



HAL
open science

La valorisation de HAL. Finalités, outils et process

Joachim Schöpfel, Hélène Prost, Bernard Jacquemin, Éric Kergosien, Florence Thiault

► To cite this version:

Joachim Schöpfel, Hélène Prost, Bernard Jacquemin, Éric Kergosien, Florence Thiault. La valorisation de HAL. Finalités, outils et process. 2024. hal-04450898v1

HAL Id: hal-04450898

<https://hal.science/hal-04450898v1>

Preprint submitted on 10 Feb 2024 (v1), last revised 25 Jul 2024 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

La valorisation de HAL

Finalités, outils et process

Joachim Schöpfel¹
Hélène Prost²
Bernard Jacquemin¹
Éric Kergosien¹
Florence Thiault³

¹ Univ. Lille, ULR 4073 – GERiiCO – Groupe d'Études et de Recherche Interdisciplinaire en Information et Communication, F-59000 Lille, France

² CNRS, ULR 4073 – GERiiCO – Groupe d'Études et de Recherche Interdisciplinaire en Information et Communication, F-59000 Lille, France

³ UR 4246 – PREFICS – Pôle de Recherche Francophonies, Interculturel, Communication, Sociolinguistique, Université Rennes 2, F-35043 Rennes, France

Résumé

L'article présente les résultats d'une analyse qualitative de l'utilisation de la plateforme HAL par les laboratoires de recherche. L'analyse s'appuie sur des entretiens semi-directifs avec des représentants de 50 laboratoires affiliés aux dix universités de recherche Udice. Elle porte sur la fonction que remplit HAL pour les laboratoires, sur sa valeur ajoutée pour leur fonctionnement et leur développement. En particulier, nous interrogeons les finalités de l'utilisation de HAL par les laboratoires, le recours à des outils internes et externes, et les trajectoires des pratiques. Nous discutons les résultats sous trois angles : les discours et les communautés de pratiques dans les laboratoires ; le périmètre et les enjeux des dispositifs mis en place par les laboratoires en amont et en aval de HAL ; et la transformation de HAL d'une plateforme d'auto-archivage et de communication scientifique directe vers une infrastructure de suivi et d'évaluation de la performance scientifique. Il s'agit du premier d'une série de trois articles issus du projet HAL/LO.

Mots-clés

Science ouverte, libre accès, archives ouvertes, voie verte, HAL, laboratoires de recherche, universités, systèmes d'information sur la recherche

Introduction

Le mouvement de la science ouverte, entendue comme « la diffusion sans entrave des résultats, des méthodes et des produits de la recherche scientifique » (MESR, 2021), modifie en profondeur les modalités d'accès et de diffusion de l'information scientifique et technique. L'ouverture des publications qui selon les termes du deuxième plan national pour la science

ouverte (loc.cit.) doit devenir « une pratique incontournable », s'appuie notamment sur le dépôt dans une archive ouverte publique comme HAL¹.

La plateforme HAL a été conçue pour « faciliter et accélérer la diffusion des publications scientifiques suivant les principes d'une communication scientifique directe entre pairs (...) La mission de HAL est de diffuser des articles, de niveau recherche, publiés ou non » (Berthaud et al., 2021). Inspiré du serveur de preprints arXiv², le projet HAL a été porté lors de sa création en 2001 par le département des sciences physiques et mathématiques (SPM) et l'Institut national de physique nucléaire et de physique des particules (IN2P3) du CNRS. Sa vocation a été, dès le début, d'être une archive ouverte multidisciplinaire, pour l'ensemble des disciplines et communautés scientifiques, au service des chercheurs mais aussi, et de plus en plus, des établissements et organismes de recherche. Aussi, depuis son lancement, « l'histoire de HAL s'inscrit dans celle des stratégies institutionnelles en matière d'accès ouvert aux publications et, plus largement, de science ouverte » (loc.cit.).

Aujourd'hui, début 2024, HAL joue un rôle central dans la politique de la science ouverte en France, avec 1,3 millions de documents scientifiques et 3,5 millions de références, dont beaucoup d'articles mais aussi des communications, rapports, ouvrages et chapitres, des thèses etc., issus de toutes les disciplines scientifiques. Ses services et fonctionnalités sont en constante évolution pour faciliter l'acceptation et répondre à la demande des autorités et des institutions (loc.cit.). Poursuivre le développement de HAL (considéré comme « archive ouverte nationale ») est l'une des mesures prioritaires du deuxième plan national de la science ouverte afin de « simplifier son appropriation par les chercheurs et par les institutions », une amélioration qui passe notamment par la simplification « du dépôt dans HAL pour les chercheurs qui publient sur d'autres plateformes en accès ouvert dans le monde » et par le développement d'un « service intégré d'auto-archivage (et) de collecte automatisée de publications (...) » (MESR, 2021).

Plusieurs études ont analysé l'appropriation de HAL par les chercheurs et les institutions, révélant notamment des différences disciplinaires et institutionnelles (Mahé & Prime-Claverie, 2017 ; Larrieu & Schöpfel, 2022). Apparemment, beaucoup de chercheurs connaissent et utilisent HAL, d'une manière ou d'une autre – moins pour chercher de l'information, davantage pour signaler et diffuser leurs travaux. Cependant ils semblent moins nombreux à utiliser HAL pour valoriser leurs publications dans leurs pages web personnelles (Schöpfel et al., 2022).

Quant aux laboratoires de recherche, plusieurs pratiques et approches de HAL ont été identifiées, conditionnées par des choix individuels et collectifs, des politiques d'établissements et des pratiques disciplinaires, des variables qui forment un écosystème de facteurs interdépendants (loc.cit.). Tous les laboratoires, ou presque, sont présents dans HAL grâce aux publications de leurs chercheurs ; plus de la moitié ont leur propre collection³, surtout en sciences et technologies et en SHS, et dans les unités mixtes (Schöpfel et al., 2023a). La création d'une telle collection permet de regrouper les publications du laboratoire au même endroit ; son objectif est de rendre la production du laboratoire visible et de mieux

¹ HAL <https://hal.science/>.

² arXiv <https://arxiv.org/>.

³ En 2021, HAL comptait 7263 collections créées depuis 2002, dont 39% de collections de laboratoire (Berthaud et al., 2021).

gérer et suivre la production en interne afin de faciliter les bilans individuels et collectifs et contribuer aux campagnes d'évaluation (Schöpfel et al., 2022). En effet, rappelons que l'évaluation des laboratoires de recherche par le Hcéres examine depuis récemment « comment les différentes dimensions de la science ouverte sont prises en compte » et que ses bilans bibliométriques incluent désormais des « indices d'accès ouvert » (Coulhon & Larrouturou, 2023).

Cependant, utiliser HAL ne veut pas nécessairement dire l'apprécier. L'avis sur HAL est partagé. Au sein des laboratoires, « on l'apprécie ou on ne l'apprécie pas (...) les uns trouvent HAL facile, pratique et performant (tandis que) les autres disent que HAL est compliqué, lourd, peu convivial » (Schöpfel et al., 2022). Sur le blog du CCSD⁴, le commentaire d'un chercheur est plus direct : « (...) il y a une inflation de ces interfaces, qui souvent ne fonctionnent pas bien, ou ne sont pas conviviales du tout (...) HAL est un très bon exemple de complexité non conviviale (...) L'urgence, pour nous chercheurs, est de simplifier tout cela. Mon seul mot d'ordre est donc : SIMPLIFIER, SIMPLIFIER, SIMPLIFIER » (Lubino, 2023).

Or, une telle appréciation ambiguë pose un problème dans la mesure où l'acceptation d'un système d'information n'est pas seulement liée à la perception de son utilité mais aussi à la perception de la facilité de son utilisation (Davis, 1989). Ainsi, depuis plusieurs années le CCSD poursuit un double objectif : améliorer la facilité d'utilisation de HAL, par exemple par la simplification des dépôts, et renforcer la perception de l'utilité de HAL, par le développement de nouveaux services et fonctionnalités. « Pour le CCSD, l'enjeu est de générer de la confiance, de faciliter l'utilisation de HAL et de proposer des services complémentaires à valeur ajoutée aux chercheurs comme aux institutions » (Berthaud et al., 2021). Retenons trois aspects de cette démarche :

- la stratégie du CCSD s'adresse aussi bien aux individus (chercheurs, enseignants-chercheurs...) qu'aux institutions (établissements, organismes, laboratoires...), ce qui reflète la transformation de HAL d'un service d'auto-archivage vers un dispositif de suivi et d'évaluation de la performance scientifique (Schöpfel et al., 2023b) ;
- ce développement concerne aussi bien l'amont que l'alimentation de HAL que l'exploitation de ses données (publications) et métadonnées (notices) ;
- le développement de nouveaux services de HAL est accompagné par la mise en place d'outils extérieurs à HAL, construits à partir des fonctionnalités et données de HAL et portés par des laboratoires ou d'autres services.

Parmi ces outils extérieurs à HAL, citons ici à titre d'exemple la suite d'outils HALUR développé à Rennes (Jonchère & Troccaz, 2023), l'outil de contrôle des publications OCdHAL développé à Grenoble (Reynier & Bellemain, 2023), les HalTools développés par l'INRIA (cf. la boîte à outils du portail inria.hal.science) ou encore le projet SoVisu à Toulon (Reymond et al., 2023). Ces outils répondent à différents besoins exprimés au niveau individuel et au niveau institutionnel. Ils remplissent plusieurs fonctions de gestion de dépôts (individu) et d'administration de HAL (institution) (tableau 1).

⁴ Centre de Communication Scientifique Directe, une unité d'appui et de recherche rattachée au CNRS, opérateur de la plateforme HAL <https://www.ccsd.cnrs.fr/>

Tableau 1. Les fonctions de quelques outils externes

Niveau	Fonction
Individu (chercheur...)	Vérifier sa liste de publications HAL (bilan CNRS)
	Créer sa page web
	Exporter des références
	Importer par lot de références
	Importer par lot de publications
Institution (laboratoire...)	Exporter la bibliographie des productions pour l'évaluation du Haut conseil d'évaluation de la recherche et de l'enseignement supérieur (Hcéres) (bilan collectif)
	Référencer la bibliographie des laboratoires dans HAL
	Ajouter et diffuser le texte intégral des articles
	Assurer la qualité des métadonnées
	Évaluer le taux d'accès ouvert aux publications

Dans un contexte plus large, on pourrait également mentionner les projets de monitoring de niveau local et national (Bracco, 2022 ; Bracco et al., 2022) du Baromètre de la science ouverte (BSO) dont la première version a été développée à partir de 2018 et qui s'appuie en partie sur les métadonnées de HAL.

