



La Théorie sur la "voie verte" de l'Open Access

03-12-2008

Jocelyne Jerdelet - Sandrine Reyes
CERN-DSU-SI



La Théorie sur la "voie verte" de l'Open Access

- L'Open Access (OA) au CERN
- L'OA et le Service d'Information Scientifique (SIS)
- L'OA et les objectifs du SIS
- La Théorie: 1ère Etude concrète d'une série OA
- Tâches et difficultés rencontrées
- Evaluation finale
- Conclusions



L'Open Access (OA) au CERN

Les Principes de l'OA soutenus par le CERN depuis sa création

- **Par sa convention** : Adoptée en 1953 par 12 Etats Membres stipule que *«les résultats de ses travaux expérimentaux et théoriques [de l'Organisation] sont publiés ou de toute autre façon rendus généralement accessibles»*

 - **Par les avancées technologiques**
 - Création du Web (1989) qui popularisa Internet
 - Serveurs de documents électroniques
- 1993 : CERN Document Server (CDS) - adapté aux besoins des chercheurs et des documentalistes



L'OA et le Service d'Information Scientifique (SIS)

- **Mission du SIS** : Acquérir, gérer et diffuser les résultats de recherche du CERN à la communauté scientifique
- **Informatisation du catalogage** :
 - 1982 - Informatisation du système de gestion des données bibliographiques du SIS
- **Numérisation** :
 - 1990 - Début de la numérisation pour les nouveaux documents
 - 1993 - Mise en ligne sur CDS de ces documents



L'OA et le Service d'Information Scientifique (SIS)

- **Substitution du catalogage par la soumission électronique**
 - 1991 - Les auteurs soumettent leurs papiers sur le serveur d'arXiv
 - 1994 - CDS propose sa propre soumission électronique pour les documents CERN.

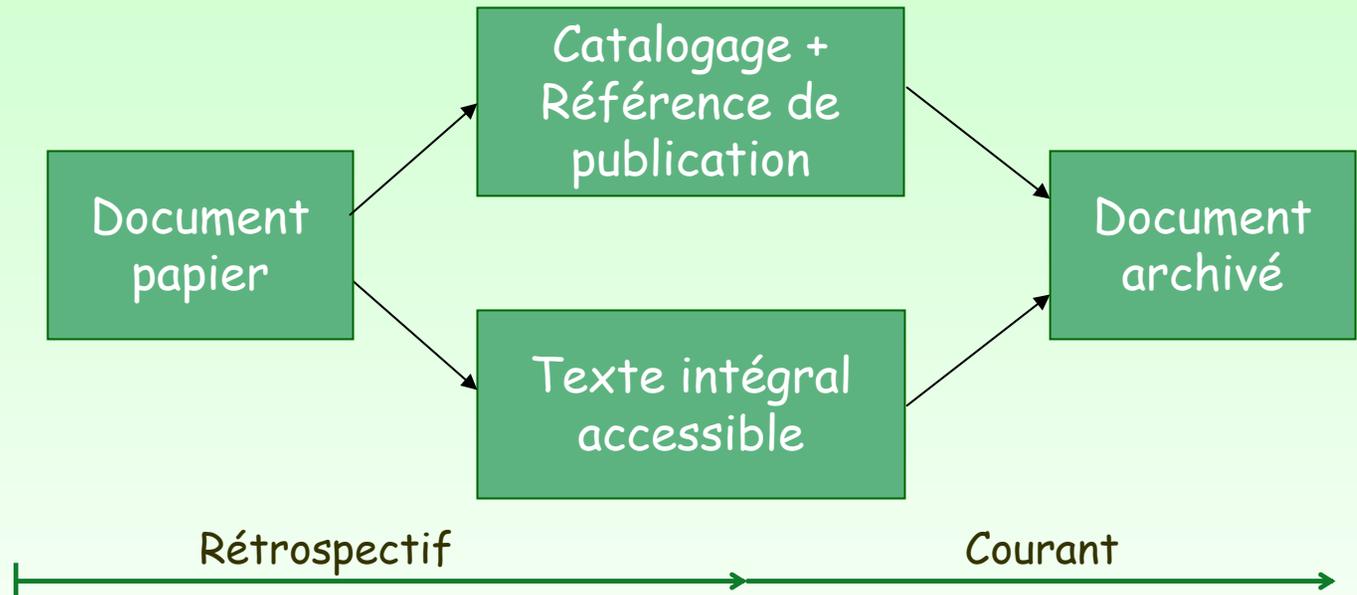
- **Importation de notices bibliographiques :**
 - 1994 - Importation des métadonnées d'arXiv sur CDS

