



PRESENTATION DE PROJETS OMEKA DE L'INIST



AVEC DES DONNÉES ISSUES DE NAKALA

Titouan BOUDART

ANF 2022 – MAÎTRISER L'EXPOSITION DES
DONNÉES ENTREPOSÉES DANS NAKALA

25 NOVEMBRE 2022

CNRS - INIST

INTRODUCTION - OMEKA À L'INIST

Notre offre de service / pourquoi
Omeka ?

Mettre à disposition des laboratoires des plateformes web collaboratives pour le partage et la valorisation de leurs données de recherche

INTRODUCTION

- Service exposition de données
Département « Valoriser les données de recherche »
 - **5 documentalistes** chargés de la gestion des projets, des relations avec les laboratoires, de la création, curation, enrichissement, modélisation des métadonnées
 - **3 informaticiens + 1 correspondant** du Service Ingénierie des Systèmes d'information (**SISI**) pour le déploiement + les Administrateurs systèmes et réseaux (**ASR**)
 - **2 techniciens** spécialisés dans la numérisation de documents (atelier numérisation)



INTRODUCTION

- Omeka - logiciel libre permettant de valoriser la production scientifique des laboratoires par la création de bibliothèques numériques
 - Réalisation du **Roy Rosenzweig Center for History and New Media de l'Université George Mason** (Virginie, États-Unis)
 - Nombreux contributeurs dans le monde et très actifs en France

INTRODUCTION

- Pourquoi Omeka ?

Logiciel Libre et Open Source

Large communauté

APIs, standards d'interopérabilité (OAI PMH, formats d'exports xml, json...)

Fonctionnalités avancées :

- Imports de données en CSV/TSV
- Moteur de recherche à facettes
- Liens entre les ressources
- Cartes

INTRODUCTION

- Omeka S

Plus récent, coexiste avec la version Classic

Utilisé à l'INIST pour les nouveaux projets (et migration des anciens sites sous Classic vers la version S)

Orienté vers le web sémantique

Des fonctionnalités en plus :

- Multi-sites
- IIIF (Images utilisables par défaut + module pour la présentation en manifeste)
- Un module qui permet d'accéder en direct à des vocabulaires tels que GeoNames, IdRef, les vocabulaires du Getty, PACTOLS, Frantiq ...

Industrialisation

Demande croissante de sites

- Nécessité de garder une homogénéité de nos sites



Plateforme « de référence », facilite la maintenance et les mises à jour

- Sélection de plugins préconfigurés (~ 30)
- Moteur d'indexation Solr
- Notre thème « générique » customisé ensuite

INTRODUCTION

Quelques chiffres

- 17 plateformes en production sous Omeka Classic
- 11 plateformes en cours de développement sous Omeka S

28

Nombre de plateformes

21

Laboratoires

3

MSH

ÉDITORIALISER DES DONNÉES

Workflow avec les laboratoires

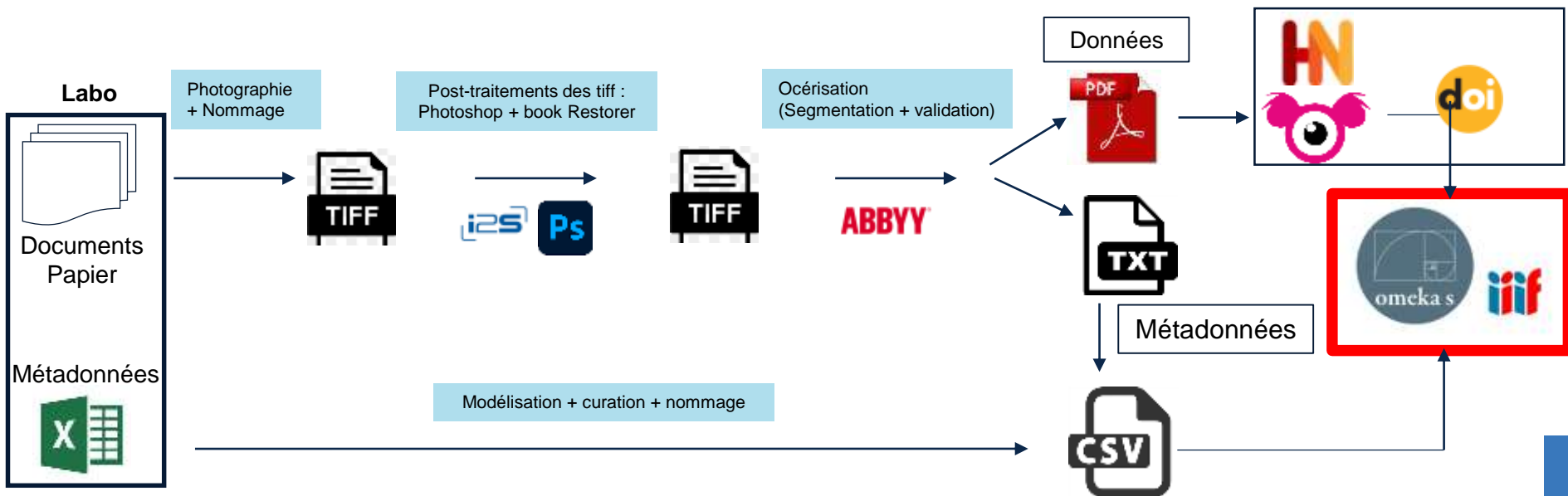
ÉDITORIALISER DES DONNÉES

- Principales étapes et opérations :
 - Demande et étude de faisabilité avec le laboratoire
 - Établissement d'un devis (fait office de convention)
 - Transmission des métadonnées conformément au format prédéfini + médias
 - Mise en place de l'infra et déploiement plateforme de Référence
 - Traitement des métadonnées : conversion, curation, enrichissement
 - Traitement des documents associés : numérisation
 - Customisation de l'apparence du site
 - Mise en production du site après validation par le laboratoire

ÉDITORIALISER DES DONNÉES

Du document papier au texte intégral

Océrisation et indexation par le moteur solR => chercher dans les documents en langage naturel



ÉDITORIALISER DES DONNÉES

- Les problématiques
 - Modèle de données beaucoup moins restreint sur un site Omeka (entièrement libre) que sur Nakala
 - Difficultés des laboratoires à déposer en lots sur Nakala (manque de connaissances, API et outils divers)

ÉDITORIALISER DES DONNÉES

- Dans le futur
 - Utilisation d'images IIIF provenant de Nakala dans nos plateformes
 - Étude (en cours) d'une procédure visant à déposer en lots sur Nakala depuis nos sites Omeka S

EXEMPLES AVEC DES DONNEES PRESENTES DANS NAKALA

EXEMPLES



Cenhtor

cenhtor-msh-lorraine.cnrs.fr



RevColEurop

revcoleurop.cnrs.fr



Architoul

architoul-msh-toulouse.cnrs.fr

Merci



inist.fr

www.cnrs.fr