ourouk.fr
CONSEIL
EN MANAGEMENT
DE L'INFORMATION

ANR-10-IDEX-0004-02

ETUDES COPIST (CATALOGUE D'OFFRES PARTAGEES D'IST)

Rapport de l'étude n° 4 : L'articulation des archives des établissements et de l'archive nationale pluridisciplinaire HAL

Version finale du 24/08/2018

Référence : CNRS_COPIST_ETUDE 424/08/2018_VF

Cette étude résulte du travail du cabinet Ourouk.
Ses commanditaires ne sont pas engagés par les recommandations qu'elle contient.

5 RUE AMBROISE THOMAS
75009 PARIS
www.ourouk.fr

TÉL : 01 44 82 09 99
FAX : 01 44 82 72 70
MÉL : ourouk@ourouk.fr

SARL AU CAPITAL DE 358 800 €
RCS B 387472160 - APE 7022Z
AGRÉMENT FORMATION : 11753245175

SOMMAIRE

Introduction ... **3**
Objet du document ... 3
Six études pour la mise en cohérence des projets numériques en IST de l'ESR ... 3
Attendus de l'étude ... 4
Eléments méthodologiques ... 4
Résumé ... **6**
Première partie : Les archives institutionnelles ... **7**
Définition ... 7
Etat des lieux ... 7
Motivations pour la création d'une archive institutionnelle ... 8
Des contextes singuliers ... 10
Les freins à la création d'une archive institutionnelle ... 12
Un dispositif articulant les archives locales et l'infrastructure nationale HAL ... 12
Deuxième partie : Les portails institutionnels dans HAL ... **13**
Genèse et développement des portails institutionnels dans HAL ... 13
Caractéristiques de l'offre portail ... 15
Ouvrir un portail dans HAL : un choix avant tout économique au regard du service délivré ... 17
Les limites des portails HAL invoquées par les établissements ... 18
Troisième partie : Attentes des établissements et solutions envisageables ... **20**
Introduction ... 20
Prendre en compte la production scientifique et technique des établissements dans son ensemble ... 21
Améliorer la qualité des référentiels ... 26
Enrichir l'offre fonctionnelle proposée aux chercheurs pour valoriser leurs dépôts ... 28
Développer des outils pour les gestionnaires qui déposent pour les chercheurs, complètent et enrichissent les métadonnées ... 30
Réexaminer les principes régissant la modération ... 31
Simplifier le reversement dans HAL des archives institutionnelles locales ... 33
Des évolutions du socle technique de HAL à moyen terme ... 34
Poursuivre, en la développant, la démarche d'écoute et de prise en compte des demandes des institutions utilisatrices de HAL ... 35
Quatrième partie : Synthèse des évolutions attendues et de leurs impacts ... **37**
Les évolutions les plus attendues ... 37
Les autres évolutions ... 38
Les moyens limités du CCSD ... 39
Annexes ... **40**
Annexe 1 : Liste des institutions et des personnes interrogées ... 40
Annexe 2 : Pondération des évolutions proposées ... 42
Annexe 3 : Commentaires post-étude des commanditaires ... 43

Introduction

Objet du document

Le présent document constitue la version finale du rapport de l'étude Copist n°4 « *L'articulation des archives des établissements et de l'archive nationale pluridisciplinaire HAL : besoins et solutions* ».

Il intègre des commentaires post-étude de ses commanditaires (CPU, ADBU, COUPERIN, EPRIST et CNRS) en annexe 3, page 43.

Six études pour la mise en cohérence des projets numériques en IST de l'ESR

L'enquête nationale COPIST[1] sur les services d'information scientifique et technique a révélé une forte intention de partage dans les différents domaines de l'IST de la part de la centaine d'établissements de l'ESR qui y ont participé.

L'analyse conjointe des résultats de l'enquête par la CPU, l'ADBU, COUPERIN, EPRIST et le CNRS a abouti à la définition de 6 thèmes d'études destinées à préciser les attentes prioritaires des institutions de l'ESR.

Les thèmes retenus s'inscrivent tous dans la perspective d'une mutualisation de grands services d'IST à l'échelle nationale, à l'instar de la démarche qui a présidé au développement du projet ISTEX :

1. Des outils mutualisés de signalement et d'accès aux ressources documentaires électroniques pour l'ESR : enjeux, conditions, faisabilité
2. Une plateforme mutualisée pluridisciplinaire de stockage, de gestion, de signalement et de partage de données de recherche : enjeux, conditions, faisabilité
3. Des modèles innovants de publication et d'édition scientifiques publiques : conditions de développement
4. L'articulation des archives des établissements et de l'archive nationale pluridisciplinaire HAL : besoins et solutions
5. Des modèles économiques et de financement des services d'IST
6. Des actions de communication et de formation destinées à assurer une meilleure visibilité et appropriation des services d'IST par les communautés de recherche et les professionnels de l'information

Ces études de cadrage ont une visée opérationnelle. Leur objectif général est de définir la portée et le périmètre des évolutions de services existants et de nouveaux services répondant aux besoins et aux usages des différentes communautés de l'ESR, structurés et dotés d'un modèle économique pérenne.

[1] Rapport disponible sur www.cnrs.fr/dist/z-outils/documents/copist_rapport-analyse-conjointe_18-05-2017.pdf

Il s'agit, *in fine*, d'aboutir dans le consensus, et à partir d'un état des lieux initial des ressources et pratiques, à un Catalogue d'offres partagées d'Information scientifique et technique (COPIST) qui s'inscrit dans la stratégie du CNRS et de ses partenaires de l'ESR de « *Mieux partager l'IST pour mieux partager les connaissances* ».

Attendus de l'étude

La création de portails institutionnels dans HAL répond aux besoins d'un nombre croissant d'institutions qui y voient le moyen de disposer de leurs archives sans qu'il soit nécessaire de développer le service *ad hoc*, le dépôt de documents non éligibles pour HAL étant autorisé dans le cadre d'un portail institutionnel. Mais nombre d'établissements estiment que le service *Portails* de HAL ne répond pas, en l'état, à l'ensemble de leurs besoins propres, conduisant certains à développer leur propre système d'archives.

Conscient de ce problème, le CCSD souhaite prendre en compte ces attentes pour les intégrer dans sa feuille de route 2016-2020.

L'objectif de l'étude est de recenser et de décrire les attentes fonctionnelles des établissements pour exposer leurs archives dans HAL et de définir les conditions et moyens de leur satisfaction.

Elle doit permettre de répondre aux principales questions suivantes :

- Quels sont les besoins précis des établissements ?
- Quelle doit être la structure optimale de collaboration entre les établissements et le CCSD pour articuler les archives ouvertes institutionnelles et HAL ?
- Quelles réponses le CCSD peut-il apporter et dans quelles conditions organisationnelles, techniques, économiques ?

Eléments méthodologiques

Une vingtaine d'entretiens ont été réalisés auprès d'institutions ayant fait le choix, d'une part, d'une archive ouverte institutionnelle locale et, d'autre part, d'un portail HAL.

Les entretiens ont porté sur la démarche et les raisons qui ont conduit à la solution retenue, sur les évolutions à apporter à l'offre de portail HAL pour qu'elle réponde davantage aux besoins actuels de l'institution, et sur les orientations que pourrait prendre HAL à moyen terme au-delà de son périmètre historique.

Deux entretiens avec le CCSD ont permis d'éclairer les demandes d'évolutions formulées et d'en apprécier les impacts sur le modèle HAL et sur les charges de développement ou d'exploitation.

Institutions interrogées

Sur les 17 institutions interrogées, 7 ont développé une archive locale et 10 ont fait le choix d'un portail HAL.

Concernant les premières, il s'agit de projets anciens parfois antérieurs au développement des portails HAL, comme de projets récents ou en cours :

- IRSTEA (2005)

- INRA (2006)
- Université de Liège (2008)
- Université Paris-Dauphine (2009)
- ENSAM (2013)
- Université de Strasbourg (2016)
- Université de Lille (en projet : décembre 2017)
- Université de Bordeaux (en projet : 1er trimestre 2018)

L'Université de Liège, bien que hors contexte, a été interrogée car son archive locale est souvent citée comme un exemple d'orientations à suivre pour HAL.

Concernant les secondes, il s'agit d'institutions actives dans le développement de leur portail et investies dans la démarche d'amélioration de l'offre du CCSD notamment dans le cadre de CasuHal, le club utilisateur de HAL :

- INRIA (2004)
- INSERM (2006)
- Université de Toulouse 2 (2011)
- Université de Rennes 1 (2012)
- Aix-Marseille Université (2013)
- Université de Grenoble Alpes (2013)
- Université de La Réunion (2015)
- Université de Lorraine (2015)
- Université de Montpellier (2017)
- Université Paris Lumières (2017)[2]

Personnes interrogées

Les personnes interrogées au sein de ces établissements (cf. annexe 1 page 40) sont directement en charge des archives ouvertes, le plus souvent dans le cadre d'une mission plus large de soutien à l'Open Access ou de services d'appui à la recherche.

Dans les universités, à l'exception de Toulouse 2, les interlocuteurs relèvent des SCD et des bibliothèques universitaires.

[2] Les universités Paris 8 Vincennes-Saint-Denis et Paris Ouest Nanterre La Défense qui forment l'université Paris Lumière disposent d'un Portail HAL, respectivement depuis 2012 et 2016)

Résumé

L'étude n° 4, *« L'articulation des archives des établissements et de l'archive nationale pluridisciplinaire HAL : besoins et solutions»*, examine les besoins spécifiques des institutions en matière d'archives institutionnelles et les évolutions de l'archive HAL qui permettraient de les satisfaire.

Les projets d'archives institutionnelles visent en premier lieu la mise à disposition en accès libre et gratuit de la production scientifique de l'établissement. Toutefois, d'autres motivations sont également à l'œuvre à des degrés variables selon les établissements, comme la réalisation d'une vitrine de la totalité des productions pédagogiques, scientifiques ou techniques de l'établissement, l'analyse de la production des chercheurs et la préservation à terme des documents produits.

L'offre de portail institutionnel proposée par HAL depuis 2003 est un sous-ensemble de l'archive nationale pluridisciplinaire qui permet aux institutions de valoriser leurs productions en les réunissant dans des espaces dédiés. Cette offre a été adoptée par plus de 120 institutions, notamment pour des motifs économiques au regard du service délivré.

Les entretiens menés auprès d'institutions ayant fait le choix d'une archive ouverte institutionnelle locale ou d'un portail HAL ont mis en évidence 25 mesures politiques, techniques ou organisationnelles qui favoriseraient la création de nouveaux portails institutionnels par des établissements qui n'en disposent pas encore, conduiraient certains disposant d'une archive locale à l'abandonner au profit d'un portail dans HAL, et amélioreraient l'articulation entre les archives institutionnelles locales et l'archive nationale. Ainsi, les principales mesures visent à :

- élargir les types de documents éligibles dans HAL,
- autoriser un embargo illimité pour certains documents,
- intégrer des données de la recherche liées à des documents déposés,
- améliorer la qualité des référentiels,
- enrichir l'offre fonctionnelle proposée aux chercheurs pour valoriser leurs dépôts,
- développer des outils pour les gestionnaires qui déposent pour les chercheurs,
- réexaminer les principes régissant la modération,
- simplifier le reversement des archives institutionnelles locales dans HAL.

Un chapitre conclusif interroge les capacités humaines et budgétaires de l'actuel CCSD pour répondre dans des délais raisonnables à tout ou partie de ces attentes.

Première partie : Les archives institutionnelles

Définition

Dans son glossaire de l'Open Access, l'INIST propose la définition suivante de l'archive institutionnelle : *« Une archive institutionnelle relève d'une institution (université, grande école, organisme de recherche, association professionnelle) et a pour objectif de contenir, valoriser et conserver* ***l'ensemble de la production scientifique*** *de celle-ci.*[3] ».

