



HAL
open science

Utiliser les archives ouvertes pour valoriser ses travaux l'exemple de HAL-SHS

Agnès Magron

► To cite this version:

Agnès Magron. Utiliser les archives ouvertes pour valoriser ses travaux l'exemple de HAL-SHS. Catherine Schnedecker, Angelina Aleksandrova. Le doctorat en France : mode(s) d'emploi, Peter Lang, pp. 227-239, 2017, 9782807606210. sic_01697164

HAL Id: sic_01697164

https://archivesic.ccsd.cnrs.fr/sic_01697164v1

Submitted on 31 Jan 2018

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.

Utiliser les archives ouvertes pour valoriser ses travaux

l'exemple de HAL-SHS

Agnès Magron

Centre pour la Communication Scientifique Directe (CCSD)

Introduction : le libre accès

La circulation des savoirs est au cœur des enjeux du mouvement du libre accès – ou open access – qui se développe depuis le début des années 1990. Il se définit comme la mise à disposition libre et sans restriction des résultats de la recherche scientifique¹.

Le chercheur est doublement concerné. En tant que lecteur tout d'abord : pouvoir consulter et utiliser les publications de ses pairs sans obstacle financier est un avantage indéniable, chacun peut l'éprouver dans sa pratique. En tant qu'auteur ensuite : en élargissant le lectorat, ses recherches ont une probabilité accrue d'être citées. Des études² démontrent en effet une augmentation du nombre de citations par la diffusion en libre accès.

Pour ce faire, il peut publier dans une revue qui est déjà en libre accès³, ou dans une revue qui l'est partiellement parce que l'éditeur le propose en option (option payante). Parallèlement à la publication, il peut en outre déposer ses travaux dans une archive ouverte.

Une archive ouverte est une base de données (on trouve aussi les termes d'entrepôt, de réservoir) dans laquelle les chercheurs mettent à disposition gratuitement les résultats de leurs travaux. L'auteur peut aussi utiliser ce type d'outil pour diffuser un texte non encore publié, qu'il propose à la communauté scientifique en vue de recueillir des avis. Cette pratique de circulation et d'échanges est à l'origine de la création de la toute première archive, arXiv⁴, créée dans le milieu des physiciens en 1991 et qui sert de modèle depuis.

L'archive ouverte française HAL

Le Centre pour la Communication Scientifique Directe (CCSD) a été créé en 2000 par le Centre National de la Recherche Scientifique (CNRS), à l'initiative du physicien Franck Laloë qui souhaitait justement développer pour la communauté des chercheurs français une archive ouverte sur le modèle d'ArXiv.

¹ Voir les textes fondateurs sur le site Libre Accès à l'Information Scientifique et technique, site qui permet aussi de suivre l'actualité dans ce domaine : <http://openaccess.inist.fr>

² Le projet [OpCit](#) (The Open Access Citation Advantage Service) recense les études sur ce thème : en septembre 2016, 46 sur les 70 études listées démontraient un bénéfice sur le nombre de citations.

³ Le [DOAJ](#) (Directory of Open Access Journals) recense les revues en libre accès.

⁴ <https://arxiv.org/>

La mission première du CCSD était donc de créer et développer les archives ouvertes HAL⁵ et TEL⁶, spécifique au dépôt des thèses. Ses missions se sont progressivement étendues à d'autres services, notamment les plateformes SciencesConf.org⁷ pour la gestion de colloque et Episciences.org⁸ pour l'hébergement d'épi-revues. En 2014, l'INRIA et l'Université de Lyon sont devenues tutelles du CCSD, au titre des organismes de recherche et des universités, concrétisant ainsi ses missions nationales. L'année précédente en effet, une convention de partenariat⁹ regroupant 25 établissements et organismes scientifiques et universitaires légitimait HAL comme l'archive commune et partagée et comme un des éléments pivot de la politique nationale en faveur du libre accès. Elle est depuis 2016 inscrite dans la feuille de route des infrastructures de recherche du Ministère de l'enseignement supérieur et de la Recherche.

