



Service de préservation à long terme des données numériques de la recherche SHS

Journées OAIS – 23 et 24 octobre 2012

- Les objectifs et les acteurs
- Les différentes étapes
de la construction d'un plan d'archivage à la diffusion des données
- Les technologies utilisées
- Bilan
- Les services à l'étude

Archivage à long terme – Les objectifs

- Résoudre les problèmes de conservation des données produites par la recherche
- Préservation à long terme de ces données ainsi que leur lisibilité
- Fournir un accès en ligne à ces données (web service)

Archivage à long terme – Les acteurs

- Les services versants:
 - les ex-CRDO de Paris et d'Aix-en-provence: producteurs de données
- Le CINES (Montpellier):
 - Archivage à long terme des données sur PAC
- Le TGE-ADONIS:
 - coordinateur du projet et diffusion des données

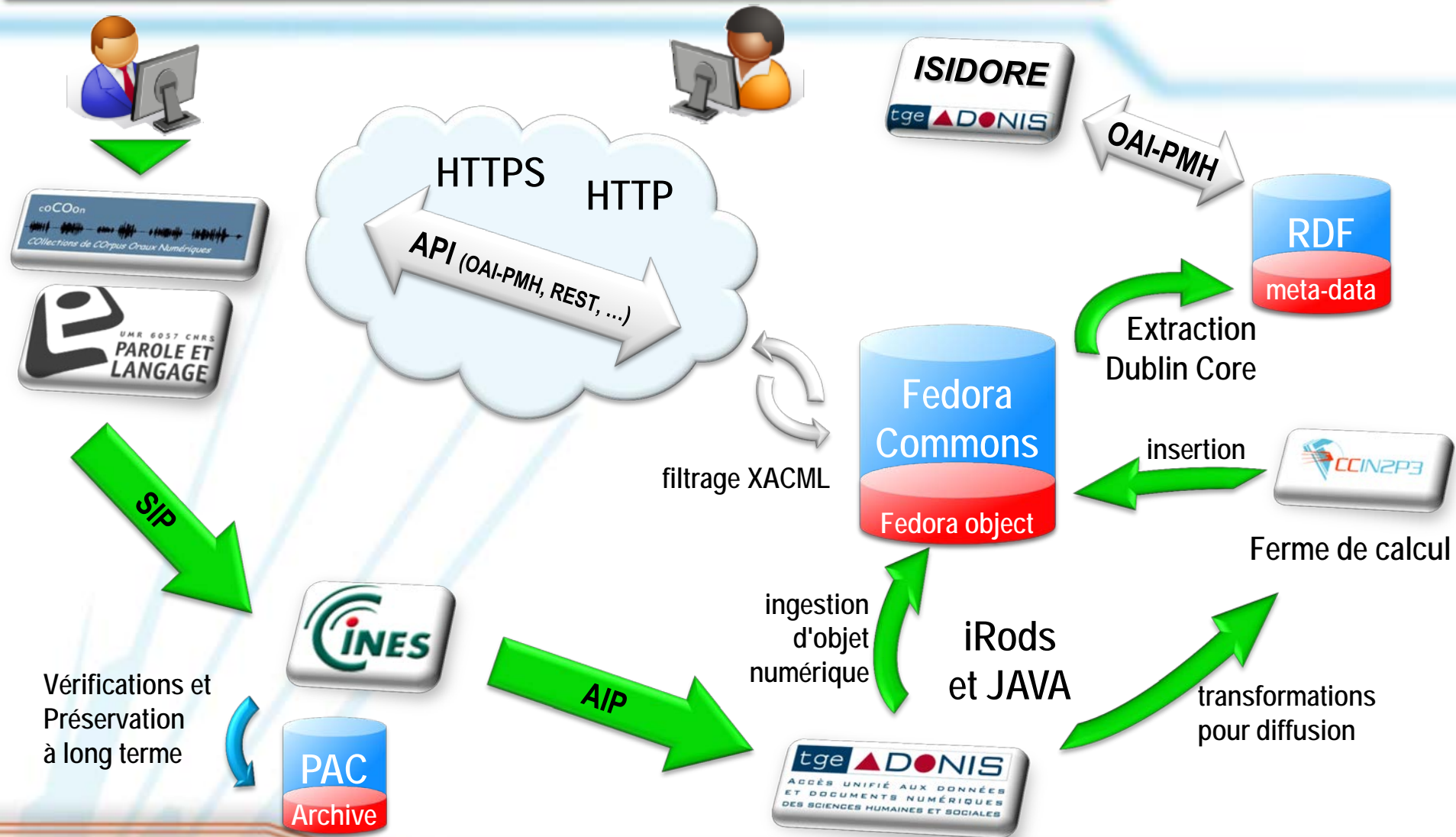
- Constitution du plan d'archivage
 - inclus dans la curation globale des données
 - "épuration" des données
 - définition de la granularité de l'archive
(collection, corpus, ...)
 - définition des relations

