

# Les archives ouvertes

## L'archive HAL

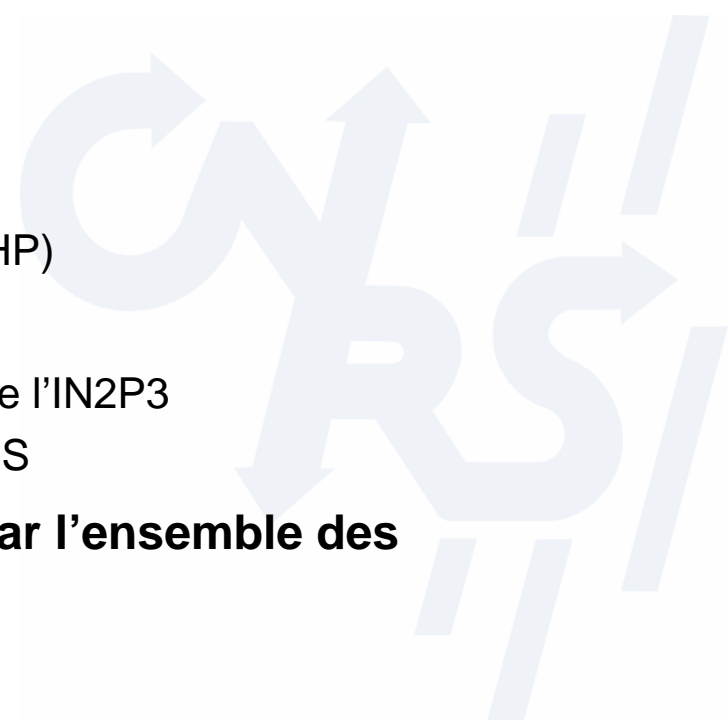


© ccscd - Cette présentation contient des animations ; elle doit être consultée en mode « diaporama », hors de ce mode, un certain nombre d'informations ne sera pas visible. Pour les mêmes raisons l'impression est déconseillée.

# L'archive ouverte HAL



- ❖ **Une initiative du CNRS**
- ❖ **Développée par le CCSD**
- ❖ **Partenariat avec l'INRIA**
- ❖ **Technologies « libres »**
  - ◆ LAMP (Linux, Apache, MySQL, PHP)
- ❖ **Environnement sécurisé**
  - ◆ Hébergé par le Centre de Calcul de l'IN2P3
  - ◆ Réplication des données au CINES
- ❖ **Un futur placé sous le pilotage par l'ensemble des établissements signataires**



# HAL, centralisation apparente, mutualisation certaine ! une plate forme commune pour,

- ❖ Garantir l'**accessibilité** au texte intégral
- ❖ Permettre l'**indexation** global du texte intégral contenu dans la base
- ❖ Préserver un **niveau scientifique** homogène et de qualité
- ❖ Offrir une **meilleure visibilité** à l'international
- ❖ Interconnecter avec les **bases mondiales** de référence
- ❖ Fournir des **URL pérennes** et ainsi permettre la citation numérique
- ❖ « **Horodater** » les dépôts et ainsi valider une antériorité scientifique
  - ◆ C'est le dépôt de la création auprès d'un tiers (on ne peut être témoin pour soi-même) qui va permettre de façon irréfutable de prouver cette antériorité, élément clé de la preuve.
- ❖ Enrichir automatiquement les **référentiels institutionnels**
- ❖ **Limiter la saisie** des références à un seul système
- ❖ Effectuer l'**archivage à long terme**
- ❖ *Et ... limiter l'effort à une seule unité spécialisée*

# HAL Quels documents ?

Normalisation en cours sur la liste du ministère

## ❖ Pré publications

### ◆ Typologie

- Publications avec comité de lecture (ACL)
    - internationales
    - nationales
  - Conférences invitées (INV)
  - Procédures de publications sans comité de lecture (SCL)
  - Publications avec actes (ACT)
    - internationales
    - nationales
  - Séminaires (SEM)
  - Communications sans actes (COM)
  - Chapitres (OS)
  - Livres (OV)
  - Directions d'ouvrages (DO)
  - Brevets (AP)
  - Thèses (AI)
  - Autres activités internationales (AI)
  - Information et culture scientifique et technique
  - Valorisation : contrats de recherche, partenariat industriel, créations d'entreprises
- Pour les brevets, certificats d'obtention végétale et logiciels.

## II.2 - Bilan quantitatif sur les quatre dernières années concernant :

II.2.1 Articles dans des revues avec comité de lecture (ACL)

- internationales
- nationales

II.2.2 Articles dans des revues sans comité de lecture (SCL)

II.2.3 Conférences invitées (INV)

II.2.4 Communications avec actes (ACT)

- internationales
- nationales

II.2.5 Communications sans actes (COM)

II.2.6 Ouvrages scientifiques (ou chapitres) (OS)

II.2.7 Ouvrages de vulgarisation (ou chapitres) (OV)

II.2.8 Directions d'ouvrages (DO)

II.2.9 Autres publications (AP)

II.2.10 Autres activités internationales (AI)

II.2.11 Information et culture scientifique et technique

II.2.12 Valorisation : contrats de recherche, partenariat industriel, créations d'entreprises

Pour les brevets, certificats d'obtention végétale et logiciels.

# Déposer dans HAL

## ❖ Modalités de dépôt

- ◆ Identification simplifiée
  - Compte auto validé (possibilité d'authentification institutionnelle)
  - Contributeurs : auteurs, documentalistes, etc.
- ◆ Niveau scientifique requis
  - « **Tout article qui, soumis à une revue à comité de lecture, serait envoyé à un rapporteur** »
  - Vérification avant mise en ligne
    - Vérification technique
    - Validation scientifique sommaire par domaine scientifique
- ◆ Format des documents libre
  - Format visualisable obligatoire (PDF, PS)
  - Fichiers sources recommandés

# Dépôt sans retrait ?

- ❖ **L'absence de retrait depuis l'interface assure la responsabilisation des dépôts, et donc le niveau scientifique de l'archive ouverte (ce n'est pas un blog...). C'est absolument essentiel pour l'existence même de l'archive.**
- ❖ **Cela évite des utilisations scientifiques frauduleuses :**
  - ◆ dépôt de résultats contradictoires pour prendre date avec retrait postérieur des dépôts faux, etc.
- ❖ **On ne peut pas citer dans une publication scientifique un texte qui risque à tout moment d'être retiré par son auteur ! La stabilité des dépôts et des adresses est donc cruciale**
- ❖ **L'alternative au dépôt sans retrait est la possibilité de déposer de nouvelles versions**
  - ◆ La dernière version est proposée par défaut au lecteur, mais celui-ci peut toujours accéder aux versions antérieures

# Déposer dans HAL, responsabilités

- ❖ La responsabilité éditoriale des éditeurs scientifiques et comités de lecture est remplacée par la responsabilité individuelle de l'auteur qui présente son travail à la communauté internationale.
- ❖ Comme pour une publication traditionnelle, un dépôt engage la responsabilité du laboratoire et des institutions dont il dépend.
- ❖ Le travail décrit doit être un travail achevé, bien rédigé, comparable à un manuscrit soumis à une revue scientifique. Une archive ouverte n'est pas un *blog* ou un cahier de laboratoire !

# Politique des éditeurs

❖ Les éditeurs, sont de plus en plus nombreux à permettre le dépôt en archives ouverte de la version auteur.