Peu de laboratoires ont publié leurs pratiques et expériences en matière de science ouverte, à l'instar de l'UMR *Ambiances Architectures Urbanités* (AAU) de Grenoble où l'obligation de déposer les publications dans HAL fait partie depuis 2020 du règlement intérieur. Ce laboratoire décrit plusieurs « leviers, incitatifs ou contraignants », facteurs favorables à l'accélération de la démarche, notamment le financement par l'ANR, le bilan annuel d'activité des chercheurs du CNRS, l'accord cadre Ministère de la Culture-CNRS et le processus d'évaluation du Hcéres mais également des événements « pour motiver les équipes » comme la participation au HALathon organisé par casuHAL⁵ et le suivi des résultats de leur engagement (Acquier et al., 2023).

⁵ casuHAL <https://www.casuhal.org/>

De son côté, l'UMR *Grenoble Image Parole Signal Automatique* (GIPSA) a développé une application (OCdHAL) qui propose un « tableau de bord permettant de voir et de contrôler sur un seul écran toutes les métadonnées des documents déposés dans HAL et de gérer les collections d'une structure » ; cet outil inclut la gestion des identifiants individuels des chercheurs (OCidHAL), partant du constat qu'un laboratoire de recherche doit « pouvoir produire la liste des publications scientifiques de ses chercheurs et disposer rapidement et facilement de quelques statistiques ou indicateurs bibliométriques, afin de répondre aux demandes de ses tutelles, de valoriser sa production scientifique et de contribuer à la politique de science ouverte » (Reynier & Bellemain, 2023).

Augmenter la visibilité et l'impact de la production scientifique et développer l'activité du laboratoire est un « argument marketing » pour utiliser HAL. D'autres arguments sont d'ordre politique (accélérer la diffusion des résultats, renforcer l'accessibilité des publications du laboratoire, contribuer au développement de la science ouverte) ou institutionnel (répondre à la demande des tutelles et des agences de financement et d'évaluation). L'enquête du CCSD sur la formation à HAL en 2023 reflète ces différents arguments et besoins : « Les formations suivies sont principalement les formations *Gestionnaire de collection/référent de structure* et *Administrer un portail* pour les professionnels de l'IST, et *Déposer dans HAL* et *Gestionnaire de collection/référent de structure* pour les chercheurs » (Lubino, 2023). Déposer, oui, mais aussi et surtout gérer et administrer ; on sort d'une simple logique d'auto-archivage vers un dispositif plus complexe, à plusieurs finalités.

La transformation de HAL d'une archive ouverte vers un dispositif de suivi et d'évaluation a été décrite ailleurs, en mettant l'accent sur les nouveaux enjeux qui en découlent, comme l'importance cruciale de la qualité et fiabilité des métadonnées, avec une dimension éthique (Schöpfel et al., 2023b). Force est de constater que le développement de la plateforme HAL et l'évolution des pratiques dans les laboratoires est un processus interdépendant, dans un environnement (écosystème) dynamique et complexe.

Afin de poursuivre l'analyse des pratiques des laboratoires et pour mieux comprendre la diversité des situations et des trajectoires constatées sur le terrain, nous avons mené une étude qualitative avec un panel de laboratoires sous forme d'entretiens semi-directifs. L'objectif de l'enquête était triple :

1. analyser les objectifs de l'utilisation de HAL par les laboratoires, ainsi que les outils et procédures mis en œuvre. ;
2. analyser les différentes approches d'incitation et d'accompagnement des pratiques au sein et par l'environnement des laboratoires ;
3. analyser quelques pratiques représentatives, y compris l'utilisation d'autres plateformes, ainsi que les avis récoltés à propos de HAL.

Notre article – le premier d'une série de trois articles – présente une partie des résultats de cette enquête : quelle fonction remplit HAL pour les laboratoires, quelle est sa valeur ajoutée pour leur fonctionnement et leur développement ?

Méthodologie

La trame de l'entretien contient 28 questions, certaines fermées, d'autres ouvertes (cf. annexe 1). Elle a été développée à partir des résultats de notre première enquête en ligne en 2021 (Schöpfel et al., 2022), et elle a été testée avec plusieurs personnes.

Nous avons constitué un panel de 50 laboratoires rattachés aux dix universités Udice (cf. annexe 2) à partir de l'échantillon des 400 laboratoires de l'enquête de 2021. Nous avons repris contact avec les 43 laboratoires qui avaient signalé, lors de l'enquête en ligne, leur disposition à participer à la suite de l'enquête ; 34 ont répondu favorablement, et nous avons ajouté 16 autres laboratoires de manière à représenter toutes les disciplines et toutes les universités.

Les personnes contactées étaient soit les répondants de l'enquête en ligne, soit la personne en charge de la collection HAL ou responsable de la documentation du laboratoire, ou la directrice ou le directeur de l'unité. Chaque personne ainsi identifiée a été contactée personnellement via son adresse électronique, avec une invitation à participer à l'étude sur l'usage de HAL sous forme d'un entretien d'environ 30 à 60 minutes.

Les entretiens ont été menés entre mai 2021 et février 2023, par cinq membres de l'équipe projet, seul ou en binôme (Aline Benvegnú dos Santos, Marine Chochoy, Bernard Jacquemin, Joachim Schöpfel, Florence Thiault). Les questions ont été communiquées plusieurs jours avant l'entretien. Les entretiens ont été menés à distance, avec Zoom, et enregistrés sur le serveur de l'université. Chaque entretien a fait l'objet d'une transcription synthétique ; la transcription a été envoyée aux interlocuteurs, avec un lien vers l'enregistrement en ligne, pour relecture et validation. La version validée a été saisie sur la plateforme SurveyMonkey pour l'analyse, tandis que l'enregistrement a été supprimé.

L'analyse des contenus des entretiens a été réalisée de juin à décembre 2023. Par la suite, les verbatim sont cités en italique. Les attributs des citations (ex. : SU-074) correspondent aux identifiants internes du projet HAL/LO⁶.

Résultats

Les répondants

50 laboratoires ont participé à cette enquête, dont 40 unités mixtes de recherche (UMR), 7 unités de recherche universitaires (UR/EA), une unité de recherche propre du CNRS (UPR), un observatoire des sciences de l'univers (OSU) et une unité mixte de service (UMS). Ils couvrent l'ensemble des dix disciplines du répertoire national des structures de recherche⁷ et

⁶ Valorisation sur HAL de la production des laboratoires dans l'environnement de la science ouverte. Projet financé par le GIS URFIST.

⁷ RNSR <https://appliweb.dgri.education.fr/rnsr/>

sont affiliés à au moins une des dix universités Udice (cf. annexe 2). 19 laboratoires relèvent de plusieurs disciplines et se décrivent, notamment en SHS, comme inter- ou pluridisciplinaire.

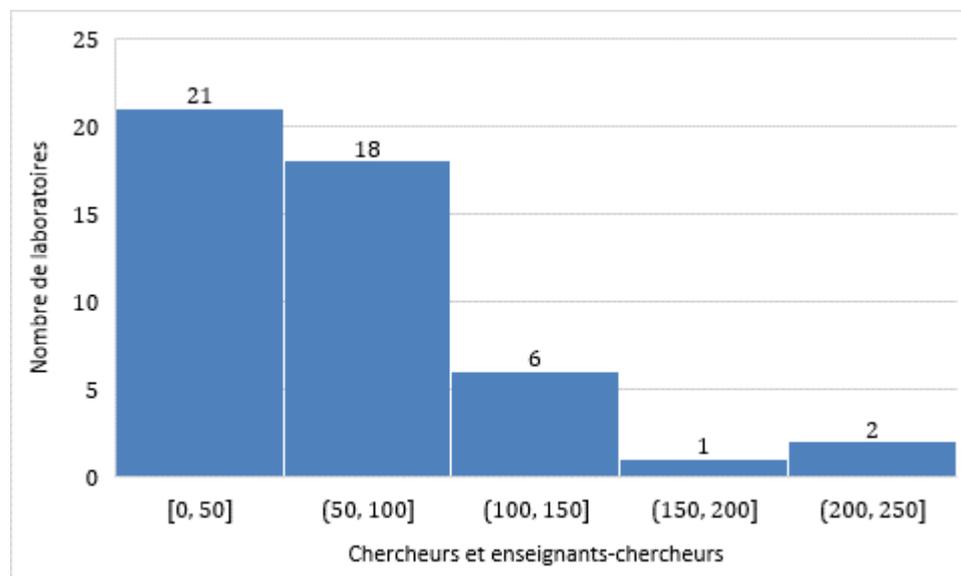
Les entretiens ont été menés avec 58 personnes (42 entretiens individuels, 8 entretiens avec deux personnes), dont 33 chercheurs ou enseignants-chercheurs et 25 ITA, avec différentes fonctions au sein de leur laboratoire (table 1) :

Tableau 2. Fonction des interlocuteurs (N=58)

Fonction	Nombre d'interlocuteurs
Direction, administration	29
Gestion HAL	15
Documentation	8
Divers (communication, science ouverte...)	6

La taille des laboratoires est très variable, avec une médiane de 60 chercheurs et enseignants-chercheurs (figure 1), auxquelles s'ajoutent jusqu'à 200 doctorants (médiane = 52) et 750 ITA (médiane = 31).

