

REUNION des CorIST	<b>Le : 21 février 2019</b> <b>Rédactrice : Isabelle DEBANO</b>
<b>Ordre du jour :</b> <ol style="list-style-type: none"> <li>1) HAL - Retours et échanges sur les exercices pratiques de dépôt de publications</li> <li>2) Plateformes d'édition</li> <li>3) Préparation réunion CORIST du 21 mars autour des données de la recherche</li> <li>4) Questions diverses</li> </ol>	

	Entités	Présents
<b>Participants CORIST</b>		
Patrice BELLOT	CNRS-INS2I	✓
Jean-Jacques BESSOULE	CNRS-INSB	✓
Joël CUGUEN	CNRS-INEE	Excusé
Stéphanie CASTEX	CNRS-INC	✓
Frédéric HELEIN	CNRS-INSMI	✓
Martina KNOOP	CNRS-INP	✓
Laurent LE GUILLOU	CNRS-IN2P3	Excusé
Lionel MAUREL	CNRS-INSHS	✓
Benoit PIER	CNRS-INSIS	✓
Pascale TALOUR	CNRS-INSU	Excusée
<b>Participants DIST</b>		
Francis ANDRE	CNRS-DIST	✓
Serge BAUIN	CNRS-DIST	✓
Christine BERTHAUD	CNRS-CCSD - Directrice	✓
Annie CORET	CNRS-DIST	✓
Isabelle DEBANO	CNRS-DIST	✓
Laurence EL KHOURI	CNRS-DIST	✓
Nathalie FARGIER	PERSEE - Directrice	Excusée
Claire FRANCOIS	CNRS-INIST - Directrice	✓
Joanna JANIK	CNRS-DIST	✓
Paolo LAI	CNRS-INIST	✓
Maxime RAGOT	CNRS-DIST	✓
Sylvie ROUSSET	CNRS-DIST	✓
Didier TORNAY	CNRS-INSHS/DIST	✓
Christine WEIL-MIKO	CNRS-INIST/DIST	Excusée
<b>Invités permanents</b>		
Volker BECKMANN	CNRS-IN2P3	✓
Odile CONTAT	CNRS-INSHS	Excusée
Michèle DASSA	CNRS-INSHS	Excusée
Stefan HOHENEGGER	CNRS-IN2P3	Excusé
<b>Invités</b>		
Bénédicte Kuntziger	CNRS-CCSD	✓

## Validation du CR 17 janvier 2019

Le Compte rendu de la réunion du 17 janvier 2019 a été validé par l'ensemble des participants.

### 1) HAL - Retours et échanges sur les exercices pratiques de dépôt de publications

Dans le cadre de la campagne à venir pour le dépôt obligatoire dans HAL<sup>1</sup>, il avait été demandé aux CORIST, lors de la réunion de janvier, de mener des exercices pratiques de dépôt d'articles dans HAL et de faire remonter les questions et les remarques.

Afin d'échanger et répondre aux questions, Bénédicte Kuntzinger – Responsable assistance utilisateurs CCSD, avait été invitée à participer.

Outre ce moment de questions/réponses autour de HAL, l'autre objectif était de parvenir à identifier des recommandations, suggestions qui pourraient alimenter la lettre qui sera envoyée par Alain Schuhl aux chercheurs afin de les aider à déposer dans HAL. Cette lettre annoncera les modifications dans CRAC et RIBAC afin que seules les publications déposées dans HAL soient référencées dans les comptes rendus des chercheurs.

### Comment accompagner et inciter le chercheur à déposer dans HAL ?

#### Quelques recommandations issues des échanges entre les CORIST et le CCSD

- L'export de HAL vers CRAC et RIBAC est déjà possible
- La première mise en route prend un peu de temps : **le chercheur doit créer son profil (l'affilier à l'orcid, créer l'id-HAL)** puis il doit faire le ménage dans ses formes auteur répertoriées dans HAL en rajouter si nécessaire puis trier les documents qui sont déjà déposés dans l'archive.