Pourquoi spécifiquement HAL ? Sa vocation multidisciplinaire¹⁰ rassemble un large spectre de la littérature savante produite. Mais une des raisons de son succès – HAL est utilisée par l'ensemble des chercheurs français – est sans doute due à son adaptation à l'organisation administrative de la recherche en France. Un chercheur exerce son activité dans un laboratoire ou une équipe de recherche, qui dépendent de plusieurs tutelles qui les financent. Par exemple, une UMR (unité mixte de recherche) a pour tutelles le CNRS et une, voire plusieurs universités. Elle a un ancrage à la fois national (avec le CNRS) et territorial avec l'université.

Dans les faits, HAL est devenue un outil commun et ouvert à l'ensemble des chercheurs, même s'il existe des archives institutionnelles¹¹ développées en-dehors de HAL.

Les publications déposées peuvent être citées dans des bibliographies et les fichiers contenant le texte doivent être accessibles. Aussi, une attention particulière est-elle portée à la stabilité des identifiants (url des dépôts notamment) et à la pérennité des documents. La date et l'heure du dépôt sont affichés ce qui assure la paternité du texte déposé.

HAL satisfait aux critères d'interopérabilité qui permet, entre autres, l'interconnexion avec les archives thématiques internationales ArXiv, PubMed, RePec et OpenAIRE. Une publication déposée dans HAL peut être automatiquement reversée dans ces bases à la demande de l'auteur, ce qui lui évite des doubles saisies.

Une fonctionnalité spécifique de HAL est celle du portail : c'est un sous-ensemble de HAL qui n'affiche que les publications correspondant aux critères de création du portail. Ces critères peuvent être institutionnels : un organisme de recherche comme l'INSERM ou une université, mais aussi thématiques et c'est le cas du portail dédié aux sciences humaines et sociales, HAL-SHS.

L'archive ouverte dédiée aux sciences humaines et sociales : HAL-SHS

IL est apparu assez vite que, pour convaincre les chercheurs en SHS de diffuser sur une archive ouverte leurs publications, il leur fallait un espace dédié. C'est ainsi que HAL-SHS¹² a été ouvert à la

⁵ <https://hal.archives-ouvertes.fr/>

⁶ <https://tel.archives-ouvertes.fr/>

⁷ <https://www.sciencesconf.org/>

⁸ <https://episciences.org/>

⁹ <http://www.enseignementsup-recherche.gouv.fr/cid71277/parteneriat-en-faveur-des-archives-ouvertes-plateforme-mutualisee-hal.html>

¹⁰ Liste des disciplines consultable : <https://hal.archives-ouvertes.fr/browse/domain>

¹¹ Par exemple : OATAO qui regroupe plusieurs établissements de l'université de Toulouse

communauté scientifique en janvier 2005. On doit la création de cette archive au travail commun mené par le CCSD, l'Institut des Sciences de l'Homme ([ISH](#)) et l'université Rennes 2 avec le soutien du département SHS du CNRS.

Que peut-on y déposer ?

Déposer, archiver ou auto-archiver sont les termes généralement utilisés pour qualifier le fait de télécharger son texte dans une archive ouverte. Les métadonnées sont les informations, formalisées selon les normes de descriptions bibliographiques, qui permettent de décrire le fichier déposé (le titre du texte, son ou ses auteurs, le résumé, le titre de la revue pour un article ou titre de l'ouvrage pour un chapitre, la date de publication, etc.).

Comme pour HAL, tous les types de publications peuvent être archivés dans HAL-SHS dès lors qu'elles sont le fruit de recherches produites dans un cadre scientifique : article publié dans une revue, ouvrage ou chapitre d'ouvrage, communication ou poster présentés lors d'un colloque. Les rapports de recherche et les manuscrits (qui n'ont pas encore été soumis à un éditeur par exemple, aussi appelés preprint ou document de travail) peuvent aussi être déposés, de même que les thèses et les Habilitation à Diriger des Recherches (HDR). Par contre, HAL-SHS ne diffuse pas les mémoires de master et les rapports de stage.

Le dépôt est constitué principalement du texte intégral de la publication mais un chercheur, s'il le souhaite, peut lui ajouter des fichiers annexes (graphiques, illustrations, etc). Cependant, les données non textuelles comme les images et les cartes, ainsi que les enregistrements sonores ou vidéo sont identifiées dans HAL-SHS comme des données de la recherche pouvant faire l'objet d'un dépôt en tant que tel.