Figure 1. La taille des laboratoires, en nombre de chercheurs et enseignants-chercheurs (N=50)

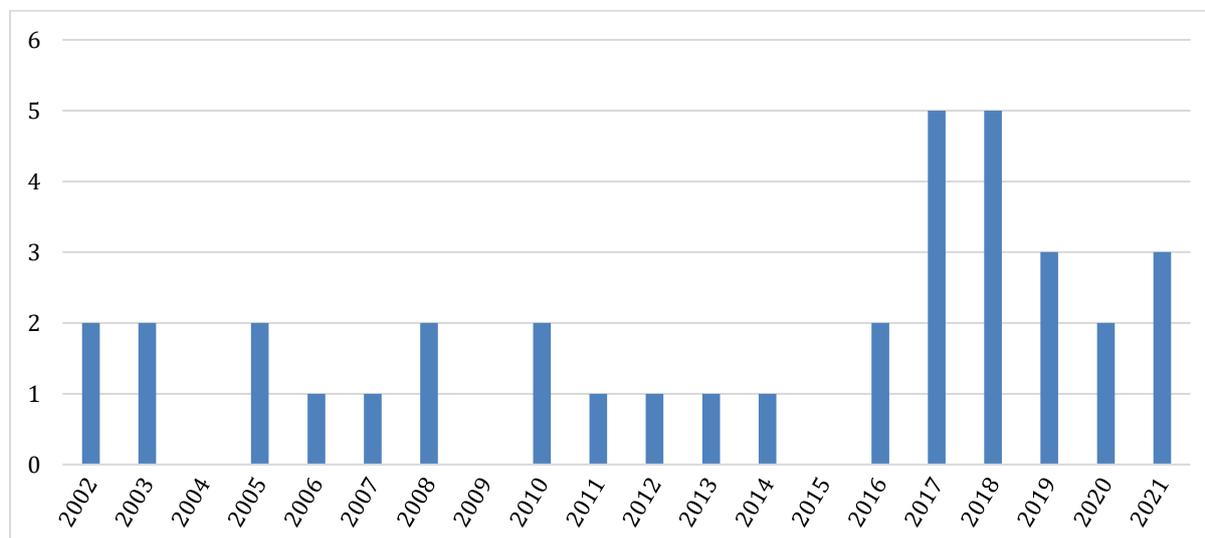


Les finalités

47 laboratoires ont une collection sur HAL. Un laboratoire *commence tout juste à s'occuper de sa collection HAL* (UCA-036) ; deux autres vont utiliser HAL pour la prochaine campagne

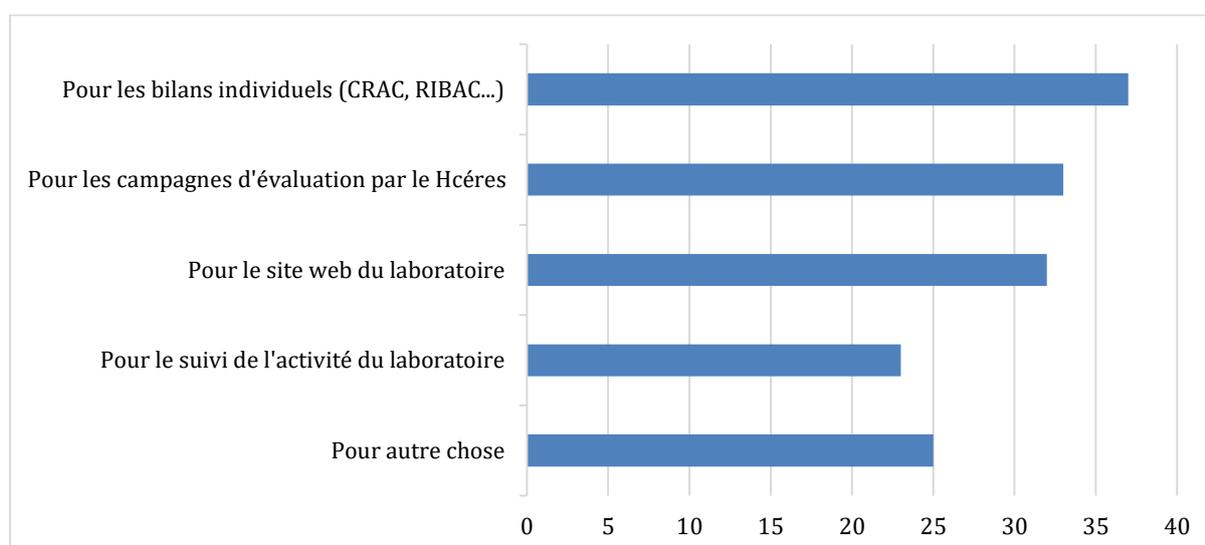
Hcéres. Tous les autres laboratoires pratiquent HAL à différents degrés, depuis plus ou moins longtemps (figure 2) et souvent pour plusieurs raisons, comme pour le suivi et les bilans des chercheurs et du laboratoire, pour l'évaluation par le Hcéres et/ou pour le site web.

Figure 2. Année de la mise en place de la collection HAL⁸ (N=36)



On peut distinguer deux types de pratiques, souvent liées : une pratique plutôt individuelle (bilans des chercheurs du CNRS, de l'INRAE ou de l'INRIA), et une pratique plutôt collective ou institutionnelle (suivi et bilan de la production du laboratoire et de ses équipes, évaluation, site web) (figure 3).

Figure 3. Utilisation de HAL par les laboratoires (N=49)



⁸ NB : 12 répondants ne savent plus quand leur collection a été mise en place.

Quelques unités voient HAL surtout comme un vecteur de diffusion ou décrivent un *usage rudimentaire* (USPC-164). Le directeur d'un laboratoire en sciences de la vie dit qu'au niveau institutionnel, *on ne l'utilise pas pour les bilans(...) On utilise HAL comme une plateforme où les publications de l'institution sont déposées en libre accès* (SU-074). Un autre directeur précise pourquoi il n'utilise pas HAL pour les campagnes d'évaluation par le Hcéres : *On n'est pas sûr que tout le monde rentre toutes les publications dans HAL, on a préféré sur 600 publications réparties sur une dizaine d'équipes (...) de gérer et de faire ce travail (de manière manuelle), même si c'est un peu fastidieux, surtout quand on a des doublons qui sont partagés sur deux équipes qui publient en même temps* (UDL-072). Même approche dans une UMR en SHS : *Certains CNRS utilisent HAL pour leur RIBAC, mais pas tous. Pour le Hcéres, on avait fait des extractions du site web du laboratoire* (SU-118). L'ergonomie est un autre argument : *On n'utilise pas HAL directement pour le Hcéres parce que ce n'est pas très pratique pour les exports* (UDL-141).

Pour d'autres laboratoires cependant, HAL est devenu l'outil central et incontournable pour les bilans et le suivi des publications : *Au conseil du laboratoire, ils étaient tous d'accord pour que HAL soit le « centre de dépôt » et que la direction du laboratoire fasse des extractions pour le bilan Hcéres (...) Le laboratoire comptait un peu sur l'intérêt de pouvoir alimenter les pages personnelles des chercheurs à partir de HAL, avec un identifiant HAL (...) HAL est également utilisé pour le suivi de l'activité, aussi bien pour l'ensemble du laboratoire que pour certains chercheurs individuellement (...) aussi pour les CV personnels, pour 'l'autopromotion'* (AMU-062). En tout, 45 laboratoires utilisent HAL pour les bilans individuels, pour le suivi de l'unité et/ou pour les campagnes d'évaluation. En d'autres mots, neuf sur dix utilisent HAL comme un dispositif d'information sur la recherche.

Régulièrement, les répondants expliquent l'utilisation de HAL avant tout par l'obligation d'établir les bilans individuels notamment pour les chercheurs du CNRS. Produire les bilans à partir de HAL est fortement recommandé par les organismes de recherche ; mais de l'autre côté, *l'utilisation par les chercheurs CNRS pour les RIBAC, avec leur identifiant IdHAL, (est) très pratique pour eux* (USPC-084). En plus, cette pratique individuelle plus ou moins obligatoire contribue à renforcer l'intérêt de HAL pour le suivi d'activité des laboratoires. L'action individuelle (dépôt) comme condition nécessaire pour l'action collective (bilan) : c'est un argumentaire qu'on retrouve dans la réponse d'une chargée de l'IST d'une UMR où l'alimentation de HAL *ne se fait pas de manière globale (...) chaque chercheur fait ses propres dépôts. Pour extraire les références de HAL, pour les rapports d'activité et autres statistiques, j'utilise le soft de l'université de Rennes* (AMU-075).

Aussi, on peut observer deux pratiques différentes pour l'usage multifonctionnel de HAL :

- alimentation de HAL pour ensuite mettre à jour le site Web du laboratoire, puis utiliser le site Web pour les besoins institutionnels de suivi et d'évaluation ;
- ou bien, alimentation de HAL, puis d'une part exporter les notices HAL pour alimenter le site Web et d'autre part, effectuer les analyses et les bilans directement dans HAL.

Les laboratoires rencontrent essentiellement trois problèmes : des difficultés techniques pour extraire les données de HAL, des problèmes avec l'identification des personnels, et des problèmes pour retrouver les publications dans HAL.

