

Science ouverte et présence numérique des chercheurs en sciences humaines et sociales

ANNAÏG MAHÉ, URFIST DE PARIS-ECOLE NATIONALE DES CHARTES
annaig.mahe@chartes.psl.eu

ET CAMILLE PRIME-CLAVERIE, UNIVERSITÉ PARIS NANTERRE
cclaveri@parisnanterre.fr

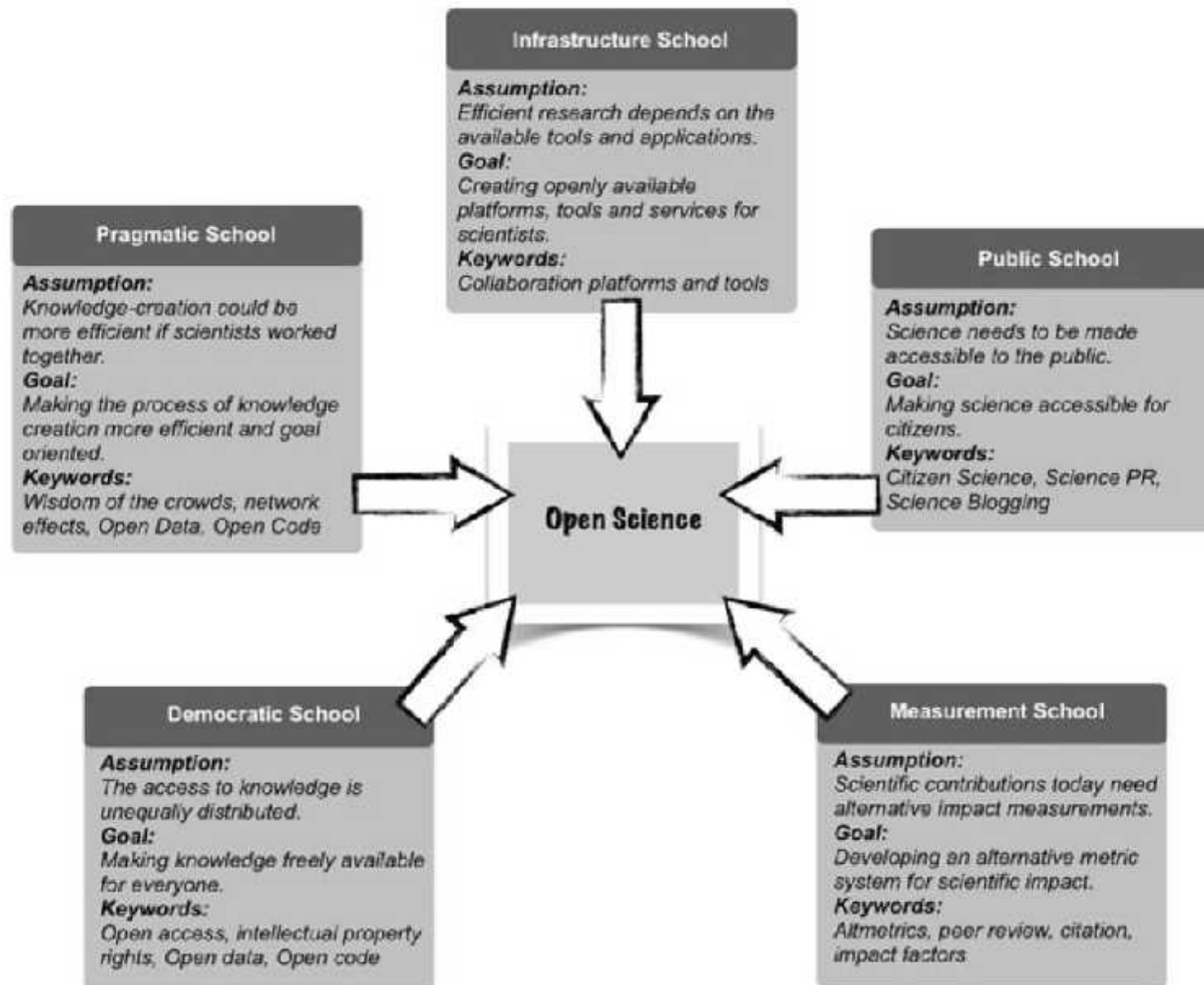
Dicen-Idf

Contexte et questions de recherche

Analyse de la visibilité en ligne de la recherche à l'heure de la « science ouverte »

Science ouverte : meilleur accès aux publications, aux résultats et aux données de la recherche, leur partage et leur réutilisation

Accès ouvert (Open access) : accès au texte intégral des publications sans barrière technique, juridique ou commerciale



Ouverture = caractéristique technique et « état d'esprit » (Weller, 2011)

Mise en œuvre de l'interopérabilité dans l'univers de la science ouverte : technique, sémantique, organisationnelle

2017 - Le défi de l'interopérabilité entre plates-formes pour la construction de savoirs augmentés en sciences humaines et sociales

Logiques d'appropriation des nouveaux modes de communication par les acteurs de la communauté scientifique (producteurs et utilisateurs de ressources)

2013 - Sites de dépôt en libre accès et formes de médiations : quelles évolutions ?