En France, les premières archives institutionnelles sont apparues au début des années 2000 quand les institutions ont commencé à rejoindre le mouvement en faveur de l'Open Access initié par les chercheurs et les professionnels de l'IST.

Etat des lieux

En 2012, Marine Rigeade établit à 82 le nombre des archives institutionnelles répondant à la définition de l'INIST : 17 établissements disposaient alors de leur propre archive ouverte institutionnelle et 65 avaient créé un portail dans HAL.[4] L'étude recensait en outre 19 archives locales dédiées aux seules thèses et mémoires de doctorants et d'étudiants inscrits dans les établissements.

En 2016, l'enquête COPIST constatait que la moitié des 105 établissements y ayant répondu déclarait disposer d'une archive institutionnelle locale (9) ou d'un portail HAL (45).

En 2017, la synthèse de l'enquête menée par le consortium Couperin indique que 82% des 121 répondants ont une archive opérationnelle ou en cours de mise en œuvre, alors qu'ils n'étaient que 62% en 2014[5]. Sur l'ensemble des répondants, les portails HAL représentent 79% des archives et 84% pour les seules universités. Une vingtaine d'établissements a opté pour la réalisation d'une archive locale.

Quant au site de HAL, il dénombrait 120 portails d'établissements en décembre 2017.

Certains établissements, enfin, disposent à la fois d'une archive locale et d'un portail HAL : Paris-Dauphine, Sciences Po et l'IRD, par exemple.

In fine, il est donc difficile d'établir un bilan quantitatif fiable des archives institutionnelles des établissements de l'ESR. On peut néanmoins avancer sans risque que la part des archives locales avoisine les 20%. Des projets récents comme ceux des universités de Strasbourg, de Bordeaux et de Lille confirment la volonté d'établissements disposant des crédits *ad hoc* de développer leurs propres archives. Dans le même temps, la dynamique des portails HAL se poursuit avec la création de 15 nouveaux portails en 2017, majoritairement par des universités.

[3] http://openaccess.inist.fr/?+-Archive-institutionnelle-+
[4] RIGEADE, Marine. Les archives ouvertes institutionnelles en France : état des lieux et perspectives (2012)
[5] http://www.couperin.org/images/stories/AO/Depouillement_enquete_AO2017.pdf

Motivations pour la création d'une archive institutionnelle

Si la mise à disposition en accès libre et gratuit de la production scientifique de l'établissement est une visée commune à tous les projets d'archives institutionnelles, les distinguant en cela d'une plateforme de gestion électronique de documents (GED) ou d'un intranet, d'autres motivations sont également à l'œuvre comme :

- la réalisation d'une vitrine de la recherche et de la diversité des productions de l'établissement ;
- le recensement de la production de ses communautés ;
- le partage de connaissances ;
- l'analyse de la production scientifique et technique de l'établissement ;
- la préservation à terme des documents produits.

S'inscrire dans le mouvement de l'Open Access

Toutes les archives institutionnelles ont comme finalité et premier argument incitatif à l'égard de leurs chercheurs, la mise à disposition au plus grand nombre des résultats de la recherche scientifique et de l'enseignement (thèses et mémoires), comme en témoignent les exemples suivants issus de pages de présentation d'archives institutionnelles :

*« Signataire en 2003 de la «Déclaration de Berlin de sur le Libre Accès à la connaissance en Sciences Exactes, Sciences de la Vie et Sciences Humaines et Sociales», l'**Inserm** s'associe en 2006 à de nombreux établissements autour d'une archive ouverte nationale multidisciplinaire: HAL ».*

*« S'inscrivant dans le mouvement de l'Open Access, Spire met à la disposition de tous, sur Internet et gratuitement, le texte intégral des publications, dans le respect des dispositions relatives au droit d'auteur, contribuant ainsi à une meilleure diffusion et à la valorisation de la production scientifique de **Sciences Po**. »*

*« **L'Université de Rennes 1** encourage depuis 2012 ses EC et C à référencer, déposer et diffuser leurs publications scientifiques dans l'archive ouverte HAL-Rennes 1. La diffusion en libre accès de nos productions scientifiques est une plus-value importante pour la science, pour la société et pour la visibilité de nos recherches, tant au bénéfice des acteurs individuels, que des laboratoires et de l'établissement. ».*

Néanmoins, l'archive institutionnelle ne s'inscrit pas toujours dans une politique globale d'établissement. L'étude Couperin réalisée en 2017 met en évidence que seule une moitié des établissements ayant répondu à l'enquête a une politique d'établissement, le plus souvent peu contraignante quant aux dépôts dans des archives ouvertes.

Constituer une vitrine de la recherche de l'établissement

L'archive institutionnelle est une vitrine de la recherche de l'établissement et un vecteur de son rayonnement sur la scène internationale. A ce titre, elle est considérée par certains comme un outil stratégique dont la maîtrise est déterminante.

Par exemple, l'archive ouverte BIRD de l'université Paris Dauphine a certainement pesé dans l'obtention de l'accréditation EQUIS (labellisation des business schools), généralement délivrée aux seules écoles de commerce.

Faciliter le partage de la production de l'établissement entre les équipes

Parce qu'elle réunit les documents scientifiques, voire techniques ou à caractère pédagogique (mémoires) produits par la communauté de l'établissement, quels qu'en soient la forme, l'état d'avancement ou la communicabilité, l'archive ouverte facilite la capitalisation et le partage des connaissances entre les équipes de l'institution.

Rarement mise en avant, cette motivation apparait, par exemple, dans les missions assignées à l'archive institutionnelle de l'IFREMER.

Recenser la production de l'établissement

De plus en plus, et pour des raisons diverses, les établissements souhaitent disposer d'une vue d'ensemble des différents types de documents et de données produits par leur communauté, qu'ils soient ou non publiés ou destinés à l'être, réservés à un public restreint…

Outre des articles scientifiques, des communications à des conférences, l'archive institutionnelle peut également intégrer des ouvrages et chapitres d'ouvrages, des rapports de recherche, des mémoires et des travaux d'étudiants, des fiches et des rapports techniques, des documents de vulgarisation des vidéos, des pages web…

Le signalement est opéré pour partie par les professionnels de l'IST et vise *a minima* à identifier la production de l'établissement. Le signalement peut constituer une première étape vers le dépôt de documents par les chercheurs. La proportion de notices orphelines *versus* celles assorties d'un document en texte intégral (dont l'accès peut être restreint) est très variable, mais elle est supérieure à 50% pour la majorité des archives institutionnelles en production.[6]

Ce signalement bibliographique, avec ou sans texte intégral, peut notamment faciliter la production des rapports d'évaluation Hcéres en automatisant l'édition des listes de publications. Il permet également de nourrir les pages web des laboratoires et des chercheurs. Enfin, des institutions s'en servent pour produire des indicateurs.

Analyser la production scientifique et technique

Pour certains établissements, une archive institutionnelle bien alimentée constitue un outil au service de l'évaluation et du pilotage de la recherche. Support de calcul d'indicateurs bibliométriques, elle peut intervenir dans l'élaboration des rapports d'activité.

Par exemple, une analyse des co-auteurs et de leurs affiliations permet d'identifier les relations entre l'institution et la communauté scientifique internationale. L'archive de l'INRA, ProdInra, propose ainsi une cartographie des collaborations de ses équipes de recherche. Dans le même esprit, l'analyse des mots-clés peut faciliter l'identification d'experts (service *Trouver un expert* de la base Archimer de l'IFREMER).

[6] Enquête Couperin 2017

La qualité des référentiels, auteurs et affiliations, est un pré requis pour fiabiliser ces analyses.

Préserver les différents documents produits

La préservation à long terme des productions scientifiques et techniques dispersées dans les multiples outils de stockage des laboratoires, des chercheurs et des enseignants-chercheurs est une fonction de l'archive institutionnelle à laquelle leurs auteurs sont généralement sensibles. L'archivage pérenne proposé par HAL *via* le CINES est un réel atout souligné par les établissements.

Pour certains projets d'archives institutionnelles les plus récents, référencer voire archiver des données de la recherche, notamment celles afférentes aux publications, est une évolution envisagée.

Des contextes singuliers

Les motivations présentées ci-dessus interviennent à des degrés divers dans les projets d'archives institutionnelles des institutions interrogées. De même, les jugements portés sur l'offre de portail HAL et sur sa capacité à répondre aux besoins sont largement impactés par le contexte des établissements.

Dès lors, les arguments avancés et les demandes formulées lors des entretiens ont pu parfois apparaitre comme une juxtaposition de situations particulières, peu de demandes faisant consensus *in fine*.

Parmi les facteurs qui éclairent la nature des demandes, on peut citer les principaux points suivants.

La politique de dépôt de l'établissement

L'existence d'une politique d'établissement encadrant le dépôt dans l'archive ouverte et *a fortiori* d'un mandat le rendant obligatoire ne concerne qu'une partie des établissements interrogés (INRIA, ENSAM, université de Bordeaux...).

L'étude de Laurent Jonchère (2013) a montré l'efficacité de ces mandats pour accompagner l'auto archivage par le chercheur[7]. A l'inverse, pour les établissements ne disposant pas d'un mandat, les professionnels de l'IST doivent faire davantage de pédagogie et peuvent être amenés, plus volontiers, à développer des services de dépôt en lieu et place du chercheur.

L'interdiction des notices orphelines

L'ENSAM, par exemple, a fait le choix de proscrire le dépôt de références bibliographiques sans texte intégral.

Pour l'université de Montpellier, la demande de la recherche est de disposer d'une vue d'ensemble des publications et de faciliter la production des listes bibliographiques des rapports d'évaluation Hcéres. Le dépôt du texte intégral est avant tout une préoccupation du SCD.

[7] Laurent Jonchère. Synthèse sur les politiques institutionnelles de libre accès à la recherche.2013, pp.1-35.

Les disciplines de recherche au sein de l'établissement

Les disciplines scientifiques représentées au sein d'un établissement influent sur sa politique en matière d'Open Access[8] et sur la propension de ses chercheurs à déposer le texte intégral de leur production, sur les types de documents scientifiques produits, sur les formes souhaitées pour l'exploitation des listes bibliographiques...

Le périmètre documentaire visé

Certains établissements visent le recensement de toute la production documentaire de leurs chercheurs, enseignants-chercheurs, ingénieurs… comme à l'IRSTEA ou à l'IFREMER.

D'autres établissements, comme l'ENSAM, s'attachent en premier lieu à la mise en ligne du texte intégral des articles de recherche et des interventions lors des congrès.

Pour les établissements d'enseignement, le recensement des mémoires des étudiants est un enjeu.

L'utilisation de l'archive comme support de pilotage de la recherche

Si la bibliométrie n'est pas une fonction attendue pour l'archive de Toulouse 2, elle est déterminante ou importante pour d'autres établissements comme l'INRIA ou l'IRSTEA, par exemple.

Les moyens dévolus à l'archive ouverte

Les moyens budgétaires et humains susceptibles d'être consacrés au développement et à l'exploitation d'une archive institutionnelle varient nettement d'un établissement à l'autre, et influencent naturellement leurs ambitions.

Il est notable que les regroupements récents d'universités soutenus par des financements exceptionnels ont pu constituer un « effet d'aubaine » pour développer une archive institutionnelle locale.

La maturité du système d'information dans lequel s'inscrit l'archive institutionnelle

L'archive institutionnelle est couramment articulée avec le catalogue de la bibliothèque et le site web de l'institution.

Certains établissements ont ou envisagent des intégrations plus complètes :

- Avec leur annuaire (LDAP) et leurs applications RH internes pour authentifier les utilisateurs, gérer des droits sur les documents, récupérer de façon automatique et fiable des affiliations…
- Avec un CRIS (Current Research Information Systems, ou système d'information sur la recherche) existant ou à venir. L'archive est alors l'outil d'exposition des publications dont le cycle de vie est géré par le CRIS.
- Avec des applications internes, de bibliométrie par exemple.