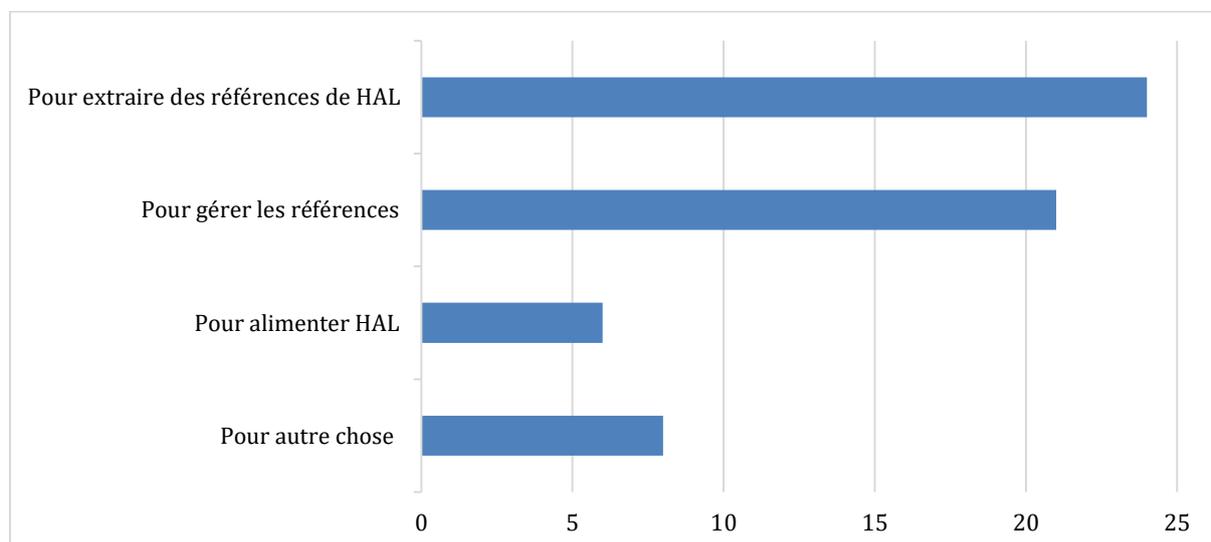
Le suivi individuel avec HAL ne se limite pas aux seuls chercheurs mais s'étend progressivement aux autres personnels scientifiques, en premier lieu aux universitaires : *Au-delà des bilans individuels CRAC/RIBAC, je m'en sers aussi pour les entretiens avec les enseignants-chercheurs qui n'ont pas les CRAC ou RIBAC (UGA-031)*. D'après les entretiens, il y a trois raisons à cela : la volonté d'un même traitement pour toutes les catégories ; la possibilité d'alimenter la vitrine du site web avec l'ensemble des publications ; et surtout, la nécessité d'utiliser HAL pour l'évaluation de la performance globale du laboratoire.

Un laboratoire utilise HAL pour valoriser le patrimoine scientifique de l'unité à travers les travaux de plusieurs membres émérites : *L'UMR a recruté des vacataires pour scanner et déposer les documents pour quatre membres émérites, en petite équipe et en étroite collaboration avec ces collègues, qui étaient tous très partants. Il s'agit de déposer les textes intégraux, pas uniquement de référencer la production (USPC-084)*.

Les outils

La grande majorité des laboratoires – quatre sur cinq – utilisent un ou plusieurs programmes, applications etc. pour alimenter HAL, gérer les références, extraire des notices ou analyser la production (figure 4).

Figure 4. Utilisation d'outils (logiciels, codes) par les laboratoires (N=42)



Parfois il s'agit d'un usage plutôt personnel. Ainsi, le directeur d'une UMR regrette que *chacun fait sa propre référence en BibTeX, donc c'est un peu le bazar dans tous les disques durs (SU-076)*. En effet, plusieurs répondants racontent que leurs collègues se servent, à titre personnel, d'un gestionnaire de références (Zotero, EndNote, Mendeley, BibTeX) ou la base

de données d'ORCID pour le suivi de leurs publications et leurs bilans. *Les chercheurs utilisent ce qu'ils veulent, souvent ils utilisent Zotero ou EndNote (AMU-058).*

Néanmoins, les autres décrivent comment ils font appel, au niveau du laboratoire, à un ou plusieurs outil(s), à l'instar d'une référente et gestionnaire de collection HAL qui évoque un *mix de plein d'outils*, dont Zotero et Bib2HAL pour l'import initial de références, un *plug-in* pour alimenter le site web, ExtrHAL pour les bilans, Publish or Perish (harzing.com) pour les citations et indicateurs, et Google Scholar, ResearchGate et Academia pour identifier les publications des chercheurs (AMU-062). En tout, nous avons compté presque 20 applications, logiciels, etc., de plusieurs catégories.

Des outils externes développés pour HAL

En amont, pour l'alimentation de HAL, les laboratoires utilisent plusieurs outils de conversion, comme Bib2HAL et X2HAL (INRIA), Ad HALstra (INSU). Pour alimenter HAL, une documentaliste fait appel à HALstra et Zotero *mais (...) utilise aussi OCdHAL qui est un outil qui permet de vérifier, corriger les notices (AMU-099).*

En complément de ces outils, un contrôle manuel peut s'avérer nécessaire : *Si les chercheurs déposent directement sur HAL, c'est facile de récupérer en un clic vers Zotero. Le tamponnage est fait à la main au fur et à mesure, car il y a une vérification à faire de ce que le chercheur a déposé (affiliations, métadonnées). Pour les imports dans l'autre sens on utilise X2HAL, l'outil qui permet les imports en format BibTeX dans HAL (UDL-141).*

En aval, pour l'extraction et l'export des notices de HAL, les laboratoires utilisent OCdHAL (Grenoble), une application web (un tableur dynamique, connecté par webservices à HAL), et ExtrHAL (Rennes) qui permet d'extraire, de mettre en forme puis d'exporter les publications HAL d'un laboratoire, d'une équipe ou d'un auteur. Les deux outils se complètent : *Personnellement j'utilise OCdHAL, qui permet de questionner plus facilement le contenu de HAL. J'utilise également ExtrHAL, qui est assez pratique pour faire les rapports d'activité (AMU-058).*

Cette dernière réponse implique par ailleurs que les laboratoires ont le choix entre plusieurs options et outils. Le documentaliste d'une UMR le confirme très clairement quand il dit que pour l'extraction, *j'utilise l'outil développé par Rennes (ExtrHAL) (...), en BibTeX. L'université développe un outil pour l'évaluation Hcéres en XML mais c'est moins facile. Je n'utilise pas l'outil de Grenoble (OCdHAL) (UCA-033).*

D'autres outils externes

Pour la veille en amont de l'alimentation de HAL, les laboratoires utilisent des bases de données multidisciplinaires (Scopus et Web of Science) ou disciplinaires (SciFinder en chimie), des archives ouvertes (arXiv), des moteurs de recherche (Google Scholar et PubMed), des réseaux sociaux (ResearchGate et Academia). Dans certains cas, ces outils sont connectés à un gestionnaire de références (EndNote et Zotero). Deux laboratoires ont installé le logiciel Publish or Perish comme une sorte de métamoteur bibliométrique de recherche ; ainsi, un ingénieur gestionnaire de la collection HAL constate qu'il a *encore une grande liste de publications à déposer, issue de la recherche avec « Publish or Perish » deux fois par an, ce qui est fait peu à peu vu la lourdeur de la tâche et de la plateforme en soi (UCA-*

036). La plupart de ces outils sont gratuits, au moins pour un usage non-commercial, tandis que l'accès aux grandes bases de données est payant et conditionné par un abonnement institutionnel.

Un laboratoire de Saclay a installé le package QC Labo de QUASAR Conseil qui contient un outil pour la gestion des publications d'un laboratoire, avec interaction bidirectionnelle avec HAL⁹. En sciences de la vie, un autre laboratoire déclare qu'ils ont *un autre outil qui s'appelle SAMPRA¹⁰ développé par le CHU de Lille pour l'ANR qui nous permet de comptabiliser nos publications sur Web of Science et PubMed et d'extraire des bilans (SU-074).*

Tous ces systèmes n'ont pas été développés spécifiquement pour HAL. Aussi, le directeur d'un laboratoire de physique raconte qu'il a développé *un outil en Python pour récupérer régulièrement une liste des publications du laboratoire dans Scopus. Cette liste est classée par équipe, et je la transmets une fois par an à (la documentaliste) pour qu'elle puisse vérifier si la liste est complète et pointer si toutes les publications se trouvent bien dans HAL (SU-101).* Un tel code interne en amont de HAL reste une exception.

Des outils développés en interne

En revanche, quelques laboratoires ont développé des scripts ou *plug-ins* en aval, pour les exports (rapports d'activité, bilans etc.), et aussi pour alimenter le site web du laboratoire, à l'instar d'un laboratoire de Paris qui utilise depuis 2017 un *plug-in* pour rapatrier les données pour la page publications du site web (USPC-084).

Deux autres exemples : *L'informaticien du laboratoire a créé un lien de HAL vers le site web du labo, il fait tourner un script pour récupérer les publications dans HAL (USPC-138).* Un tel script peut aussi augmenter la granularité et enrichir les métadonnées, en ajoutant le lien avec les équipes du laboratoire : *Dans HAL, les chercheurs ne rentrent pas les équipes. Comme nous avons (besoin) des équipes pour les bilans, nous avons développé un script en interne pour faire dans notre site web le lien avec HAL. Le script a été développé par un ingénieur et s'appuie sur une liste des équipes et des chercheurs (UCA-034).*

Dix laboratoires au moins – un sur cinq – ont mis en place une base de données en interne. Dans une unité mixte par exemple, *une chercheuse CNRS du laboratoire fait le bilan (annuel) des publications à travers BibCNRS (avec en particulier le Web of Science) et alimente un outil interne, rebase, et les pages web du laboratoire (UGA-043).*

Il y a plusieurs raisons pour un tel choix. Souvent, on entend que *c'est des choses qui en partie prédatent l'usage de HAL, mises en place quand il n'y avait pas la possibilité d'exporter des données de HAL (UGA-086).* Pour d'autres, c'est une question de qualité : *Je travaille aussi dans une base locale pour le référencement, dans laquelle j'ajoute beaucoup de choses qui ne sont pas dans HAL. Dans cette base il y a une homogénéisation de données qui est beaucoup plus forte que dans HAL – et cela car sur HAL il y a beaucoup d'intervenants sur*

⁹ QUASAR Conseil <https://www.quasarconseil.fr/>

¹⁰ SAMPRA se décrit comme un logiciel d'analyse bibliométrique, une « brique d'un système d'information sur la recherche » pour des chercheurs et des structures, interopérable avec HAL <https://www.reseau-chu.org/article/apres-sigaps-voici-sampra-le-nouveau-logiciel-danalyse-bibliometrique-made-in-lille/>

une référence (les auteurs mais aussi les différents services documentaires liés aux tutelles), et si ces différents acteurs n'ont pas les mêmes besoins, des choses différentes sont mises en avant, donc les données ne sont pas homogènes (UDL-141).

Le même laboratoire ajoute que *c'est beaucoup plus facile de faire des extractions fines et selon nos besoins dans Zotero (UDL-141)*. Un autre préfère la base interne à HAL car *on ne fait pas une énorme confiance technique dans HAL, en termes d'export et des choses comme ça (UGA-086)*.