2017 - Qui dépose quoi sur Hal-SHS ? Pratiques de dépôts en libre accès en sciences humaines et sociales

2017 - Science ouverte et présence numérique des chercheurs en sciences humaines et sociales

Développement de nouvelles métriques (état de l'art)

2015 - Altmetrics - Nouvelles mesures de la visibilité des résultats de la recherche

Objectifs des études sur HAL/SDV et HAL-SHS

- ✓ Identifier et évaluer la contribution des différentes catégories d'acteurs
- ✓ Définir les logiques de dépôts
- ✓ Comparer les logiques disciplinaires et/ou d'acteurs

Niveaux de contribution sur HAL

Plus ou moins éloignés de l'objectif de de communication scientifique directe

Qui dépose ?	Auteur	Chercheur non auteur du document	Intermédiaire (non chercheur)	Transfert entre plateformes
Niveau de contribution	Communication scientifique directe entre chercheur Niveau 1	Communication scientifique entre chercheur Niveau 2	au service de leurs communautés Niveau 3	au service des institutions Niveau 4

Logiques de contribution sur HAL (SDV) et HAL-SHS

	Dépôt récent	Dépôt ancien
Dépôt avec fichier	Communication Scientifique immédiate	Archivage
Dépôt sans fichier	Visibilité de front de recherche (référencement)	Visibilité des recherches antérieures (recensement)

Nos hypothèses

		Auteur	chercheur non auteur du document	Intermédiaire (non chercheur)	Transfert entre plateformes
Notice avec texte intégral	Document récent	CSD immédiate			Diffusion du texte intégral
	Document ancien				
Notice seule	Document récent			Référencement / recensement	
	Document ancien				



disciplines

- Nos données
- *Sciences de la vie*
 - *Sciences humaines et sociales= HAL-Shs*
 - *etc.*

portails

- *Universités*
- *Centres de recherche*
- *Images = MediHAL*
- *Theses = TEL*
- *etc.*