[8] Si les mathématiciens et les physiciens constituent des communautés en faveur du développement de l'Open Access, l'intérêt des autres communautés est très variable, voire limité pour certaines, les juristes par exemple.

L'utilisation d'une plateforme locale, *versus* un portail HAL, est plus propice à ces objectifs d'intégration. En outre, la même infrastructure pourrait également permettre l'archivage des données de la recherche.

Les freins à la création d'une archive institutionnelle

L'absence de projet d'archive institutionnelle au sein d'un établissement peut traduire une sensibilité limitée de sa communauté à l'Open Access, voire une faible appétence à s'associer à un projet centralisateur, comme une insuffisance de moyens humains et financiers mobilisables.

Certains établissements ont ainsi créé une *Collection* dans HAL qui permet de regrouper les dépôts réalisés par leurs chercheurs dans HAL sans passer le cap de la création d'un *Portail* qui engage davantage l'institution dans le cadre d'une convention et rend nécessaire la mobilisation de ressources pour l'administrer[9].

Le développement de l'usage des réseaux sociaux scientifiques peut également venir satisfaire les besoins d'échange des chercheurs avec leur communauté de travail au détriment du développement d'une archive institutionnelle.

Un dispositif articulant les archives locales et l'infrastructure nationale HAL

Dès 2006, le « *Protocole d'accord en vue d'une approche coordonnée, au niveau national, pour l'archivage ouvert de la production scientifique* » signé par le CEMAGREF (devenu IRSTEA), le CIRAD, le CNRS, la CPU, l'INRA, l'INRIA, l'INSERM, l'Institut PASTEUR, l'IRD et la Conférence des Grandes Ecoles, acte que la plateforme HAL sera accessible soit par dépôt direct, par l'intermédiaire ou non d'une interface spécialisée, soit, le cas échéant par dépôt indirect à partir d'un système propre à l'établissement, permettant le transfert des données sur la plateforme commune en respectant le modèle des données de celle-ci.

En 2013, la « *Convention de partenariat en faveur des archives ouvertes et de la plateforme mutualisée HAL* » réaffirme les deux modalités possibles de mutualisation des archives ouvertes, l'utilisation d'un portail HAL ou le développement par le partenaire d'une archive institutionnelle locale qui duplique son contenu (texte intégral et signalement) dans l'archive HAL, sous réserve de se conformer aux conditions scientifiques et techniques d'éligibilité à la plateforme.

[9] HAL intègre 6 collections de Comue, 134 d'institutions, 45 d'universités et 55 d'écoles. Ces collections peuvent être également un regroupement destiné à l'export des références vers leur propre archive institutionnelle locale (HEC, Sciences Po, INRA…).

Deuxième partie : Les portails institutionnels dans HAL

Les portails HAL, qu'ils soient institutionnels ou thématiques, constituent un sous-ensemble de l'archive nationale pluridisciplinaire dont ils partagent les principes de fonctionnement, le réservoir, les outils et les référentiels.

Un portail dispose d'une entrée propre qui lui permet d'apparaître comme une archive à part entière et de possibilités de paramétrage pour réaliser quelques adaptations en relation avec le contexte de l'institution ou de la communauté thématique (type de documents spécifiques, métadonnées supplémentaires…).

Genèse et développement des portails institutionnels dans HAL

La genèse des portails institutionnels dans HAL

A son lancement en 2001, HAL est positionnée comme une plateforme de communication directe entre chercheurs. Elle élargit son offre dès 2003 en permettant à des communautés et à des institutions de valoriser les productions déposées en les réunissant dans des espaces dédiés : les *Collections* et les *Portails*.

Avant HAL, le CCSD avait déjà développé une archive dédiée aux thèses, Thèse-en-ligne (TEL), une archive disciplinaire, @rchiveSIC, et l'archive institutionnelle de l'Institut Jean Nicod, les deux dernières devenant rapidement des portails HAL.

Avec ces premières initiatives, le CCSD offre une réponse aux projets d'archives ouvertes institutionnelles et stimule la dynamique de dépôts en s'appuyant sur les initiatives individuelles des chercheurs, sur les projets d'établissement et sur les professionnels de l'IST qui en sont le plus souvent les gestionnaires.

Depuis, le CCSD fait évoluer HAL et son offre *Portail* avec la constante préoccupation de conserver un juste équilibre entre les évolutions de l'archive ouverte nationale nécessaires pour les chercheurs, et les besoins propres des institutions.

« *L'offre portail est un service spécifiquement dédié aux institutions et établissements souhaitant valoriser les publications et données scientifiques de leurs chercheurs et enseignants-chercheurs. L'offre portail est une solution technique économique, permettant à l'organisme de déployer ses ressources humaines dans l'animation autour de son archive et l'aide aux chercheurs.* » [10]

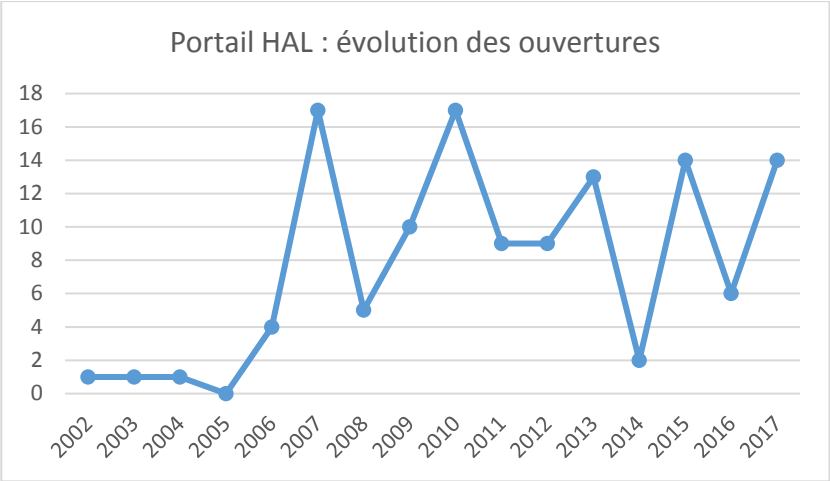
L'offre de portail HAL est donc présentée avant tout comme un outil de valorisation, économique pour les institutions.

Evolution des ouvertures de portails institutionnels dans HAL

Fin 2002, le portail de l'institut Jean Nicod suivi par celui de l'IN2P3 en 2003 marquent les débuts des archives institutionnelles dans HAL.

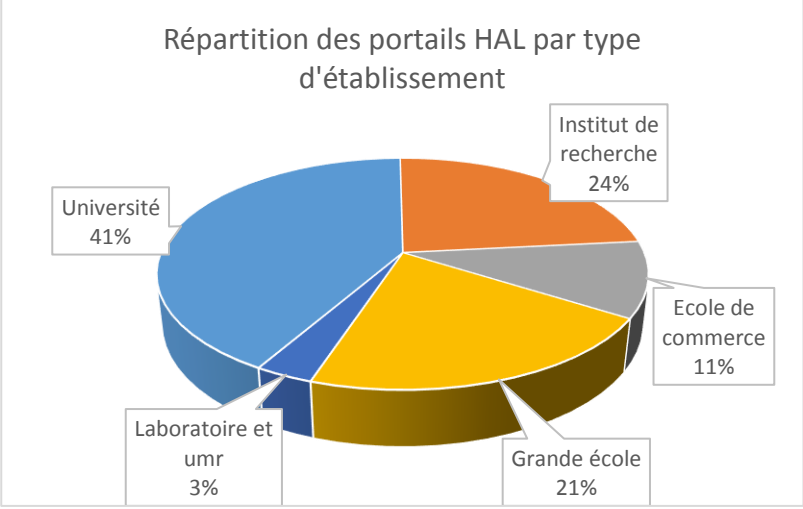
A ce jour (décembre 2017) HAL compte 141 portails dont 123 sont des archives institutionnelles, d'une façon plus ou moins marquée.

[10] https://www.ccsd.cnrs.fr/portails-institutionnels/



Après le lancement en 2004 du portail de l'INRIA, une première vague d'ouverture de portails HAL intervient suite à la signature du « *Protocole d'accord en vue d'une approche coordonnée, au niveau national, pour l'archivage ouvert de la production scientifique* » en juillet 2006. Il s'agit pour l'essentiel d'organismes de recherche. Les grandes écoles et les écoles de commerce ouvrent leur portail à partir de 2010-2011.

Les universités qui représentent 41 % des portails HAL aujourd'hui sont présentes dès 2007, mais leur part augmente nettement à partir 2010 pour constituer l'essentiel des ouvertures récentes.



La contribution des archives locales à HAL

Les institutions ayant fait le choix d'une archive locale sont articulées avec HAL pour importer et exporter des dépôts. En 2015, 7 archives locales reversaient leur contenu dans HAL. Ces dépôts représentaient 3% du texte intégral de HAL, Sciences Po étant le plus gros déposant.[11] Dans notre étude, les sept institutions interrogées ayant fait le choix d'une archive locale déposent ou ont intégré dans le projet en cours la duplication de leur contenu dans HAL.

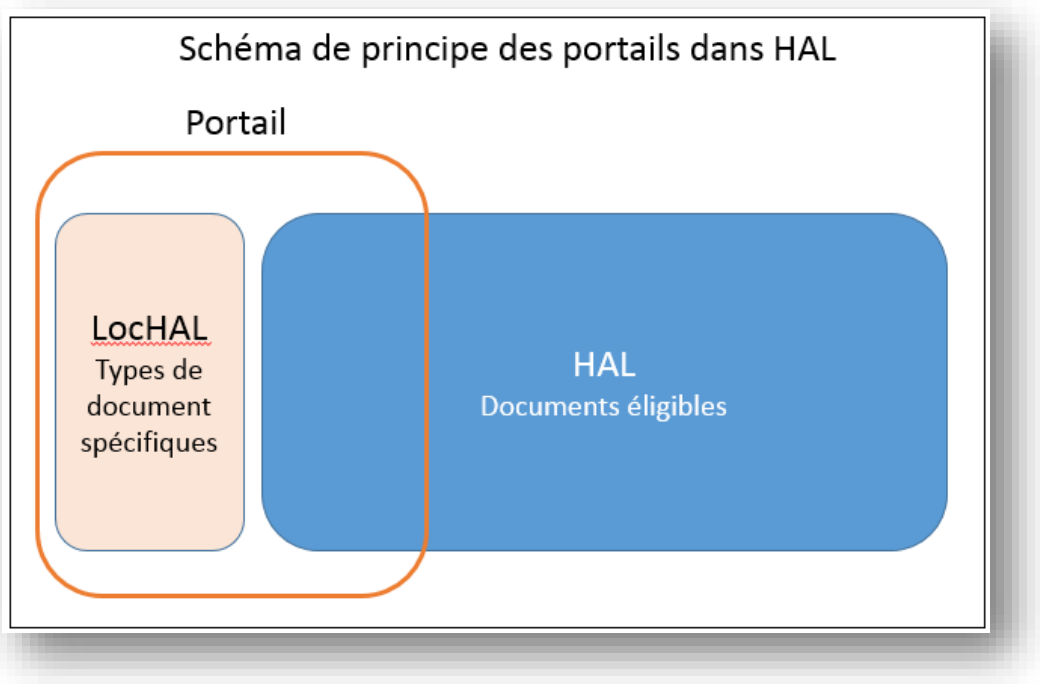
[11] https://www.ccsd.cnrs.fr/2015/11/2014-2015-retour-sur-deux-annees-riches-en-evolutions-pour-le-ccsd-et-hal/

Caractéristiques de l'offre portail

Principe

Le portail est une entrée particulière permettant de déposer dans HAL.

Selon le paramétrage retenu par l'institution, tout ou partie des dépôts opérés dans le portail sont vus aussi bien dans celui-ci que dans HAL. De même, les dépôts visibles dans le portail ont été déposés dans celui-ci et/ou résultent d'un regroupement de dépôts réalisés dans HAL.