- **Importation de fichiers électroniques :**
 - Travail considérable de numérisation de pré-tirages dont ceux du CERN par l'Institut KEK au Japon.
 - Autorisation de KEK de lier leurs numérisations à nos notices bibliographiques



L'OA et les objectifs du SIS

Privilégier les collections de documents produits par le CERN



- Pré-tirage catalogué
- Recherche de la référence de publication et mise à jour de la notice bibliographique
- Texte intégral en accès libre sur CDS
- Copie papier conservée aux Archives



La Théorie

1ère Etude concrète d'une série OA

➤ Historique administratif :

2006 - 12000 documents de la Division Théorie versés aux Archives du CERN.

Choix du SIS : Collection + Prolifique et + ancienne du CERN pour appliquer les principes de l'OA.



- **Numérotation** : Numéro de rapport unique pour l'identification « CERN-TH-Numéro »
- **Archivage du Secrétariat** : 50 ans de production de pré-tirages répertoriés et identifiés par années, numéros de rapports, titres, auteurs (Base de travail)



La Théorie

1ère Etude concrète d'une série OA

➤ Historique bibliographique :

- **Catalogage** : Catalogage rétrospectif antérieur à 1982 à faire (notamment les 2000 premiers documents)
- **Numérisation** : Numérisation antérieure à 1990 à faire
- **Références de Publication** : Recherche dans différentes bases de données extérieures (INSPEC et SPIRES)



Tâches et difficultés rencontrées

➤ **Catalogage :**

- Standardisation et homogénéisation des champs (extraction de métadonnées)
- Récupération des données perdues lors des différentes conversions du système.

➤ **Numérisation :**

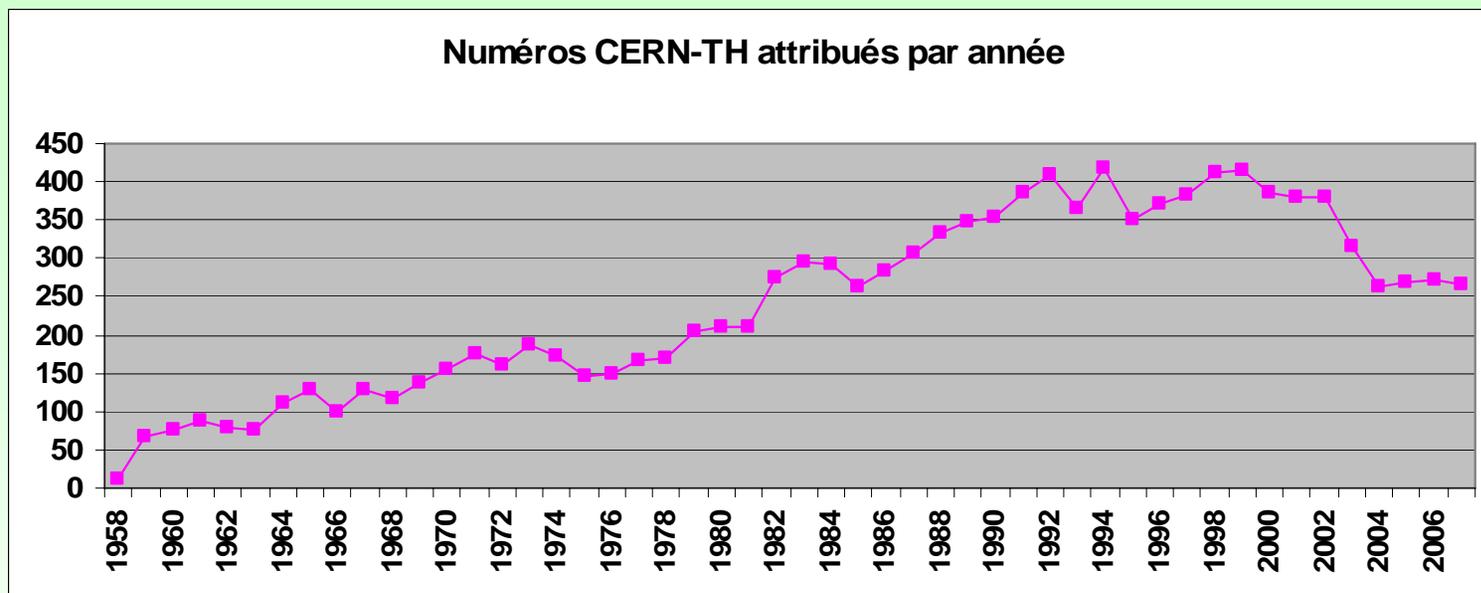
- Contrôle des liens hypertextes
- Détection des fichiers corrompus

