

7. HAL & les archives ouvertes

**On appelle archives ouvertes,
alias *open archives*,
les serveurs informatiques sur lesquels
sont déposés des documents numériques
issus de la recherche scientifique
(on parle aussi d'entrepôt, ou de *repository*).
Ces réservoirs sont intimement
liés au mouvement de
la science ouverte.**

Les documents déposés peuvent être de natures différentes : manuscrits scientifiques avant processus de publication (version auteur ou *preprint*), publications scientifiques ayant fait l'objet d'un processus éditorial (articles de revue, livres, chapitres d'ouvrages collectifs, actes de colloques) dans leur version relue avant publication (*postprint*) ou dans leur version publiée (version éditeur), travaux de la *littérature grise*, c'est-à-dire diffusés sans être issus du processus de l'édition, (mémoires, thèses, Habilitation à Diriger des Recherches, projets de recherche, posters & communications), mais aussi cours, données de la recherche, logiciels, brevets qui disposent le plus souvent d'entrepôts spécifiques.

Les dépôts sont effectués par les auteurs et titulaires des droits (auto-archivage), ou par leur institution (généralement les bibliothèques et les centres de documentation), avec trois objectifs :

- **identifier et certifier** la production scientifique, protéger les droits des auteurs et des éditeurs, préciser les autorisations qu'ils accordent : la date de dépôt atteste de la paternité de la production scientifique. Les identifiants numériques uniques attribués aux documents (DOI) et aux auteurs (ORCID, IdHAL) permettent de les identifier sans ambiguïté. Les licences explicites de type *Creative Commons* permettent aux auteurs d'indiquer les utilisations qu'ils autorisent pour leurs documents, la gestion des différentes versions (*preprint*, *postprint*, version éditeur) et des embargos permettent de garder l'historique d'un travail ainsi que sa confidentialité.

- **diffuser le plus largement possible** : les archives ouvertes permettent d'augmenter la visibilité des publications en donnant accès, au plus grand nombre, sans barrière financière, juridique ou technique, aux résultats publiés de la recherche. Le dépôt en archive ouverte constitue ainsi, sous le nom de voie verte (*Green Open Access*) l'un des principaux moyens de mise en œuvre du libre accès aux publications scientifiques. Leurs contenus sont bien référencés par les moteurs de recherche et des liens d'interopérabilité sont établis d'une archive à l'autre. De nombreuses études démontrent et mesurent le gain de visibilité et d'impact potentiel — en termes de citations — obtenu par le dépôt en archive ouverte¹. Le dépôt en archive ouverte permet de respecter les obligations de diffusion en *open access* faites par les financeurs (ANR, Commission Européenne, Wellcome Trust Company).

- **conserver sur le long terme** : en adossant les serveurs à des institutions publiques, en attribuant des liens d'accès pérennes.

Une grande diversité d'archives ouvertes

Devant la multiplicité des archives ouvertes, il n'est pas évident de savoir où déposer. Il convient de privilégier les entrepôts offrant les meilleures garanties pour les objectifs d'identification, de diffusion, de conservation. La simplicité d'utilisation, les politiques d'accompagnement des établissements et les services à valeur ajoutée constituent des critères complémentaires.

On distingue :

Les archives ouvertes institutionnelles : mises en place par et pour une institution (organisme de recherche, université, école, etc.), elles ont pour objectif de repérer, conserver, valoriser, rassembler l'ensemble de la production scientifique de l'établissement en garantissant une intégration optimale à leur système d'information et la meilleure adaptation possible aux besoins spécifiques des usagers de l'établissement : *Spire* pour Sciences Po Paris, *Archimer* pour l'Ifremer, *Bird* pour l'Université Paris-Dauphine, *SAM* pour les Arts et Métiers et Paristech, *Univoak* pour les établissements alsaciens, *Lilloa* pour l'Université de Lille²...

Les archives ouvertes nationales : Elles regroupent la production scientifique d'un pays.

Par exemple : HAL (Hyper Articles en Ligne) en France, qui propose des portails par université ou organisme (HAL-Université de Paris, HAL-INSERM) ou par thématique (HAL -SHS) ou par support (MédiHAL pour les images, TEL pour les thèses).

