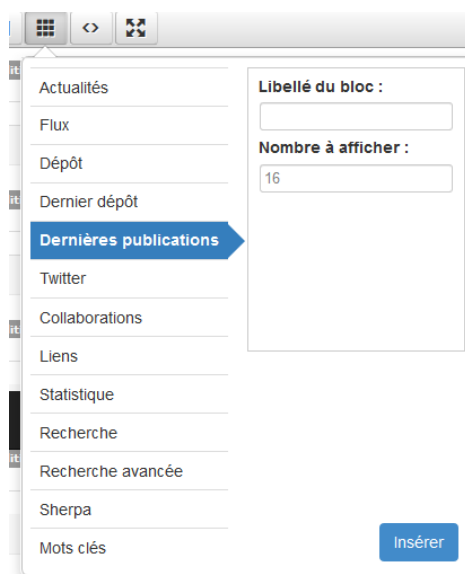


Synthèse CasuHAL

Janvier - 2021

Questions soulevées

- **Financement du CCSD** (Bibliothèques Ulm-LSH et Maths-Info): comment le nombre d'enseignants-chercheurs est comptabilisé par établissement (en fonction de leurs affiliations ? mais elles peuvent multiples, allons-nous payer plusieurs fois pour un même chercheur ? Et qu'en est-il des chercheurs associés ? Des doctorants ? Ensuite que se passe-t-il lorsque nous nous retrouvons avec des portails en "poupées russes" suite à des créations ou des fusions d'établissements ? →
 - Je souhaite transformer pour une COMUE une collection HAL en portail et j'ai le même type de questions : nous avons plusieurs établissements, avec des statuts d'enseignants et chercheurs différents. Certains sont également affiliés à plusieurs établissements, membres ou pas de notre regroupement... La réponse du CCSD faisait mention, en plus de la grille tarifaire, du fait que les `enseignant.es/chercheur.es` des établissements membres de notre COMUE ayant déjà un portail seraient exclus du décompte. Je suppose que la réponse serait la même pour des établissements qui fusionnent (HESAM Université)
- **Affichage des dernières publications sur une page de collection** (Institut de Physique des 2 Infinis de Lyon): création d'une page pour afficher les dernières publications d'une collection http://hal.in2p3.fr/NEUTRINOS_IP2I/page/index. Il n'y a que deux publications affichées alors que la collection en compte 16. Les deux qui s'affichent ont un fichier associé, les autres non. Savez-vous comment on peut afficher tous les dépôts de la collection ? →
 - Le widget qui permet d'afficher les dernières publications affichent que les "10 derniers dépôts en texte intégral". Peut-être y-a-t-il moyen de modifier le widget pour qu'il n'exclue pas les notices (Bibliothèque de l'Université de Perpignan).
 - Il faut paramétrer le widget avant son insertion (Service de la Documentation et d'Archives - École des hautes études en santé publique)