collections

- *Laboratoires*
- *Revue*
- *etc.*

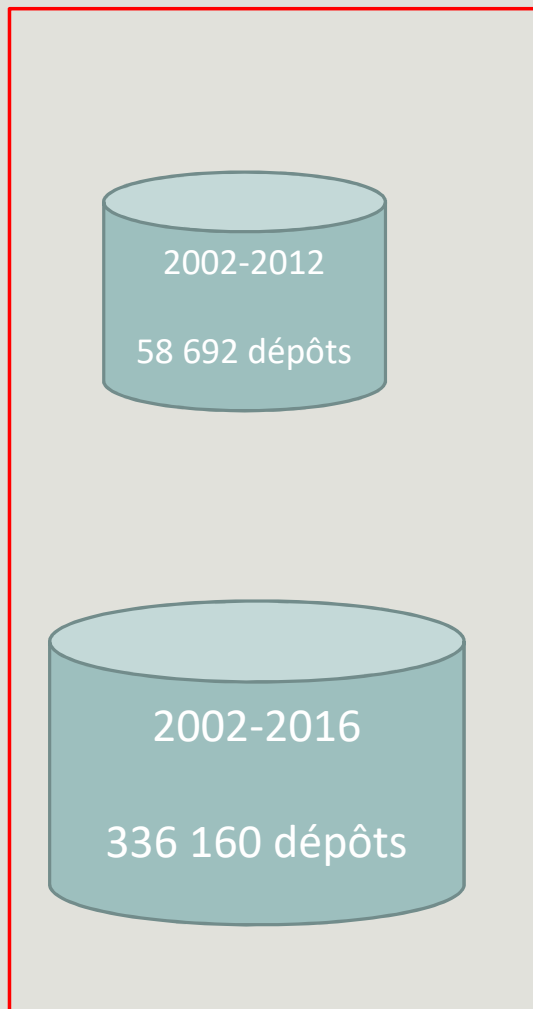


1 600 000 dépôts
1 000 000 notices
600 000 documents

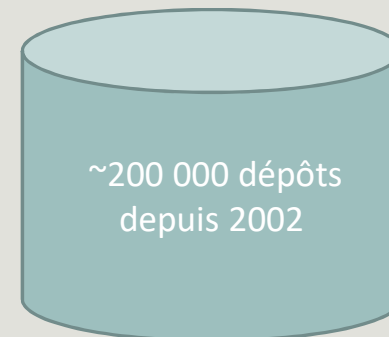
Sciences de
la vie

HAL-SHS

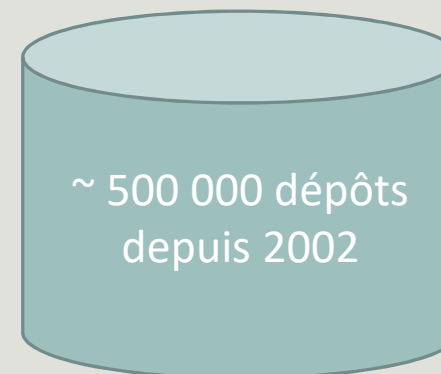
Nos données



Aujourd'hui



+200% depuis 2012

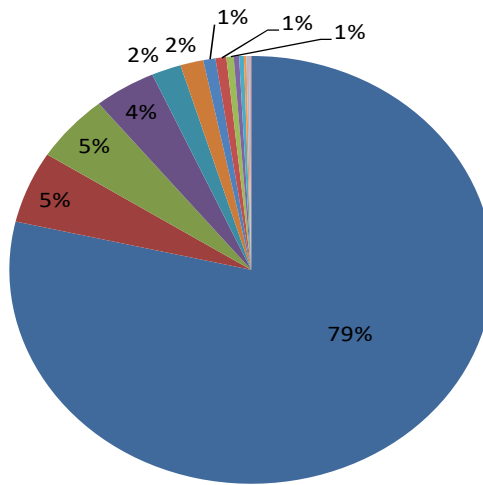


+40% depuis 2016

Catégorie de dépôts

Sciences de la vie

Articles = 4/5 dépôts

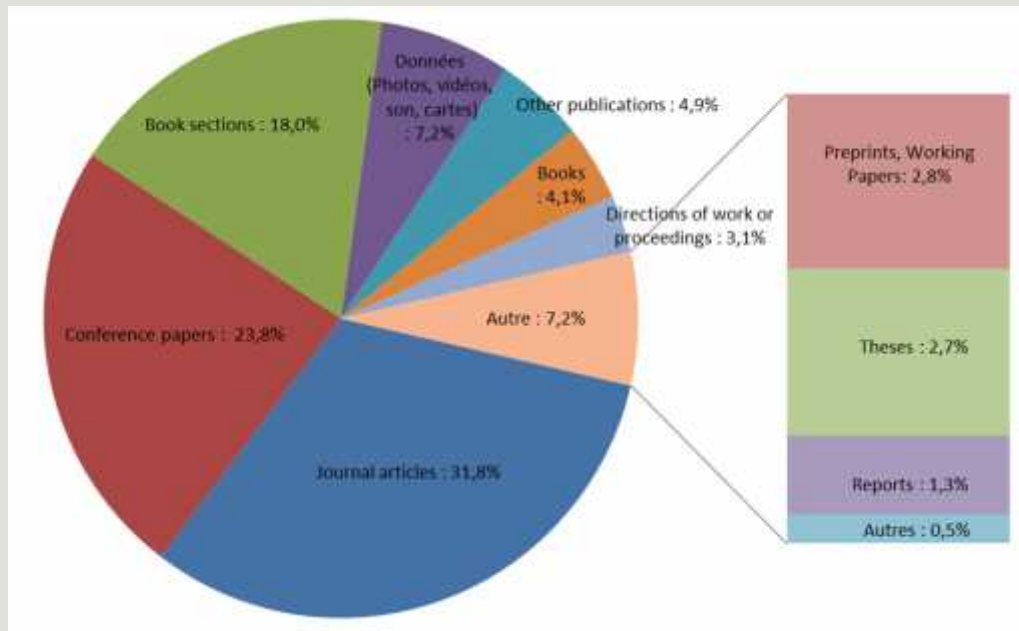


- Articles (journals with peer evaluation)
- Communications (with proceedings)
- Theses
- Communications (without proceedings)
- Book chapters
- Invited conferences
- Articles (journals without peer evaluation)
- Undefined

SHS

3 catégories majeures de dépôts

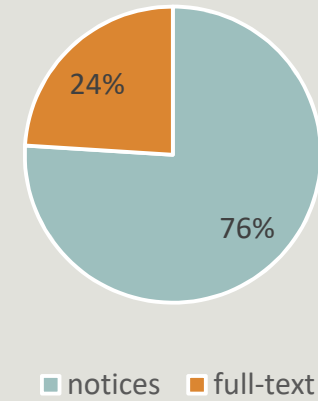
- Articles de revues
- Articles de conférences
- Chapitres d'ouvrage



Texte intégral ou simple notice ?

		Total
Full-text	Life sciences	24%
	HSS	24,5%
Notice only	Life sciences	76%
	HSS	75,5%

Sciences de la vie



SHS

Plus de texte intégral pour :

- preprints, working papers
- Rapports

= littérature grise



Texte intégral ou simple notice ?

+ dépôt récent* ou ancien ?

* Dépôt récent = texte intégral ajouté moins d'un an après la date de création de la notice

		Recent deposit	Older deposit	Total
Full-text	Life sciences	16%	8,5%	24%
	HSS	12%	12,5%	24,5%
Notice only	Life sciences	41%	34,5%	76%
	HSS	33%	43%	75,5%

Plus de texte intégral ancien en SHS :

- Embargos plus longs ?
- Plus de valeur pour les textes plus anciens ?

	Dépôt récent	Dépôt ancien
Full-text	Communication scientifique directe	Archivage
Notice only	Recensement	Référencement

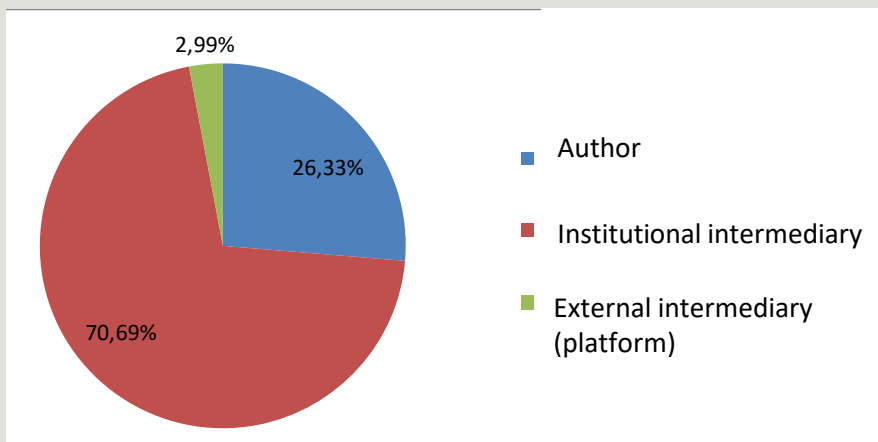
HAL-SHS : Logiques de dépôts

Test de khi-2 : Les « logiques de dépôts » varient en fonction des types de déposants

- Les auteurs déposent plus de texte intégral (récent ou ancien)
- Les intermédiaires non chercheurs déposent en grande majorité des références sans texte intégral associé, plutôt des références récentes (logique de recensement) ;
- Les plateformes déposent quasiment exclusivement des fichiers, à répartition presque égale entre fichiers récents et anciens.