❖ Depuis mai 2004 : Elsevier

*« An author may post his version of the final paper on his personal web site and on his institution's web site (including his institutional repository). Each posting should include the article's citation and link to the journal homepage (or the article DOI). The author does not need our permission to do this, but any other posting (e.g to a repository elsewhere) would require our permission. By his version we are referring to his word or tex file, not a pdf or HTML download from science direct, but the author can update his version to reflect changes made during the refereeing and editing process. »*

K. Hunter senior vice president strategy Elsevier, may 2004



# Respect des éditeurs

## ❖ Certaines conditions doivent être respectées pour déposer sur une archive ouverte comme Hal :

- ◆ On déposera généralement le « fichier auteur » au sein du laboratoire, sauf deux exceptions :
  - il est légal de déposer le « fichier éditeur » de certaines revues en « open access »
  - on peut le faire également si l'éditeur (qui est très souvent commercial) a accepté de le déposer
- ◆ De même un « fichier auteur » ne doit pas être déposé si l'auteur a cédé les droits exclusifs (y compris de diffusion électronique) à un éditeur par un contrat explicite qu'il a signé.

N'oublions pas : les éditeurs commerciaux n'ont que les droits qui leur sont cédés par les auteurs ! À ceux-ci de ne signer que ce qui est strictement nécessaire!

### ➡ Droits sur les fichiers

Fichiers "version auteur" :

- Les fichiers que je dépose<sup>1</sup> sont des fichiers auteur

Fichiers "version éditeur" :

- Cet éditeur fait partie de ceux qui autorisent le dépôt des "fichiers éditeur" sur une archive ouverte<sup>1</sup>
- J'ai obtenu l'accord explicite de cet éditeur pour ce dépôt particulier
- Le ou les auteurs n'ont cédé par écrit aucun droit à cet éditeur

La page suivante vous permettra, le cas échéant, de régler la durée d'embargo que vous impose votre éditeur.

<sup>1</sup>- Pour voir la liste des journaux dont l'éditeur autorise le dépôt d'articles publiés ou à publier, voir le site SHERPA/RoMEO.

est recommandé d'ajouter dans le titre, et le lien DOI vers la

ter la période d'embargo de

Mettre en ligne le fichier de ce dépôt :

immédiatement	▼
immédiatement	
dans 15 jours	
dans 1 mois	
dans 3 mois	
dans 6 mois	
dans 1 an	
dans 2 ans	

# HAL, usages

## ❖ Scénario

- ◆ Le chercheur dépose sa *pré publication* dans l'archive
- ◆ Des retours rapides de sa communauté scientifique vont peut être initier de sa part le dépôt d'une nouvelle version
- ◆ C'est « presque » avec l'assentiment de sa communauté qu'il dépose sa « pré publication » chez un éditeur
- ◆ L'article est publié, le chercheur met à jour les références bibliographique dans la base, l'article devient une *publication*
- ◆ Le chercheur souhaite faire de nouvelles mise à jour de cet article ? Il en dépose une nouvelle version ! *post publication*
- ◆ Le cycle de vie d'un article va donc au-delà de sa publication dans une revue : e-publication ? *Les archives ouvertes bouleversent la publication : post publications et pré publications deviennent des notions plus floues, voir différentes !*

# HAL, le choix des méta données « à l'œil »

Conductivity and Fourier's law for a system of harmonic oscillators perturbed by a noise conserving energy and momentum.

Giada Basile,<sup>1</sup> Cédric Bernardin,<sup>2</sup> and Stefano Olla<sup>3</sup>

<sup>1</sup> *Università di Firenze, Viale Morgagni 67a, 50134 Firenze - Italy.*

<sup>2</sup> *UMPA, UMR-CNRS 5669, ENS-Lyon, 46, allée d'Italie, 69364 Lyon Cedex 07 - France.*

<sup>3</sup> *Ceremade, UMR-CNRS 7534, Université de Paris Dauphine, 75775 Paris Cedex 16 - France.*

(Dated: September 28, 2005)

We consider a system of harmonic oscillators perturbed by a stochastic dynamics conserving momentum and energy. We compute the effective thermal conductivity via Green-Kubo formula. In the limit as the size  $N$  of the system goes to infinity, conductivity remains finite if a pinning (on site potential) is present or in dimension  $d \geq 3$ . In the unpinned case conductivity diverges like  $N$  in dimension 1 and like  $\ln N$  in dimension 2. Then we consider the open system in contact with 2 heat bath at different temperature in the stationary state. We prove that the conductivity of the open system coincides with the Green-Kubo formula, and a corresponding Fourier's law in the cases of finite conductivity. Mathematical complete proofs of these results are in reference [1].

PACS numbers: 44.10+i, 05.80.k, 63.10.+a

Keywords: Thermal conductivity, Fourier's law, non-equilibrium systems, Green-Kubo formula, anomalous heat transport

When a small gradient of temperature  $\nabla T$  is applied to a material, we expect that, in the steady state, the heat current satisfies the Fourier's law

$$J = -\kappa \nabla T$$

where  $\kappa$  is the conductivity of the material. So if the system has length  $N$  and the left and right ends are in contact with heat baths at temperature  $T_L$  and  $T_R$ , respectively, the steady state current should be proportional to  $(T_L - T_R)/N$ .

The derivation of the macroscopic Fourier's law for thermal conduction from a microscopic deterministic Hamiltonian dynamics is one of the main problems in non-equilibrium statistical mechanics (cf. [2] and [3] for recent reviews on the subject). Even the existence of the transport coefficients (conductivity), defined by the Green-Kubo formula, cannot be proven, because of the (poorly understood) slow decay of the space-time correlations in deterministic systems. Furthermore in Hamiltonian mechanical system is not proven that the conductivity  $\kappa$  appearing in the Fourier's law is connected to the Green-Kubo definition through the space time integral of equilibrium correlation of the energy current [4].

In some low dimensional systems ( $d \leq 2$ ) anomalous thermal conductivity is observed numerically [5] and experimentally in nanotubes technology. The anomalous

(i.e. Fourier law is not satisfied) [6].

In non-linear cases, as in Fermi-Pasta-Ulam systems without pinning (on-site) potential, numerical evidence shows a conductivity diverging like  $N^\alpha$ , with  $\alpha < 1$  in one dimension, and like  $\log N$  in dimension 2 (cf. [7] and references in there). Finite conductivity is observed numerically in all pinned cases, as long as some non-linearity is present in the interaction.

We study here a system of harmonic (linear) coupled oscillators where the hamiltonian dynamics is perturbed by a random exchange of momentum between nearest neighbors atoms. Each atom can be pinned or not by an harmonic on site potential. The random exchange of momentum conserves the total momentum and the total energy. In dimension 1 we consider a random exchange of momentum between nearest and second nearest neighbors atoms that has the same conservation laws. Because of the conservation laws, this noise introduces a certain non-linearity in the model. In some sense this stochastic perturbation of the dynamics simulates (qualitatively) the long time (chaotic) effect of the deterministic non-linear model.