- Vous pourriez récupérer le flux RSS de la page qui présente tous les dépôts et utiliser le widget "flux" pour un affichage sur la page d'accueil (Université de Normandie). →
 - Je viens de tester sur une collection. Par contre, c'est dommage car chaque entrée est considérée comme "un billet" et malgré le flux qui devrait afficher toutes les publications et la spécification dans le widget pour que les 16 publications apparaissent, il n'y en a que 6. Voir la copie d'écran ou le lien vers la collection: <https://hal.archives-ouvertes.fr/GREENH-CITY/page/index>. J'ai fait une recherche avancée que j'ai ensuite ajoutée dans l'onglet "consultation". Ça ne répondait pas exactement à la demande de l'équipe de recherche qui aurait préféré voir toutes les publications sur la page d'accueil, notices incluses (Ladyss - Université Paris Nanterre).
 - Afin d'afficher l'exhaustivité des publications, création dans le menu d'une page avec un lien web et avec le suffixe suivant : /search/index/q*/rows/30/. Ce qui donnerait pour votre collection : http://hal.in2p3.fr/NEUTRINOS_IP2I/search/index/q*/rows/30/ Ensuite vous pouvez copier cet hyperlien sur la page d'accueil de votre collection ce qui permet de renvoyer vers une page de recherche à facettes avec des outils d'exports très utiles pour les chercheurs (Bibliothèque Universitaire Maurice Agulhon).
 - Cette restriction aux notices avec pdf est gênante. Il faudrait à minima laisser le choix à l'utilisateur. Cette restriction perd son sens avec les notices en AO sur le site de l'éditeur et non dans HAL. Une évolution de HAL là-dessus serait très souhaitable
 - Pourquoi ne pas intégrer, dans un nouvel onglet "page web" classique, l'un des outils "[pages web dynamiques alimentées par les dépôts HAL](#)" ? Je crois que certains permettent d'afficher aussi les notices sans texte intégral (URFIST de Paris - École nationale des chartes).
- **Crac-ribac/HAL** (Bibliothèque universitaire Savoie Mont Blanc) : La période du CRAC incite à mettre son HAL à jour même à la dernière minute. Que répondez vous aux auteurs dont le dépôt a été refusé et qui disent ne pas avoir les docs d'origine.... →
- Leur conseiller d'utiliser la rubrique "remarques/commentaires" du RIBAC (cf page 22 du manuel RIBAC ([lien vers PDF](#)) : "Cette rubrique vous permet de nous faire part de vos remarques à la fois sur l'outil RIBAC (activités à développer ou à réduire, avis sur le contenu et l'ergonomie, etc., **et sur vos activités**".) La même rubrique existe peut-être dans le CRAC ? Cela ne résoud pas directement le problème (qui demande aux auteurs une recherche de leur version postprint ou une autorisation de l'éditeur, ce qui peut prendre du temps), mais cela permet de communiquer une partie de l'info.... En attendant un développement de l'interopérabilité entre les 2 systèmes qui permettrait la prise en compte des métadonnées des dépôts en attente de vérification (UMR Triangle - ENS de Lyon).
 - Il y a un onglet d'aide à la saisie sur le site du CRAC (<https://crac.dsi.cnrs.fr>) mais cet onglet n'est disponible que pendant la période d'ouverture (en 2020, du 3/11 au 15/12). Il n'est donc pas possible actuellement de copier cette page. La page suivante contient une FAC : https://intranet.cnrs.fr/Cnrs_pratique/recruter/carriere/chercheurs/Pages/CRAC.aspx (Université de Montpellier)
 - Le document est peut-être accessible depuis l'interface CRAC, en quel cas il faut être connecté et donc avoir des identifiants chercheurs pour y accéder. Il faut voir avec les collègues chercheurs du laboratoire (Bibliothèque universitaire Savoie Mont Blanc).
 - Quand le cas se présente, inviter l'auteur : 1) s'il est auteur correspondant : à se connecter à la plateforme de soumission de la revue (en lui donnant le lien direct et la démarche à suivre, d'après les indications du site Direct2AAM), 2) s'il est co-auteur, à contacter l'auteur correspondant de l'article pour obtenir le manuscrit pre-proof (Université de Rennes 1).
- **Champ vide** (MSH Dijon) : Dans une collection HAL, quelle serait la requête pour obtenir toutes les références dont le champ "référence interne" est vide ?

- **Affichage publications équipes suivant template du CV HAL** (Service DIST - LIRMM - CNRS UMR5506 / Univ de Montpellier) : le laboratoire procède à une refonte de son site internet et aimerait afficher les publications des équipes de recherche (ayant un identifiant de structure dans Aurehal) sous la forme du CV Hal du chercheur (avec la même mise en page). Quelqu'un l'a-t-il déjà fait, ou sinon avez-vous des pistes pour m'aider à y parvenir ? →
 - Tout dépend du CMS que vous utilisez et des moyens dont vous disposez. Dans notre laboratoire, le site web est sous SPIP. Les pages des membres du laboratoire se composent de deux parties. Une partie rédigée et mise à jour "à la main" où on trouve les activités de recherche, les responsabilités académiques et les enseignements et une partie dynamique avec les publications. Nous utilisons Haltools qui permet de créer des pages même pour les chercheurs qui non pas encore d'IdHal. Voici un exemple: <https://www.ladyss.com/lamarche-thomas> (Ladyss - Université Paris Nanterre).