Il est possible d'associer entre eux plusieurs dépôts : une illustration avec le ou les articles dans lesquels elle a été utilisée par exemple, ou bien lier deux chapitres publiés dans un même ouvrage collectif. A l'affichage du dépôt, l'internaute se verra proposer les liens vers les autres ressources pour compléter sa consultation.

Parallèlement à l'archivage de fichiers, HAL-SHS tout comme HAL sont aussi utilisées comme outils de signalement : le chercheur ou le documentaliste ne renseignent que la référence bibliographique de la publication. Ce n'est par contre pas possible pour les thèses ou les HDR, ni pour les données de la recherche identifiées comme telles (image, vidéo, son, carte).

Qui dépose

Le chercheur est au cœur du dispositif : publier est une activité inhérente à son métier et il tient à jour la liste de ses publications. Il est le mieux à même d'en décrire le contenu et il dispose systématiquement d'une copie du texte, quelle qu'en soit la version : soumise, acceptée, publiée.

Lorsque c'est l'auteur lui-même qui dépose, on utilise le terme d'auto-archivage. Ce travail peut bien entendu être aussi effectué par toute personne en charge des publications dans un laboratoire, une bibliothèque, etc. On constate par exemple que, dans certaines équipes, l'activité de dépôt est répartie : les métadonnées sont saisies par un documentaliste et le fichier ajouté par le chercheur.

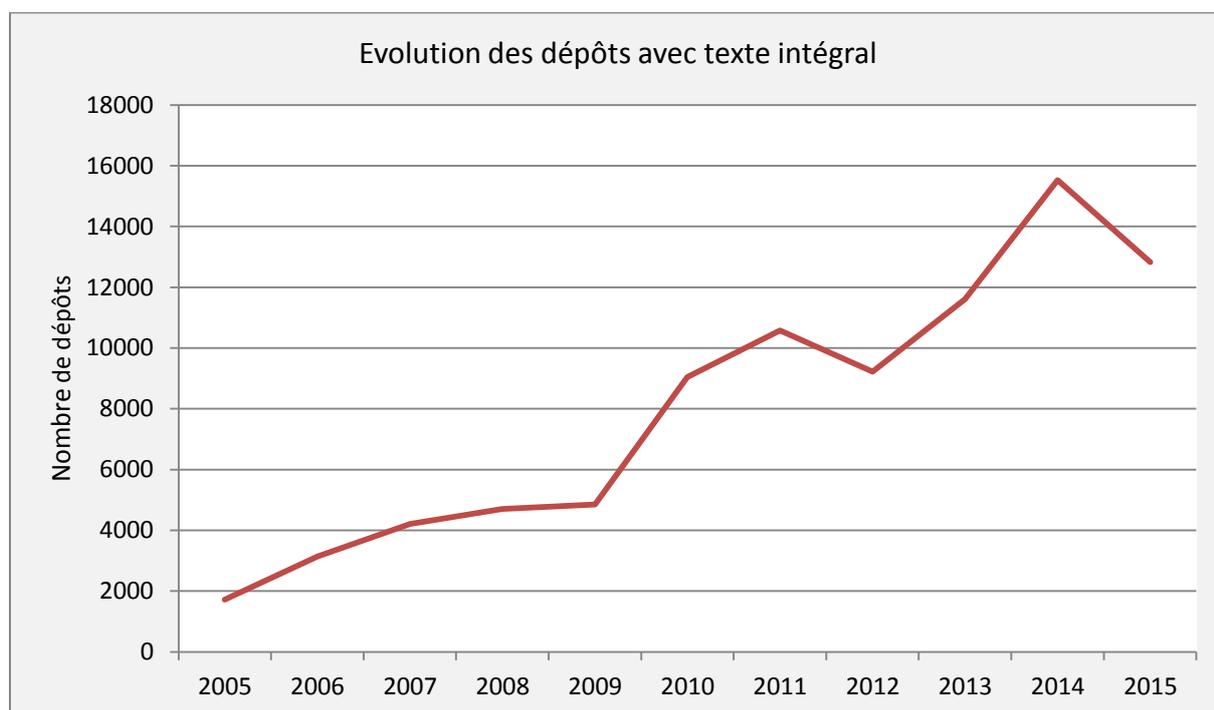
¹² <https://halshs.archives-ouvertes.fr/>