Ces outils nécessitent des compétences informatiques (voire d'un informaticien) et servent à produire le rapport d'activité et le rapport Hcéres et aussi les références pour les pages web du laboratoire. Apparemment, l'enjeu majeur est la synchronisation avec HAL, comme avec l'outil de conversion Ad HALstra (INSU) pour exporter les références de la base interne vers HAL. D'autres outils ont été mis en place pour la synchronisation inverse ; une unité mixte en STIC a développé un outil bibliométrique (LabMetry) qui *aspire toutes les nuits la production scientifique sur HAL pour mettre à jour la base de données et pour alimenter d'une manière automatique et par programmation dynamique le site web du laboratoire et les pages web des chercheurs* mais permet aussi de gérer les doublons. *Il y a une dizaine d'années, j'ai lancé au laboratoire un programme de développement d'un outil de bibliométrie qui nous sert très souvent, notamment pour les évaluations Hcéres mais pas que. L'objectif était d'avoir un outil qui permet d'avoir très rapidement un état de la production scientifique du laboratoire, d'avoir très facilement des graphiques de la production scientifique – pas seulement des histogrammes mais aussi les courbes d'évolution de l'impact factor des revues (...) Cet outil nous permet aussi d'alimenter les documents de bilan, comme par exemple pour le Hcéres (...) C'est un outil extrêmement utile pour nous et qui repose sur HAL. Dans LabMetry, on a l'organigramme du laboratoire, avec l'ensemble des individus par équipe, et ainsi nous avons une bibliométrie à tous les niveaux de granularité du laboratoire (UDL-113)*. Mais un tel développement sur mesure reste également l'exception, comme le code en amont de HAL (cf. *supra*).

Les dynamiques

L'utilisation de HAL n'est pas figée dans le temps. Souvent, les répondants racontent des trajectoires, une situation dynamique avec un avant et un après. Ainsi, la responsable administrative de laboratoire explique que *avant, on alimentait HAL automatiquement à partir de la base de données interne de publications (...) En 2008, chargement massif des anciennes publications dans HAL (...) À partir de ce moment-là, le nombre de documents (...) a fortement augmenté (UGA-018)*. Une autre trajectoire, rapportée par le directeur d'une UMR : *Aujourd'hui le transfert de HAL vers arXiv fonctionne bien mais avant, les chercheurs utilisaient beaucoup arXiv, et notre institut a développé un outil qui devait déverser toutes les publications d'arXiv dans HAL (UCA-034)*.

Parfois, le changement est intervenu après le départ d'une personne : *Il y avait une base de données développée en interne pour gérer les publications du labo mais on l'a arrêtée quand HAL est arrivé et quand l'informaticien qui s'en occupait est parti à la retraite (UCA-033)*. Ici, HAL est décrit comme une solution pour parer à la perte d'une compétence-clé.

Certains changements sont d'ordre institutionnel, comme la fusion de plusieurs laboratoires ou la disparition d'un établissement : (...) *avant on réalisait une exportation via Madis (...) la base institutionnelle (...) Sauf que cette année cela va s'arrêter, car c'était historique IFSTTAR (UDL-096).* Ces changements institutionnels peuvent créer des problèmes techniques : *Notre laboratoire a été créé en 2018 après la fusion de deux laboratoires, et ça ajoute un peu de confusion (...) En début d'année, on a remarqué qu'il manquait beaucoup de prépublications dans notre page web parce que le script récupérait seulement un identifiant et pas l'autre (USPC-138).* Même genre de soucis dans un autre laboratoire : *Après la création du laboratoire (suite à la fusion de deux autres laboratoires), il y a un versement dans HAL d'anciennes publications de l'un des deux laboratoires, jusqu'en 2016 (...) Suite à la fusion, on a encore des publications qui traînent un peu à gauche et à droite (UCA-034).*

D'autres changements sont uniquement d'ordre technique : *Il y a eu un problème il y a quelques années avec le transfert vers arXiv (...) Je me souviens que dans le passé le flux était dans les deux sens, entre HAL et arXiv. Après, le flux d'arXiv vers HAL n'était plus possible mais je ne me souviens plus des détails (USPC-138).* Ce n'est pas uniquement un problème avec HAL : *La difficulté, c'est que les API changent, et nous avons un ingénieur d'étude (...) au laboratoire qui s'occupe du suivi de l'outil. Dès qu'il y a un changement sur le Web of Science, sur CoRR¹¹ ou sur HAL, il met à jour l'outil pour pouvoir continuer d'alimenter la base de données (UDL-113).*

Parfois, le changement est en cours ou programmé, par exemple à l'occasion de la prochaine campagne d'évaluation : *Auparavant le laboratoire avait une base de données bibliographiques en interne, sur un serveur qui sera bientôt désactivé (...) Nous allons passer sur Wordpress (USPC-164).* Même situation dans une UMR qui utilisera HAL pour la prochaine évaluation Hcéres, *car avant on réalisait une exportation via Madis (...) Madis va disparaître cette année, je vais donc verser directement dans HAL au nom de mes collègues (UDL-096).* À la question de savoir si le laboratoire utilisait des outils, le directeur d'une autre UMR répond par *non, aucun outil (...) on verra lors de la prochaine campagne d'évaluation dans deux ans si on va mettre en place ce genre d'outils (Saclay-156).* Dans un petit laboratoire universitaire en SHS, *pour l'instant, il n'y a pas de liaison entre HAL et les pages personnelles des chercheurs sur le site web du labo (...) mais c'est un projet à venir (Saclay-081).* Un dernier exemple : *Il y a un projet pour le site de « plugger » directement avec HAL, pour que chaque notice déposée sur HAL soit automatiquement importée dans le site, sur la page de chaque chercheur (SU-118).*

Changement ne veut pas dire discontinuité ; la documentaliste d'une UMR témoigne ainsi que *avec le directeur précédent, dans le cadre de la dernière évaluation, il y avait une personne en informatique, avec qui j'avais travaillé pour qu'elle puisse faire certains rapports (...) Cette personne est partie et nous n'avons pas le même directeur ; l'actuel a repris la continuité, donc il joue la carte de HAL. Un stagiaire en informatique va développer un outil spécifique (UCA-013).*

Plusieurs réponses font le lien explicite avec l'environnement institutionnel. Une UMR constate qu'il n'y a pas d'outil au niveau du laboratoire et que les collègues utilisent des outils

¹¹ Computing Research Repository <https://dblp.org/db/journals/corr/index.html>

comme Zotero ou Mendeley pour ajouter que cela va changer car *Paris-Saclay nous incite à avoir un compte ORCID alimenté par HAL* (Saclay-233). Une autre UMR explique que les collègues ont commencé à utiliser Zotero car *il y a eu une grosse campagne de formation sur Zotero il y a quelque temps* (SU-089). D'une manière similaire, un certain nombre de laboratoires raconte la mise en place d'une collection sur HAL pour le Hcéres, à la demande d'une tutelle et/ou avec l'aide d'un SCD.

D'une manière plus générale, un directeur d'UMR admet que *depuis que le CNRS a signé la Déclaration de San Francisco (DORA), les critères de promotion et de recrutement ont beaucoup changé (et que) la bibliométrie est passée au plan secondaire. (Néanmoins) on maintient ces indicateurs d'impact basés sur les citations tant que nos instances ne proposent pas d'autres procédures ou d'autres outils pour une autre forme d'évaluation. On est conscient de l'importance du qualitatif mais pour l'instant, on reste avec les indicateurs traditionnels* (UB-033).

En résumé, les entretiens révèlent des pratiques en évolution, dans des situations dynamiques dont les variables indépendantes se situent à plusieurs niveaux :

- organisation : création ou fusion de structures (unités, établissements...);
- institution : incitation et/ou soutien (formation...);
- Hcéres : nouvelle campagne d'évaluation ;
- ressources humaines : départ ou arrivée d'une nouvelle personne (direction, informatique, documentation) ;
- dispositifs : création ou disparition d'outils, de fonctionnalités ou de services.

Les entretiens montrent que ces différents facteurs sont liés, et que les pratiques dans les laboratoires se développent et évoluent sur des trajectoires plus ou moins longues. Mais ce que les entretiens révèlent surtout, c'est la capacité d'adaptation des laboratoires à faire face aux défis et enjeux internes et externes.

Discussion

Les études antérieures ont mis en lumière la diversité des pratiques individuelles et collectives de HAL, dans le cadre d'un écosystème dynamique constitué de variables politiques, disciplinaires, institutionnelles, organisationnelles et individuelles. L'analyse qualitative avec un panel de 50 entretiens confirme cette diversité mais révèle surtout la capacité d'adaptation des laboratoires, leur flexibilité face à un environnement complexe et en constante évolution. D'une certaine manière, les réponses esquissent un modèle d'évolution où le laboratoire, l'espace de travail et l'environnement principal et immédiat pour les activités scientifiques des chercheurs, est en perpétuelle transformation, subissant au fil du temps quelques modifications structurelles et fonctionnelles majeures. Par la suite, nous allons discuter les résultats sous trois aspects : les discours et les communautés au sein des laboratoires, le périmètre et les enjeux des dispositifs mis en place par les laboratoires, et la transformation de HAL vers une infrastructure de suivi et d'évaluation de la recherche.

Discours et communautés

Tous les laboratoires ont au moins une partie de leurs publications sur HAL, presque tous utilisent la plateforme ou ont l'intention de le faire. En revanche, ils ne le font pas de la même manière et pour les mêmes raisons. Si certains voient HAL avant tout comme un moyen efficace pour la diffusion des publications, d'autres expliquent l'utilisation de HAL surtout par l'obligation d'établir les bilans individuels des chercheurs ; dans quelques laboratoires, HAL est devenu l'outil central et incontournable pour le suivi et l'évaluation de la performance scientifique.