- Construction des SIP
 - <http://www.cines.fr/pac/sip.xsd>
 - Dublin Core
 - autres métadonnées de l'archive
(relations de version, parenté ou mise à jour; identifiant; ...)
 - liste des fichiers et leurs métadonnées
(empreinte, ...)
- AIP: SIP + métadonnées d'archivage
 - <http://www.cines.fr/pac/aip.xsd>

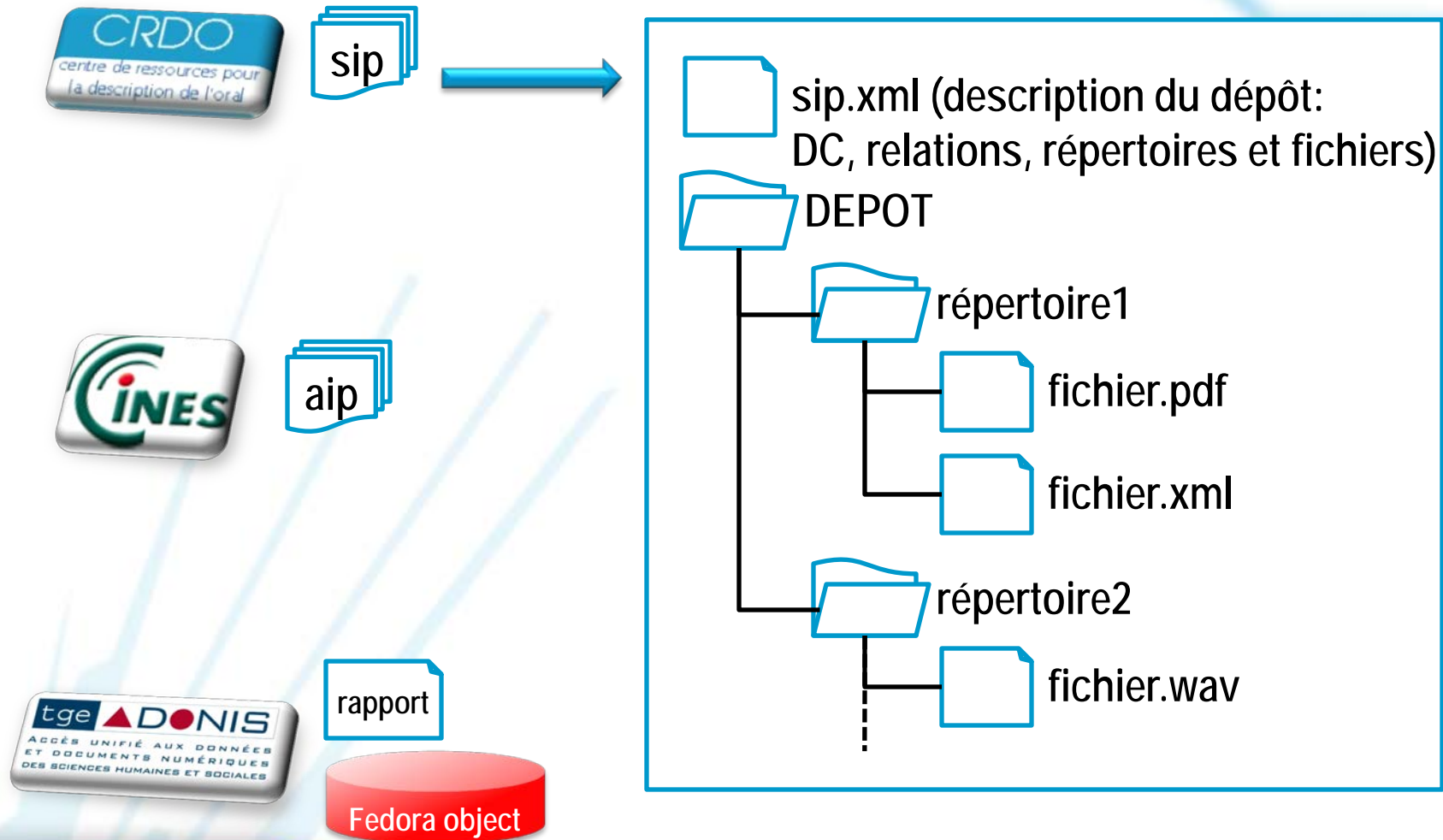
- Validation des formats (CINES)
 - liste réduite mais évolutive
 - sélection de format pérenne
 - migration si format obsolète
 - tests unitaires sur <http://facile.cines.fr/>