We compute explicitly and rigorously the conductivity  $\kappa_N$  of the finite closed system of linear size  $N$ , by using Green-Kubo formula for the dynamics in the microcanonical ensemble. We find that it has a finite explicit limit

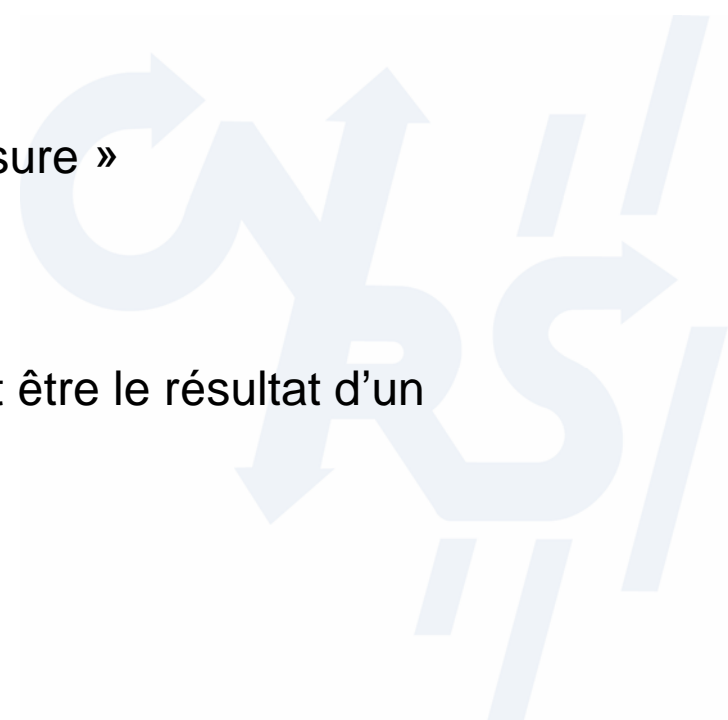
# Cohérence des méta données

## ❖ Chaque fois que cela est possible les méta données

- ◆ Seront issues de listes d'autorité pouvant provenir
  - Des référentiels institutionnels
    - Listes de laboratoires, d'universités, base LDAP, etc.
  - De listes normalisées
    - Listes des pays, des langues
  - De listes constituées au « fur et à mesure »
    - Listes de journaux

## ❖ S'il existe un référentiel

- ◆ L'alimentation de la liste d'autorité devrait être le résultat d'un automatisme
  - Mise à jour tous les ...





Accueil Déposer

..: Consulter les laboratoires con

Sigle  
(information transférée à arXiv)

Intitulé

Institution(s)

Université(s)

Grande-Ecole(s)

URL du laboratoire

Adresse postale

Pays du laboratoire

**Identifiant interne du laboratoire** 1804

**Sigle (info à ar)**

**Inti**

- Physique [physics]
- Mathématiques [math]
- Informatique [cs]
- Science non linéaire [nlin]
- Chimie
- Sciences de l'Homme et Société

**Ins**

- Anthropologie sociale et ethnologie
- Archeologie
- Anthropologie biologique

**Uni**

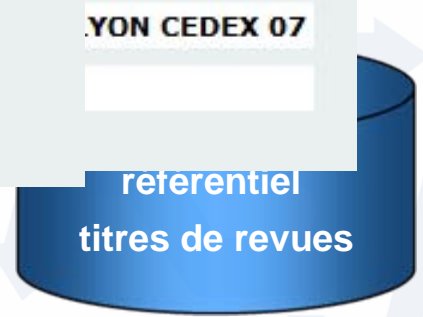
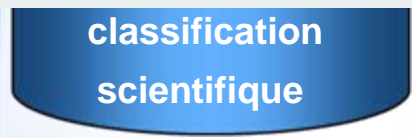
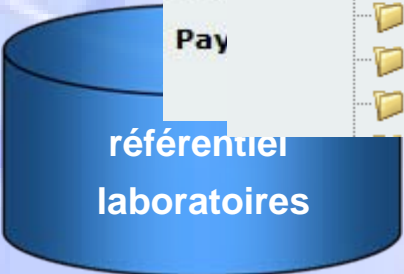
- Art et histoire de l'art
- Demographie
- CRB

**Gra**

- Economie et finances *RePEc*
- Education *Temalice*
- Etudes classiques
- Etudes de l'environnement

La classification scientifique est faite par les chercheurs et pour les chercheurs.

Le but recherché est d'avoir un maximum de 2 niveaux arborescents contenant une trentaine d'item au maximum pour une discipline.



Les principaux référentiels de HAL

# Stabilité

## ❖ Stabilité des identifiants

- ◆ La référence d'un article en base d'archives ouvertes doit pouvoir être consigné dans le texte.
- ◆ Il doit être possible de citer un article en donnant son adresse dans une archive ouverte

## ❖ Des URL courtes et pratiques sont fournies

- ◆ <http://hal.ccsd.cnrs.fr/LAAS>
- ◆ <http://hal.ccsd.cnrs.fr/autlab/nomauteur/nomlabo/>

## ❖ Il importe que l'URL de l'archive donc des documents contenus soit préservée coûte que coûte !

STATISTICS OF I

LAURIE SHAW<sup>1</sup>

<sup>1</sup>Institute of Astronor

<sup>2</sup>NASA/Fermilab Astrophysic

<sup>3</sup>Department of Physics and

<sup>4</sup>Princetn

We have identified over 20 from the output of 1024<sup>3</sup> par ing length  $3.2h^{-1}\text{kpc}$ ). We p into dynamical equilibrium,  $\xi$  all the halos in our sample. T axis. We find there to be nc that the higher mass halos ar the orbital angular momentu: momentum of the 'host' halo halos to their host. Overall, by the fraction of mass cont on the ratio of halo mass ( $M$  self-similarity is in the expec substructure than 'young hal  
*Subject headings:* cosmology: tions

### 1. INTRODUCTION

N-body dark matter simulations ha of our most valuable tools for testing the standard model of hierarchical s Indeed, it is largely through studying form in such simulations that we are proposed models that describe the r bution of halos and the subhalos wi as their physical characteristics such concentration.

Until the end of the last decade, ble reach the required mass resolu identification of galaxy mass halos

arXiv:astro-ph/0509856 v1 28 Sep 2005

# Identifiant et citabilité électronique

## ❖ Un « document » dans l'archive ouverte HAL :

- ◆ une ou plusieurs versions d'un même article comportant chacune
  - la notice descriptive
    - Incluant éventuellement des références bibliographiques
  - un ou plusieurs formats du même fichier (PDF, TeX, XML, DOC, etc.)
  - éventuellement des fichiers annexes
    - images, sons, vidéos, posters, présentation « ppt », etc.

## ❖ L'identifiant pérenne, « citable », référence un « container » qui abrite tous les composants de l'article scientifique

- ◆ une URL simplifiée qui garantit au lecteur d'arriver toujours sur la version la plus récente de l'article et lui permet l'accès à l'historique
- ◆ une URL stable ne dépendant pas du stockage physique du document

# Identifiants HAL

- ❖ **Identifiant HAL**      **hal-00021931**
- ❖ **URL pérenne**      <http://hal.archives-ouvertes.fr/hal-00021931>
- ❖ **Identifiant(s) OAI**
  - ◆ oai:hal.archives-ouvertes.fr:hal-00021931\_v1
  - ◆ oai:hal.archives-ouvertes.fr:hal-00021931\_v2
  - ◆ oai:hal.archives-ouvertes.fr:hal-00021931\_v3
  - etc.
- ❖ **Par exemple pour oai:...21931\_v1 l'URL « physique » est**
  - ◆ <http://hal.archives-ouvertes.fr/docs/00/06/41/63/PDF/photongravit.pdf>
  - ◆ cette URL n'offre pas de garantie absolue de pérennité
  - ◆ elle est fournie lors d'une requête OAI-PMH avec le metadataFormat AO.FR à des fins de téléchargement « immédiat » pour stockage dans le SI local



# Interopérabilité, interconnexions, imports/exports

## ❖ OAI-PMH

- ◆ multi formats dont OAI\_DC (Dublin core non qualifié)

## ❖ REDIF

- ◆ pour le moissonnage par RePeC (économistes)