Chaque déposant peut partager la propriété de ses dépôts, avec ses co-auteurs notamment : cela signifie que chaque utilisateur qui partage ce dépôt a les droits pour mettre à jour les données et/ou le fichier. Cette fonctionnalité a le double avantage de limiter la création de doublons et de mutualiser le travail de saisie et/ou de complétude des informations.

Une croissance régulière des dépôts

HAL-SHS connaît une croissance régulière depuis sa création. 2010 a vu le pallier de 10 000 dépôts par an franchi, tous types de documents compris. Depuis 2010, on compte une moyenne de près de 1000 dépôts par mois.

En juin 2016, le seuil des 100 000 documents contenus dans HAL-SHS a été dépassé. La part de HAL-SHS dans HAL varie quant à elle entre 20 à 25%.



Graphique 1 : évolution du nombre de dépôts avec texte intégral

Les statistiques de dépôts témoignent de l'adhésion des chercheurs en sciences humaines aux nouvelles technologies, contrairement aux idées reçues. Néanmoins, les sciences humaines et sociales couvrent un large champ et on relève des disparités de pratiques selon les disciplines.

Une appropriation différente selon les disciplines

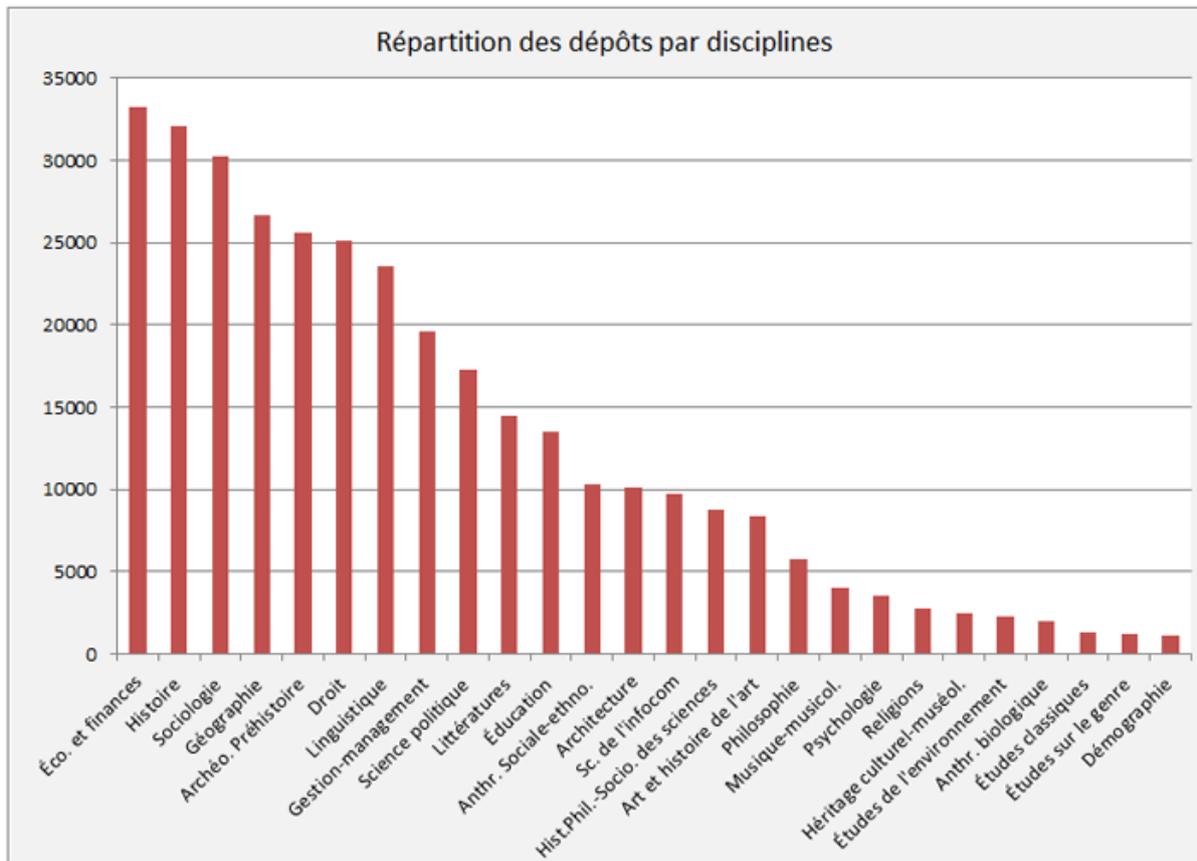
Pour qualifier son dépôt, le déposant doit obligatoirement choisir comme métadonnée au moins une discipline parmi les 27 proposées dans HAL-SHS :