- CINES: système PAC
- TGE-ADONIS
 - iRods
 - FedoraCommons
 - Java, XML, RDF
 - OAI-PMH, SOAP, REST
 - ferme de calcul du CC-IN2P3

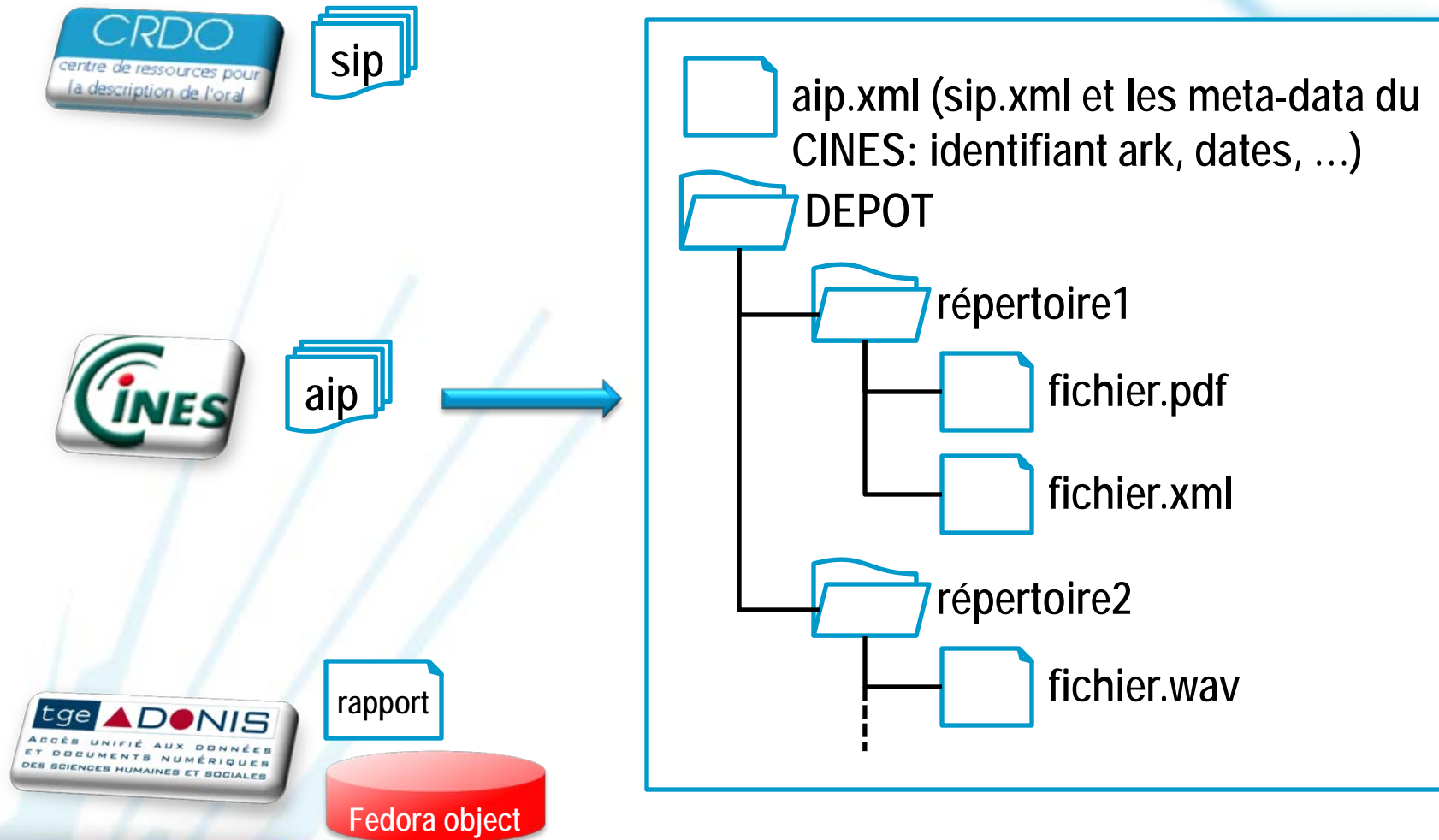
Archivage à long terme – Le flux des données