Les types de documents éligibles dans HAL / dans les Portails

Les types de documents éligibles dans HAL sont les suivants : articles de revue, communications dans un congrès, posters, ouvrages (y compris éditions critiques et traductions), chapitres d'ouvrage, directions d'ouvrage et proceeding, brevets, prépublications et documents de travail, rapports (de recherche, techniques), thèses, HDR, cours, images, vidéos, sons, cartes, autres publications (billets de blog, comptes rendus d'ouvrage, notices de dictionnaire ou d'encyclopédie).

Les types de document spécifiques créés à la demande des établissements dans les Portails concernent à ce jour[12] :

- Les mémoires d'étudiants. Un peu plus de 500 ont été déposés dans des portails institutionnels, dont 60% pour l'INRIA. Mais le portail DUMAS (pour *Dépôt Universitaire de Mémoires Après Soutenance*)[13] qui est consacré à l'archivage ouvert de travaux d'étudiants de niveaux bac+4 et bac+5, validés par un jury, dans toutes les disciplines, en compte plus de 16 000, et le portail mémSIC[14] dédié à

[12] Liste et comptage établis en décembre 2017
[13] https://dumas.ccsd.cnrs.fr
[14] https://memsic.ccsd.cnrs.fr

l'auto-archivage des mémoires de master en Sciences de l'information et de la communication, un peu moins de 500.

- Les documents associés à des manifestations scientifiques (663)
- Les notes de lecture (122)
- Autres rapport, séminaire et workshop (60)
- Les notes de synthèse (18)
- Les comptes rendus de table ronde (2)

Fonctionnalités et services

Aujourd'hui, les principales fonctionnalités spécifiques aux portails HAL sont les suivantes :

- Un nom de domaine spécifique
- Une interface de dépôt dédiée (i.e. un point d'entrée particulier dans HAL)
- La personnalisation de la charte graphique
- L'extension possible du modèle de données (métadonnées supplémentaires)
- Le « raffinage » ou l'extension de la nomenclature scientifique
- L'ajout de typologies particulières de documents scientifiques et de mémoires. Ces types de documents doivent être déposés via le portail et la consultation en est limitée aux portails.
- Deux vues possibles au choix de l'institution depuis la page du portail :
  - La consultation est limitée aux seuls documents qui ont été déposés depuis ce portail.
  - Comme dans une collection, la consultation agrège les dépôts sur un critère d'appartenance (ex : au moins un auteur appartient à un laboratoire affilié à l'institution) que le dépôt ait été réalisé dans le portail, dans HAL ou sur un autre portail. Dans le vocabulaire HAL, ce dispositif est appelé un « tampon automatique ».
- Des outils de pilotage : listes de publication, statistiques de dépôt, de consultation et de téléchargement
- Un profil d'administrateur permettant d'intervenir sur les utilisateurs, les dépôts et les référentiels.

En septembre 2017, une version majeure de HAL a répondu en premier lieu à une attente de simplification du dépôt par le chercheur, par une amélioration de l'ergonomie et une limitation des métadonnées obligatoires à saisir, les affiliations notamment (une seule affiliation est dorénavant obligatoire). Deux autres évolutions intéressent plus particulièrement les institutions ayant un portail HAL :

- L'introduction de nouveaux types de documents éligibles dans HAL : billets de blog, comptes rendus de lecture, notices de dictionnaire et traductions.
- L'ajout dans AuréHAL, le référentiel de HAL, de deux niveaux supplémentaires de regroupement des structures pour décrire la hiérarchie administrative : regroupement de laboratoires et regroupement d'institutions (ex : COMUE).

Tarifs

Aujourd'hui, le coût d'accès à l'offre de portail consiste en un « droit d'entrée » forfaitaire de 1 200 € HT ou de 2 400 € HT si l'institution souhaite ajouter des types de documents spécifiques, non éligibles dans HAL.

Une nouvelle grille tarifaire reposant sur un abonnement annuel dont le montant varie en fonction du nombre de chercheurs de l'établissement devrait entrer en vigueur prochainement.

Ouvrir un portail dans HAL : un choix avant tout économique au regard du service délivré

La principale motivation du recours à un portail HAL par les institutions souhaitant se doter d'une archive institutionnelle réside dans son « rapport qualité / prix ».

Pour les **universités de La Réunion et de Toulouse 2**, l'option d'une archive locale n'a jamais été envisagée compte tenu de son coût de développement et d'exploitation.

En 2014, un groupe de travail de **l'université Aix-Marseille** souligne dans son rapport d'étude « *Vers une archive ouverte pour Aix-Marseille Université* »[15] des fonctionnalités jugées importantes que HAL ne propose pas. Il s'agissait :

- Du dépôt en accès restreint pour certains articles.
- De l'envoi de tirés à part pour les articles placés en accès restreint.
- De la gestion des notices orphelines.
- De statistiques de fréquentation attractives et lisibles, permettant au chercheur de prendre la mesure des usages de sa production.
- De l'édition de listes de publications des chercheurs proposant plusieurs formats de sortie afin d'être réutilisables dans différents contextes[16].

Par ailleurs, une analyse du coût complet (investissement et personnel) d'acquisition et de gestion d'une solution locale estimait le budget à 135 K€ en année 1 et à 85K€ en année 2, *versus* à 60 K€ / an avec HAL.

Aussi, au vu des améliorations apportées à HAL dans la version de 2014, du constat d'une démarche de prise en compte des besoins utilisateurs et de considérations économiques, le groupe de travail recommande d'opter pour HAL : « *Il nous semble stratégique d'adopter HAL dans un premier temps, en espérant que les fonctionnalités qui manquent actuellement soient développées à terme. Si ce n'était pas le cas fin 2016, il faudrait alors réfléchir à la mise en place d'une instance locale pour AMU en 2017-2018*. »

[15] Denis Bertin, Marin Dacos, Marlène Delhaye, Michaël Hug, Marie Masclet de Barbarin, et al. Vers une archive ouverte pour Aix-Marseille Université. Une démarche en faveur de l'Open Access : Conclusions du groupe de travail sur le référencement des articles scientifiques produits par AMU. [Rapport Technique] Aix Marseille Université. 2014. <hal-01226882>

[16] Depuis, l'envoi de tirés à part est possible dans HAL et des améliorations ont été apportées aux statistiques.

L'université de Lorraine a conduit en 2014 un benchmark des solutions pour son projet d'archive ouverte qui intégrait notamment DSpace, Okina[17] et HAL, d'où il ressortait que les deux dernières solutions couvraient 80% des besoins prioritaires de l'établissement. L'Université qui envisageait initialement la mise en place d'une archive locale tranche en 2015 en faveur de HAL.

Récemment, **l'INRA** a fait le choix d'un portail HAL en remplacement de son archive locale, ProdINRA, essentiellement pour des raisons économiques. A l'instar de l'INRIA, l'INRA devrait s'associer prochainement au CCSD en rejoignant l'UMS, et dédier une personne au développement et à l'amélioration des fonctionnalités de HAL.

L'IRSTEA qui a développé sa propre archive ouverte à une époque où HAL disposait d'une couverture fonctionnelle moindre et de performances jugées peu satisfaisantes, n'exclut pas de la migrer à terme dans HAL sous réserve, notamment, d'un élargissement du périmètre des productions archivables (les documents à caractère technique, principalement) et d'une évolution des règles relatives à l'embargo de l'accès à certains documents.

Les limites des portails HAL invoquées par les établissements

Avec des intensités variables, qu'ils aient ou non ouvert un portail dans HAL, les établissements souhaiteraient que le service *Portail* offre de nouvelles possibilités destinées à permettre une meilleure adaptation à leur contexte et à leurs ambitions.

Pour les établissements ayant fait le choix de développer leur propre archive locale, HAL ne permet pas une maîtrise totale de leur politique de dépôt et de valorisation de leurs productions, et ne leur offre pas les outils et services adaptés à leurs besoins locaux.

Les principales limites invoquées du service *Portail* sont brièvement présentées ici. Elles feront l'objet d'une analyse détaillée et de propositions de solutions dans la partie suivante.

Les types de documents acceptés

HAL est une archive ouverte des productions scientifiques répondant à des critères d'éligibilité précis. Les portails HAL peuvent intégrer des types de documents spécifiques, comme les mémoires par exemple, mais ils ne permettent pas à ce jour de déposer tous les types de documents techniques.

L'introduction de nouveaux types de document est soumise à l'arbitrage des instances de décision de HAL.

L'accès aux documents

Dans HAL, la période d'embargo pendant laquelle un document déposé n'est visible que par ses auteurs est limitée à 2 ans. Or, des institutions souhaiteraient pouvoir déposer des documents avec une période d'embargo illimitée. Cette demande peut être associée à la possibilité de créer un ou plusieurs espaces réservés au sein du portail destinés à tout ou partie des membres de l'établissement.

[17] Okina est un développement réalisé par l'Université d'Angers sur un socle Drupal.

Un workflow de dépôt et de validation adapté à l'organisation interne de l'établissement

HAL et les portails partagent un seul et même workflow de dépôt (définition des rôles et des étapes). En l'état, il n'est pas prévu d'en permettre un paramétrage spécifique pour un établissement.

Des référentiels de qualité insuffisante

Les établissements qui souhaitent réaliser des analyses bibliométriques à partir des dépôts faits dans HAL, jugent la qualité des référentiels des auteurs et des structures de HAL insuffisante.

Pour les auteurs et les structures des établissements, les archives locales utilisent leurs annuaires internes nécessairement plus fiables mais, hors de ce périmètre, un référentiel reste nécessaire.

Le dépôt de données de recherche en relation avec des publications

Ce besoin émerge du fait notamment des exigences du programme Horizon 2020 de la Commission Européenne.

Troisième partie : Attentes des établissements et solutions envisageables

Introduction

Les attentes décrites ci-après ont, pour une large part, été mentionnées dans des rapports antérieurs. Les plus prégnantes ont été inscrites dans le programme d'évolutions de HAL à l'issue de la première réunion de son Comité Scientifique et Technique en 2014[18], et reprises par Serge Bauin dans son rapport[19]. Elles ont également été discutées dans le cadre du club des utilisateurs de HAL, CasuHAL, créé en 2016.

Elles portaient principalement sur l'élargissement des types de documents éligibles dans HAL, de l'articulation de l'archive nationale avec les archives institutionnelles, de la question de l'embargo illimité, de la simplification et de l'ergonomie du dépôt.

L'objet de cette troisième partie du rapport est d'exposer et d'analyser les attentes exprimées par les institutions interrogées, dont celles précédemment citées font partie, et de proposer des solutions :

- La prise en compte de la production scientifique et technique de l'établissement dans son ensemble.
- L'amélioration de la qualité des référentiels.
- L'enrichissement de l'offre fonctionnelle proposée aux chercheurs pour valoriser leurs dépôts.
- Le développement d'outils pour les gestionnaires qui déposent pour les chercheurs, complètent et enrichissent les métadonnées.
- Les principes régissant la modération.
- La simplification du reversement des archives institutionnelles locales dans HAL.
- La gouvernance de HAL.

Pour éclairer le lecteur, les solutions proposées sont pondérées par des critères visant à en préciser l'importance pour les institutions interrogées et les impacts pour HAL et le CCSD[20].

[18] https://www.ccsd.cnrs.fr/2014/06/la-premiere-reunion-du-comite-scientifique-et-technique/
[19] www.cnrs.fr/dist/z-outils/documents/CNRS DIST Rapport Bauin sur CCSD et HAL septembre 2014.pdf
[20] Voir les critères en annexe 2

Prendre en compte la production scientifique et technique des établissements dans son ensemble

Elargir les types de documents éligibles dans HAL

Problème posé

Les gestionnaires de portail souhaitent répondre aux différentes attentes de dépôt des chercheurs, qu'il s'agisse de les convaincre de déposer ou d'éviter des refus qui alimentent les réticences de certaines communautés. Ils ont également besoin d'adapter la typologie des documents pour répondre aux exigences de production de rapports et d'indicateurs.