- **Extraction de la littérature citée** (Université Sorbonne Nouvelle) : comment exporter les informations de plusieurs articles (par API, OCDHAL ou tout autre moyen) qui ont des informations de citations (sur la page publique de l'article, "Littérature citée"). On peut "télécharger", ce qui renvoie à un fichier HTML assez plat. Je ne trouve aucun champ dans l'API HAL qui me permette de les exporter, pas non plus de solution trouvée dans OCDHAL. Est-ce que quelqu'un-e d'autre a rencontré ce problème ? →
 - On ne trouve pas de champs dans l'API de HAL, que ce soit pour avoir la liste ou bien pour savoir si une liste est présente (sauf erreur cette liste n'est présente que pour les notices avec fichier). Peut être est-ce possible en OAI MPH : on voit dans le format *oai_dcterms* le champs *bibliographicCitation*. <https://api.archives-ouvertes.fr/oai/hal/?verb=ListMetadataFormats>. En tout cas que g. Scholar remontait bien ces infos de citation

- **Statut juridique des métadonnées HAL** (SCD Université Bordeaux Montaigne) : Quel est le statut juridique des métadonnées des notices de HAL ? J'ai lu la page "données personnelles" de HAL mais qui, me semble-t-il, concerne surtout les données du compte HAL (adresse mail, nom prénom, etc.). Pour les fichiers en texte intégral c'est aussi relativement clair : l'auteur cède un droit de diffusion non-exclusif en déposant, en acceptant l'impossibilité de la rétraction. Mais qu'en est-il des MD ? Je donne un exemple pour rendre ma question plus concrète : mettons qu'une entreprise privée dont le modèle économique repose sur l'accès payant à des métadonnées de publications scientifiques (réseau social de recherche, logiciels de bibliométrie, base de connaissance) moissonne HAL et en commercialise les MD dans le cadre de son activité. Peut-il légalement le faire sans autorisation ? Plus prosaïquement : un chercheur qui souhaiterait faire de la fouille de données sur les données de HAL, avec mise à disposition ultérieure de son corpus. A-t-il besoin de demander une autorisation, et si oui à qui ? →
 - Pas mal de fournisseurs de données indiquent assez clairement le statut de leurs métadonnées, comme l'INHA qui les met sous licence CC. Pour les métadonnées de HAL, je n'ai trouvé que cette information : "Moissonnage : conditions d'utilisation des données. Les métadonnées de HAL peuvent être consultées de façon totale ou partielle par moissonnage dans le respect du code de la propriété intellectuelle. Pas d'utilisation commerciale des données extraites. Obligation de citer la source (exemple : hal.archives-ouvertes.fr/hal-00000001)." Ce qui équivaut en quelque sorte à une licence CC-BY-NC. L'information est disponible dans la documentation de HAL. Cela empêcherait donc une entreprise privée de faire une réutilisation onéreuse de ces informations, mais qui rendrait possible leur utilisation par un chercheur (Université Sorbonne Nouvelle).
 - - les extractions de métadonnées de HAL sont-elles considérées comme des documents administratifs, définis dans l'[article L300-2](#) du Code des relations entre le public et l'administration

("documents produits ou reçus, dans le cadre de leur mission de service public, par l'Etat, les collectivités territoriales ainsi que par les autres personnes de droit public ou les personnes de droit privé chargées d'une telle mission") ? Si non, pourquoi ?

- si oui, les données qui figurent dans ces extractions sont-elles considérées comme des informations publiques, évoquées dans l'[article L321-1](#) du même code ("Les informations publiques figurant dans des documents communiqués ou publiés par les administrations mentionnées au premier alinéa de l'article [L. 300-2](#) peuvent être utilisées par toute personne qui le souhaite à d'autres fins que celles de la mission de service public pour les besoins de laquelle les documents ont été produits ou reçus.") ? Si non, pourquoi ? J'imagine qu'on peut considérer que sur le champ "résumé" en particulier (et peut-être aussi sur le champ "titre") "des tiers détiennent des droits de propriété intellectuelle" ce qui justifierait que ces champs ne soient pas considérés comme des "informations publiques" selon l'[article L321-2](#) du même code.