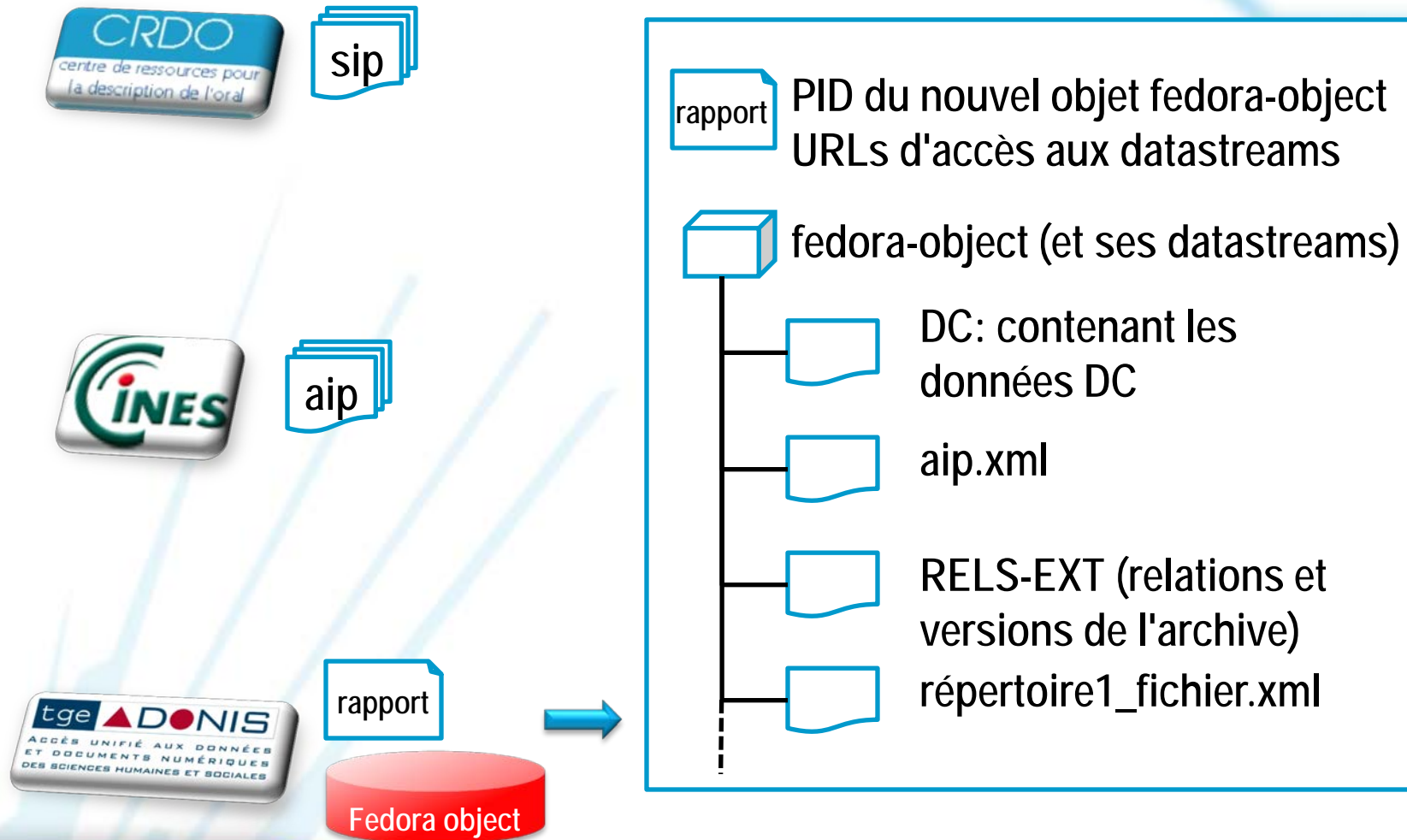
Archivage à long terme – Le flux des données



