

Journée Utilisateurs CINES – 22 juin 2016

L'archivage pérenne des thèses et mémoires : un partenariat Lille 2 - CINES

Groupe PEPITE de l'Université de Lille Droit et Santé :

Anne Flahaut-Steiner (SCD)

Sophie Quenton (DSI)

Anne-Sophie Guilbert (SCD)

Antoine Cordier (DSI)

Solenn Bihan (SCD)

Contexte et originalité du projet

- Plateforme locale de diffusion **PEPITE** (thèses de doctorat, **thèses d'exercice, mémoires**) utilisant **Nuxeo** et **ORI-OAI**
Thèse/mémoire = fichier(s) PDF + MD de référencement
- Charte de diffusion des thèses et mémoires (01/01/2014) :
Version électronique = version de référence => **conservation pérenne**
- Lille 2 est le **premier établissement français** à archiver ce **type de documents** (fichiers + métadonnées)



DURÉE :
1 année complète
De janvier 2015 à
janvier 2016

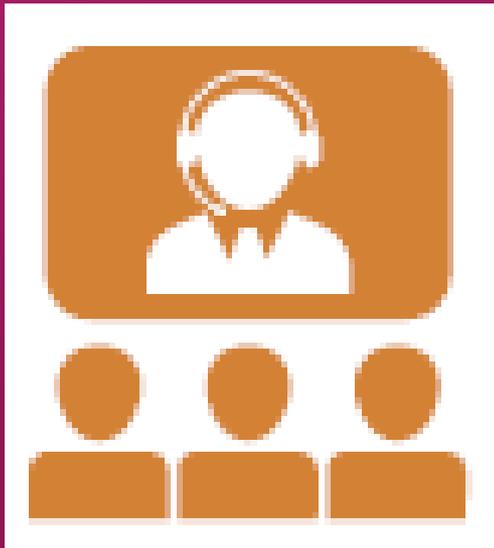


ATOUT : 1 équipe projet de **5** personnes qui ont suivi **PEPITE** depuis le début avec des compétences complémentaires :

- **3** bibliothécaires
- **2** informaticiens
- **20** semaines *temps homme*

Plan

1. Documenter le projet et se comprendre
2. Préparer et conduire des tests
3. Automatiser les versements: une vraie valeur ajoutée de Lille 2
4. Adapter le workflow
5. Archiver la production courante
6. Des pistes d'amélioration



1. Documenter le projet et se comprendre

Janvier - Juillet



Documenter le projet et se comprendre

- Échanger des courriers officiels entre le CINES et la direction du SCD
- Se réunir pour lancer et **planifier** le projet (visio-conférence)
- Comprendre et renseigner les **nombreux documents** envoyés par le CINES.

Ex: renseigner le ppdi

PROBLEMES : Culture métier différente entre le CINES (archivistes, informaticiens) et Lille 2 (bibliothécaires, informaticiens)
Temps d'adaptation réciproque

SOLUTIONS : Très bon accompagnement au niveau du CINES
Réunions mensuelles avec CR pour garder le rythme de travail



2. Préparer et conduire des tests

Juin - Octobre

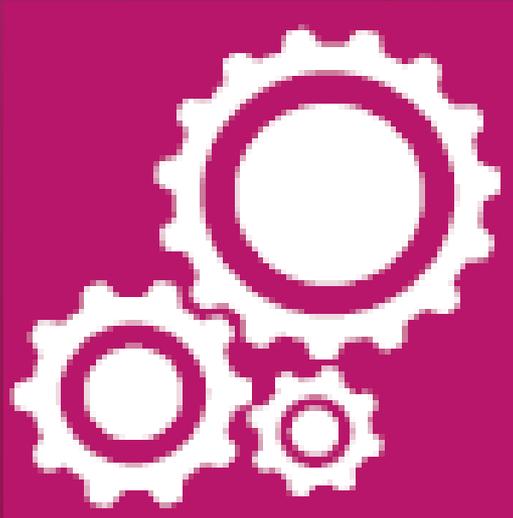


Préparer et conduire les tests

- Repérer les **fichiers PDF** archivables : mise en place d'un **programme de tri** pour séparer les fichiers PDF valides et ceux à retraiter
- Faire en sorte que les documents versés soient **conformes** aux caractéristiques demandées par le CINES : écrire un **schéma de métadonnées** pour les mémoires

PROBLEMES : Ecrire un schéma de métadonnées
Modifier par lot toutes les notices pour les rendre conformes au schéma retenu

SOLUTIONS : Autoformation sur les schémas Dublin Core.
Aide importante du CINES pour valider le schéma et corriger les erreurs.