## ❖ RSS

## ❖ Connexions

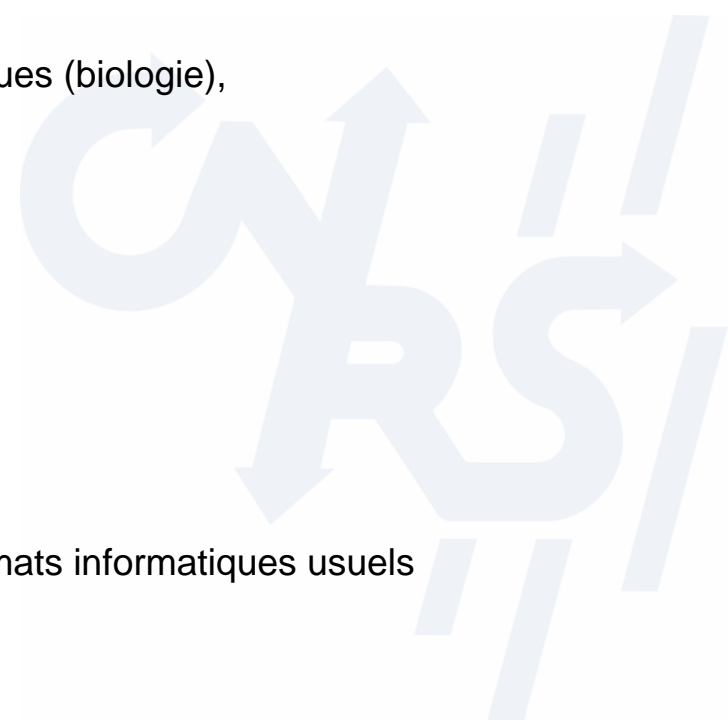
- ◆ **arXiv** dans les domaines de physique et mathématiques (biologie), uniquement de HAL **vers** arXiv
- ◆ **Pubmed** (HAL-INSERM)
  - Facilité d'import de la notice **depuis** Pubmed
  - Transfert **vers** Pubmed Central

## ❖ Imports

- ◆ Fichiers XML, méta données et texte intégral
- ◆ Web services

## ❖ Exports

- ◆ Listes selon critères de recherches dans tous les formats informatiques usuels
- ◆ « Home page » chercheur
- ◆ Exports institutionnels (Crac, Labintel, Graal, etc.)
- ◆ Web services