Archivage à long terme – Le flux des données



Archivage à long terme – Le flux des données

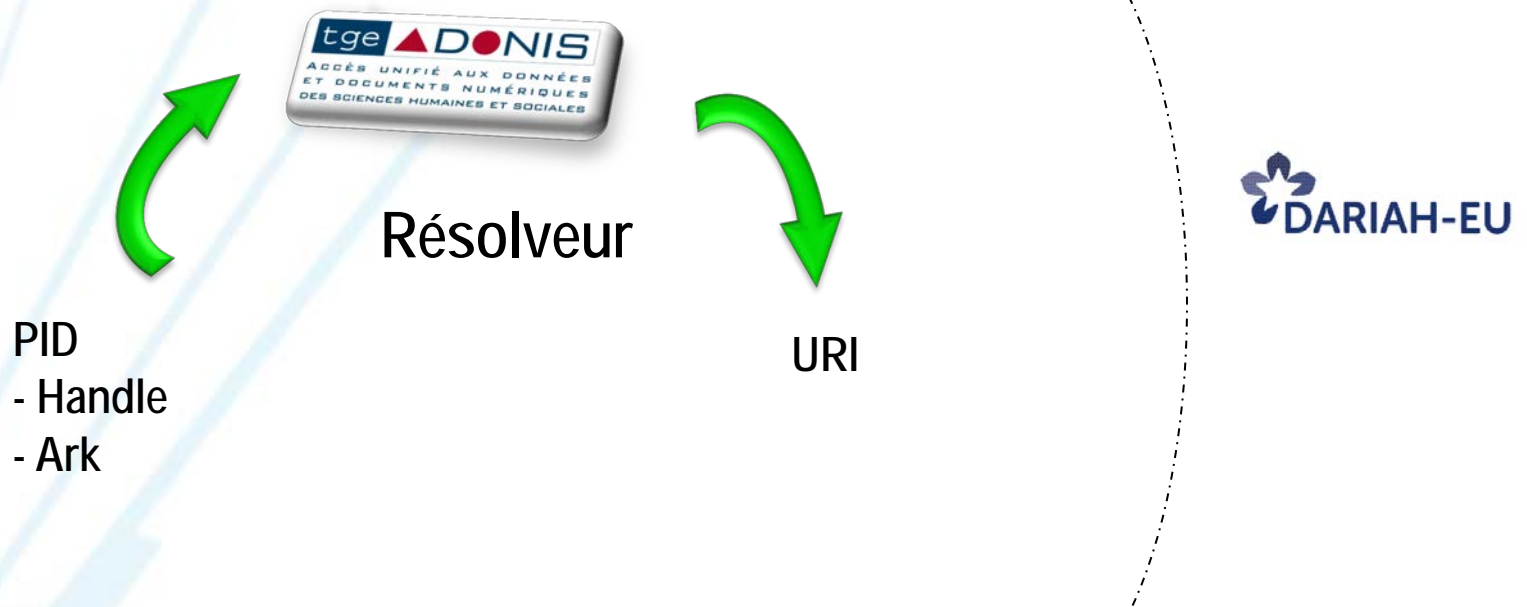


Archivage à long terme – Bilan

- En production depuis mars 2010
- Plateforme évolutive
- Demande beaucoup de préparation au service versant (plan d'archivage)