Aussi, l'analyse des réponses concernant l'intérêt de HAL révèle trois types de discours :

1. un discours administratif qui met l'accent sur l'incitation institutionnelle, sur la demande de la part de la tutelle : *cela devient une obligation pour les chercheurs CNRS (SU-089)* ;
2. un discours politique voire idéologique qui souligne l'intérêt d'utiliser HAL pour une recherche meilleure, plus responsable et plus ouverte à la société : *une démarche très romantique, dans le sens où c'est un bel idéal (...) un peu utopique pour l'instant (UCA-036)* ;
3. un discours utilitaire et pragmatique qui met en exergue l'usage de HAL pour le bon fonctionnement et le développement de l'unité : *Pour les bilans annuels du laboratoire. Nous avons besoin de l'information sur le nombre de publiants car notre budget en dépend en partie. On utilise HAL aussi pour assurer le suivi de l'utilisation d'un certain nombre de ressources (Saclay-106).*

Le discours pragmatique paraît majoritaire, en compétition parfois avec un discours plus politique sur l'intérêt de la science ouverte. Parfois, une perception plutôt mitigée de l'utilité de HAL et de sa facilité d'utilisation semble en contradiction avec l'opinion favorable de la science ouverte : *De manière générale, HAL c'est génial ; je pense que la connaissance n'appartient à personne, surtout quand elle est financée en très grande partie par l'État (...) Mais, techniquement, cette plateforme ne marche pas comme elle devrait (...) Un concept très positif, mais quelques soucis techniques (AMU-074).* Cette expression d'une dissonance cognitive entre l'opinion et l'expérience n'est pas un cas isolé.

Les entretiens révèlent par ailleurs plusieurs manières de vivre avec, voire diminuer cette dissonance. Se convaincre et/ou convaincre les collègues de l'utilité de HAL est une option, tout comme insister sur l'obligation d'avoir recours à HAL pour les bilans et l'évaluation, en se rassurant de ne pas avoir le choix. Miser sur l'amélioration progressive de HAL en est une autre.

De leur côté, les outils et les dispositifs au sein des laboratoires contribuent à leur manière à réduire cette dissonance, dans la mesure où ils augmentent l'utilité de HAL (par exemple, par une plus grande granularité des bilans), tout en facilitant son utilisation (par exemple, par l'import de plusieurs références et/ou de documents). Mais la mise en place de ces dispositifs témoigne d'un autre élément caractéristique de l'écosystème HAL, le développement d'une voire de plusieurs communauté(s) de pratique qui dépassent les laboratoires.

Ces communautés de pratique et réseaux métier ont été décrits par Granger et Janik (2021), y compris par rapport à leur articulation et leur complémentarité avec les services

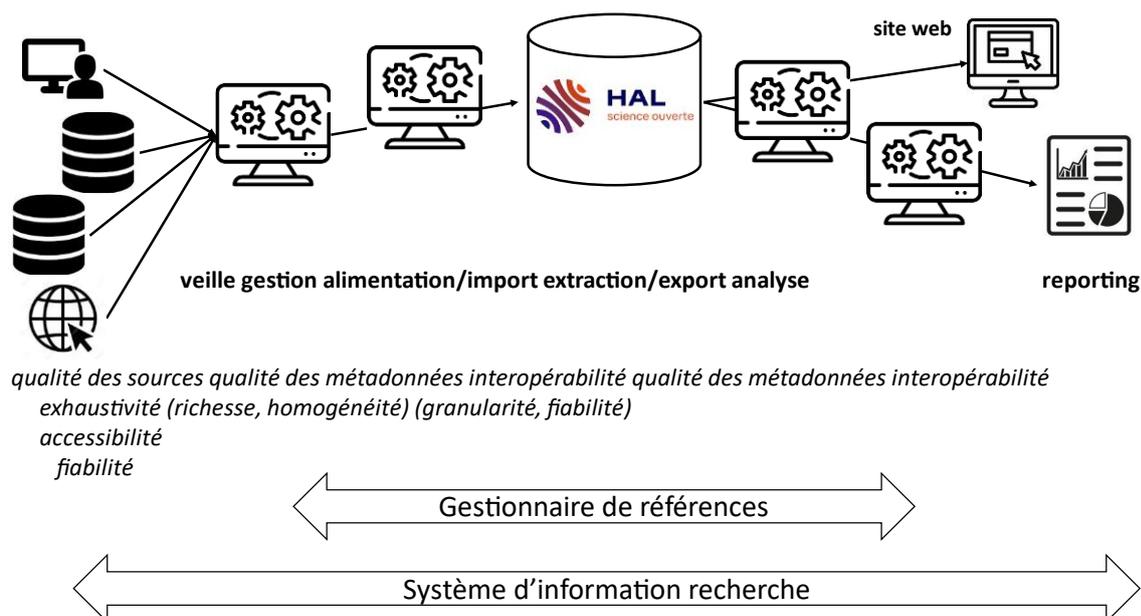
institutionnels, tels que les SCD : « (...) la création en 2016 d'un réseau d'utilisateurs tel que CasuHAL a répondu quant à elle au besoin des utilisateurs professionnels de mutualiser leurs connaissances à l'échelle inter-établissements afin de faire évoluer leurs compétences. Cet acteur national offre ainsi un cadre d'échanges complémentaire au cadre local » (Granger et Janik, 2021, §26 ; cf. aussi Alarcon et al., 2023). L'utilisation d'outils développés par d'autres laboratoires et d'autres services, le partage d'une solution interne, la demande et l'échange de conseil, l'entraide et la formation sont un thème récurrent des entretiens et reflètent la réalité de ces communautés. Un exemple parmi d'autres : *L'outil est open source et potentiellement utilisable par d'autres (...) Nous avons participé à des groupes de travail à l'université de Lyon, à l'époque il y avait un référent HAL, un ingénieur, qui rassemblait tous les directeurs de laboratoire pour échanger sur la politique des laboratoires vis-à-vis de HAL et pour partager les outils des uns et des autres. L'École des Mines de Saint-Étienne avait demandé de pouvoir utiliser l'outil et on leur avait fourni tous les éléments mais je ne sais pas s'ils l'utilisent réellement (UDL-113).*

D'après les entretiens, ces communautés de pratique sont souvent du domaine de la documentation mais incluent aussi, plus rarement, l'informatique, la direction ou bien l'administration des laboratoires.

Périmètre et enjeux des dispositifs

Les entretiens permettent de dresser un bilan détaillé des outils et dispositifs mis en place par les laboratoires dans le cadre de leurs pratiques de HAL. Certains sont positionnés en amont de HAL, pour assurer la veille de la production scientifique, gérer les références voire alimenter HAL. D'autres se situent en aval et servent à mettre à jour une base interne, alimenter des pages Web, analyser la production et/ou produire des bilans. La figure 5 montre l'articulation de ces outils – et de leurs principales fonctions (en gras) – avec HAL.

Figure 5. Enjeux et périmètre des outils et dispositifs mis en place par les laboratoires



Aux fonctions des outils mis en place correspondent des enjeux spécifiques (en italique dans la figure 5), en particulier :

- la veille : la qualité des sources exploitées et la fiabilité de leurs données, leur exhaustivité et aussi leur accessibilité ;
- la gestion : la qualité des métadonnées, leur richesse et aussi leur homogénéité, par exemple quand elles proviennent de sources différentes ;
- l'import/export : l'enjeu majeur est l'interopérabilité des formats, par exemple la typologie des documents, la nomenclature des domaines scientifiques, les identifiants des personnes et des organisations etc. ;
- l'analyse : comme pour la gestion, l'enjeu principal est la qualité des métadonnées, leur fiabilité et aussi leur granularité (axes, projets, équipes...) ;
- le reporting : le problème est surtout l'interopérabilité avec les systèmes de suivi et d'évaluation des établissements et organismes et du Hcéres.

Les différents outils proposent des solutions mais, d'après les entretiens, ne sont pas toujours à la hauteur des enjeux. Aussi, plusieurs actions documentaires manuelles sont-elles nécessaires pour assurer une certaine qualité de ces dispositifs, à différents niveaux – il faut sélectionner des publications et compléter les listes, dédoublonner et trier des notices, corriger et enrichir les références, sans parler des tâches habituelles d'administration d'une collection sur HAL (gestion et suivi des identifiants, tamponnage automatique et manuel, etc.). Ces actions sont généralement prises en charge par la ou le documentaliste du laboratoire ; mais tous les laboratoires ne disposent pas d'une fonction documentaire, et dans ce cas, elles reviennent à la direction ou l'administration du laboratoire, voire à un ou plusieurs collègues désignés, un(e) doctorant(e) ou un(e) vacataire (ou stagiaire). Ou bien, alternativement, la qualité est simplement considérée comme affaire de tous.

Certains laboratoires sont outillés, d'autres pas ; certains ont les ressources nécessaires pour faire ce travail correctement (postes, compétences, équipement) tandis que d'autres constatent que *honnêtement, la seule chose dont on aurait besoin, c'est la liste de nos publications dans HAL en format BibTex (ou n'importe quel format bien documenté) et regrettent que on n'a pas les outils d'INRIA. On aurait besoin d'une interface de programmation avec HAL, un truc qu'on peut scripter et qui nous sort des choses dans un format qu'on peut traiter facilement* (UGA-086).

Ajoutons deux observations.

Le périmètre des dispositifs. Les laboratoires ont mis en œuvre des dispositifs à périmètre variable. Certains laboratoires se limitent à la gestion des références de leur production scientifique, avec des outils de gestion, d'import et d'export, où HAL joue essentiellement le rôle d'une base de données bibliographiques, parfois en remplacement d'une base interne (cf. figure 5). D'autres laboratoires vont plus loin, avec des dispositifs qui, à l'instar de LabMetry, couvrent au moins partiellement les fonctions d'un système d'information sur la recherche, notamment pour assurer le suivi des projets scientifiques et fournir des informations nécessaires aux bilans, aux rapports d'activité et aux campagnes d'évaluation (De Castro, 2018), même si très peu de laboratoires semblent utiliser des logiciels fonctionnels et adaptés, comme SAMPRA ou QUASAR.

La qualité des outils. Sont-ils provisoires ou temporaires (rustines), ou représentent-ils une solution plus durable, avec une certaine stabilité ? Qui en assure la maintenance et le développement ? Dans la mesure où certains sont complémentaires tandis que d'autres remplissent des fonctions similaires, plus ou moins utiles et adaptés et plus ou moins faciles à utiliser, existe-t-il une sorte de marché, une offre clairement identifiée ? Et pour aller plus loin, à quel degré sont-ils ouverts, standardisés, interopérables avec d'autres systèmes ; ou pour utiliser un terme actuel, à quel degré sont-ils FAIR ?