- si oui, je suis un peu surpris des "conditions d'utilisation des données" prévues sur la page <https://doc.archives-ouvertes.fr/questions-juridiques/> puisque l'[article L323-2](#) du même code prévoit que "Les administrations qui élaborent ou détiennent des documents contenant des informations publiques pouvant être réutilisées dans les conditions prévues au présent chapitre sont tenues de mettre préalablement des licences types, par voie électronique, à la disposition des personnes intéressées par la réutilisation de ces informations." et que "Lorsque la réutilisation à titre gratuit donne lieu à l'établissement d'une licence, cette licence est choisie parmi celles figurant sur une liste fixée par décret, qui est révisée tous les cinq ans, après concertation avec les collectivités territoriales et leurs groupements. Lorsqu'une administration souhaite recourir à une licence ne figurant pas sur cette liste, cette licence doit être préalablement homologuée par l'Etat, dans des conditions fixées par décret." ; la "liste fixée par décret" apparaît à l'[article D323-2-1](#) du même code et je n'y vois aucune licence portant sur les données qui limiterait un usage commercial, et je ne vois pas la licence d'utilisation des données de HAL sur la page <https://www.data.gouv.fr/fr/licences> qui recense les "licence homologuées".

- Les métadonnées de HAL peuvent être consultées de façon totale ou partielle par moissonnage OAI-PMH dans le respect du code de la propriété intellectuelle. Pas d'utilisation commerciale des données extraites. Je précise que cette mention doit sans doute être ancienne et ne me paraît plus conforme à l'évolution du droit, car en vertu de la Loi République numérique, ce sont des données publiques qui devraient être librement réutilisables. Ce qui rend les choses un peu plus compliquées, c'est que les métadonnées de HAL contiennent aussi des données personnelles. Mais en tout cas, s'il y a des restrictions à la réutilisation, elles ne peuvent certainement plus être basées sur la propriété intellectuelle.

➤ **Thèses confidentielles et sous embargo sur HAL** (SCD – Université Polytechnique Hauts-de-France) : comment d'autres établissements gèrent leurs thèses confidentielles et sous embargo sur HAL. Pour les thèses confidentielles, existe-il des notices sur HAL ? Pour celles sous embargo, les documents sont-ils déposés avec une visibilité après l'embargo ? Vous êtes-vous posés la question de la visibilité interne de vos thèses sous embargo ?

- A l'UVSQ, les thèses confidentielles n'ont pas de notices HAL. Concernant celles sous embargo, elles sont signalées sur HAL via TEL puis accessibles en texte intégral après la fin de l'embargo. N'ayant pas d'Intranet, l'auteur de la thèse sous embargo nous fournit un exemplaire imprimé, mis à disposition de la communauté universitaire. (*BU de St-Quentin - Direction des Bibliothèques et de l'IST*).
- Les thèses sous embargo sont obligatoirement diffusées à minima dans l'enceinte de l'établissement de soutenance pendant la période d'embargo. Cette diffusion peut se faire, en fonction des moyens techniques de l'établissement : sous forme d'une copie imprimée ou sous forme électronique, à partir d'un serveur de diffusion local, avec accès authentifié. Le mode d'accès des thèses sous embargo (papier ou accès électronique restreint) est signalé sur theses.fr. Les thèses sous embargo

sont donc bien mises à la disposition des membres de l'établissement concerné. Cette information n'est pas disponible sur HAL, qui est une plateforme d'archives ouvertes, mais sur theses.fr, qui est le portail national des thèses de doctorat (ABES)

- Nous sommes également en demande d'une visibilité de TOUTE la production scientifique de notre École. Le contrat est rempli, sauf pour les thèses sous embargo ou confidentiel, voir même des thèses plus anciennes, datant d'avant notre autorisation à délivrer, que nous souhaitons valoriser lorsque l'un de nos EC étaient Co-Directeur de thèse.

Par ailleurs, sur le portail INRAE, les thèses (sans fichier) sont signalées. Peut-être d'autres portails ? (Centre de documentation – IMT Mines Albi-Carmaux).