➤ **Enquêtes auprès des auteurs pour 267 documents manquants :**

- Recherche d'email (Phone-Book du CERN, HEPNAME)
- Envoi de message en vue d'obtenir une référence de publication et un fichier ou une copie papier du document (167 emails envoyés)



Evaluation finale



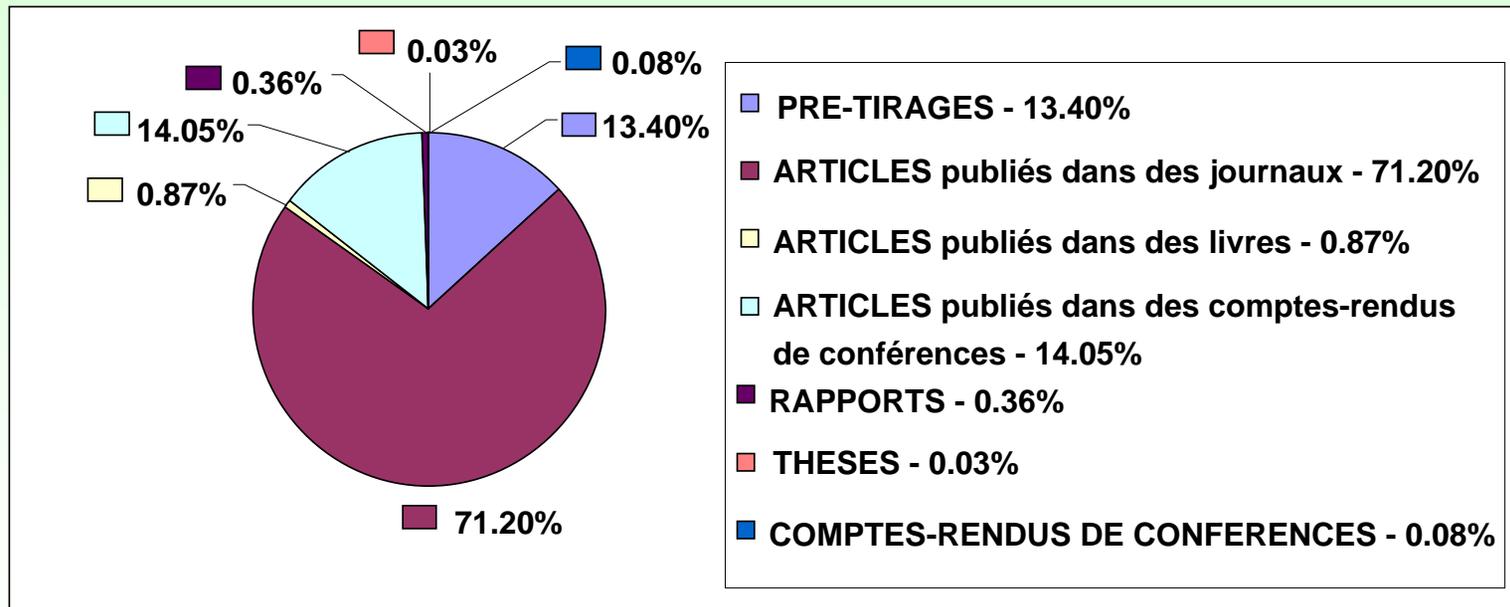
Production des pré-tirages par année

- **Fin de l'analyse : Collection de 11780 notices**
 - 9 457 initialement cataloguées dans la base
 - 928 nouvellement cataloguées
 - 1 395 identifiées grâce à une recherche par titre
 - 2 527 documents numérisés par le service du scanning du CERN

Evaluation finale

➤ Type de documents :

- 86% articles publiés dont 71,2% dans des journaux scientifiques et 14% dans des comptes-rendus de conférences
- 13% jamais publiés
- **restant** : rapports, des comptes-rendus de conférences et des thèses.