#### Motivation pour les chercheurs

- **Un archivage pérenne, ouvert, accessible à tous et en accord avec la politique de science ouverte** qui se met en place, qui peut leur permettre de regrouper la production de leurs unités en temps réel (pour faciliter le recensement des productions lors des évaluations par exemple...)
- Le soutien de professionnels, pour un bon nombre d'établissements possédant des sous-portails HAL, les **bibliothécaires ont été formés et peuvent aider les chercheurs** dans leurs démarches.
- La possibilité d'utiliser **HAL pour déposer le texte intégral d'une soumission aux éditeurs** qui peuvent y accéder via un hyperlien, tout comme peut le proposer ArXiv.
- HAL propose une large visibilité au niveau national et international.
- HAL est très bien référencé sur Google et Google Scholar.
- HAL est une voie complémentaire au dépôt chez un éditeur classique ou prestigieux.

### Echanges autour du caractère obligatoire du dépôt dans le rapport CRAC

Il a été évoqué par certains correspondants le risque que cette obligation du dépôt dans HAL ne soit pas appréciée, et que si cette volonté est confirmée, elle doit être accompagnée par des moyens supplémentaires. Il serait préférable d'inciter le dépôt en précisant l'utilité et l'avantage pour le

---

<sup>1</sup> Pour la prochaine campagne CRAC 2019 et, probablement RIBAC, le CNRS va demander aux chercheurs de présenter seulement les articles qui sont importés de HAL.

chercheur, tout en proposant des solutions pour faciliter, aider le dépôt avec des options d'accompagnement dans les établissements par exemple.

Maintenant, cette contrainte pourrait inciter les chercheurs à plus déposer systématiquement et les habituer à la pratique du dépôt, même si l'autre crainte serait qu'il ne soit déposé que des notices et non le texte intégral.

Sylvie Rousset rappelle qu'il est attesté que la seule motivation personnelle sans obligation, ne permet que d'atteindre 20% d'articles déposés dans les archives ouvertes. Afin de se rapprocher des 70 %, il est nécessaire de passer par une demande à caractère plus coercitif.

Dans ce registre, l'université de Marseille a adopté un système de bonus-malus sur la subvention d'état, par exemple obtention de 80% de la dotation si 80% de dépôts de publications dans HAL.

Pour le CNRS, il est privilégié de passer par l'obligation de déposer, au niveau du compte rendu des chercheurs, ce qui représente un niveau intermédiaire de contrainte, et de considérer que les chercheurs, en règle générale plutôt sérieux, verront en cette action un intérêt à terme.

**>> Les CORIST précisent que le dépôt individuel est plutôt aisé surtout par le biais du DOI.**

La question se pose pour inciter le dépôt du texte intégral, et pas uniquement les notices, car HAL est une archive ouverte et non une base seulement de signalement de notices. Il est rappelé le plan national pour la Science Ouverte qui est une orientation très volontariste de passer de l'ère de la bonne volonté à l'ère de la politique publique, et l'ANR en implémentant son plan 2019 impose cette contrainte de dépôt en open access.

Enfin, les pratiques et habitudes des chercheurs en matière de dépôt dans les archives ouvertes varient selon les disciplines.

### Réponses techniques du CCSD suite aux retours des exercices pratiqués par les CORIST

#### ▪ **Dépôt individuel :**

Le dépôt individuel est facilité en utilisant le **DOI** de l'article. Si le **DOI** est mentionné dans le fichier ou si le chercheur l'a ajouté, le système récupère les métadonnées qui sont associées au DOI dans la base [CrossRef](#). Le DOI est conservé dans les métadonnées.

Une méthode alternative consiste à déposer un **fichier PDF** (version déposée : éditeurs ou MAA), la récupération des métadonnées s'opère via **GROBID**<sup>2</sup> (logiciel proposé par l'INRIA). Pour information, si le DOI est disponible dans le PDF, il est récupéré également via GROBID, et ensuite CROSSREF est interrogé afin de compléter les métadonnées qui n'auraient pas été récupérés avant.