Qui dépose ?

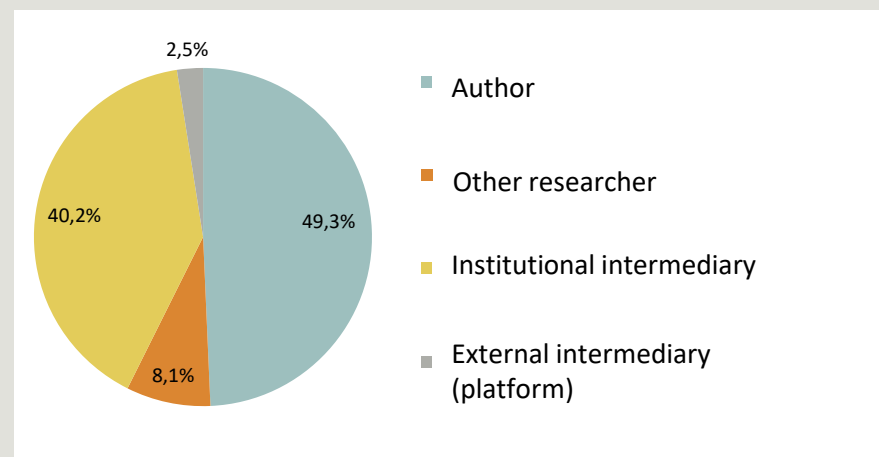
Sciences de la vie



✓ Intermédiaire de l'institution en sciences de la vie (70%)

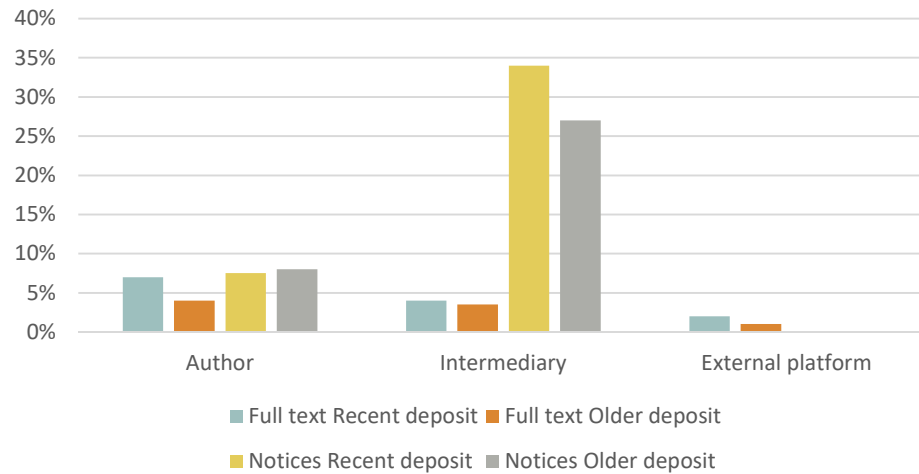
✓ Chercheur (57%) ou auteur en SHS (49%)

HSS



Qui dépose quoi ?