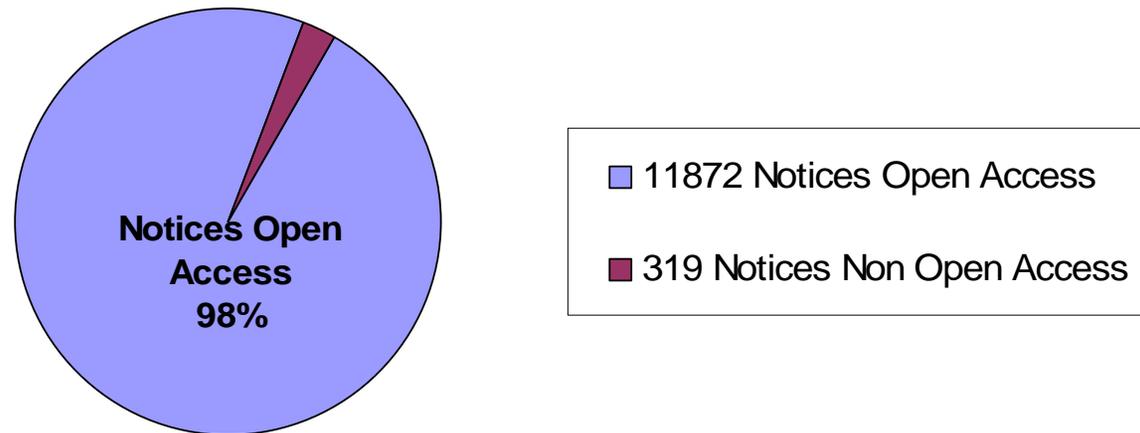
Série Théorie par type de document



Evaluation finale

- 98% de la collection disponible sur CDS en Open Access

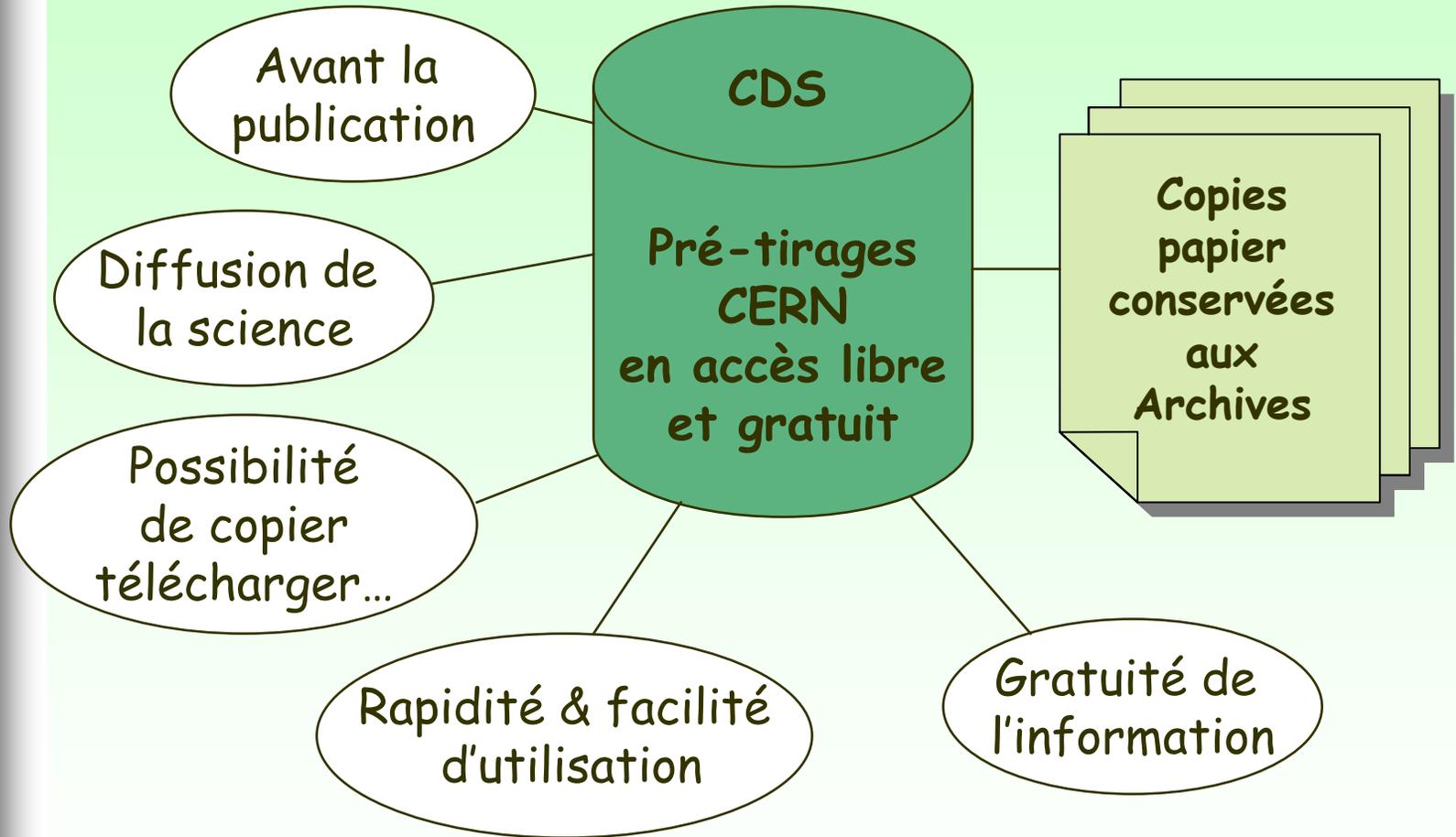