- Anthropologie biologique
- Anthropologie sociale et ethnologie
- Archéologie et Préhistoire
- Architecture, aménagement de l'espace
- Art et histoire de l'art
- Études classiques

- Démographie
- Droit
- Économies et finances
- Éducation
- Études de l'environnement
- Études sur le genre
- Géographie
- Gestion et management
- Histoire, Philosophie et Sociologie des sciences
- Histoire
- Sciences de l'information et de la communication
- Linguistique
- Littératures
- Héritage culturel et muséologie
- Musique, musicologie et arts de la scène
- Philosophie
- Psychologie
- Religions
- Science politique
- Sociologie
- Méthodes et statistiques

Cette liste n'a pas vocation à être exhaustive : pour faciliter le dépôt, le choix a délibérément été fait d'un seul niveau hiérarchique, l'ajout de mots-clés permettant d'apporter des précisions sur le contenu.

Le déposant peut bien entendu choisir plusieurs disciplines, de même qu'il a la possibilité d'ajouter des domaines hors SHS s'il le souhaite. Néanmoins, les statistiques de dépôts que nous présentons ici sont calculées sur la 1^{ère} discipline sélectionnée.



Graphique 2 : répartition des dépôts par discipline SHS

La discipline Economie et finances est la discipline qui est représentée par le plus grand nombre de dépôts dans HAL-SHS, et ce, quasiment depuis le début de l'archive. Il faudrait certainement une étude spécifique pour en cerner les raisons. L'une d'elles est vraisemblablement l'interconnexion de HAL-SHS avec RePec (Research Papers in Economics)¹³ la base internationale de référence dans cette discipline. A partir de 2007 en effet, tout dépôt en économie¹⁴ dans HAL-SHS est reversé dans RePec, augmentant ainsi la visibilité internationale de la publication.

De nombreux facteurs entrent en ligne de compte qui bloquent ou qui incitent les chercheurs à déposer leurs publications, sans toutefois relever de pratiques disciplinaires. Parmi les freins, les questions de droit d'auteur sont fréquemment avancées, notamment celles en lien avec les politiques des éditeurs vis-à-vis de l'auto-archivage: ces derniers voient en effet la diffusion en libre accès sur une archive ouverte comme une concurrence. Les sites Sherpa/Romeo¹⁵ et Héloïse¹⁶ sont mis à la disposition des éditeurs de revues pour publier leurs politiques: ils y précisent leurs conditions, déclinées selon la version de l'article (soumis, validé ou version publiée). Ces sites constituent des sources d'information précieuses pour les auteurs avant de déposer.

Par contre, une politique d'incitation et d'animation autour du dépôt au sein d'une équipe de recherche est facteur de succès: en témoignent par exemple les interventions des membres des