- A l'UL, nous signalons les thèses confidentielles sous forme de notices (avec mention de la date de sortie de confidentialité en commentaire). Pour les thèses sous embargo, nous avons un lien vers notre serveur local : - avec authentification UL pour les thèses antérieures à 2016 (ex. <https://hal.univ-lorraine.fr/THESES-UL/tel-01754574v1>) ; - avec authentification Renater pour les thèses après 2016 (ex. <https://hal.univ-lorraine.fr/THESES-UL/tel-03116859v1>). Le dépôt de ce type de lien n'était dans un 1er temps permis que sur quelques portails mais est désormais accessible à tous (Champ URL du fichier dans le formulaire de dépôt de thèse, avec ce commentaire "*Url de la thèse en accès restreint sur un site ouvert à l'ESR*"). Ce champ n'est pas répétable. Pour les thèses en plusieurs volumes, le CCSD préconise le dépôt d'un zip.

Nous effectuons une double saisie, notamment parce que le problème de protection des champs n'est pas réglé côté CCSD (ce qui implique un écrasement des données dans HAL à chaque revalidation Star). Le signalement et le suivi dans STAR et dans HAL sont effectués par les mêmes personnes qui ont une vision d'ensemble sur toutes les opérations concernant les thèses (du dépôt à la modération). A priori, seules les thèses avec fichier peuvent actuellement être transférées depuis Star vers HAL. (BU – Université de Lorraine)

- **Widget HAL derniers dépôts pour une page d'accueil d'un site web Typo3** (Laboratoire Territoires, Villes, Environnement & Société – Université de Lille) : Avez-vous une piste pour intégrer un widget qui afficherait les derniers dépôts d'une collection HAL sur la page d'accueil d'un site web avec typo3 ? →

- Nous essayons de faire développer un plugin Typo3 permettant d'afficher la liste des publications d'une collection particulière ou d'un IDHAL spécifique. Il est ensuite possible de trier les publis selon l'année de parution et/ou le type de publication. Il ne permet pas d'afficher les publications selon la date de dépôt, si c'est bien cela que vous recherchez. Le plugin est encore en développement. On espère une mise en production d'ici la fin du semestre (Université de la Réunion).

Outils présentés

- **Nouvel outil Shareyourpaper.org** (URFIST de Paris - École nationale des chartes) : Je serais curieuse de savoir si certains établissements l'implémentent dans leurs sites internet institutionnels, et quelle pourrait être l'articulation avec HAL.
- **Portail Open Dissertations d'Ebsco** : Dans le but de rendre les Thèses et Mémoires Électroniques plus facilement accessibles aux chercheurs du monde entier, EBSCO et BiblioLabs ont créé EBSCO Open Dissertations. Ce portail gratuit est destiné tant aux étudiants qu'aux universités et a pour but de faciliter l'accès aux Thèses et Mémoires Électroniques. Les auteurs peuvent y télécharger leurs Thèses et leurs Mémoires Électroniques pour les partager avec d'autres. En retour, les universités

voient augmenter le trafic vers leurs dépôts institutionnels. Plus de 800 000 thèses et mémoires électroniques, provenant de 26 universités à travers le monde, sont référencées dans Open Dissertations. Pour y accéder, aller [sur BibCnrs](#) puis cliquer sur l'onglet "*Une base de données*"

Varia

- **ActuHAL : une newsletter pour connaître les dernières actualités de HAL** : En complément du blog du CCSD, est mis en place une newsletter, ActuHAL, qui paraîtra tous les 2 mois. N'hésitez pas à vous abonner et à relayer l'information : <https://www.ccsd.cnrs.fr/newsletter-actuhal/>
Le 1er n° de ActuHAL : <https://bit.ly/2N9irVl>

- **[Nouveau billet sur casuhal.org] Catherine Cardon, interview de la doyenne de HAL** :
<https://www.casuhal.org/2021/01/25/catherine-cardon-interview-de-la-doyenne-de-hal/>

- **Save the date - Journées Casuhal 7-11 juin 2021** : Afin de tenir compte des incertitudes liées à la situation sanitaire, le comité d'organisation, en accord avec le bureau Casuhal, a fait le choix de proposer un format distanciel, étalé sur la semaine complète.

- **Blog du CCSD - Formations en ligne : vous avez été au rendez-vous en 2020 !** : Le dernier billet du blog du CCSD revient sur le succès des webinaires organisés par l'équipe support et assistance <https://www.ccsd.cnrs.fr/2021/01/formations-en-ligne-vous-avez-ete-au-rendez-vous-en-2020/>