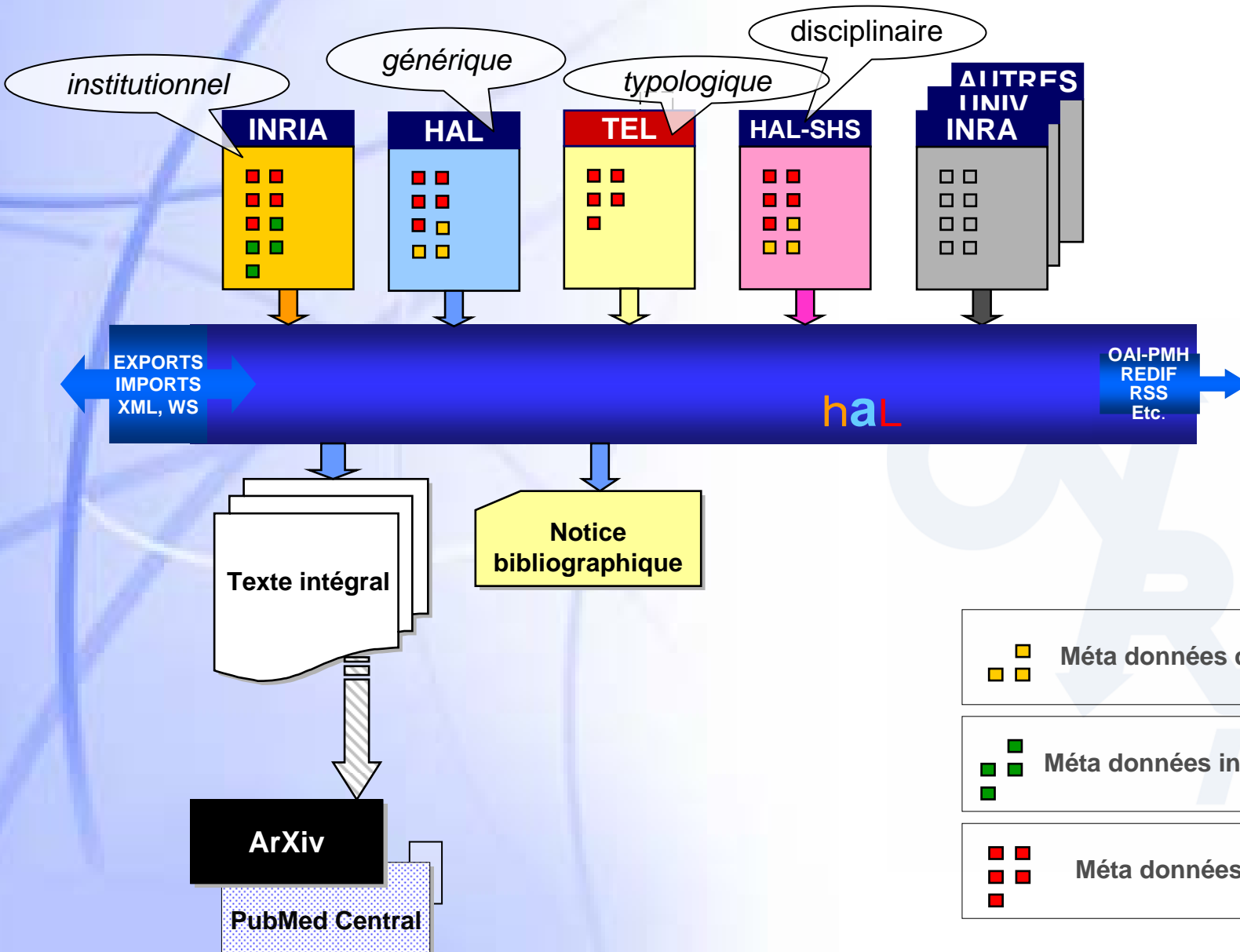


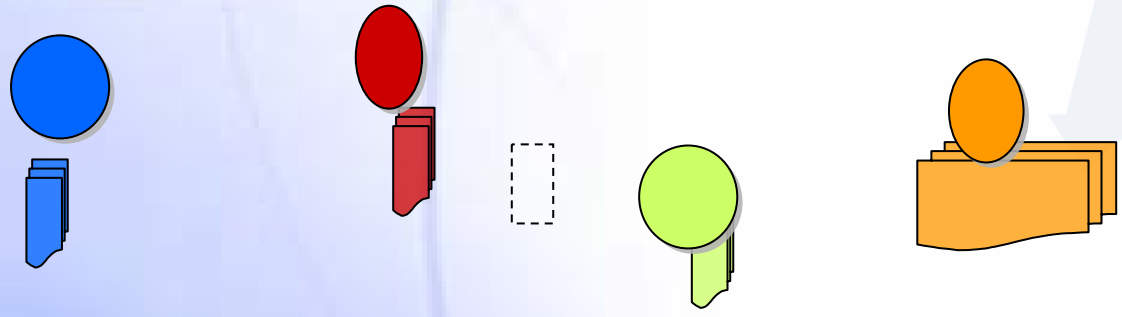
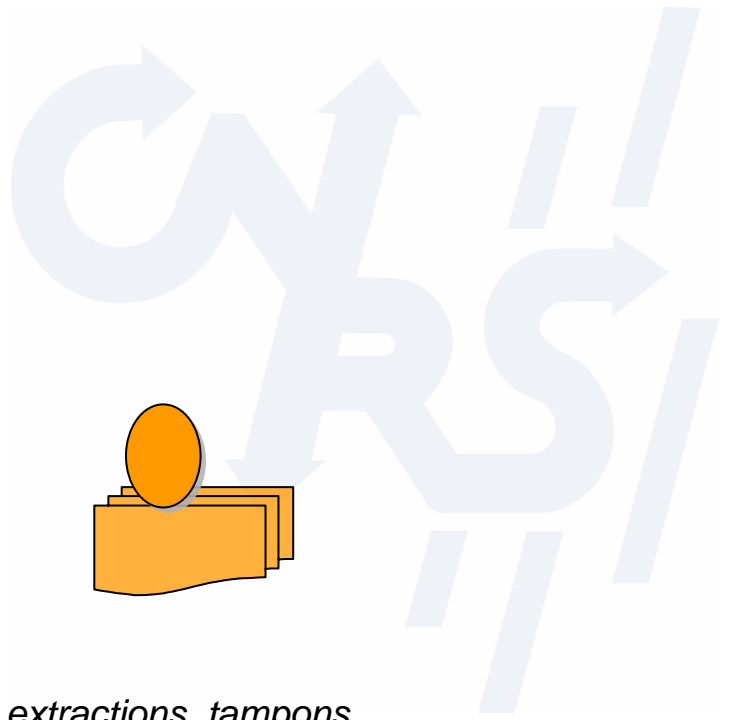
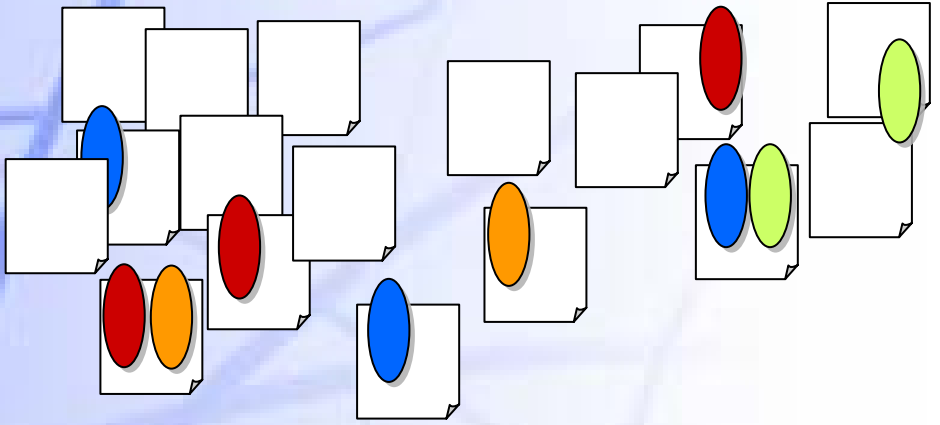
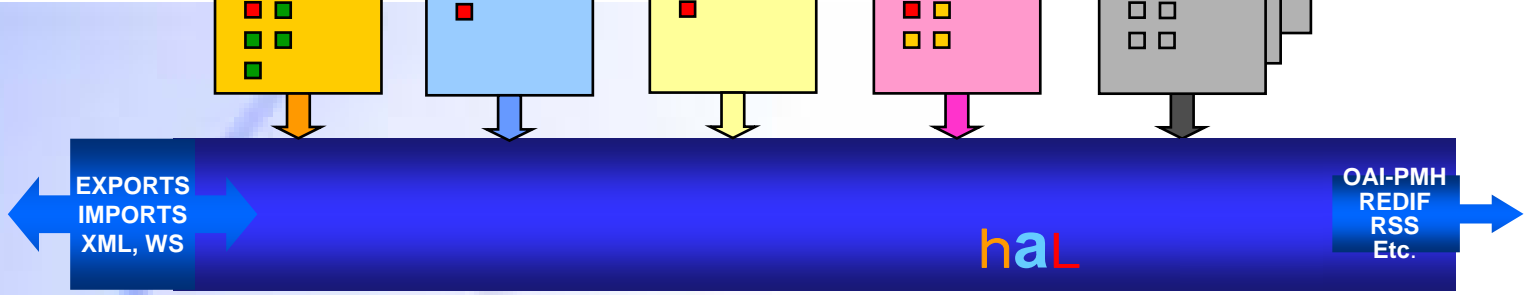
## Structure de HAL

Quels services peut-on attendre de l'archive ?



Portails de dépôts





*Collections personnalisables, extractions, tampons*

# Portails et collections

## ❖ Une base de données commune, HAL

- ◆ des portails thématiques ou typologiques
  - thèses (TEL), sciences humaines et sociales, sciences de l'environnement, etc.
- ◆ des portails institutionnels personnels
  - préservation de l'identité
  - règles spécifiques
- ◆ des collections pour
  - des épi journaux
  - des actes de congrès
  - des pages personnelles
  - des thèmes scientifiques
  - etc.

HAL - Hyper Article en Ligne

Le serveur HAL permet de déposer et de rendre publics des **documents scientifiques de toutes les disciplines**. Le dépôt doit être fait, soit par un des auteurs du document ayant l'accord des autres auteurs, soit par une personne autorisée (documentaliste du laboratoire par exemple). Avant de déposer dans HAL, il est nécessaire de se créer un compte de déposant.

HAL est un **outil de communication scientifique directe entre chercheurs**. Un texte déposé sur HAL doit décrire un travail achevé de recherche, conforme aux usages scientifiques dans la discipline ; le contenu doit être comparable aux manuscrits que les chercheurs soumettent pour publication aux comités de lecture de revues scientifiques, d'actes de colloque, etc. Les documents déposés sur HAL ne sont soumis à aucune évaluation scientifique détaillée. Seul un examen rapide permet de s'assurer que le document proposé entre bien dans la catégorie définie ci-dessus – le CCSD restant libre de mettre en ligne ou pas le document sans justification.

HAL ne se limite pas aux documents publiés, ou même destinés à être publiés, mais le contenu scientifique doit être approprié pour un affichage dans HAL. En cas de publication, au moment de la parution les contributeurs sont invités à enrichir le dépôt des références de publication et du DOI.

Les **documents déposés dans HAL** bénéficient de la pérennité de cette base, y **restent de façon permanente** et reçoivent une **adresse web stable**; ainsi, comme une publication dans un journal scientifique habituel, ils peuvent faire l'objet d'une citation dans un autre travail.

La mise à disposition gratuite en ligne de ces documents est destinée à la meilleure diffusion des travaux de recherche ; leur propriété intellectuelle reste entièrement celle du ou des auteurs. Les utilisateurs sont soumis aux règles du bon usage habituelles des publications scientifiques : respect des travaux originaux, citation des travaux antérieurs, pas de pillage intellectuel, etc. L'ensemble des règles de bon usage de HAL, en particulier celles relatives au respect des lois sur le droit d'auteur, sont résumées dans le « **guide du dépôt et du bon usage de HAL** ». Pour plus de détails, en particulier sur les types de dépôts (documents récents ou anciens), métadonnées, usage des collections et des tampons, les rôles spécifiques disponibles dans HAL, etc. voir le « **Manihal** ». Tout dépôt dans HAL, dans une discipline scientifique existante chez arxiv, est automatiquement dupliqué dans cette base (sauf si le déposant spécifie explicitement le contraire).

Enfin, HAL peut également être utilisé pour fournir des services de nature plus administrative, comme la constitution de listes de publication (par auteur, laboratoire, institution, etc.) dans des formats multiples. Pour que ces listes puissent être complétées dans le cas où le fichier du plein texte n'est pas disponible, il est possible de déposer sur HAL des « **notices bibliographiques** » ne contenant que des références de publication.