3. Automatiser les versements : **Une vraie valeur ajoutée de Lille 2 !**

Septembre - Décembre



Automatiser les tests

- Simuler des **envois** de documents sur le serveur de test du CINES
- **Repérer les erreurs** et les corriger avant le passage en production

Le travail sur l'automatisation des échanges avec le CINES est une forte valeur ajoutée de la part de Lille 2

- Export mensuel des fichiers par **requêtes dans Nuxeo** (doc en .pdf) **et ORI-OAI** (notices)
- **Création des bordereaux de versement** (sip.xml) à partir des notices, et création des structures de paquets contenant à la fois le fichier .pdf et la notice .xml
- **Traitement automatique des retours CINES** pour pointage, suivi et corrections



4. Adapter le workflow

Septembre - Décembre



Adapter le workflow

- Atteindre un niveau acceptable de **documents rejetés** par le CINES (<5%)

PROBLEMES :

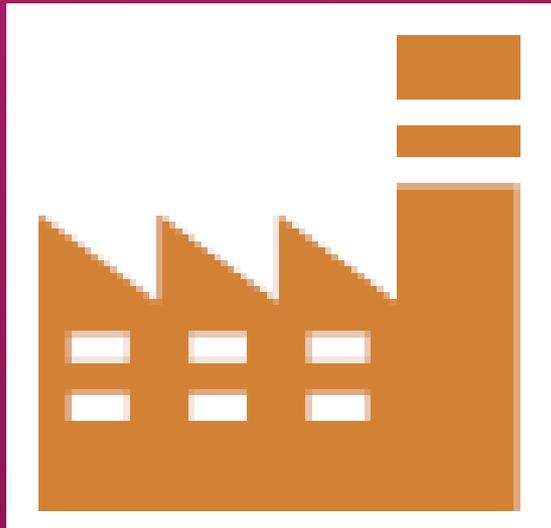
PDF non valides : 60% de rejet lors des premiers tests !

Nommage non conforme aux règles définies dans la documentation du projet

SOLUTIONS : Implémentation d'un **module de contrôle de validité des PDF (FACILE)** lors du dépôt (Nuxeo)

Optimisation des fichiers PDF non valides avec Acrobat Pro

Correction manuelle des noms de fichiers et consignes plus précises pour les nouveaux fichiers