A noter que GROBID est toujours en cours de développement, et s'il est constaté des problèmes de récupération ou d'erreurs de métadonnées, le chercheur est invité à en informer les équipes de HAL qui peuvent relayer auprès des équipes de l'INRIA qui développent GROBID et permettre l'amélioration du service rendu. Cette méthode est cependant intrinsèquement plus source d'erreur que la méthode via le DOI.

---

<sup>2</sup> GROBID Documentation : <https://grobid.readthedocs.io/en/latest/>

- **Dépôt en masse :**

A ce jour, l'INRIA a développé deux applications « BIP to HAL » et « X to HAL » pour aider au dépôt massif ou via le protocole SWORD<sup>3</sup>, qui utilise un format XML TEI. Des universités (Paris Nanterre par ex.) ont déjà mis en place ces protocoles.

D'autres universités ont également utilisé l'option des listes HCERES versées dans ZOTERO et récupérées ensuite dans HAL.

- **Accompagnement des chercheurs par des équipes de professionnels de l'IST**

En SHS, des documentalistes sont formés pour aider les chercheurs à déposer. Dans d'autres disciplines, dans les universités avec un sous-portail HAL, les professionnels de la BU organisent des formations courtes d'une heure très régulièrement pour apprendre à déposer dans HAL et sont désireux de rencontrer les chercheurs pour les former.

Au sein des réseaux métiers CNRS, par exemple RENATIS, les documentalistes sont prêts à monter des formations, comme par exemple dans les Maisons des Sciences de l'Homme.

>> L'un des objectifs des CORIST pourrait être de repérer, recenser les personnes et/ou équipes qui viendraient soutenir le chercheur dans sa démarche et aussi mieux diffuser au sein des établissements l'information autour de HAL et de son fonctionnement.

- **Dépôt de la notice seule ou avec le texte intégral :**

Le dépôt du texte intégral est bien entendu idéal. Cependant dans l'exemple d'ArXiv, il ne faut pas oublier le fait que le fichier du texte intégral n'est pas dans HAL mais qu'il est accessible via ArXiv.

Christine Berthaud informe que des discussions sont en cours avec le CERN pour qu'il développe des outils permettant de passer par HAL pour aller vers ArXiv et non l'inverse comme c'est le cas aujourd'hui. HAL apporte une diffusion nationale avec l'assurance de la pérennité de l'archivage, qui peut permettre de répondre aux besoins institutionnels, et être très visible dans sa communauté au niveau international.

Sylvie Rousset indique que cela vient à l'encontre d'une idée reçue que déposer dans HAL coupe le chercheur de leur diffusion habituelle de sa production scientifique (ex BioArXiv, etc).

Cette archive ouverte institutionnelle française est probablement enviée, d'autant qu'elle présente une modération probablement inégalée dans les autres archives, par exemple dans la gestion de la problématique des affiliations.

Didier Torny pense que cette nouvelle pratique pourrait amener une hausse très visible du nombre de dépôts de fichier de texte intégral dans HAL, car il rappelle que la première action demandée est de déposer un fichier. Cependant, il serait intéressant de savoir comment GROBID fait l'extraction de métadonnées du fichier afin d'améliorer la pratique. Car il est plus compliqué de rentrer les éléments de la notice à la main que d'utiliser le DOI. Par contre, cela risque de mettre sous tension le processus de modération en décembre.

>> Modération déjà sous tension comme le précise Bénédicte Kuntziger.

- **Point sur Dissemin.in | <https://dissem.in/>**

Ceux qui ont un ORCID suggère parfois de passer par Dissemin.in. Cela ne marche pas bien et cette application ne sera pas forcément maintenue. Cette application n'a pas été développée par le CCSD, mais par une équipe d'étudiants jeunes chercheurs. Il a été demandé de la connecter à HAL. Cette

---

<sup>3</sup> <https://api.archives-ouvertes.fr/docs/sword>

application est très peu utilisée, environ 200 dépôts en 18 mois. Il est à noter que cette application avait été développée avant la nouvelle procédure de dépôt.