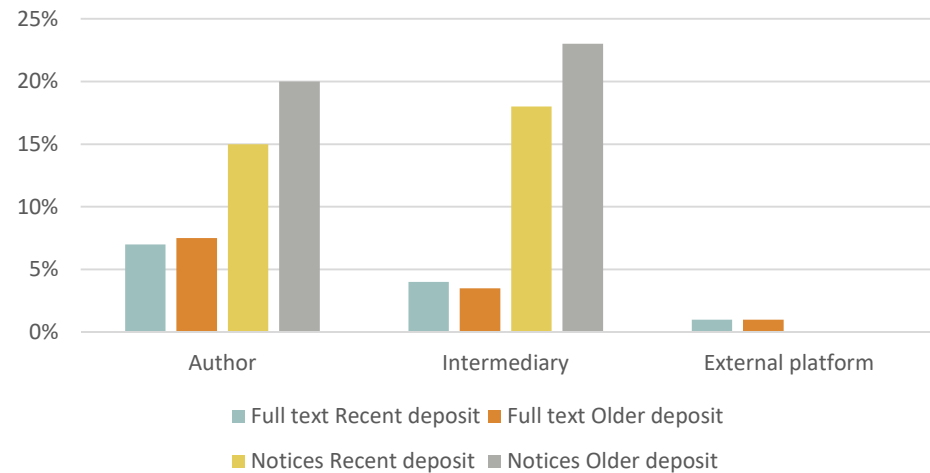
Life sciences

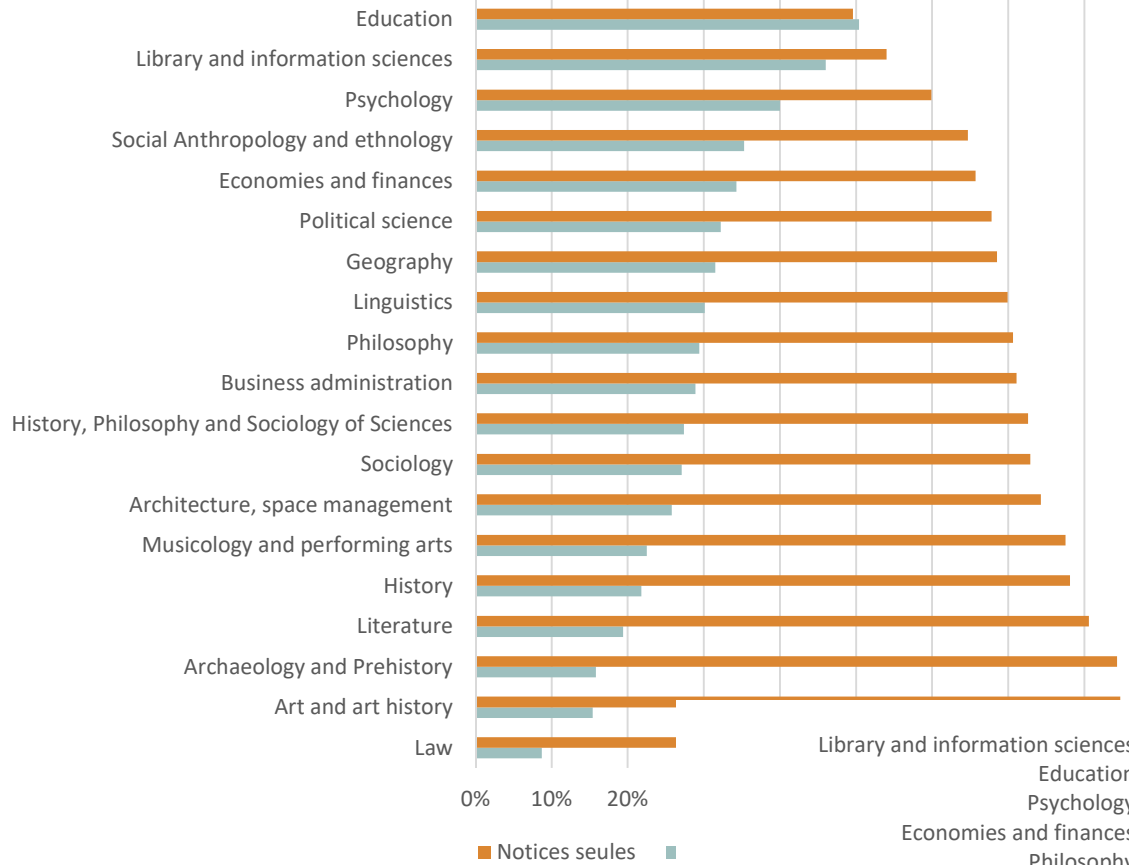


✓ En sciences de la vie, les intermédiaires déposent beaucoup de notices sans texte intégral

✓ En SHS aussi, mais les auteurs déposent autant de notices que les intermédiaires et plus de texte intégral

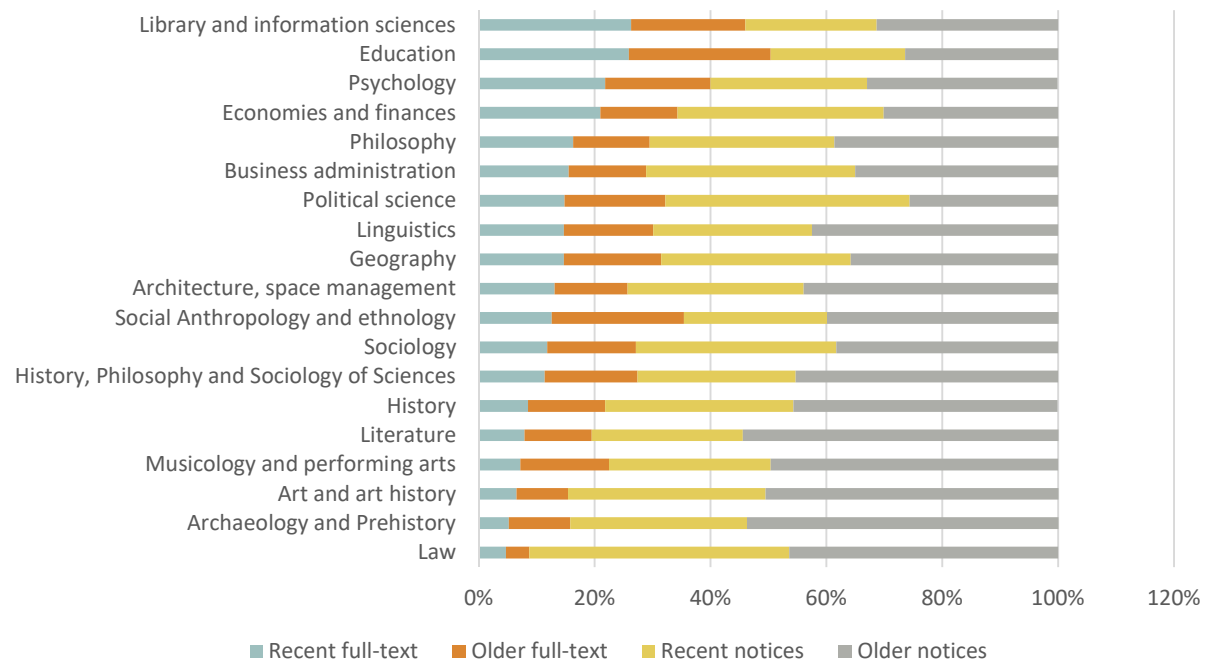
HSS





- ✓ **Articles de revues**
 - ✓ Law
 - ✓ Psychology
 - ✓ Economy
- ✓ **Actes de conférences**
 - ✓ business administration
 - ✓ education sciences
 - ✓ library and information sciences
 - ✓ linguistics
- ✓ **Chapitres d'ouvrage**
 - ✓ History
 - ✓ Art and art history
 - ✓ Literature
 - ✓ History and philosophy of sciences