Aujourd'hui, l'institution peut ajouter à son portail des types de documents qui lui sont propres et en personnaliser les métadonnées. Ces types de documents doivent être déposés via les portails et ne sont accessibles pour la consultation qu'au sein des portails. Ils représentent un très faible volume comparé aux articles et aux conférences qui totalisent 75 % des dépôts. A ce jour, deux types de documents spécifiques dépassent les 500 dépôts dans les portails : les mémoires d'étudiant et les documents associés à des manifestations scientifiques.

Mais cette possibilité est peu utilisée car ces documents ne sont pas exposés dans HAL ce qui en limite la visibilité, qu'ils doivent être déposés dans le portail ce qui n'est pas nécessairement la pratique des chercheurs de l'établissement et, enfin, qu'il s'agit d'une prestation payante[21].

S'agissant de la visibilité des types de document spécifiques, on notera toutefois qu'ils apparaissent dans les recherches réalisées via Google et Google Scholar qui sont les outils les plus largement utilisés pour rechercher dans HAL.

L'introduction de nouveaux types de documents dans HAL est étudiée au cas par cas par son Conseil Scientifique et Technique (CST), et la liste s'enrichit progressivement. Ainsi, certains types de documents initialement non éligibles dans HAL sont désormais acceptés, comme les articles de vulgarisation scientifique et les posters, les cours ou les rapports techniques, par exemple.

Les institutions interrogées souhaiteraient que la liste des documents éligibles soit élargie notamment à des documents à caractère technique (IRSTEA, IFREMER) aux mémoires d'étudiant, mais aussi, pêle-mêle à des éditions critiques, des commentaires d'arrêt, des revues de livre, des maquettes d'architecture, des notes de fouille…

Pour le CCSD, ce point est sensible car il questionne le périmètre de l'archive HAL.

[21] Seuls les nouveaux types de document spécifiques font l'objet d'une facturation. Un nouveau portail peut bénéficier sans surcoût des types de document spécifiques déjà existants.

Solutions envisageables

1/ Adopter une procédure de proposition de types de documents, clarifier les modalités d'arbitrage et définir une périodicité de mise à jour de la liste des documents éligibles

Rejoignant la question de la gouvernance de HAL, cette simple mesure organisationnelle, courante pour maintenir les outils terminologiques partagés par une communauté, constitue un premier niveau de réponse pour les administrateurs de portail.

Intensité de la demande	Impact sur le modèle HAL	Impact sur les charges	Impact sur les charges d'exploitation
Forte	Faible	Faible	Faible

2/ S'accorder sur un élargissement des types de productions scientifiques éligibles

Plusieurs institutions ont recommandé de s'appuyer sur les travaux du COAR[22] qui définit et maintient différents référentiels, dont l'un est dédié aux types de productions scientifiques, pour définir la liste de celles éligibles dans HAL. Un chantier devrait être créé sur ce sujet dans le cadre de la feuille de route 2016-2020.

Chaque institution pourrait choisir dans son portail de limiter cette liste en fonction de sa politique de dépôt.

Intensité de la demande	Impact sur le modèle HAL	Impact sur les charges	Impact sur les charges d'exploitation
Forte	Faible	Moyen	Faible

3/ Rendre éligibles certaines catégories de documents techniques

Cette évolution, qui constitue un infléchissement fort du périmètre de HAL, est un pré requis important pour les institutions ayant une double mission de recherche fondamentale et finalisée. A ce jour, seuls les « rapports techniques » sont éligibles.

Intensité de la demande	Impact sur le modèle HAL	Impact sur les charges	Impact sur les charges d'exploitation
Faible (mais forte pour les établissements concernés)	Fort	Moyen	Faible

[22] https://www.coar-repositories.org/about/

Autoriser un embargo illimité pour certains documents

Problème posé

Dans HAL générique et dans les portails HAL, la période d'embargo pendant laquelle une publication peut ne pas être visible est limitée à deux ans.

La possibilité de déposer un document avec un embargo illimité, une fonction généralement intégrée dans les archives locales, répond aux besoins suivants :

- Permettre le dépôt, la gestion et l'archivage pérenne de documents éligibles mais soumis à des clauses de confidentialité ou à des contrats d'édition particuliers.
- Laisser au chercheur la latitude de décider si sa production peut être exposée en accès libre, sachant qu'il peut communiquer ou non le document sur demande via la fonction « *Demander une copie à l'auteur* ».[23]
- Proscrire les notices sans document même si le document n'a pas vocation à être exposé en accès libre.
- S'adapter plus facilement aux évolutions réglementaires en faveur de l'Open Access, les documents étant déjà présents dans l'archive.

Cette demande est exprimée par une partie des institutions interrogées. Toutefois, si le besoin se justifie peu désormais pour les articles, compte tenu de la limitation de la période d'embargo fixée par la Loi numérique, elle peut avoir du sens pour certains rapports ou d'autres types de documents.

La position de l'archive ouverte des universités de Liège, ORBI, est souvent citée en exemple sur ce point. Soutenue par un mandat imposant le dépôt du texte intégral pour tous les articles dès leur acceptation, ORBI est au premier rang mondial pour la complétude des articles dans son archive, 64% des dépôts comportant le texte intégral. Mais, en 2015, 46 % de ces documents étaient en accès restreint[24].

Solutions envisageables

L'embargo illimité est une concession importante au principe de l'Open Access, mais il n'est pas avéré qu'il représente un volume important de documents.

Pour en encadrer les impacts, l'embargo illimité pourrait être assorti des conditions suivantes :

1/ Limitation de la fonctionnalité aux seuls portails HAL.

2/ Modération par les administrateurs des portails des dépôts assortis d'un embargo illimité.

3/ Adressage périodique (tous les deux ans, par exemple) d'un mail à l'auteur et/ou à l'administrateur du portail demandant de confirmer le maintien en embargo illimité du document.

4/ Exclusion ou facturation de l'archivage pérenne *via* le CINES des documents en embargo illimité.

[23] Cette fonction est, de fait, très peu utilisée. HAL compte actuellement quelque 17 500 documents sous embargo : 559 demandes de copies ont été effectuées, dont 221 sont restées sans réponse.
[24] L'open science, une révolution au service de l'innovation par Paul Thirion | LIEGE CREATIVE, 05.03.15

Intensité de la demande	Impact sur le modèle HAL	Impact sur les charges	Impact sur les charges d'exploitation
Forte	Moyen	Faible	Potentiellement fort

Limiter l'accès à certains documents à tout ou partie des membres de l'institution

Problème posé

Il s'agit ici d'une demande de type « intranet » consistant à pouvoir archiver dans un portail des documents dont l'accès est limité à tout ou partie des membres de l'institution.

Deux niveaux d'accès sont envisagés : le document est visible par l'ensemble des membres de l'établissement ou par un sous-ensemble. Cette possibilité est offerte par certaines archives locales, mais ne l'est pas par le service Portail de HAL.

Il s'agit d'un besoin de faible intensité, mais qui devrait être satisfait pour les thèses. En effet, l'arrêté du 25 mai 2016 fixant le cadre national de la formation et les modalités conduisant à la délivrance du diplôme national de doctorat, définit les conditions de la diffusion des thèses : « *Sauf si la thèse présente un caractère de confidentialité avéré, sa diffusion est assurée dans l'établissement de soutenance et au sein de l'ensemble de la communauté universitaire. La diffusion en ligne de la thèse au-delà de ce périmètre est subordonnée à l'autorisation de son auteur, sous réserve de l'absence de clause de confidentialité.* »

L'Université de Lorraine, par exemple, a choisi d'intégrer dans son portail les 9 000 thèses de son établissement gérées jusqu'ici sur sa plateforme locale PETALE. Or près d'un tiers de celles-ci relèveraient d'une diffusion restreinte à la communauté universitaire.

Aujourd'hui, la réponse apportée par HAL est la suivante :

- Un type de document spécifique au portail, « *Thèse à diffusion intranet* », a été créé.
- La notice dans HAL indique un lien pour consulter la thèse sur le serveur local de l'établissement, dont l'accès est subordonné à une authentification et à une vérification des droits via la fédération d'identité Renater.

Cette solution est jugée peu satisfaisante puisqu'elle oblige l'établissement à conserver localement un système de gestion des thèses.

Solutions envisageables

La gestion de documents en accès restreint suppose qu'HAL soit en mesure de vérifier les droits du demandeur qui peuvent varier selon l'institution à laquelle il appartient.

1/ Le CCSD étudie la possibilité d'associer au compte utilisateur IdHal son identification dans la fédération d'identité RENATER.

Cette mesure permettrait d'éviter à l'utilisateur déjà connecté dans son système local de saisir à nouveau un identifiant et un mot de passe pour se connecter à HAL (fonctionnalité SSO Single Sign-On). La création du compte IdHal et son alignement sur

son identité RENATER suppose néanmoins que l'auteur a déposé au moins une publication dans HAL.

2/ Associer des droits spécifiques d'accès au document en fonction de l'établissement de rattachement de l'utilisateur.

Pour les thèses, il s'agirait de permettre le dépôt du texte en embargo illimité et d'en permettre l'accès aux seuls membres de l'établissement.

La même logique permettrait de gérer une zone « intranet » au sein du portail de l'établissement.

Intensité de la demande	Impact sur le modèle HAL	Impact sur les charges	Impact sur les charges d'exploitation
Faible	Fort	Fort	Moyen

Pouvoir intégrer des données de la recherche

Problème posé

Le signalement et l'archivage de données de la recherche, en particulier celles liées à des documents déposés, est un enjeu mis en avant par les institutions ayant récemment opté pour une archive locale, mais aussi par certains administrateurs de portails HAL existants.

La possibilité d'attacher un fichier annexe à un dépôt HAL ou de déposer un média visuel ou sonore dans MediHAL ne suffit pas à couvrir le besoin. *A minima*, le modèle de métadonnées de HAL devrait permettre d'établir un lien avec un jeu de données signalé et archivé dans une plateforme annexe interne ou externe au CCSD.

Dans cet esprit, l'INRIA et le CCSD collaborent autour d'une archive ouverte des logiciels (https://www.softwareheritage.org) d'une part, en permettant le signalement dans HAL des logiciels et, d'autre part, en encourageant l'alimentation de l'archive par un transfert automatique des logiciels déposé dans HAL vers SoftwareHeritage.

Solution envisageable

Le sujet est déjà pris en compte dans la feuille de route du CCSD 2016-2020.

Il fait l'objet en outre d'une étude COPIST consacrée aux enjeux, conditions et faisabilité d'une plateforme mutualisée pluridisciplinaire de stockage, de gestion, de signalement et de partage de données de recherche.

Intensité de la demande	Impact sur le modèle HAL	Impact sur les charges	Impact sur les charges d'exploitation
Forte	Faible	Fort	Fort

Améliorer la qualité des référentiels

Problème posé

Pour les institutions interrogées, les référentiels auteurs et structures de HAL constituent un obstacle à l'exploitation des dépôts et à la production de statistiques fiables pour la réalisation d'études bibliométriques.

Par exemple, une demande de l'IDEX « Université Grenoble Alpes » à l'administratrice de son portail HAL d'identifier les collaborations internationales des chercheurs de l'université n'a pu être satisfaite compte tenu de la faible qualité des affiliations des laboratoires étrangers.

Dans HAL, un déposant peut créer un auteur et une affiliation s'il ne trouve pas les entrées qui sont nécessaires à son dépôt. Comment le rappelle la feuille de route 2016-2020 du CCSD, « *faciliter le dépôt est une priorité, la qualité des métadonnées passant par d'autres flux*. ».

En parallèle, les gestionnaires de portail peuvent intervenir sur les référentiels de leur structure pour créer, modifier et fusionner les structures. Les structures qui ont été vérifiées apparaissent en vert dans les référentiels de HAL et les déposants doivent les privilégier[25].