Transferring data from hal.archives-ouvertes.fr...

# HAL une archive ouverte « ouverte »

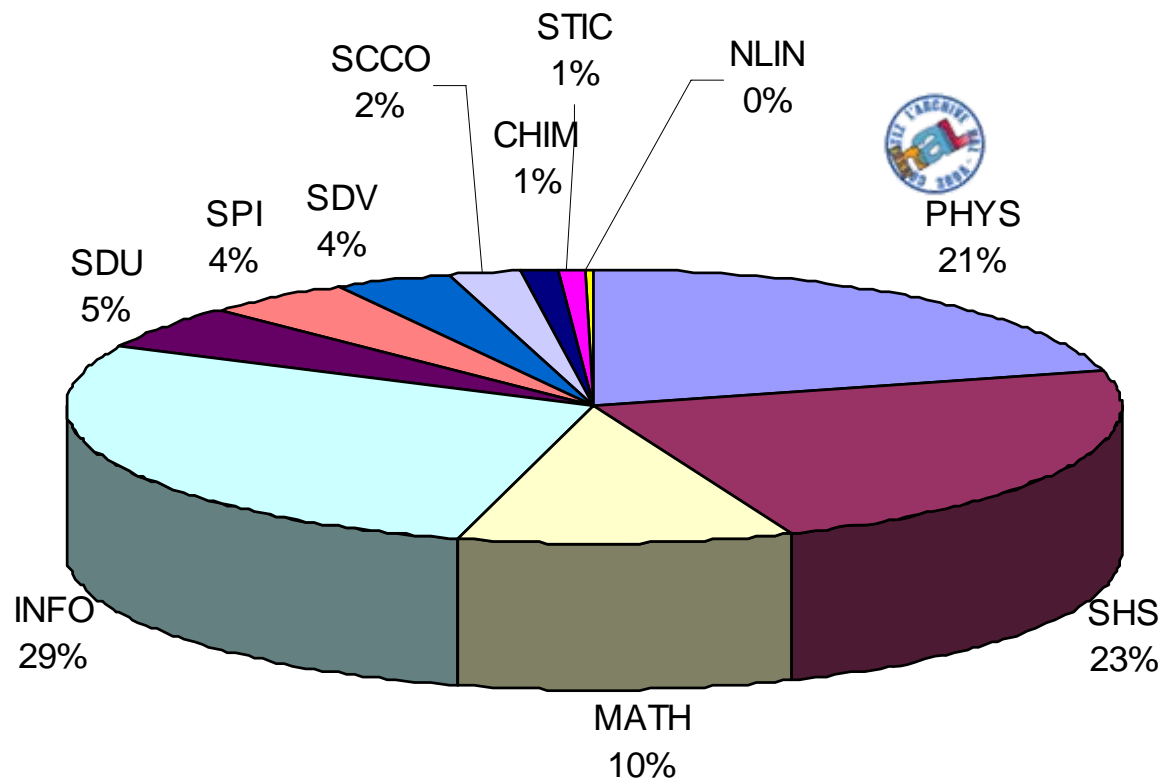
- ❖ **Accessible depuis des applications diverses via des services Web**
  - ◆ logiciels documentaires commerciaux
  - ◆ services de gestion de congrès
  - ◆ environnements numériques de travail
  - ◆ moteurs de recherches fédérés
  - ◆ GED
  - ◆ Etc.
- ❖ **En adhérant au schéma de données de HAL, il est possible**
  - ◆ de connecter son propre système documentaire
  - ◆ de créer sa propre interface de dépôt



# Statistiques de dépôt

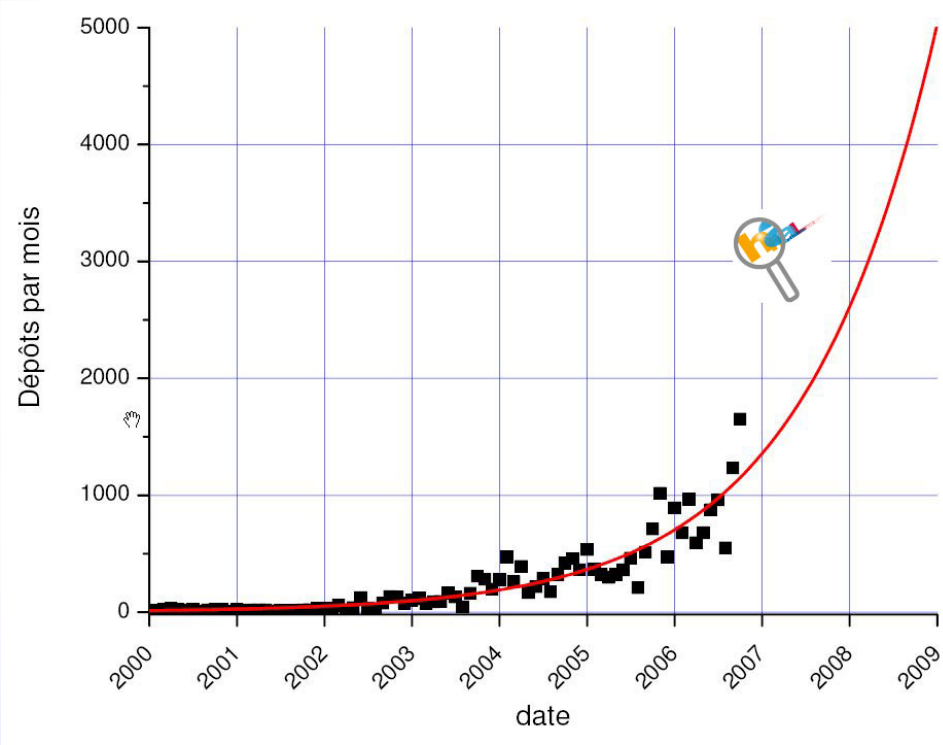
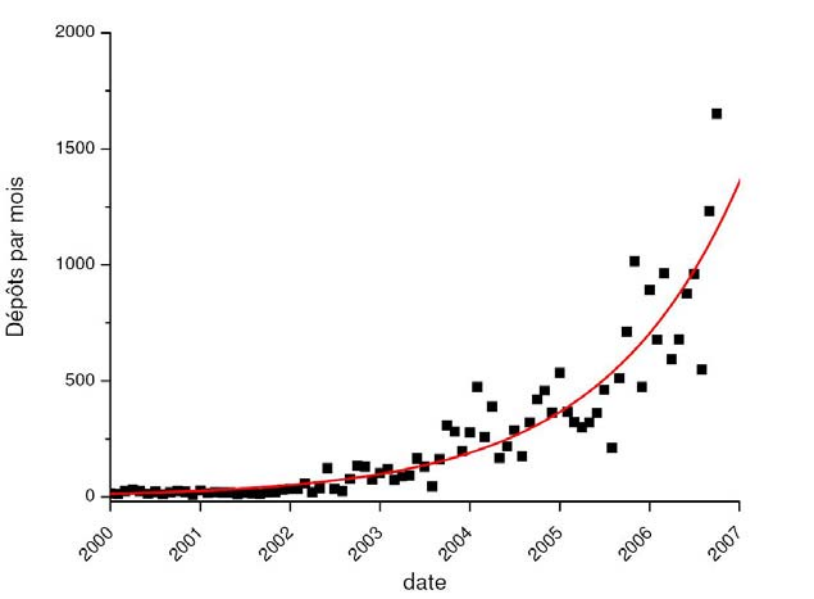