Logiques disciplinaires



Présence numérique des chercheurs : étude exploratoire à partir de deux plateformes en ligne : Hal-Shs et Hypotheses.org

- ✓ Etude centrée sur la présence numérique chercheurs et non sur les dépôts

	chercheurs	intermédiaires	plateformes
HAL-SHS	13 557	809	2
	(179 017 dépôts)	(125 337dépôts)	(7 711 dépôts)

	chercheurs		
Hypotheses.org	415 auteurs	256 blogs	16 196 billets

Présence numérique des chercheurs : étude exploratoire à partir de deux plateformes en ligne : Hal-Shs et Hypotheses.org

Objectifs

- ✓ Comparaison disciplinaire des contributeurs-chercheurs de HAL-SHS
- ✓ Évolution de la participation
- ✓ Proposition d'une typologie des contributeurs-chercheurs de HAL-SHS
- ✓ Les auteurs des carnets d'Hypotheses.org sont-ils également contributeurs de HAL-SHS ?

Comparaison disciplinaire des contributeurs-chercheurs de HAL-SHS

- ✓ Pour chaque auteur : calcul de son indice de spécialisation dans une discipline A en fonction de ses dépôts

$$P(A) = \frac{\sum_{i=0}^N p_i(A)}{N}$$

N : nombre de dépôts effectués par ce contributeur en tant qu'auteur,

$p_i(A)$, le poids de la discipline A pour l'article i, $p_i(A) = \frac{1}{n}$

- où n est le nombre de disciplines mentionnées lors du dépôt de l'article i .

Comparaison disciplinaire des contributeurs-chercheurs de HAL-SHS

Poids max discipline	1	0,95	0,85	0,75	0,65	0,55	0,45	0,35	0,25	0,15
% contributeurs	65,4	68,7	74,8	78,7	83,6	86,5	95,0	96,9	99,1	100,0
	Une discipline de spécialisation						Interdisciplinarité			

Comparaison disciplinaire des contributeurs-chercheurs de HAL-SHS

- ✓ Comparaison HAL-SHS (2002-2016) et population des enseignants-chercheurs dans les sections CNU SHS en France (2010/2016)

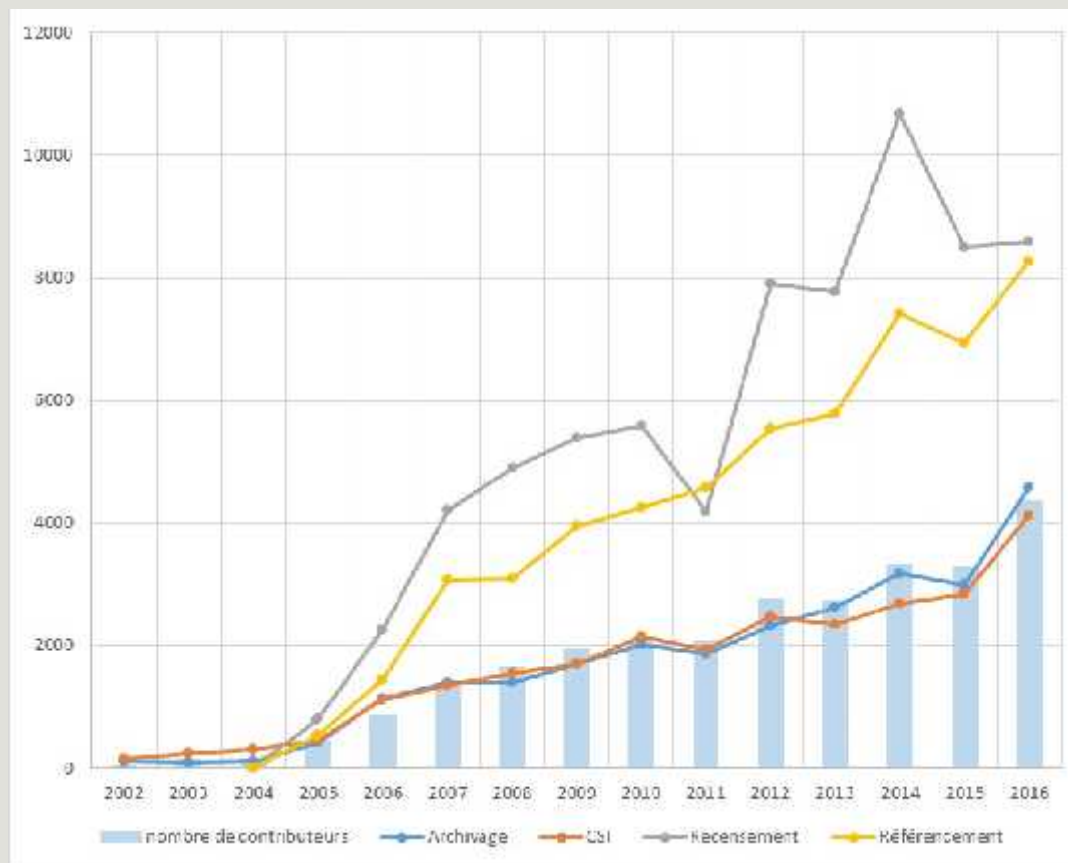
Sujets HAL-SHS	HAL-SHS Proportion de Contributeurs-chercheurs	Code Section	Section CNU	Proportion d'enseignants-chercheurs- SHS en France (2010/2016)
Linguistics	9,7%	7	Sciences du langage : linguistique et phonétique générales	3,4%
Geography	9,1%	23 et 24	Géographie	5,0%
History	9,0%	21 et 22	Histoire	8,3%
Economies and finances	7,6%	5	Sciences économiques	8,0%
Sociology	7,3%	19	Sociologie, démographie	3,9%
Demography	0,3%			
Business administration	6,7%	6	Sciences de gestion	8,8%
Education	6,6%	70	Sciences de l'éducation	2,9%
Library and information sciences	6,3%	71	Sciences de l'information et de la communication	3,5%
Law	6,0%	80 à 85	Droit	27,8%