Tous les établissements n'ayant pas un portail ou une collection dans HAL, ce travail de correction est donc partiel et, qui plus est, doit être régulièrement renouvelé pour éliminer les doublons qui surgissent à l'occasion de nouveaux dépôts des chercheurs ou lors des versements automatiques via l'API SWORD.

Enfin, le dépôt rapide récemment mis en place qui affranchit le déposant du renseignement de toutes les affiliations est une source d'inquiétude au regard de la charge incombant aux gestionnaires de portail pour compléter les métadonnées.

En l'absence de référentiels internationaux auteurs et structures uniques et exhaustifs, certaines institutions interrogées suggèrent des améliorations possibles :

- Pour le référentiel des structures françaises, il devrait être possible de fiabiliser la liste en s'appuyant sur le Répertoire National des Structures de Recherche (RNSR). Néanmoins, le répertoire n'intègre pas nécessairement tous les laboratoires.
- Le réseau des professionnels de l'IST qui contrôlent et corrigent les structures dans HAL pourrait être étendu à leurs collègues d'établissements n'ayant pas de portail qui pourraient alors réaliser les mêmes tâches pour leurs structures.
- Le développement d'algorithmes devrait permettre le nettoyage automatique ou semi-automatique des doublons les plus évidents.
- Le référentiel partagé IdRef produit par l'ABES, enrichi et géré par les seuls professionnels de l'IST, et qui a été retenu comme référentiel pour le projet Conditor, constitue la voie à privilégier en général et pour HAL.

[25] Exemple : une recherche dans le référentiel AuréHAL sur le laboratoire LEDA remonte une structure validée (vert) et 15 versions non validées (rose)

Solutions envisageables

Aujourd'hui, l'amélioration des référentiels auteurs et structures est un des axes de la feuille de route 2016-2020 du CCSD, et fait d'ores et déjà l'objet de chantiers techniques :

- **Le référentiel auteur de HAL est désormais aligné avec les référentiels ORCID et IdRef.** Le travail commun engagé entre l'ABES et le CCSD sur l'articulation des référentiels de HAL avec le projet IdRef doit être poursuivi.
- **Une API affiliation a été mise en service pour faciliter et fiabiliser le dépôt par le chercheur.** Elle repose sur un algorithme développé par l'INRIA et déjà utilisé par les institutions (Bib2HAL / X2HAL). L'algorithme opère des rapprochements entre les données auteur, affiliation et date présentes dans le document déposé, et les affiliations existantes dans HAL. Un projet commun INRIA/CCSD est prévu pour améliorer cet algorithme et en ouvrir le code.

En outre, trois améliorations pourraient être étudiées :

1/ Mettre en place une liste fermée des affiliations françaises reposant sur le RNSR

Aujourd'hui ce projet se heurte aux limites du socle technologique du Répertoire national des structures de recherche (RNSR) qui ne permet pas une récupération automatique des données via une API.

Lorsque ce frein sera levé, il sera possible de limiter les affiliations à celles présentes dans la liste ou *a minima* d'aligner les structures HAL avec celles présentes dans le RNSR.

Intensité de la demande	Impact sur le modèle HAL	Impact sur les charges	Impact sur les charges d'exploitation
Forte	Moyen	Moyen	Faible

2/ Permettre aux institutions ne disposant pas d'un portail HAL de procéder au contrôle et à la correction de leur structure dans les dépôts

Ce travail pourrait être réalisé, par exemple, par les gestionnaires d'archives locales dont le contenu est reversé dans HAL. Pour le CCSD, cette évolution constituera une charge supplémentaire de formation.

Intensité de la demande	Impact sur le modèle HAL	Impact sur les charges	Impact sur les charges d'exploitation
Faible	Faible	Faible	Moyenne

3/ Améliorer la qualité des référentiels en développant des outils de back-office pour nettoyer les structures étrangères

Des outils destinés à éliminer les aberrations et les doublons relatifs aux structures étrangères pourraient être mis à la disposition des administrateurs de portail. Néanmoins, il est vraisemblable qu'une partie de la charge reposera sur les équipes du CCSD.

Intensité de la demande	Impact sur le modèle HAL	Impact sur les charges	Impact sur les charges d'exploitation
Faible	Faible	Moyen	Moyen

Enrichir l'offre fonctionnelle proposée aux chercheurs pour valoriser leurs dépôts

Problème posé

Pointant les limites de la pédagogie pour convaincre les chercheurs, les administrateurs attendent des services « simples » et adaptés aux besoins des disciplines pour stimuler les dépôts dans HAL.

HAL offre d'ores et déjà différents services destinés à favoriser l'exploitation de ses dépôts par le chercheur :

- Le CV du chercheur qui permet d'afficher dans une page et de façon dynamique la liste de ses publications.
- Des statistiques sur le nombre de consultations et de téléchargements des documents déposés dans HAL, et les activités liées sur les réseaux sociaux si le dépôt est identifié par un DOI.
- Un plugin permettant l'affichage de ses publications sur un site web développé avec WordPress.
- Différents formats d'export des publications.

En l'état, les services de création de pages web et d'export de listes de publications sont jugés par les administrateurs de portail interrogés, difficilement utilisables sans compétence technique et trop génériques pour satisfaire les différentes communautés.

Par ailleurs, des outils périphériques à HAL ont été développés par des établissements. Leur ergonomie a été pensée pour être directement utilisables par les chercheurs sans recours à des informaticiens.

L'outil le plus souvent cité est ExtrHAL[26], développé par l'université Rennes 1, qui permet d'extraire et de mettre en forme des listes de publications. ExtrHAL offre un large potentiel de paramétrage des options d'édition et de tri des publications pour générer des listes conformes aux besoins des disciplines et aux règles définies pour les rapports Hcéres.

L'INRIA et l'université de Grenoble Alpes ont également développé des connecteurs pour faciliter l'édition des listes dans les pages chercheurs.

Le CCSD assure la promotion de ces outils, mais leur maintenance, la formation et le support restent à la charge des établissements qui les ont développés. Leur intégration dans l'environnement de HAL est une demande forte des personnes interrogées que le CCSD n'a pas pu réaliser faute de moyens suffisants.

[26] https://halur1.univ-rennes1.fr/ExtractionHAL.php

Au-delà, quelques personnes militent pour une ergonomie et une offre de services destinés aux chercheurs susceptibles de rivaliser avec ceux des réseaux sociaux comme Researchgate ou Academia.

Solutions envisageables

Les outils destinés à améliorer l'utilisation de HAL développés par certains établissements sont l'expression de besoins que l'équipe du CCSD, réduites, n'a pas pu satisfaire jusqu'à présent. Le principe d'une telle collaboration entre utilisateurs de HAL et le CCSD est positif mais, dès lors que ces services sont réputés utiles au-delà des institutions qui les ont réalisés, il est nécessaire de les intégrer pleinement dans l'environnement technique de HAL pour favoriser leur usage par le plus grand nombre.

Les solutions envisageables sont les suivantes :

1/ Intégrer un outil de paramétrage des éditions sur le modèle d'ExtrHAL.

L'outil pourra offrir des modèles types étudiés pour les besoins des disciplines, et des fonctions de paramétrage avancées.

Intensité de la demande	Impact sur le modèle HAL	Impact sur les charges	Impact sur les charges d'exploitation
Forte	Faible	Moyen	Faible

2/ Définir une procédure d'évaluation des outils développés par des établissements (usage et coût de développement pour une reprise des fonctionnalités dans HAL) et d'éligibilité à l'intégration dans la feuille de route de HAL.

Intensité de la demande	Impact sur le modèle HAL	Impact sur les charges	Impact sur les charges d'exploitation
Moyenne	Faible	Moyen	Moyen

3/ Réunir les conditions techniques pour intégrer facilement les développements réalisés par les établissements utilisateurs de HAL.

Pour cela, il est probablement nécessaire de revoir le socle technique de HAL pour permettre le développement et l'activation de modules périphériques (architecture, élaboration de standards et de recommandations à l'usage des développeurs) et de définir un processus de validation technique des développements réalisés.

Intensité de la demande	Impact sur le modèle HAL	Impact sur les charges	Impact sur les charges d'exploitation
Faible	Faible	Fort	Fort

Ces deux dernières recommandations valent également pour le chapitre suivant.

Développer des outils pour les gestionnaires qui déposent pour les chercheurs, complètent et enrichissent les métadonnées

Problème posé

Les administrateurs des portails HAL, dans des proportions variables selon le modèle organisationnel retenu, peuvent être chargés par leurs établissements de s'assurer de la complétude des signalements, de la qualité des référentiels et des dépôts, et de produire les analyses statistiques attendues par les directions de la recherche.

Or les outils de back-office sont souvent jugés insuffisants pour assurer une administration efficace du portail. Les besoins exprimés concernent en particulier :

- Un outil de comparaison des références présentes dans WOS et SCOPUS et dans HAL pour identifier les publications à déposer.
- La préparation des dépôts de notices en local à partir de fichiers de publication, et leur versement par lots dans HAL via l'API SWORD.
- Un outil permettant de faciliter la correction des métadonnées avec une présentation en mode tableau.
- Une bibliothèque de requêtes pour produire les statistiques sur les dépôts serait nécessaire. Il devrait aussi être possible de produire des statistiques sur un ensemble supérieur à 25 000 dépôts.
- La création de workflows permettant d'aménager le circuit de dépôt pour tenir compte des circuits mis en place entre les administrateurs de l'archive, les professionnels de l'IST et les chercheurs. Ce manque de flexibilité est un des arguments avancés par les institutions ayant fait le choix d'une archive locale.
- La création de rôles supplémentaires permettant de répartir la charge de dédoublonnage sur certains référents structures allègerait la charge des administrateurs.

Des institutions ont développé des outils externes pour répondre à ces besoins dont elles souhaiteraient qu'ils soient intégrés dans HAL. On citera notamment Bib2HAL[27] réalisé par l'INRIA, qui est utilisé pour préparer des versements dans HAL par lot.

Solutions envisageables

Outre la définition d'une procédure d'évaluation des outils développés par des établissements et la mise en œuvre de conditions techniques pour les intégrer facilement dans HAL, mentionnés dans le chapitre précédent, les solutions complémentaires envisageables sont les suivantes.

1/ Intégrer un outil de préparation des versements par lot sur le modèle de Bib2HAL.

Cet outil semble largement utilisé, *a minima* lors des phases de mise en place des portails.

Intensité de la demande	Impact sur le modèle HAL	Impact sur les charges	Impact sur les charges d'exploitation
Moyenne	Faible	Moyen	Faible

[27] https://doc.archives-ouvertes.fr/bib2hal/

2/ Créer une bibliothèque de workflows permettant de s'adapter aux pratiques locales.

Intensité de la demande	Impact sur le modèle HAL	Impact sur les charges	Impact sur les charges d'exploitation
Faible	Fort	Moyen	Faible

3/ Créer un rôle supplémentaire autorisant le dédoublonnage des dépôts.

Cette évolution est déjà identifiée par le CCSD et sera proposée dans l'offre *Portail* en 2018. Des charges supplémentaires de formation sont à prévoir.

Intensité de la demande	Impact sur le modèle HAL	Impact sur les charges	Impact sur les charges d'exploitation
Faible	Faible	Faible	Moyen

4/ Poursuivre le développement du module statistique.

La refonte du module statistique constitue le principal chantier technique en cours au CCSD.

Intensité de la demande	Impact sur le modèle HAL	Impact sur les charges	Impact sur les charges d'exploitation
Fort	Faible	Moyen	Moyen

Réexaminer les principes régissant la modération

Problème posé

La modération est l'examen de la conformité des dépôts au regard des critères de HAL avant leur mise en ligne. On distingue deux types de modération :

- La modération technique qui consiste à vérifier, voire corriger et compléter des métadonnées produites lors du dépôt.
- La modération éditoriale qui consiste à vérifier l'éligibilité du type du document déposé et sa version.