# Répartition par domaines scientifiques des dépôts avec texte intégral



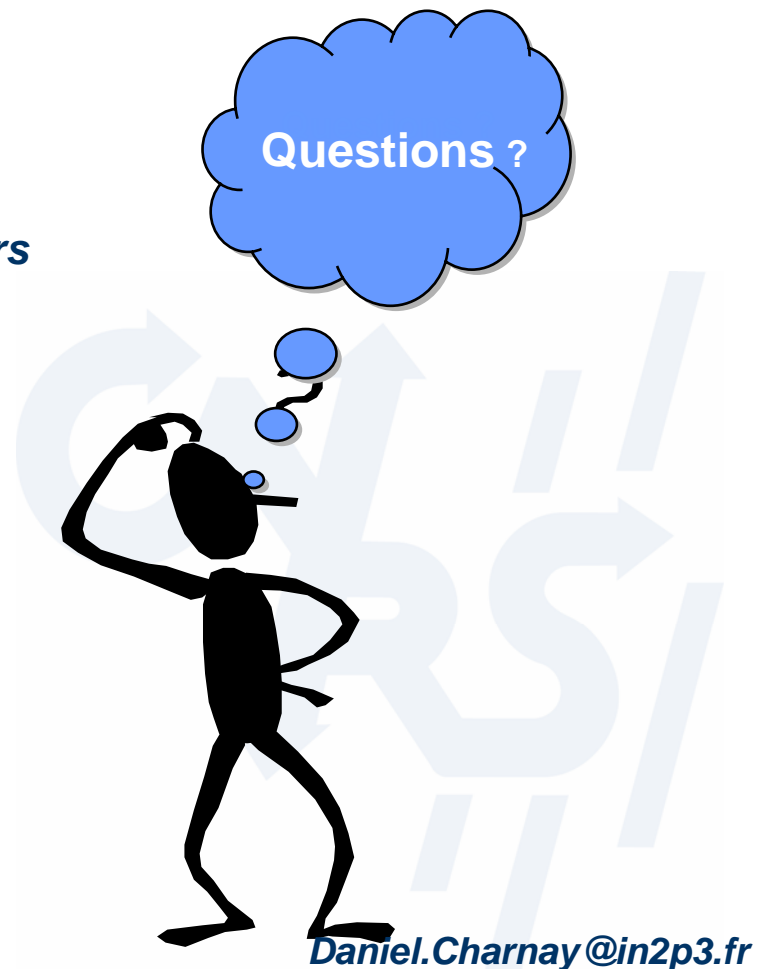


# Evolution des dépôts



# Adresses et URL

- ❖ <http://hal.archives-ouvertes.fr>
- ❖ <http://tel.archives-ouvertes.fr>
- ❖ <http://ccsd.cnrs.fr>
  - ◆ *documentations utilisateurs et développeurs*
  - ◆ *tutoriels (vidéos)*
  - ◆ *manuels*
  - ◆ *accès aux différents portails*
  - ◆ *news, etc.*
- ❖ **Support utilisateurs**
  - ◆ *mailto : [hal.support@ccsd.cnrs.fr](mailto:hal.support@ccsd.cnrs.fr)*
  - ◆ *<http://support.ccsd.cnrs.fr>*
- ❖ **Les sites archives-ouvertes.fr**
  - ◆ *<http://www.archives-ouvertes.fr>*
  - ◆ *<http://openaccess.inist.fr>*



*Daniel.Charnay@in2p3.fr*



Interconnexion entre HAL,  
plate-forme partagée d'archives ouvertes  
et les systèmes institutionnels locaux



# Les systèmes d'informations locaux

## ❖ Divers systèmes

- ◆ Base de données bibliographique des publications du laboratoire
  - Sans texte intégral (notice). Le plus fréquent
  - Avec le texte intégral (archive institutionnelle)

## ❖ Leurs raisons d'être

- ◆ Historique : le système est antérieur à l'arrivée des archives ouvertes
- ◆ Typologie de documents plus large que celle de l'archive ouverte HAL (rapports, littérature grise, contrats, etc.)
- ◆ Méta données plus riches
- ◆ Workflow de validation des dépôts
- ◆ Etc.

# Modèles de dépôt dans HAL, les acteurs

❖ La plate-forme commune archives-ouvertes.fr : HAL



❖ Le système d'information local : SI



❖ Le dépôt : fichier et méta données



❖ Le contributeur



# Interconnexion entre HAL et SI : les choix

- A.** HAL propose un portail institutionnel, ou le portail générique, le contributeur dépose directement dans un de ces portails
- B.** L'institution dispose de son propre système d'archive (SI)
  - 1. Le contributeur dépose directement dans le SI
    - Le SI se connecte à HAL et effectue le dépôt
  - 2. Le contributeur dépose dans HAL
    - Le SI vient moissonner HAL




# Alimentation directe - alimentation assistée

## ❖ Le contributeur alimente directement HAL

- ◆ en se connectant sur le site Web
- ◆ en remplissant les différents formulaires de dépôt

## ❖ Mais ...

- ◆ si les méta données (réf. biblio par exemple) pré existent
  - dans un système local reconnu
  - dans une autre base (Pubmed, ADS, ArXiv...)
- ◆ si l'identificateur externe peut être fourni
- ◆ **la saisie peut être assistée**  HAL collecte alors automatiquement l'information sur le système distant et pré-rempli le maximum de champs en fonction des données qu'il a pu recueillir
- ◆ le contributeur complète les données manquantes et joint le texte intégral de la publication

# Alimentation assistée : contraintes et solutions

- ❖ **HAL devrait développer autant de connecteurs que de bases pouvant fournir de l'assistance ?**
  - ◆ un connecteur pour Pumed, un connecteur pour ADS, un connecteur pour l'IRD, pour l'université X, l'établissement Y, etc. ?
- ❖ **Un schéma de méta données, désigné sous le nom de *AO.fr*, est défini pour permettre les échanges de données (SI<->HAL)**
  - ◆ la déclaration d'un partenariat entre un SI et HAL permet d'associer un identifiant avec les paramètres d'accès vers la base partenaire
  - ◆ IRD002732, USE0067, permettront à HAL de savoir qu'il faut récupérer les méta données normalisées du document 002732 sur la base de l'IRD ou le document 0067 à l'université de Saint Etienne
  - ◆ HAL stockera l'identifiant du dépôt d'une base partenaire, cet identifiant servira de lien interne entre les deux bases (il a essentiellement une vocation technique).