Comparaison disciplinaire des contributeurs-chercheurs de HAL-SHS

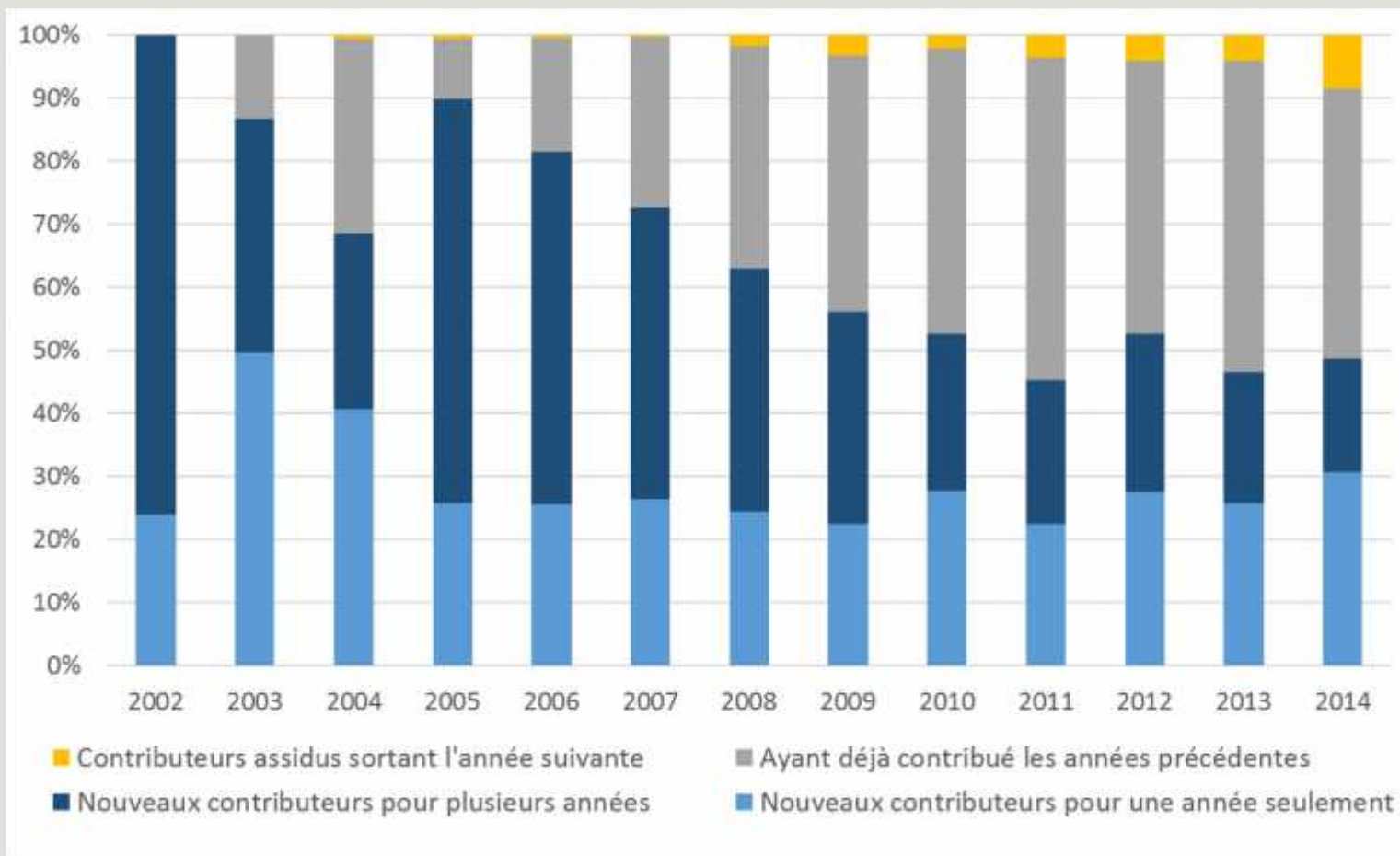
- ✓ Comparaison HAL-SHS (2002-2016) et population des enseignants-chercheurs dans les sections CNU SHS en France (2010/2016)

Présence sur HAL-SHS	Disciplines
Plus importante	linguistique, la géographie, la sociologie, les sciences de l'éducation ou les sciences de l'information et de la communication
équivalente	histoire, économie, philosophie
Plus faible	gestion, littérature, droit, psychologie