La mise en place d'une équipe de modération à l'INIST venant en appui du CCSD depuis environ un an, a permis de réduire les délais de mise en ligne des dépôts, et les corrections mineures qui relèvent de la modération technique sont désormais réalisées par les modérateurs, diminuant ainsi la part des demandes de modification. Certains administrateurs de portail estiment que cette évolution, assortie d'une « *plus grande souplesse* » des modérateurs dans la validation éditoriale des dépôts, a réglé en partie le problème du trop grand nombre de rejets.

Toutefois, les règles régissant la modération éditoriale de HAL sont jugées insuffisamment partagées et explicites.

Le contrôle opéré sur la légalité du dépôt peut être parfois en porte-à-faux avec les pratiques actuelles de l'édition scientifique. Ainsi, les auteurs peuvent être amenés à rédiger leurs publications directement sur les plateformes des éditeurs et ne disposent pas de la version auteur, sans mise en forme de l'éditeur, acceptée par HAL. L'auteur doit alors « dégrader » la version qui lui est fournie par l'éditeur pour pouvoir la déposer. Néanmoins, au regard de la Loi numérique, il n'apparait pas envisageable de procéder autrement aujourd'hui.

Le CCSD est ouvert au principe d'une modération par l'institution elle-même sous réserve qu'elle soit en mesure de respecter les engagements de délais d'examen des dépôts. Ainsi, l'INRIA assure la modération des dépôts réalisés par ses chercheurs et une expérimentation est en cours à l'université d'Aix-Marseille. Mais la plupart des administrateurs de portail interrogés ne peuvent prendre en charge la modération faute de ressources suffisantes.

Il est à noter que certaines archives locales, ORBI (universités de Liège) par exemple, délègue au chercheur l'entière responsabilité de son dépôt. La publication du dépôt est alors immédiate, seule une modération technique (contrôle des métadonnées et des formats de fichier) est assurée par sondage *a posteriori*. Cette option est jugée intéressante par certains interlocuteurs.

Solutions envisageables

La modération technique relève de la responsabilité de HAL car elle conditionne la qualité des métadonnées.

La modération éditoriale doit garantir que les dépôts répondent aux contraintes légales et aux critères d'éligibilité des documents[28].

S'agissant de la modération éditoriale, sauf à opter pour un désengagement total de HAL au profit de la responsabilité du chercheur sur le modèle d'ORBI, la réponse aux attentes des institutions passe principalement par des mesures organisationnelles.

1/ Redéfinir collectivement le fonctionnement de la modération.

- Etablir un suivi statistique des demandes de modification : part des dépôts faisant l'objet d'une demande de modification, part de ces dépôts qui ne seront pas déposés à nouveau par les chercheurs, répartition des demandes de modification par motif...
- Expliciter et partager les règles de modération
- Définir des conditions d'instruction et d'arbitrage s'agissant de l'éligibilité des documents déposés

Intensité de la demande	Impact sur le modèle HAL	Impact sur les charges	Impact sur les charges d'exploitation
Forte	Faible	Faible	Faible

[28] Par exemple, un résumé d'ouvrage ne peut pas être accepté lors d'un dépôt d'ouvrage comme le souhaiteraient parfois des chercheurs pour assurer la promotion de leur publication.

2/ Confier aux administrateurs de portail l'instruction des rejets et les arbitrages *in fine*.

Le risque de voir diverger la définition de l'éligibilité des documents en fonction des établissements devra être pris en compte.

Intensité de la demande	Impact sur le modèle HAL	Impact sur les charges	Impact sur les charges d'exploitation
Faible	Moyen	Faible	Faible

Simplifier le reversement dans HAL des archives institutionnelles locales

Problème posé

Les établissements signataires de la convention HAL de 2013 qui ont fait le choix de créer une archive locale, sont engagés à reverser son contenu, métadonnées et texte intégral, dans HAL. Mais ils souhaiteraient que le reversement soit réalisé par moissonnage de leur archive. *A minima*, ils attendent que les conditions du reversement par l'API spécifique SWORD[29] utilisé par HAL soient améliorées.

Les archives locales sont le plus souvent moissonnables, le protocole OAI/PMH étant largement utilisé par les archives institutionnelles pour importer et exporter les dépôts.

Mais pour le CCSD, moissonner les archives locales, c'est prendre le risque de devoir assurer le tri entre les bons et les mauvais dépôts et d'avoir à redresser des métadonnées pour les rendre conformes au standard HAL, le Dublin Core 15 champs utilisé pour le protocole OAI/PMH étant trop limité, en l'état, pour porter le schéma de métadonnées de HAL. Le CCSD est rejoint sur ce point par des administrateurs de portail HAL des universités qui jugent que le moissonnage viendrait fragiliser les référentiels et les métadonnées.

Par ailleurs, le reversement via l'API SWORD est largement critiqué par les institutions interrogées :

- La mise en place et la maintenance du connecteur SWORD suppose l'intervention d'une compétence informatique, externe le plus souvent à l'institution.
- L'API s'appuie sur des standards, mais intègre une part de développement spécifique qui complique son utilisation.
- Les institutions ne sont pas informées suffisamment en amont et dans le détail des évolutions de l'API pour faire réaliser les développements nécessaires localement. Dès lors, le reversement dans HAL peut être interrompu.

[29] Une API spécifique SWORD est utilisée pour verser les métadonnées et les documents dans HAL. Le protocole SWORD définit un ensemble de services Web basé sur Atom Publishing Protocol (APP) et le format des métadonnées est basé sur le format TEI.

Solutions envisageables

Deux solutions peuvent être envisagées :

1/ Créer un profil d'application du Dublin Core dédié à HAL pour le moissonnage en OAI-PMH

Ce profil préciserait les conventions d'écriture et les métadonnées supplémentaires attendues. Une étude complémentaire devra valider qu'il pourrait bien intégrer l'ensemble des métadonnées attendues par HAL et éviterait au CCSD d'avoir à redresser des métadonnées.

Les établissements souhaitant être moissonnés seraient alors engagés par convention à exposer le contenu de leur archive dans le format attendu. Si des post-traitements restaient nécessaires, le CCSD pourrait demander aux institutions de les réaliser.

Intensité de la demande	Impact sur le modèle HAL	Impact sur les charges	Impact sur les charges d'exploitation
Forte pour les institutions ayant une archive locale	Faible	Moyenne	Faible (à valider)

2/ Améliorer la communication sur les évolutions de l'API SWORD auprès des institutions versantes.

Il s'agit d'informer chaque institution sur le planning et le contenu des évolutions, et de produire une documentation adaptée des évolutions pour les prestataires en charge des évolutions.

Intensité de la demande	Impact sur le modèle HAL	Impact sur les charges	Impact sur les charges d'exploitation
Forte pour les institutions ayant une archive locale	Faible	Faible	Faible

Des évolutions du socle technique de HAL à moyen terme

Problème posé

Plateforme interopérable, HAL permet aux institutions qui le souhaitent de récupérer leurs dépôts pour les exploiter dans leur propre système d'information (site web, catalogue des bibliothèques, logiciel de bibliométrie), mais aussi de verser dans HAL des documents archivés dans des applications locales comme un CRIS, par exemple.

Les développements nécessaires à une intégration de HAL dans les systèmes d'information locaux des établissements sont à leur charge et ne bénéficient pas à l'ensemble de la communauté des utilisateurs.

Certains établissements ayant développé leur propre archive locale, ont suggéré d'étudier les évolutions techniques du socle HAL qui permettraient d'offrir une interaction plus forte entre HAL et les SI locaux et de mieux mutualiser les développements.

Solution envisageable

Faire évoluer HAL vers une solution open source

HAL repose sur un logiciel développé par le CCSD qui en assure seul la maintenance et l'évolution. Comme DSpace, par exemple, orienter HAL vers une solution open source répondrait à un double objectif :

- Favoriser la mutualisation des développements réalisés localement et notamment des connecteurs avec les logiciels les plus fréquemment utilisés par les établissements (CMS pour les sites web, CRIS…)
- Permettre aux institutions souhaitant disposer de leurs propres archives ouvertes d'installer des instances locales de HAL qui présenteraient l'avantage d'une solution nativement compatible avec l'archive nationale.

Ce projet, coûteux en développement et qui nécessitera des charges d'animation de la communauté Open source est au programme 2016-2020 de HAL, sans qu'à ce stade les orientations de mise en œuvre aient été arrêtées.

Intensité de la demande	Impact sur le modèle HAL	Impact sur les charges	Impact sur les charges d'exploitation
Faible	Faible	Fort	Moyen

Poursuivre, en la développant, la démarche d'écoute et de prise en compte des demandes des institutions utilisatrices de HAL

En dernier lieu, plusieurs des personnes interrogées dans le cadre de l'étude, en particulier des membres du club des utilisateurs de HAL, CasuHAL, ont insisté sur la nécessité de poursuivre, en la développant, la démarche de transparence, d'écoute et de prise en compte des demandes des institutions utilisatrices de HAL par le comité scientifique et technique (CST).

A ce titre, il convient de rappeler les points suivants qui s'inscrivent dans cette perspective :

- « *Le développement de l'animation des communautés utilisatrices de HAL* » est l'objectif n°1 de l'axe communication de la feuille de route 2016-2020 du CCSD[30].
- Le club des utilisateurs de HAL, CasuHAL, qui s'est réuni pour la première fois le 20 septembre 2016, réunit des administrateurs de portails et de collections, et des représentants des institutions qui déposent dans HAL.

[30] Page 5 de la feuille de route : https://www.ccsd.cnrs.fr/wp-content/uploads/2016/06/Feuille_de_route_CCSD_2016-2020.pdf

- Cinq groupes de travail ont été constitués sur les évolutions à proposer : politique, référentiels, imports, services aux chercheurs et services aux administrateurs. Le CCSD a mis un wiki à leur disposition pour faciliter le partage des bonnes pratiques.
- Une réunion mensuelle est organisée entre le CCSD et les membres de CasuHal.

Quatrième partie : Synthèse des évolutions attendues et de leurs impacts

Cette dernière partie présente deux synthèses des attentes exprimées avec plus ou moins d'intensité par les institutions interrogées.

Un chapitre conclusif interroge les capacités humaines et budgétaires de l'actuel CCSD pour répondre dans des délais raisonnables à tout ou partie de ces attentes.

Les évolutions les plus attendues

Les évolutions répertoriées dans le tableau ci-dessous nous sont apparues comme les plus largement attendues par les institutions interrogées dans le cadre de l'étude. Certaines figurent, en totalité ou en partie, dans la feuille de route 20106-20020 du CCSD.

Leur satisfaction devrait favoriser la création de nouveaux portails institutionnels par des établissements qui n'en disposent pas encore, d'en conduire certains disposant d'une archive locale à l'abandonner au profit d'un portail dans HAL, et d'améliorer l'articulation entre les archives institutionnelles locales et l'archive nationale.

Trois types d'impact majeur lié à leur mise en œuvre ont été distingués :

- Politique. L'évolution impose de revoir des principes structurants de HAL et de ses services.
- Organisationnel. L'évolution peut nécessiter de redéfinir des relations entre le CCSD et les institutions ou d'en créer de nouvelles. Elle peut également supposer la mobilisation de ressources humaines pour formaliser des processus de fonctionnement et animer la communauté des utilisateurs.
- Technique. L'évolution est subordonnée à des développements plus ou moins importants.