Bénédicte Kuntziger a constaté beaucoup d'erreurs lors de tests.

Serge Bauin pose la question stratégique qui pourrait être de récupérer le code utilisé, car il n'est pas possible de laisser cette application active sans savoir qui la maintient.

Patrice Bellot signale le PV d'AG de dissemin.in datant de septembre 2018 qui signale le maintien de cette application.

>> Faudrait-il ôter Dissemin.in de HAL ?

#### ▪ **Possibilité de moissonnage dans HAL ?**

Christine Berthaud explique que normalement cela n'est pas prévu dans une archive ouverte vs un moissonneur. La difficulté majeure est de pouvoir moissonner des bases avec des fichiers en conformité avec HAL et de savoir qui fait le travail pour la mise en conformité avec ce que HAL peut récupérer (qui génère autant de travail du côté moissonné que moissonneur).

La question posée est quelle qualité souhaite-t-on des données dans HAL et, de fait, comment faire pour récupérer ce haut niveau d'informations ?

Le moissonnage de notices sur HAL pourrait être considéré sous conditions que les chercheurs aient un ORCID, de récupérer la notice et de demander au chercheur de déposer dans HAL.

>> Les CORIST signalent que cela serait bien de pouvoir moissonner dans HAL.

A noter, une réflexion est en cours au CCSD sur une collecte ciblée en identifiant des publications en accès libre d'un auteur identifié dans HAL et lui proposer de la rapatrier sur son compte HAL.

#### ▪ **Problème avec les affiliations auteurs**

Le challenge actuel pour le CCSD est de faire « matcher » la récupération d'une « forme auteur » soit un nom d'auteur et prénom ou un une initiale de prénom et un nom, avec un auteur présent dans HAL. Par exemple, s'il y a un nom, prénom d'un côté et un nom et initiale du prénom de l'autre côté, cela n'est pas possible de « matcher » les deux, et il faut donc recréer une forme auteur.

Si le match est possible avec une forme auteur dans HAL, il est recherché dans HAL l'affiliation qui n'est pas extraite par GROBID en utilisant le même algorithme que celui de « X to HAL ». Cette opération est perfectible et peut expliquer la récupération d'affiliations incorrectes ou anciennes sans oublier les cas d'homonymie. Par ailleurs, les auteurs qui déposent ne contrôlent pas toujours si les affiliations sont correctes.

A ce jour, il est impossible pour les équipes de modérateurs HAL de vérifier toutes les données déposées dans HAL, car il y a plus de 250 dépôts par jour de full text. L'une des options pourraient être, par exemple, de donner une délégation de modération aux universités qui ont un portail et qui en font la demande.

#### ▪ **Interface pour un dépôt simplifié**

Le CCSD a développé une interface de dépôt simplifié pour le dépôt, par contre, il est difficile de savoir quels boutons mettre en avant pour faciliter la navigation, sans rendre la page d'accueil trop chargée par trop d'informations et d'options.

Par exemple, pour ceux qui importent un fichier « LATEX », il n'est pas aisé de savoir qu'il faut cliquer sur « afficher la vue détaillée » pour déposer des fichiers et la possibilité d' « ajouter un embargo ».

>> **Ergonomie du dépôt** : audit en cours pour améliorer l'interface et fonctionnalités de dépôts dans HAL

- **Lien vers PCI (Peer Community In)**

Question posée : « Si une publication fait l'objet d'une recommandation sur Peer Community In, est-il possible de faire figurer sur la page HAL correspondant à la publication un lien vers la recommandation sur PCI ? »

A ce jour, il n'existe pas de lien spécifique seul une section commentaires, toujours la difficulté de rajouter ou pas des boutons supplémentaires qui ensuite peuvent perdre les déposants si trop d'options.