- Finaliser le projet pilote
- Ouverture vers d'autres services versants
- Proposer d'autres services

- Mettre en œuvre une architecture d'identifiant pérenne



- Affiner la gestion des droits associés au données



XACML



- Prévoir un dispositif d'authentification



SSO

■ Ouvrir à d'autres services versants

Intégrer les pratiques de la communauté

Nouveaux formats à prendre en compte



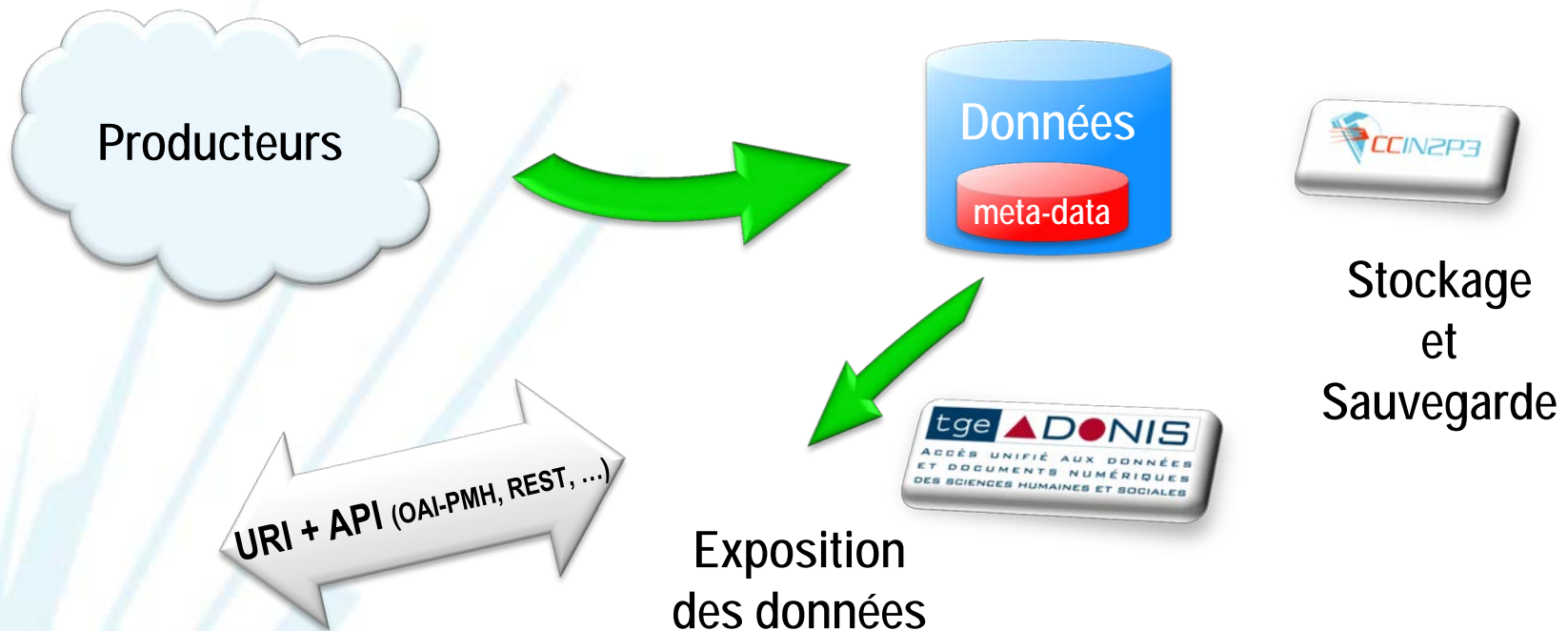
- Proposer un ensemble de services aux chercheurs
 - préservation de données
 - exposition des données