Attentes	Solutions envisagées	Impact majeur
Elargir les types de documents éligibles dans HAL	Adopter une procédure de proposition de types de documents, clarifier les modalités d'arbitrage et définir une périodicité de mise à jour des documents éligibles S'accorder sur un élargissement des types de productions éligibles	Politique Technique
Autoriser un embargo illimité pour certains documents	Limitation de la fonctionnalité aux seuls portails HAL. Modération par les administrateurs des portails des dépôts assortis d'un embargo illimité. Adressage périodique (tous les deux ans, par exemple) d'un mail à l'auteur et/ou à l'administrateur du portail demandant de confirmer le maintien en embargo illimité du document. Exclusion ou facturation de l'archivage pérenne via le CINES des documents en embargo illimité.	Politique Organisationnel

Intégrer des données de la recherche liées à des documents déposés	Evolution intégrée à la feuille de route 2016-2020 du CCSD	Technique
Améliorer la qualité des référentiels	Mettre en place une liste fermée des affiliations françaises reposant sur le RNSR ou a minima aligner les référentiels	Technique
Enrichir l'offre fonctionnelle proposée aux chercheurs pour valoriser leurs dépôts	Intégrer un outil de paramétrage des éditions sur le modèle d'ExtrHAL	Technique
Développer des outils pour les gestionnaires qui déposent pour les chercheurs, complètent et enrichissent les métadonnées	Poursuivre le développement du module statistique	Technique
Réexaminer les principes régissant la modération	Redéfinir collectivement le fonctionnement de la modération	Politique
Simplifier le reversement des archives institutionnelles locales dans HAL	Créer un profil d'application du Dublin Core dédié à HAL pour le moissonnage en OAI-PMH des archives locales Améliorer la communication sur les évolutions de l'API SWORD auprès des institutions versantes	Technique Organisationnel
Poursuivre, en la développant, la démarche d'écoute et de prise en compte des demandes des institutions utilisatrices de HAL	Assurer le « *développement de l'animation des communautés utilisatrices de HAL* » (objectif n°1 de l'axe communication de la feuille de route 2016-2020 du CCSD)	Politique

Les autres évolutions

Les autres évolutions attendues, dont certaines sont importantes, voire déterminantes pour certains établissements, sont répertoriées dans le tableau ci-dessous.

Attentes	Solutions envisagées	Types d'impact
Elargir les types de documents éligibles dans HAL	Rendre éligibles certaines catégories de documents techniques	Politique Technique
Limiter l'accès à certains documents à tout ou partie des membres de l'institution	Associer des droits spécifiques d'accès au document (dont thèses) en fonction de l'établissement de rattachement de l'utilisateur	Politique Technique
Améliorer la qualité des référentiels	Permettre aux institutions ne disposant pas d'un portail HAL de procéder au contrôle et à la correction de leur structure dans les dépôts Améliorer la qualité des référentiels en développant des outils de back-office pour nettoyer les structures étrangères	Politique Organisationnel Technique
Enrichir l'offre fonctionnelle proposée aux chercheurs pour valoriser leurs dépôts	Définir une procédure d'évaluation des outils développés par des établissements et d'éligibilité à l'intégration dans la feuille de route de HAL Réunir les conditions techniques pour intégrer facilement les développements réalisés par les établissements utilisateurs de HAL	Organisationnel Technique

Développer des outils pour les gestionnaires qui déposent pour les chercheurs, complètent et enrichissent les métadonnées	Intégrer un outil de préparation des versements par lot sur le modèle de Bib2HAL. Créer une bibliothèque de workflows permettant de s'adapter aux pratiques locales. Créer un rôle supplémentaire autorisant le dédoublonnage des dépôts.	Technique Technique Technique
Réexaminer les principes régissant la modération	Confier aux administrateurs de portails l'instruction des rejets et les arbitrages in fine	Politique Organisationnel
Des évolutions du socle technique de HAL à moyen terme	Faire évoluer HAL vers une solution open source	Technique

Les moyens limités du CCSD

Les réponses aux demandes formulées par les institutions pour faire évoluer HAL vers un modèle technique et organisationnel plus proche de leurs besoins nécessitent, pour certaines d'entre elles, des moyens budgétaires et humains pour les mettre en œuvre. Or la plupart des personnes interrogées dans le cadre de l'étude jugent que le CCSD ne dispose pas à ce jour de moyens suffisants pour réaliser ces évolutions dans des délais raisonnables.

Le CCSD compte à ce jour 16 collaborateurs dont 9 titulaires et 7 contrats à durée déterminée. Son budget annuel est de l'ordre d'1 M€ dont 85% environ sont dédiés aux salaires et 15% au fonctionnement. Outre l'archive ouverte pluridisciplinaire HAL et l'ensemble des services qui y sont rattachés, l'équipe du CCSD opère d'autres services, tous liés à HAL, comme Episciences, la plateforme de publication et d'hébergement de revues électroniques en libre accès, dites « épi-revues », et SciencesConf, la plateforme de gestion et de publication de congrès scientifiques.

L'INRA devrait rejoindre prochainement l'UMS aux côtés de l'INRIA, du CNRS et de l'université de Lyon et contribuer à l'effort de développement. En outre, la mise en œuvre de la nouvelle grille tarifaire pour les portails HAL permettra au CCSD de dégager des revenus complémentaires pour financer la réalisation des évolutions attendues. Il n'est toutefois pas acquis que cela suffise.

Une fois les résultats de cette étude partagés et discutés, et la liste des évolutions prioritaires définie, une mesure des charges et délais de celles ayant un impact technique devra être réalisée par le CCSD. Il s'agira ensuite de vérifier si leur intégration est envisageable dans la feuille de route 2016-2020.

Annexes

Annexe 1 : Liste des institutions et des personnes interrogées

Etablissement	Personne interrogée	Portail HAL	Archive locale
Université de Lorraine	Jean-François Lutz, responsable de la bibliothèque numérique Célia Lentretien,	Oui – 09/2015	Non
Aix-Marseille Université (AMU)	Marlène Delhaye, responsable des ressources numériques et OA Animatrice commission CasuHal	Oui – 12/2013	Non
INRA	Odile Hologne, directrice déléguée à l'IST	En projet	Oui – 2006
Université de Bordeaux	Jérôme Poumeyrol, conservateur des bibliothèques et chef de projet "Archive ouverte institutionnelle".	Non	En projet
Université Paris-Dauphine	Christine Okret-Manville, directrice adjointe de la bibliothèque, administratrice du portail HAL Isabelle Boudet, conservatrice	Oui – 07/2015	Oui
ENSAM	Christine Ollendorff, Directrice de la Documentation et de la Prospective – Animatrice du GTAO Couperin	Non	Oui 2013
Université de Lille	Julien Roche, directeur du SCD de Lille de Lille 1 et responsable du learning center, vice-président ADBU Marie-Madeleine Geroudet, chef de projet AOI	Non – Portail HAL Lille 3 02/2013	En projet
Université de Strasbourg	Adeline Rege, responsable du département documentation scientifique numérique et chef de projet archives ouvertes de la connaissance	Non	Oui 10/2016
Université de la Réunion	Nicolas Alarcon, Président de CasuHal	Oui 03/2015	Non
Université Paris Lumières (ComUE)	Lionel Maurel, responsable de la valorisation de l'IST Membre du CST de HAL	En projet – Paris 8 – 07/2012 Nanterre – 07/2016	Non
INSERM	Nathalie Duchange, DISC pôle IST	Oui – 01/2006	Non
Université de Liège	Paul Thirion, directeur général du réseau des bibliothèques	Sans objet	Oui – 11/2008
Université Grenoble Alpes	Lucie Albaret, responsable des services à la recherche à la BU	Oui – 07/2013	Non

Etablissement	Personne interrogée	Portail HAL	Archive locale
INRIA	Jacques Millet, délégué à l'information et l'édition scientifiques Alain Monteil, coordinateur du pôle archives ouvertes	Oui – 04/2004	Non
Université de Montpellier	Laure Lefrançois, responsable du pôle sciences et techniques au sein du SCD	Oui – 05/2017	Non
Université de Toulouse 2	Françoise Gouzi, chargée d'Information Scientifique & Technique (IST) dans le service Direction en Appui à la Recherche	Oui – 05/2011	Non
Université de Rennes 1	Laurent Jonchère, Responsable de la documentation électronique et de l'Open Access Thierry Fournier, responsable de la documentation électronique et de l'Open Access	Oui – 12/2012	Non
IRSTEA	Emmanuelle Jannès-Ober, Chargée de mission Intelligence économique / Responsable Information scientifique et technique	Non	Oui – 03/2005
CNRS-DIST	Renaud Fabre, ancien directeur de la Direction de l'Information Scientifique et Technique (DIST), interrogé lors des phases ayant précédé l'étude		

Annexe 2 : Pondération des évolutions proposées

Pour pondérer les différentes solutions proposées, les critères suivants ont été retenus :

- **L'intensité de la demande :**
  - Forte, la demande est récurrente lors de nos entretiens et présentée parmi les premières évolutions demandées.
  - Moyenne, la demande a été formulée par plusieurs personnes interrogées et, lorsqu'elle est suggérée, elle est généralement bien reçue.
  - Faible, le sujet est évoqué par de rares interlocuteurs.
- **L'impact sur le modèle de HAL :**
  - Fort, la solution envisagée remet en cause des orientations structurantes du modèle HAL (auto dépôt, Open Access, fluidité du dépôt pour le chercheur...)
  - Moyen, la solution infléchit certains principes sans néanmoins avoir des incidences pour la dynamique de dépôt par le chercheur.
  - Faible, pas d'impact sur le modèle de HAL.
- **L'impact sur les charges de développement :**

  La quantification est ici indicative et des études techniques devront être menées pour établir les charges nécessaires au développement des solutions proposées.
  - Fort, la solution envisagée suppose des développements significatifs.
  - Moyen, des développements sont nécessaires mais la solution semble peu coûteuse en première analyse.
  - Faible, pas développement ou marginal.
- **L'impact sur les charges d'exploitation :**
  - Fort, la solution envisagée suppose de mobiliser des ressources humaines ou techniques récurrentes et significatives.
  - Moyen, des ressources sont nécessaires.
  - Faible, pas d'impact sur les charges d'exploitation.

Annexe 3 : Commentaires post-étude des commanditaires

Introduction

Après relecture de l'étude, ses commanditaires ont recensé des bonnes pratiques et des actions susceptibles d'être menées dans un cadre institutionnel et dans des instances extérieures. Il s'agit d'éléments qui sont de près ou de loin abordés par l'étude, mais qu'ils jugent essentiels et veulent préciser.

En outre, ils ont dressé une liste de points de réflexion et d'éléments qu'ils auraient souhaité faire apparaître dans l'étude ou dans un travail ultérieur.

Actions à inscrire dans un cadre institutionnel

- Présenter HAL et ses services aux établissements qui souhaitent créer une archive.
- Définir un périmètre pour la création des indicateurs.
- Mesurer l'impact de l'abandon de services auprès des établissements qui souhaitent les mutualiser.
- Etudier la possibilité pour les archives institutionnelles d'être administrateur de leurs contenus dans HAL.
- Préciser, décrire, étudier les interactions possibles avec le GTAO de Couperin et CasuHal.
- Insister sur la nécessaire articulation de HAL avec OpenAIRE à tous les niveaux, pour répondre à une demande récurrente et de longue date du CCSD.
- Se référer aux études précédentes, par exemple celle IRSTEA sur les archives, qui avait déjà souligné l'importance de l'utilisation de systèmes CRIS.

Compléments de l'étude souhaités

- Etoffer la notion d'archives institutionnelles et les besoins des établissements en la matière.
- Alerte sur l'outil unique capable de répondre à plusieurs fonctions différentes.
- Préciser les périmètres d'une archive ouverte nationale.
- La migration de Prodinra dans HAL pourra être présentée car ce sera une expérience d'archive institutionnelle avec son propre fonctionnement et ses propres indicateurs.
- Donner une portée plus pratique à la partie « recommandations ».
- Etudier de manière plus approfondie le business model de HAL pour mieux nourrir les décisions à prendre.
- L'identification des documents/chercheurs/organismes pourrait faire l'objet d'une étude car c'est un enjeu important pour coordonner les archives et les plateformes entre elles.
- PURE et CRIS sont des outils de plus en plus utilisés par les établissements et les universités de recherche intensive. Il serait utile de détailler leur utilité et leur utilisation.
- L'approche comparative de l'étude gagnerait à inclure une dimension internationale.