¹³ <http://repec.org/>

¹⁴ Première discipline choisie: Economie et finances ou Gestion et management

¹⁵ <http://www.sherpa.ac.uk/romeo/index.php>

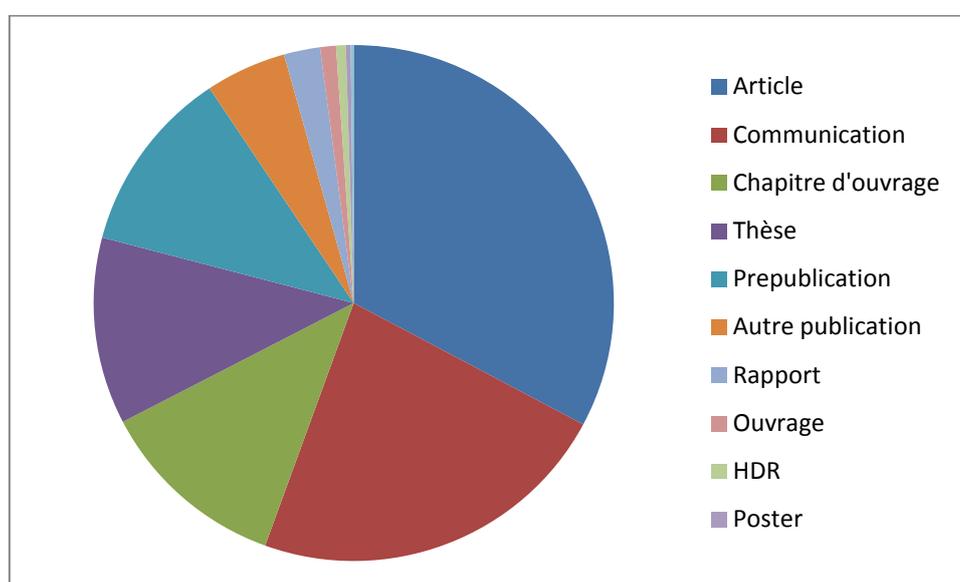
¹⁶ <http://heloise.ccsd.cnrs.fr/>

deux laboratoires Triangle et CRESSON (Centre de recherche sur l'espace sonore & l'environnement urbain) lors de la journée consacrée aux 10 ans de HAL-SHS¹⁷. Un accompagnement personnalisé des chercheurs a été mis en place par les professionnelles de l'information, soutenues par la direction des laboratoires.

Les types de documents dans HAL-SHS

La première étape du dépôt consiste à renseigner le type de publication : article, chapitre etc. Ce qui permet ensuite d'adapter l'affichage des métadonnées à compléter : par exemple, le titre d'une revue pour un article, celui d'un ouvrage pour un chapitre, la date de soutenance et l'université pour une thèse.

Le graphique suivant présente la répartition du contenu déposé dans HAL-SHS en fonction du type de document :



Graphique 3 : répartition des types de documents dans HAL-SHS

L'article est le vecteur de diffusion le plus massivement utilisé : c'est le cas en SHS comme pour toutes les disciplines. Les archives ouvertes permettent de diffuser de manière pérenne les communications dans des colloques, ce qui n'est pas forcément garanti avec les sites web créés lors des manifestations scientifiques. De même, elles permettent de valoriser un chapitre d'ouvrage, notamment si l'auteur a publié dans un ouvrage collectif, contrairement aux catalogues de bibliothèque dans lesquels ne sont recensés que les auteurs ayant dirigé ou coordonné le livre.

Près de 50% des prépublications de HAL-SHS sont en Economie et finances : le « working paper » ou document de travail, texte non définitif diffusé pour recueillir les avis et commentaires avant soumission à une revue, est en effet un mode de circulation des recherches courant dans cette discipline. Des laboratoires valorisent les recherches de leurs membres par des collections de