- ❖ **AO.FR** est un des metadataFormat OAI-PMH que « devrait /doit » exposer chacune des archives participant à **archives-ouvertes.fr**
  - ◆ il fournit le **tronc commun** des méta données tel que défini par le COST/COSTRAO
  - ◆ il définit l'ensemble des clés, vocabulaire XML, qui qualifient chacune des méta données
  - ◆ il est utilisé (cette liste n'est pas exhaustive)
    - par HAL pour l'assistance à la saisie
    - par le SI pour moissonner HAL
  - ◆ il fournit de façon dynamique (non pérenne) l'URL « physique » du fichier à des fins de téléchargement

# Dépôt via un SI équipé d'un connecteur HAL : 1 - principe

## ❖ Le contributeur alimente le système d'information local

- ◆ Les informations saisies correspondent
  - exactement au cœur des méta données de HAL ([AO.fr](http://AO.fr))
  - à un sur ensemble de ce tronc commun
- ◆ Le système d'information local
  - transmet le dépôt vers HAL
    - de façon synchrone ou asynchrone (traitement par lots)
    - il inclut l'identifiant local de l'article et l'email du contributeur
  - Conserve localement
    - la totalité de l'information recueillie (notices et texte intégral)
    - uniquement la notice
  - Il reçoit (sauf retour d'erreur)
    - de façon synchrone : l'identifiant HAL
    - de façon asynchrone : le statut du dépôt
      - ▶ en ligne - en attente de modification de l'auteur - refusé
      - ▶ dans les deux derniers cas les raisons sont indiquées dans le message de retour

# Dépôt via un SI équipé d'un connecteur HAL :

## 2 - référentiels

### ❖ Les référentiels utilisés par HAL

1. laboratoires et institutions
  - référentiel interne à HAL
  - inclut EPST, Université, Grandes écoles (multi évalué)
  - évoluera vers un référentiel national (INSE, Ministère, ?)
2. Nomenclature des domaines scientifiques
  - référentiel « maison », aligné sur ArXiv pour les domaines concernés
  - classification orientée « chercheur »
3. Typologie des documents
  - aligné sur la liste du ministère
4. Liste des revues
  - ~25000 titres
  - structuré
  - inclut la classification Sherpa

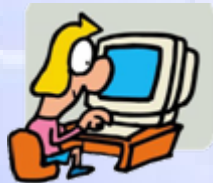
# Dépôt via un SI équipé d'un connecteur HAL :

## 3 - échanges

### ❖ Les Web-services (SOAP) sont le protocole informatique utilisé

1. le SI récupère les référentiels de HAL ; il peut les garder dans un cache réactualisé régulièrement
2. le SI collecte les méta données du tronc commun et qualifie celles-ci par rapport à ces référentiels
  - En proposant directement les référentiels de HAL
  - En utilisant ses propres référentiels et en effectuant une équivalence avec les référentiels de HAL
3. le SI effectue le dépôt sur HAL
4. le SI met à jour de façon synchrone dans sa base l'identificateur HAL
5. le SI reçoit plus tard l'information de mise en ligne, de refus ou de demande de modification
  - il appartient au SI de gérer les deux dernières situations

# Dépôt direct



1 - Connexion à HAL



3- Ajout du texte intégral

2- Saisie des méta données



# Dépôt direct



## 4 - Dépôt dans HAL



Retour

# Dépôt direct assisté

3- HAL demande les méta données  
au SI partenaire en fournissant l'identifiant



2- fourniture de l'identifiant externe



1 - connexion à HAL



# Dépôt direct assisté



4-via OAI-PMH Hal récupère les données (notice)

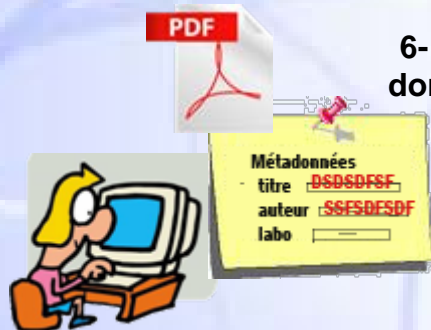


5- HAL utilise ces données pour pré remplir le formulaire





# Dépôt direct assisté



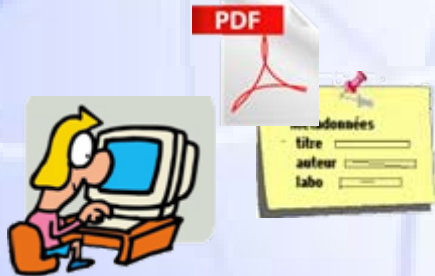
6- finalisation des méta  
données et ajout du texte  
intégral



# Dépôt direct assisté



Les méta données sont à jour dans les deux systèmes



7 - dépôt dans HAL



Retour

# Dépôt via le SI

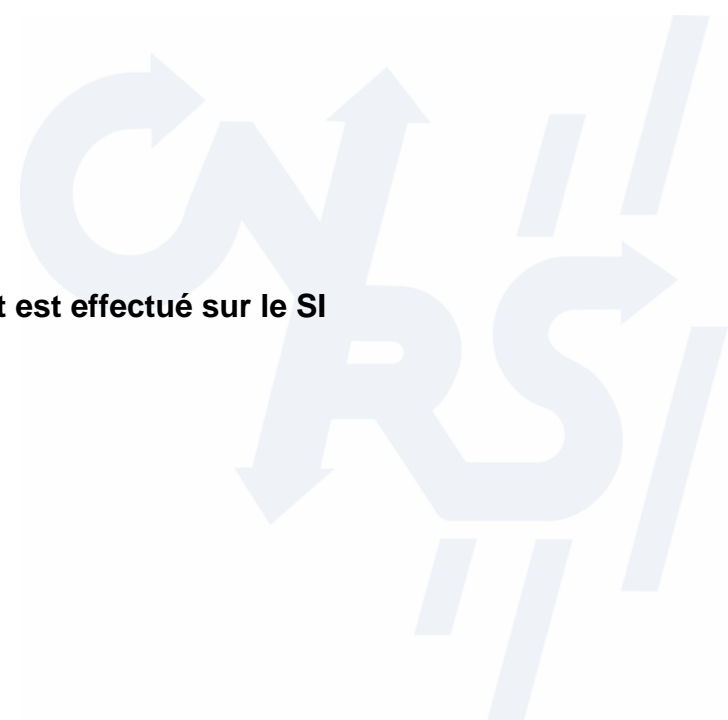
Le SI tient à jour une copie des référentiels HAL



Le dépôt est effectué sur le SI



Les méta données saisies sont calibrées sur les référentiels



# Dépôt via le SI

Plus tard



Le SI transfère le dépôt sur HAL

Le SI peut ne conserver  
que la notice



HAL envoie  
l'information de mise en  
ligne ou de refus



Retour

# Post alimentation du SI

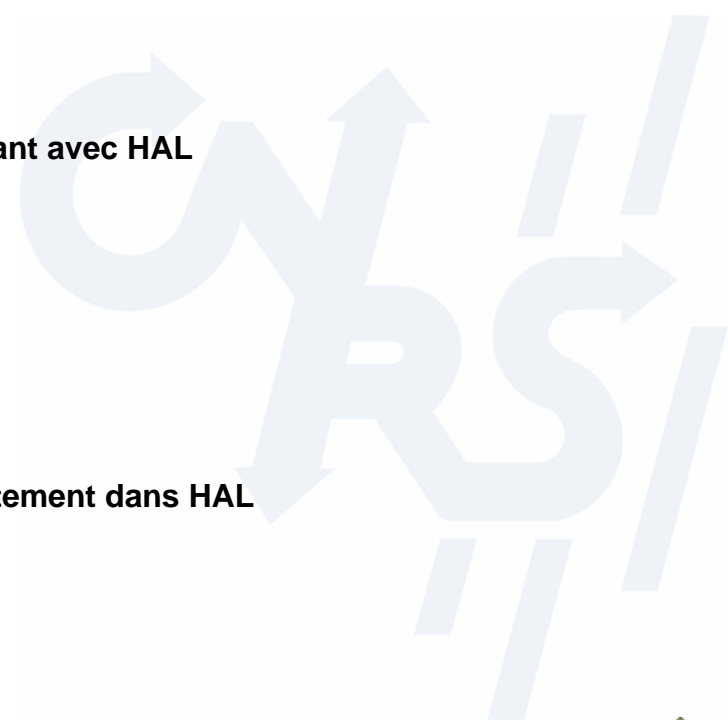
2- Le SI moissonne HAL



3- Le SI synchronise son identifiant avec HAL



1- Le dépôts est fait en standard, directement dans HAL



Retour