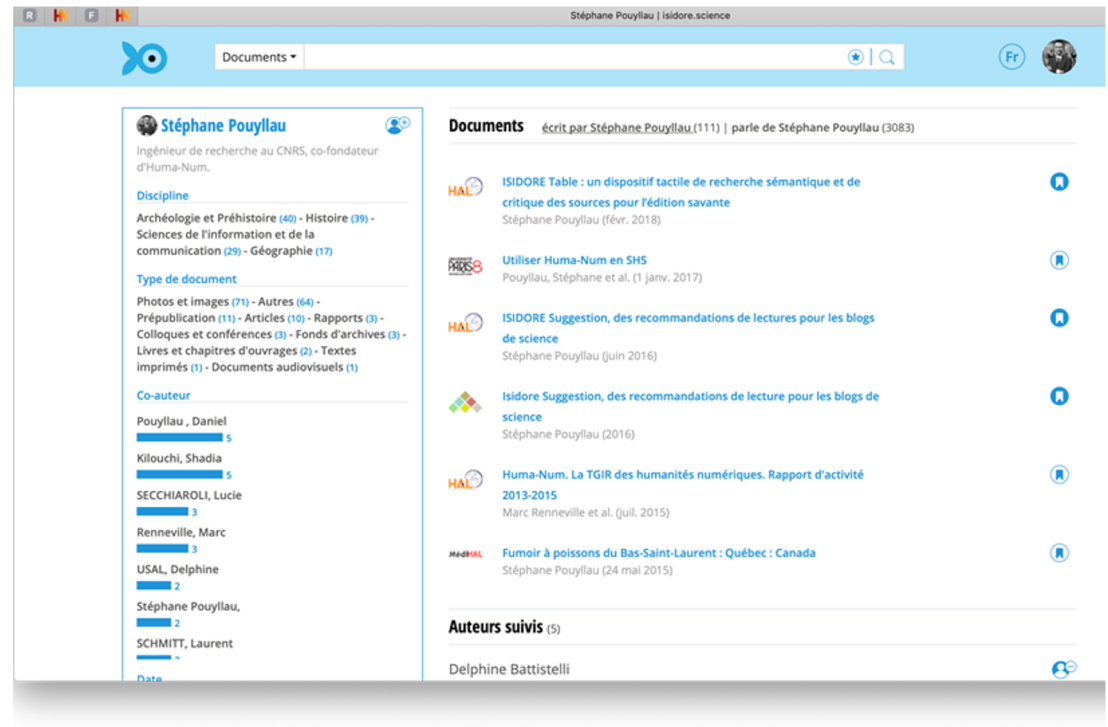
L'IFRA a constitué un fonds d'archives de presse (plus de 50



# isidore.science



- Un “assistant” de recherche
- un réseau social de scientifiques SHS
- un outil de veille scientifique et de diffusion dans le cadre de la science ouverte



https://isidore.science



Votre assistant de recherche  
en Sciences Humaines et Sociales

Collections ▾ Rechercher dans les 7 368 collections de ISIDORE... | Q

[i démonstration](#)

## Préserver et archiver les données

Huma-Num accompagne les projets de préservation à long terme

Liens entre les producteurs de données et le CINES

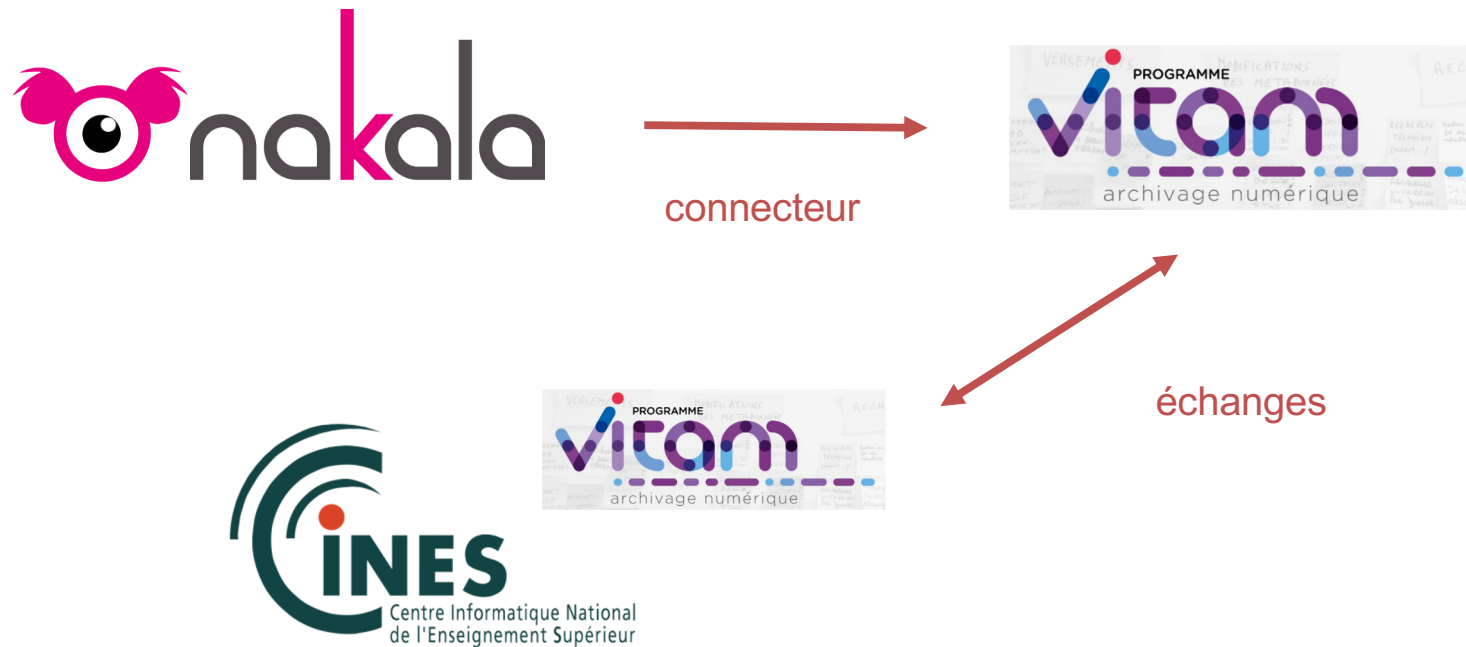
Suggestion de nouveaux formats et prise en compte (e.g. formats pour la 3D, la TEI)

Liens avec la communautés des archivistes et les institutions



# Préserver et archiver les données

Test du logiciel Vitam avec NAKALA



Définition d'une politique d'archivage