La transformation de HAL

Nous avons analysé ailleurs l'évolution de HAL d'une plateforme d'auto-archivage vers une infrastructure avec des fonctionnalités d'un système d'information sur la recherche (Schöpfel et al., 2023a et 2023b), en insistant sur le fait qu'il ne s'agit pas d'un phénomène spécifiquement français mais d'une transformation globale du paysage des archives ouvertes (Schöpfel & Azeroual, 2021). Plusieurs laboratoires ont parfaitement intégré cette transformation, à l'instar d'une UMR de Grenoble dont le directeur reconnaît que *chaque fois qu'on a besoin d'un indicateur, on utilise HAL* (UGA-031).

Le CCSD, opérateur technique de HAL, est conscient du risque de cette évolution qui risque « de brouiller l'ambition initiale de communication scientifique directe et de privilégier le signalement au détriment de l'accès » (Berthaud et al., 2021). Cependant, au moment où, dans le cadre du projet HAL+, les procédures de HAL commencent à changer (Barborini et al., 2023), posons la question de la transformation d'une autre façon, du point de vue des laboratoires. Quelle stratégie faut-il privilégier ? Le projet en cours prévoit l'intégration des nouvelles fonctionnalités dans l'offre de services de HAL, au même titre que le dépôt, la recherche ou la consultation. En d'autres mots, les nouvelles fonctionnalités remplaceront à terme les outils mis en place dans les laboratoires, ce qui correspond à externaliser certaines fonctions des laboratoires vers le CCSD et les SCD.

Ce processus répond au constat de la diversité des situations dans les laboratoires et à une demande de la part des établissements et organismes tutelles. Cependant, il faut être conscient qu'une telle externalisation ou délégation peut être vécue comme une sorte de perte de contrôle et de compétences par les laboratoires concernés. Deux enjeux : Que deviennent les outils « *on top of HAL* », au moins partiellement développés et maintenus comme logiciels libres par une communauté de pratique ? Et quid de la qualité des métadonnées, de leur richesse, des identifiants etc., qui est d'une importance cruciale pour la perception de l'utilité de ces nouvelles fonctionnalités par les chercheurs et surtout, par leurs laboratoires ?

Conclusion

La science ouverte modifie en profondeur les modalités d'accès et de diffusion de l'information scientifique et technique. L'ouverture des publications s'appuie en partie sur le dépôt dans une archive ouverte publique comme HAL. Initialement conçue pour « faciliter et accélérer la diffusion des publications scientifiques suivant les principes d'une communication scientifique directe entre pairs » (Berthaud et al., 2021), HAL joue aujourd'hui un rôle central dans la politique de la science ouverte en France.

Dans le cadre du projet HAL/LO, nous avons essayé de mieux comprendre l'utilisation de HAL par les laboratoires de recherche, à travers des approches quantitatives (Schöpfel et al., 2023a) et qualitatives (Schöpfel et al., 2022). Nous avons ainsi démontré la transformation de HAL vers une infrastructure de suivi et d'évaluation de la performance scientifique des établissements et organismes, des laboratoires et des personnels (Schöpfel et al., 2023b).

L'analyse des contenus de 50 entretiens avec des représentants de laboratoires confirme ces résultats, avec une évidence empirique riche et détaillée du terrain. Les entretiens reflètent la coexistence et parfois la dissonance de trois types de discours sur HAL au sein des laboratoires (administratif, politique, pragmatique) et témoignent de la création de communautés de pratique autour de HAL et autour des dispositifs mis en place en amont et en aval de HAL dans une partie des laboratoires. Ces dispositifs ont différents périmètres fonctionnels ; dans certains laboratoires, ils correspondent à des gestionnaires de références tandis qu'ailleurs, ils ont des similitudes avec des systèmes d'information sur la recherche.

Les entretiens montrent surtout une grande diversité de situations et d'approches de la plateforme HAL qui ne sont pas stables mais qui évoluent sur des trajectoires plus ou moins longues, révélant la capacité d'adaptation des laboratoires pour faire face aux défis et enjeux internes et externes.

Par la suite, nous allons poursuivre l'analyse qualitative des pratiques des laboratoires dans leurs dimensions systémique et fonctionnelle. Dans un premier temps, nous allons analyser quelques pratiques représentatives de HAL par les laboratoires, y compris l'utilisation d'autres plateformes, ainsi que les avis récoltés à propos de HAL. Cette approche permettra de mieux comprendre les facteurs de l'acceptation et de l'acculturation de HAL dans les laboratoires, les différents types de services d'accompagnement mis en œuvre par les laboratoires, et le positionnement complexe et unique de HAL à travers une analyse de ses équivalents fonctionnels.

Dans un deuxième temps, nous allons élargir la perspective afin d'étudier, toujours à partir des 50 entretiens, les différentes approches d'incitation et d'accompagnement des pratiques collectives et individuelles au sein et par l'environnement des laboratoires, notamment au niveau des universités. Nous allons également ajouter un regard sur les sites Web des laboratoires pour savoir dans quelle mesure ils font le lien avec HAL et valorisent les publications des chercheurs.

HAL est sorti de sa phase « de jeunesse » pour devenir un partenaire obligé et essentiel des chercheurs et de leurs unités de recherche (Schöpfel et al., 2022). Ainsi, il est devenu un objet privilégié de la science sur la science, de l'étude des pratiques des chercheurs et de leurs communautés et institutions.

Pour cette raison, les enjeux à venir pour HAL seront également les pistes pour de futurs projets de recherche. Actuellement, nous pouvons anticiper trois perspectives.

Directement lié à la transformation en cours de HAL, l'étude de ses contenus et services par rapport aux systèmes d'information sur la recherche s'impose. Il s'agit là en premier lieu de l'exhaustivité, de la pertinence et de la qualité des données et métadonnées de HAL, y compris par rapport aux identifiants pérennes, avec des questions concernant l'alimentation,

de la curation, de la gestion et de l'usage de HAL, en tenant compte également de la dimension éthique et politique (gouvernance).

L'évaluation de la recherche est en train de changer, vers davantage de diversité, d'équité et d'ouverture vers la société. Si la *Déclaration de San Francisco sur l'évaluation de la recherche* (DORA)¹² a tracé le chemin, la coalition européenne CoARA¹³ commence à publier des recommandations pour la mise en œuvre de cette réforme. Quel sera l'impact sur HAL ? HAL contient déjà beaucoup de ressources issues de la bibliodiversité, et HAL est ouvert.

Le troisième enjeu est sans doute le développement rapide des outils de l'intelligence artificielle (IA). Le potentiel de l'IA pour HAL est évident, aussi bien pour la qualité des données et métadonnées (enrichissement) et la découvrabilité des dépôts (recommandations) que pour l'analyse des contenus. L'analyse de ce potentiel mais aussi de ses risques sera certainement l'une des pistes prioritaires pour la science sur la science en général et pour l'analyse du dispositif de HAL en particulier.

Remerciements

Nous tenons à remercier chaleureusement tous nos interlocuteurs pour l'intérêt qu'ils ont porté à notre projet de recherche, pour leur disponibilité et pour leur franchise.

Nous remercions également les étudiants stagiaires à Lille et à Rennes pour leur contribution à la préparation, la réalisation et l'analyse des entretiens (Aline Benvegnú dos Santos, Marine Chochoy, Chamseddine Hammouda, Odélin Jousset).

Le projet HAL/LO (Valorisation sur HAL de la production des laboratoires dans l'environnement de la science ouverte) a reçu un financement dans le cadre d'un appel à projet du GIS « Réseau URFIST », de 2020 à 2022.

Bibliographie

Acquier, F., Bizien, L., & Charles, L.-A. (2023). Apprendre à penser science ouverte dans un laboratoire de recherche en architecture : l'UMR CNRS Ambiances Architectures Urbanités (AAU). *Culture et Recherche*, (144), 20–22. Disponible sur : <https://www.culture.gouv.fr/fr/Thematiques/Enseignement-superieur-et-Recherche/La-revue-Culture-et-Recherche/La-science-ouverte> (Consulté le 07/02/2024)

Alarcon, N., Bidégaray-Fesquet, B., Boileau, N., Chancerel, G., Filatre, A., Gibier, F., ... Schmitt, A. (2023). *Enquête Adhérents 2023 : synthèse des résultats*. CasuHAL. https://hal.science/sic_04397787v1

¹² DORA <https://sfdora.org/read/read-the-declaration-french/>

¹³ CoARA <https://coara.eu/>

Barborini, Y., Kuntziger, B., & Chan, P. (2023). A new way for researchers to deposit files in HAL, the French Open Archive. *Open Science FAIR 2023*, September 26, 2023. <https://doi.org/10.5281/zenodo.8402880>

Berthaud, C., Charnay, D., & Fargier, N. (2021). Diffuser et pérenniser le savoir scientifique : 20 ans d'histoire de HAL. *Histoire de La Recherche Contemporaine*, 10(2), 165-176. <https://doi.org/10.4000/hrc.6330>

Bracco, L. (2022). Mesurer l'ouverture de la science : le cas de l'Université de Lorraine. *Revue française des sciences de l'information et de la communication*, (24). <https://doi.org/10.4000/rfsic.12474>

Bracco, L., L'Hôte, A., Jeangirard, E., & Torny, D. (2022). *Extending the open monitoring of open science: A new framework for the French Open Science Monitor (BSO)*. Preprint. <https://hal.science/hal-03651518>

Coulhon, T., & Larrourou, B. (2023). La science ouverte et l'évaluation de la recherche. *Culture et Recherche*, (144), 99–100. Disponible sur : <https://www.culture.gouv.fr/fr/Thematiques/Enseignement-superieur-et-Recherche/La-revue-Culture-et-Recherche/La-science-ouverte> (Consulté le 07/02/2024)

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, 13(3), 319. <https://doi.org/10.2307/249008>

De Castro, P. (2018). Mapping the European CRIS infrastructure and its potential applications. *Antwerp ECOOM Workshop "Working with National Bibliographic Databases for Research Output"*, 10-11 September 2018. <https://dspacecris.eurocris.org/handle/11366/705>

Granger, S., & Janik, J. (2021). Se fédérer pour évoluer : les communautés de pratique de l'information et de l'édition scientifiques. *Histoire de La Recherche Contemporaine*, 10(2), 156–164. <https://doi.org/10.4000/hrc.6192>

Jonchère, L., & Troccaz, O. (2023). HALUR : des outils connectés à HAL au service des chercheurs. *Culture et Recherche*, (144), 40–41. Disponible sur : <https://www.culture.gouv.fr/fr/Thematiques/Enseignement-superieur-et-Recherche/La-revue-Culture-et-Recherche/La-science-ouverte>

Larrieu, M., & Schöpfel, J. (2022). Éclairer les différences disciplinaires dans le contexte de la science ouverte. Analyse avec les publications de l'archive ouverte HAL. *DocSoc 2022. 8e Conférence Document Numérique & Société*. 23-24 Juin 2022, Liège (Belgique).