■ Stockage sécurisé simple

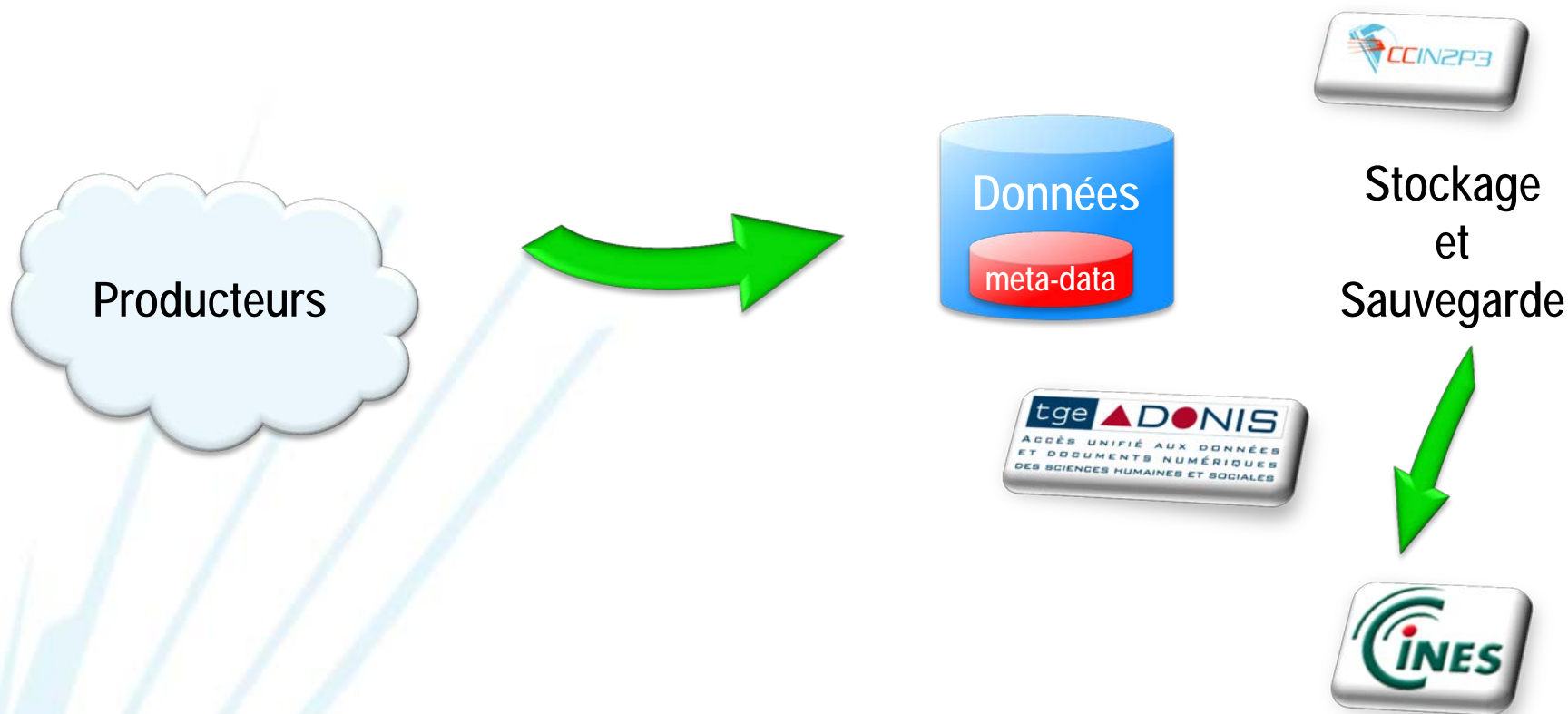


Archivage à long terme – Services complémentaires

■ Stockage sécurisé & Exposition des données



■ Stockage sécurisé & Pérennisation



Archivage à long terme – Services complémentaires

■ Stockage sécurisé & Pérennisation & Exposition des données