Les archives ouvertes thématiques : elles réunissent les productions de plusieurs institutions dans un domaine scientifique donné. Par exemple : arXiv pour la physique, les mathématiques et l'informatique, RePEC pour les sciences économiques, SSOAR pour les sciences sociales, PubMed Central pour le biomédical et les sciences de la vie.

Au plan technique, elles doivent respecter le protocole OAI-PMH (*Open Access Initiative - Protocol for Metadata Harvesting*), pour leur permettre d'être interopérables, c'est-à-dire de communiquer entre elles. Il est donc possible de déposer dans une archive ouverte et de faire en sorte que les documents soient visibles dans une autre (transferts de HAL vers arXiv ou vers PubMed Central)

Focus sur HAL (Hyper Articles en Ligne) : l'archive ouverte nationale française

HAL et le CCSD (Centre pour la Communication Scientifique Directe), unité mixte de service du CNRS, de l'INRIA, de l'INRA et de l'université de Lyon représentant l'ensemble des universités françaises, ont été créés par le CNRS en 2001 à l'initiative du physicien Franck Laloë.

HAL a été conçu sur le modèle d'arXiv (base d'archive ouverte pionnière pour la physique et les mathématiques) avec pour objectif d'offrir aux chercheurs un outil permettant l'auto-archivage des publications quelles que soient leurs disciplines et leurs institutions de rattachement. HAL permet la conservation pérenne des documents qui y sont déposés grâce à l'archivage au C.I.N.E.S. (Centre Informatique National de l'Enseignement Supérieur) et sur *Software heritage* pour les logiciels.

Au 1er juillet 2019, le nombre de documents en texte intégral déposés dans HAL était de 600 000 et le nombre de références bibliographiques de 1 900 000. En 2018, HAL a enregistré 143 574 755 consultations ou téléchargements.

HAL propose un certain nombre de services qui permettent de valoriser tant l'activité des chercheurs que celle des laboratoires ou des institutions.

Au niveau de chaque article, HAL permet d'obtenir des chiffres de consultation, de téléchargement et dans certains cas de mentions sur les réseaux sociaux (altmetrics). Ces chiffres complètent les indices bibliométriques traditionnels à l'aune desquels on mesure habituellement les revues (et par dérivation les articles qui y sont publiés).

Le chercheur peut se créer un identifiant propre à HAL, **l'IdHAL**, qui lui permet de rassembler l'ensemble des documents déposés par lui-même ou des coauteurs et de créer ainsi un CV dynamique qui s'actualise automatiquement à chaque dépôt. Ce CV peut ensuite être intégré dans un site autre.

HAL propose également la création de collections spécifiques pour les laboratoires. Ainsi les laboratoires ont une page dédiée dans laquelle sont rassemblées toutes les publications des membres du laboratoire. Cette page peut être reliée au site internet du laboratoire.

Une fois le document déposé dans le portail HAL de l'institution, celui-ci est visible dans la collection du laboratoire, mais aussi dans le portail HAL général rassemblant l'ensemble de la production française et dans les portails HAL institutionnels des différents co-auteurs de la publication.

1 - Gentil-Beccot, A., Mele, S. & Brooks, T.C. *Scientometrics* (2010) 84: 345. <https://doi.org/10.1007/s11192-009-0111-1> : sur 4839 articles en physique publiés dans 5 revues de référence, les articles déposés dans ArXiv reçoivent en moyenne 14 citations contre 4 pour les autres.

Rentier, Bernard. Étude de l'université de Liège : <https://fr.slideshare.net/MyScienceWork/oa-luniversit-de-lige-le-pari-dorbi-par-bernard-rentier> : nombre de citations multiplié par deux pour les articles référencés dans le Web of Science ou Scopus et déposés dans l'archive institutionnelle Orbi.

Bacache-Beauvallet, Maya, Benhamou, Françoise, Bourreau, Marc. RAPPORT IPP N°11 – JUILLET 2015

Les revues de sciences humaines et sociales en France: libre accès et audience : <http://www.ipp.eu/wp-content/uploads/2015/07/revues-shs-rapport-IPP-juillet2015.pdf>

2- Forum des archives ouvertes institutionnelles françaises, 2018. <https://www.sciencespo.fr/bibliotheque/fr/nous-connaître/partena-riats/forumAOI>