>> A noter que HAL a un service support<sup>4</sup> qui peut répondre aux questions spécifiques.

- **Question de l'embargo, à quel moment déposé ?**

Bénédicte Kuntziger suggère de déposer de suite la version auteur finale en lui ajoutant un embargo si nécessaire. Si l'éditeur ne veut pas laisser la consultation de suite, mais avec un embargo de 6 mois à un an selon les disciplines, l'auteur n'a plus à s'en préoccuper.

>> A noter que HAL n'a jamais été contacté par un éditeur pour un problème de délai d'embargo.

## 2) Plateformes d'édition

Sylvie Rousset : peut-on encourager des modèles de plateformes plus vertueux ?

Les CORIST avaient déjà fait un premier état des lieux des plateformes utilisées par les différentes disciplines. Pour soutenir une ou plusieurs plateformes, il faudrait connaître le coût et le modèle économique des plateformes.

### **Organisation d'une journée plateformes d'édition pour les CORIST**

Objectif : soutenir un modèle et à quel coût ? intégrer une diversité de modèles (plateformes, épi journaux, groupements de commandes, etc.)

Une discussion entre les CORIST sur ces 3 plateformes : E-life, Plos One et Frontiers, qui ont le même niveau d'APC, a permis de choisir E-Life qui semble la plus innovante.

### **Plateformes retenues à rencontrer lors de cette journée**

- Openedition, Copernicus, Centre mersenne, F1000, E-Life, Episciences, EDP Sciences, Open library of humanities, Scipost

Il reste à :

- Définir une date de cette journée.
- Définir des questions types pour les responsables des plateformes.
- Inviter les responsables des plateformes sélectionnées.

## 3) Préparation réunion CORIST du 21 mars autour des données de la recherche

Invitations à faire :

- Correspondants données des instituts en plus des CORIST

---

<sup>4</sup> [hal.support\[@\]ccsd.cnrs.fr](mailto:hal.support[@]ccsd.cnrs.fr)

Interventions pressenties :

- MICADO
- HUMANUM
- RGPD

**Objectif : quelle va être la feuille de route CNRS sur les données de la recherche ?**

**Avoir une présentation :**

- OPIDOR
- MPR du CNRS
- RDA - Francis André
- Finalisation du programme avec Francis André

#### 4) Questions diverses

- **Demande d'élargissement d'accès à BibCNRS :**

Ces demandes sont à valider par les CORIST. Pour information, s'il s'agit de l'institut secondaire du laboratoire, l'INIST le fait automatiquement. La question se pose au-delà de 2 instituts, donner des accès à tout un labo ou seulement pour une équipe.

>> Rappel de la procédure :

Au-delà deux instituts, la demande d'élargissement doit passer par les CORIST pour validation.

Afin de pouvoir identifier le coût supplémentaire d'accès à BibCNRS au-delà de 2 instituts, qui semble peu important, Claire François va nous proposer un bilan financier.

- Organisation d'une **Journée Science Ouverte** | Mardi 08 octobre 2019

Campus Gérard Mégie – Amphithéâtre Marie Curie

Cette journée, organisée dans le cadre des 80 ans du CNRS, est à destination des instituts (DI, DAS et CSI).

- Enquête BibCNRS :

L'INIST a initié en février 2019, dans le cadre d'une démarche UX (Expérience Utilisateur) mobilisant les outils de la conception centrée utilisateurs, une **enquête pour mieux connaître les pratiques documentaires et les attentes de ses utilisateurs**, ainsi que leurs attentes.

Claire François nous indique que 4 000 répondants ont rempli le questionnaire.

---

**Prochaine réunion : Jeudi 21 mars | 13h30 – 16h30 | Université Paris Diderot**

**Rappel des prochaines réunions**

- 18 avril – K024
- 16 mai – K024
- 20 juin – K024