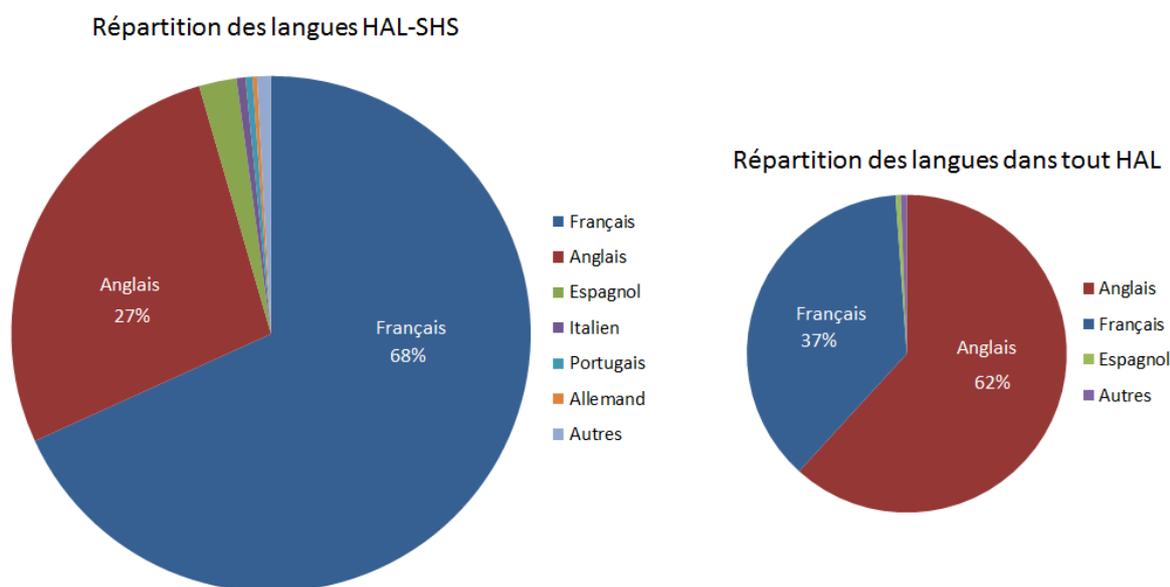
¹⁷ <https://10anshalshs.sciencesconf.org/>

working paper numérotées. Une personne référente, souvent le/la documentaliste, les déposent sur HAL-SHS : c'est le cas par exemple de de la Paris Jourdan School of Economics ou du CERDI¹⁸.

Croiser ces données de HAL-SHS avec celles plus globales de HAL permet aussi de mettre en lumière les spécificités des SHS qui se distinguent par la production d'ouvrages : 70% des dépôts de chapitres d'ouvrages et d'ouvrages dans HAL proviennent de HAL-SHS.

Les langues : richesse des langues d'expression

Publier en anglais est indispensable pour les chercheurs tant cette langue s'est imposée pour les échanges internationaux. Le français domine toutefois pour les SHS comme on peut le relever en comparant HAL-SHS avec HAL :



Graphique4 : répartition des langues – comparaison entre HAL-SHS et HAL

On recense 37 langues, parmi lesquelles on relève notamment le basque et le breton qui sont des langues régionales françaises.

Des services adaptés aux besoins du chercheur

Le contenu riche et varié de HAL-SHS démontre la vitalité de la recherche en SHS et constitue une ressource documentaire incontournable. HAL-SHS est indexée et très bien référencée par les moteurs de recherche, notamment Google Scholar, le plus utilisé par la communauté scientifique.

Mais en plus d'être un outil permettant au chercheur de diffuser en accès ouvert ses publications, qu'apporte HAL-SHS en terme de services qui peut le soutenir dans ses activités ?

Accès aux statistiques de consultation et téléchargement

Voir son article cité dans l'article d'un de ses pairs est un élément tangible de l'impact de ses recherches pour un chercheur. Les index de citations, comme le Web of science¹⁹, sont notamment

¹⁸ Collections HAL : https://halshs.archives-ouvertes.fr/PJSE_WP et <https://halshs.archives-ouvertes.fr/CERDI>

utilisés pour les repérer. Dans ses pages de résultats, Scholar Google mentionne le nombre de fois où une référence est citée et permet à l'internaute de rebondir sur les publications.

Cependant, la citation ne traduit qu'un type d'usage : lire un article n'implique pas qu'il sera cité. On voit de plus en plus apparaître sur les sites des éditeurs des métriques qui traduisent d'autres usages : consultations en ligne, téléchargements des fichiers, mentions sur les réseaux sociaux. Ces métriques sont proposées à tout internaute, indices de l'audience de l'article.

HAL-SHS présente aussi ces informations à l'affichage d'un dépôt : le nombre de consultations de la notice, le nombre de téléchargement du fichier (nombre de clics sur le fichier). Les métriques associées aux mentions sur les réseaux sociaux ne sont affichées, si elles existent bien sûr, que si un identifiant tel que le DOI²⁰ est présent dans les métadonnées du dépôt : c'est en effet à partir de ce type d'identifiant que l'application²¹ calcule ces statistiques.