Lubino, P. (2023). Vos besoins de formation à HAL : retour sur l'enquête. *Blog du CCSD*, 21 septembre 2023. Disponible sur : <https://www.ccsd.cnrs.fr/2023/09/vos-besoins-de-formation-a-hal-retour-sur-l-enquete/> (Consulté le 07/02/2024)

Mahé, A., & Prime-Claverie, C. (2017). Qui dépose quoi sur Hal-SHS ? Pratiques de dépôts en libre accès en sciences humaines et sociales. *Revue française des sciences de l'information et de la communication*, 11. <https://doi.org/10.4000/rfsic.3315>

MESR (2021). *Deuxième Plan national pour la science ouverte*. Paris : Ministère de l'Enseignement supérieur et de la Recherche. Disponible sur : <https://www.enseignementsup-recherche.gouv.fr/fr/le-plan-national-pour-la-science-ouverte-2021-2024-vers-une-generalisation-de-la-science-ouverte-en-48525> (Consulté le 07/02/2024)

Reymond, D., Tabariès, A., & Bara, L. (2023). *Accompagnement et visualisations de la Science Ouverte (SoVisu). Version 1.0.2*. Toulon : IMSIC. <https://hal.science/hal-04103412/>

Reynier, P., & Bellemain, P. (2023). OCdHAL, un outil de contrôle des publications déposées dans HAL. *Culture et Recherche*, (144), 42–43. Disponible sur : <https://www.culture.gouv.fr/fr/Thematiques/Enseignement-superieur-et-Recherche/La-revue-Culture-et-Recherche/La-science-ouverte> (Consulté le 07/02/2024)

Schöpfel, J., & Azeroual, O. (2021). Current research information systems and institutional repositories: From data ingestion to convergence and merger. In D. Baker & L. Ellis (Eds.), *Future Directions in Digital Information* (pp. 19–37). Oxford : Chandos. <https://doi.org/10.1016/B978-0-12-822144-0.00002-1>

Schöpfel, J., Kergosien, E., Prost, H., & Barrié, J. (2022). « Pas si simple que ça... » : Une enquête sur l'usage de HAL par les unités de recherche des universités IdEx. *I2D - Information, Données & Documents*, n° 2(2), 150–183. <https://doi.org/10.3917/i2d.222.0150>

Schöpfel, J., Thiault, F., Prost, H., Jacquemin, B., & Kergosien, E. (2023a). L'utilisation de HAL par les laboratoires de recherche. Une étude quantitative. *Balisages*, 6. <https://dx.doi.org/10.35562/balisages.1166>

Schöpfel, J., Chaudiron, S., Jacquemin, B., Kergosien, E., Prost, H., & Thiault, F. (2023b). The Transformation of the Green Road to Open Access. *Publications*, 11(2), 29. <https://doi.org/10.3390/publications11020029>

Annexes

Annexe 1 : La trame de l'entretien

Informations préalables

1. Quand a lieu l'entretien ?

2. Prénom et nom de la personne contact

3. Fonction

- Directrice/Directeur du laboratoire
- Responsable de l'administration du laboratoire
- Documentaliste
- Autre (veuillez préciser)

4. Participation à l'enquête en ligne ?

- Oui
- Non

5. Laboratoire : identifiant du projet HAL/LO

6. Laboratoire : acronyme

7. Laboratoire : nom

8. Laboratoire : statut

- Une unité de recherche universitaire (EA, UR ou ULR)
- Une unité mixte de recherche (UMR)
- Une unité propre de recherche d'un organisme de recherche - CNRS, INSERM etc. (UPR, U)
- Autre (veuillez préciser)

9. Laboratoire : université de rattachement

- Aix-Marseille Université
- Sorbonne Université
- Université de Lyon
- Université Côte d'Azur (UCA)
- Université de Bordeaux
- Université de Paris
- Université de Strasbourg
- Université Grenoble-Alpes
- Université Paris-Saclay
- Université Paris Sciences et Lettres (PSL)

- Autre (veuillez préciser)

10. Laboratoire : domaine scientifique (RNSR)

- 1 Mathématiques et leurs interactions
- 2 Physique
- 3 Sciences de la terre et de l'univers, espace
- 4 Chimie
- 5 Biologie, médecine et santé
- 6 Sciences humaines et humanités
- 7 Sciences de la société
- 8 Sciences pour l'ingénieur
- 9 Sciences et technologies de l'information et de la communication
- 10 Sciences agronomiques et écologiques
- Autre (veuillez préciser)

Collection et dépôts HAL

Les questions suivantes concernent la collection HAL du laboratoire (la plupart de ces laboratoires en possèdent une) et les dépôts sur HAL. Y a-t-il quelqu'un qui dépose systématiquement les publications du labo sur HAL ? Y a-t-il un autre soutien ? Pour situer cette information dans le contexte, la section commence par une estimation du nombre de chercheurs et du nombre des publications.

11. Combien de chercheurs ?

12. Nombre de publications par an ?

13. Existe-t-il une collection sur HAL ?

- Oui
- Non
- Projet en cours
- Commentaires, remarques

14. Si oui, quand a-t-elle été créée ?

15. Les dépôts dans HAL sont-ils effectués de manière systématique ?

- Oui
- Non
- Si oui, par qui et comment ?

16. Existe-t-il un accompagnement, une aide, des formations pour guider les dépôts ?

- Oui
- Non
- Si oui, comment ?

Utilisation et valorisation de HAL

La troisième section porte sur l'usage de HAL fait par le laboratoire. A quoi sert HAL ? Y a-t-il d'autres outils ? Est-ce que tout le monde fait pareil, ou y a-t-il des différences au sein du personnel du laboratoire ? La dernière question est plus large ; elle concerne l'évaluation par le Hcéres mais permet de parler de la science ouverte d'une manière plus générale.

17. Comment utilisez-vous HAL dans votre laboratoire ?

- Pour les bilans individuels (CRAC, RIBAC...)
- Pour les campagnes d'évaluation par le Hcéres
- Pour le site web du laboratoire
- Pour le suivi de l'activité du laboratoire
- Pour autre chose (veuillez préciser)

18. Avez-vous des outils particuliers pour préparer les bilans (logiciels, codes...) ?

- Pour alimenter HAL
- Pour extraire des références de HAL
- Pour gérer les références (Zotero, RefWorks, EndNote...)
- Pour autre chose
- Commentaires, détails...

19. Tout le monde fait pareil ? Ou y a-t-il des différences entre les équipes, axes etc. ?

20. Et entre les chercheurs et les enseignants-chercheurs (dans les UMR) ?

21. Dans votre domaine, à part HAL, y a-t-il d'autres sites ou réseaux pour les publications ?

- Je ne sais pas
- Non, je ne crois pas
- Oui (veuillez préciser)

22. Lors de la dernière campagne d'évaluation, est-ce qu'il a été question de la science ouverte ?

L'environnement

Les dernières questions cherchent à mieux comprendre quelques facteurs-clés pour l'usage de HAL : l'environnement institutionnel, l'impact sur le laboratoire, mais aussi l'opinion qu'ont les chercheurs eux-mêmes de HAL.

23. Votre université a-t-elle une archive institutionnelle ?

- Oui, sur HAL (portail HAL)
- Oui, en dehors de HAL
- Non, mais il y a un projet
- Non
- Commentaire, remarques...

24. Quelle est la politique de science ouverte de votre université ?

- Il existe une charte en faveur du libre accès
- Incitation à diffuser les publications en libre accès
- Obligation à diffuser les publications en libre accès
- Autre chose
- Commentaire, remarques...

25. Comment décrire le rôle du laboratoire dans cet environnement ?

26. Et par rapport au CNRS ou aux autres organismes (surtout pour les UMR) ?

27. D'une manière générale, que pensez-vous de HAL ? Plutôt satisfait, plutôt critique ?

28. Et les autres membres du laboratoire, qu'en pensent-ils ?

Annexe 2 : L'échantillon de l'enquête

Nombre de laboratoires par université, avec la discipline dominante (classification RNSR).

Discipline Université	1	2	3	4	5	6	7	8	9	10	Total
	Aix-Marseille				1		2		3		1
Bordeaux				1			2			2	5
Côte d'Azur	1	1	1			2					5
Grenoble Alpes			1	1	1	1	1		2		7
Lyon			1	1			2	2	1	1	8
Paris Cité	1					1	2				4
Paris Sciences & Lettres	1										1
Paris-Saclay		1				1		1	1	1	5
Sorbonne Université		2	1		2	1		1			7

Strasbourg			1								1
Total	3	4	5	4	3	8	7	7	4	5	50

Discipline : Classification RNSR

- 1 Mathématiques et leurs interactions
- 2 Physique
- 3 Sciences de la terre et de l'univers, espace
- 4 Chimie
- 5 Biologie, médecine et santé
- 6 Sciences humaines et humanités
- 7 Sciences de la société
- 8 Sciences pour l'ingénieur
- 9 Sciences et technologies de l'information et de la communication
- 10 Sciences agronomiques et écologiques