5. Archiver la production courante

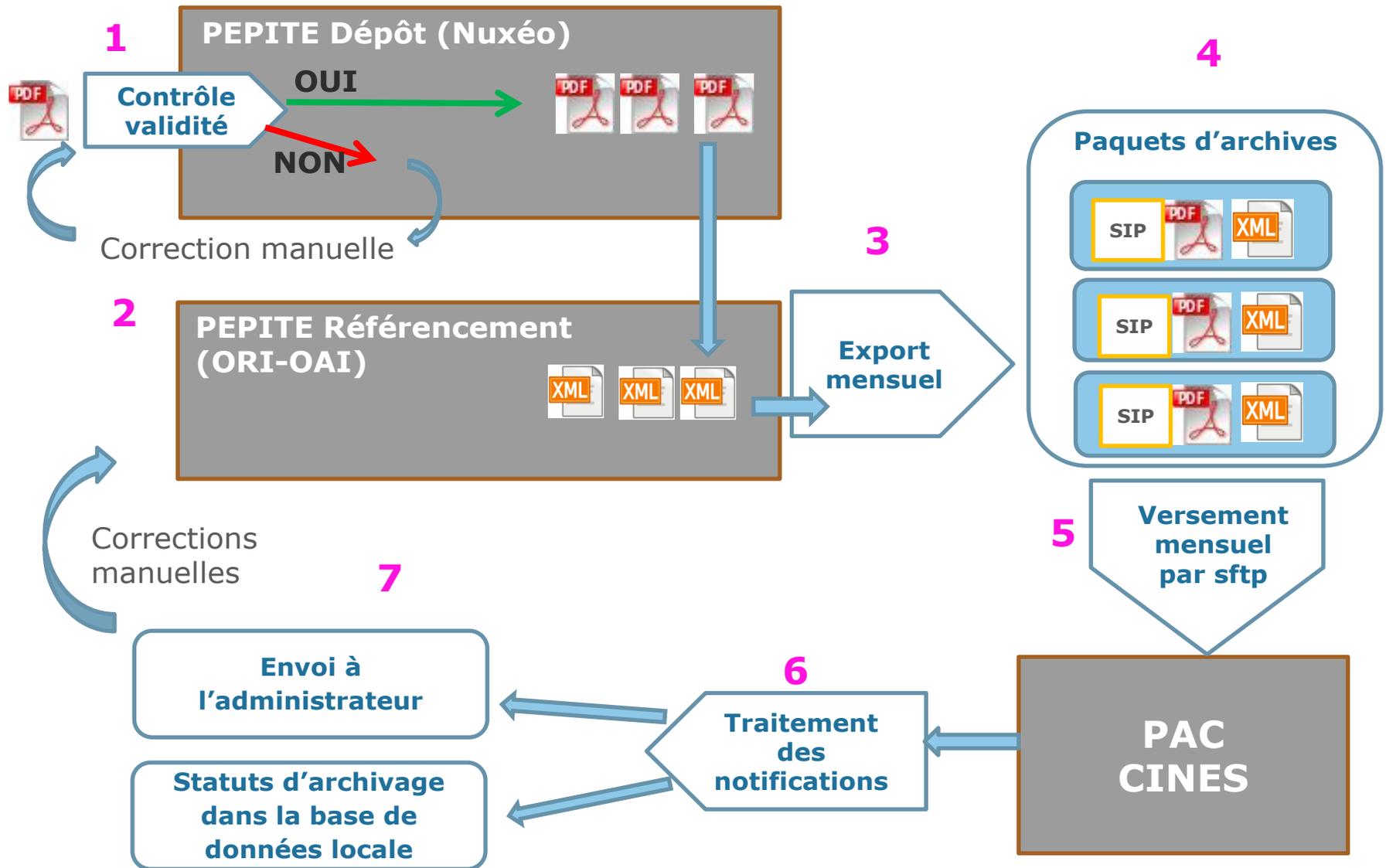
Janvier...

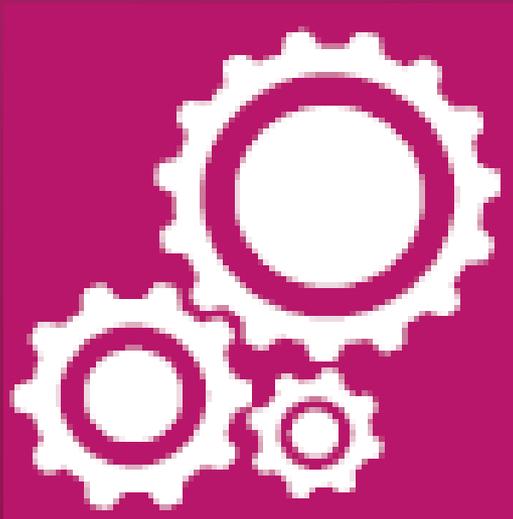


Archiver la production courante

- Passer en production (**versement régulier**) sans surcharge de travail
- Mettre en place un **contrôle qualité** pour gérer de manière fluide les rejets résiduels
- Conservation des **certificats d'archivage** renvoyés par le CINES
- Pas de souci technique dans la mise en place des versements
- Grâce à l'automatisation des échanges et à l'adaptation du workflow, il n'y a **pas de surcharge de travail** pour la gestion courante et les rejets sont minimales
- Il reste à traiter manuellement le **rétrospectif** (depuis l'adoption de la Charte de Diffusion des Thèses au 1^{er} janvier 2014)

Schéma de versement des documents au CINES





6. Des pistes d'amélioration

Pour les établissements : + de rigueur

Des PDF valides (FACILE)

Des règles de nommage des fichiers pour tous les cas particuliers

Un schéma de métadonnées pour chaque type de document

Pour le CINES : + de pédagogie

Diffuser une culture d'archivage dans les établissements

Simplifier les documents fournis aux utilisateurs

Former les établissements à la gestion des formats PDF

**Merci de votre
attention**

pepite@univ-lille2.fr