Série Théorie 12191 Notices



Collection CERN-TH-Open-Access



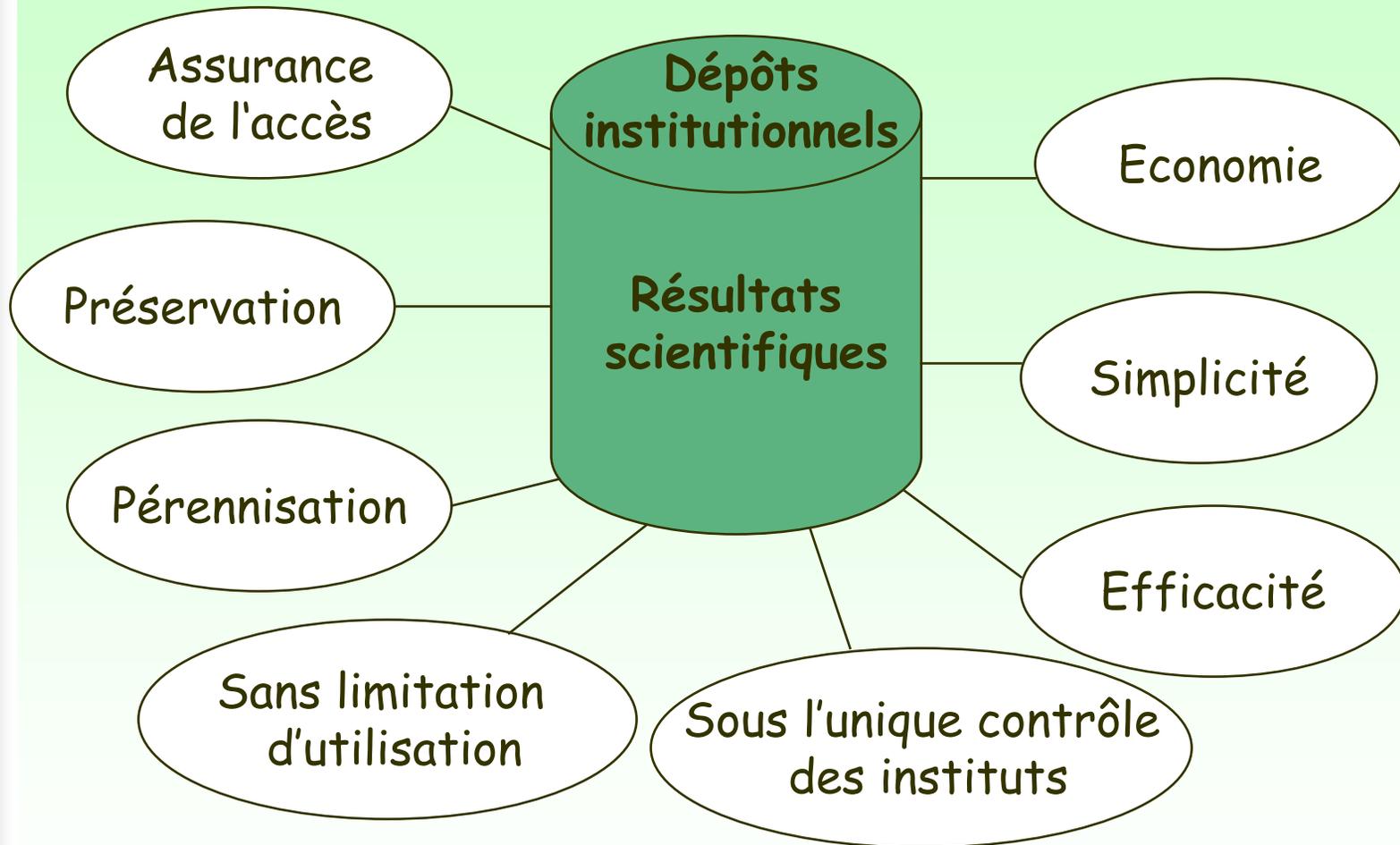
Conclusion



Mission du SIS et la Voie Verte OA



Conclusion



Généralisation de la Voie Verte OA