En revanche, sur le site de l'éditeur, l'auteur n'a pas plus d'informations que celles proposées à tous. Sur HAL-SHS, tout auteur d'une publication ayant un compte utilisateur dispose de statistiques plus fines et détaillées, qu'il ait effectué lui-même le dépôt ou une tierce personne. Il peut les consulter comme les exploiter dans un tableur après les avoir exportées. Il peut suivre l'évolution dans le temps des consultations et/ou des téléchargements d'une ou de plusieurs de ses publications, visualiser la provenance de ces consultations sous forme de carte. C'est souvent une surprise pour un auteur de constater que son dépôt a été consulté en Chine ou en Afrique du Sud.

Si toutes ses publications sont dans HAL-SHS, il peut aisément obtenir une répartition par type de documents, par année, par titre de revue. Toutes les métadonnées présentes dans les dépôts peuvent être potentiellement utilisées pour satisfaire la curiosité ou produire des rapports.

Identifiant auteur avec l'idHAL et CV en ligne

Le chercheur est de plus en plus incité à gérer sa présence en ligne et à assurer la visibilité de ses recherches. Le succès de réseaux sociaux académiques comme Researchgate ou Academia témoigne de ce qui est devenu un impératif. HAL-SHS propose deux fonctionnalités qui lui permettent très facilement de prendre ces questions en main.

Le premier est l'idHAL : c'est un identifiant unique que le chercheur configure pour rassembler toutes ses publications, quelle que soit la façon dont son nom d'auteur apparaît dans la base de données. On résout ainsi les problèmes liés aux différences de saisie (prénom sous forme d'initiale ou en entier), les changements de nom (notamment pour les femmes qui au cours de leur carrière peuvent signer sous leur nom d'usage ou d'épouse) et bien sûr les homonymies. Les publications sont ainsi associées au bon auteur qui est identifié de façon unique, ce qui est précieux pour toute recherche documentaire notamment. Si l'auteur ajoute à son idHAL ses comptes sur les réseaux sociaux (twitter par ex ou Researchgate), les identifiants auteur utilisés dans d'autres contextes (ORCID, IdRef, etc.) ou l'adresse url de son blog, l'internaute qui consulte ses dépôts pourra aussi visiter les autres sites en cliquant sur des liens, et compléter ainsi ses connaissances sur l'auteur.

¹⁹ Le Web of science est produit par la société ISI – Institute for Scientific Information de Thomson Scientific. Il regroupe des bases de données bibliographiques

²⁰ Digital Object Identifier

²¹ HAL-SHS utilise l'application commercialisée par la société Altmetric

Ces informations sont aussi accessibles via le CV, qui est la seconde fonctionnalité proposée : le chercheur qui a créé son idHAL peut ensuite générer une page web pour y afficher automatiquement toutes ses publications présentes dans HAL-SHS. Il complète cette page avec des informations libres : il peut par exemple y présenter ses recherches ou ses enseignements. L'avantage de cette page web dynamique pour le chercheur est qu'il n'a pas à se préoccuper des aspects techniques liés à la maintenance technique d'un site. De plus l'adresse url est unique et pérenne : elle est composée sur le format <https://cv.archives-ouvertes.fr/idHAL>. L'idHAL étant le plus souvent formé à partir du nom et prénom, cela contribue à la présence en ligne du chercheur.

Conclusion : pourquoi déposer dans HAL-SHS

La vitalité de HAL-SHS et la diversité de son contenu sont des signes notables de sa place dans la circulation des résultats de recherche en SHS. Un chercheur, qu'il soit en début de carrière comme plus confirmé, y trouve un outil pour valoriser ses travaux et des services adaptés à son activité.

HAL-SHS constitue, et c'est un atout précieux, un espace où il pourra centraliser ses publications – y compris ses premières communications et sa thèse. Quelles que soient ses affectations tout au long de son parcours et au gré de celles-ci, il peut utiliser HAL-SHS. Les données étant exportables sous des formats variés, il pourra les exploiter aisément²².

²² Voir par exemple le témoignage de N. Fressengeas <http://factuel.univ-lorraine.fr/node/4803>