Evolution de la participation



Evolution de la participation

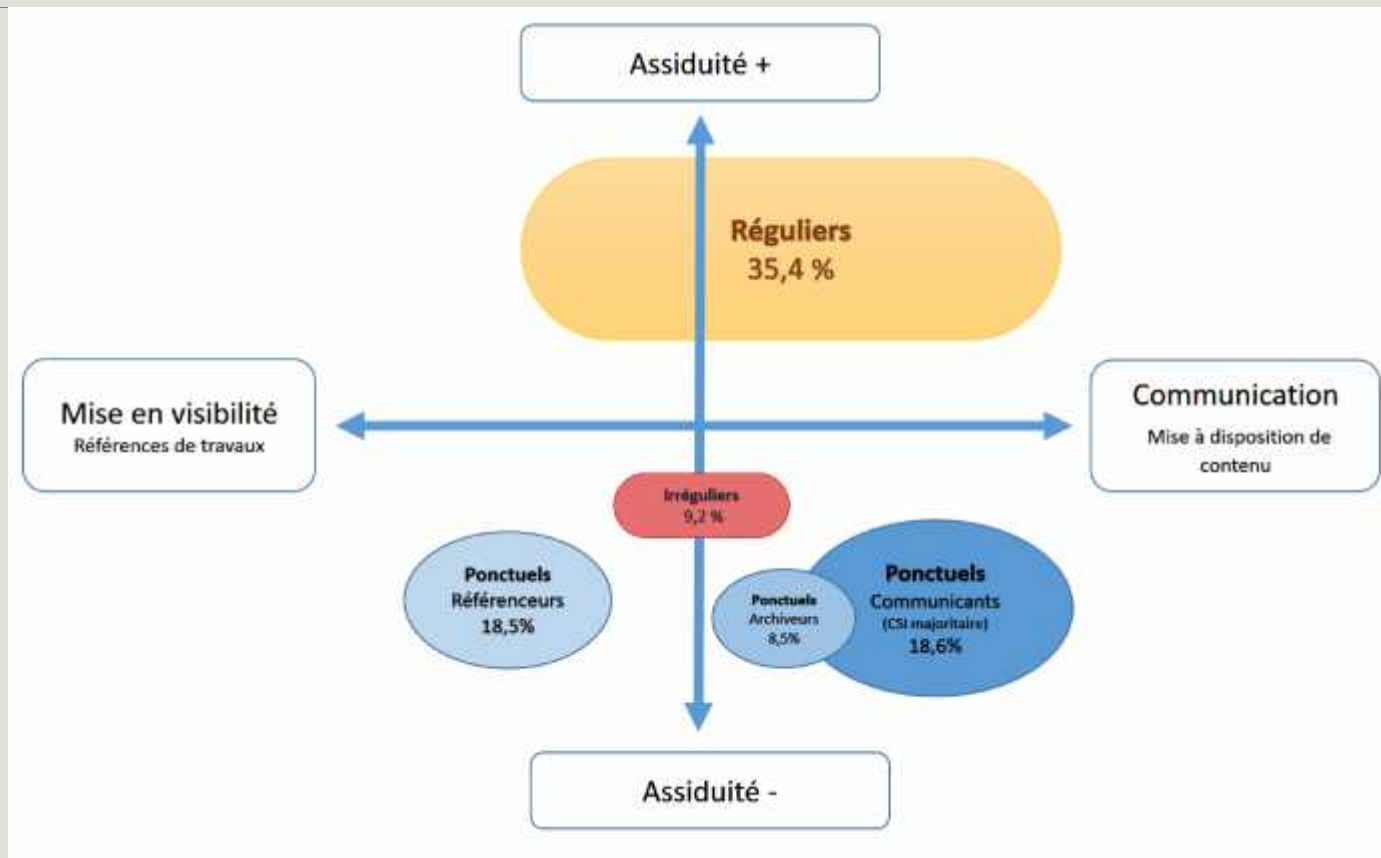


Typologie des contributeurs-chercheurs de HAL-SHS

La typologie est fondée sur les critères suivants

- ✓ **durée des dépôts** : période en année pendant laquelle un contributeur effectue des dépôts ;
- ✓ **assiduité** : nombre d'années de dépôt sur la période/durée de dépôts. L'assiduité ne peut être évaluée que pour les déposants présents au moins trois ans dans le corpus ;
- ✓ **type de dépôts** : CSI, Archivage, etc. ;
- ✓ **volume** : nombre de dépôts.

Typologie des contributeurs-chercheurs de HAL-SHS



Connexion entre les plateformes

Des difficultés techniques

- Sur Hypotheses.org
 - Usage de pseudonymes
 - Manque de structuration des données (champs multivalués)
 - Pas d'identifiant pour distinguer les contributeurs

- **Identifiants sur HAL-SHS**

- IDHAL : 5013
- ORCID : 825

Connexion entre les plateformes

- 57 auteurs d'hypotheses.org / 415 ont nécessité recherches pour identification
- 15 : nom du blog
 - « mes » pour le blog « mémoire et société »)
- 11 non identifiés
 - Pseudonyme utilisé pour représenter une autorité morale et ne correspond pas à une personne en particulier
 - soit parce que l'auteur souhaite rester anonyme (par exemple : « *phdelirium* » sur le blog « les aspects concrets de la thèse »)

Premiers résultats

	HAL-SHS	Présence commune	Hypotheses.org (carnets de recherche)
Contributeurs-chercheurs	13 557	111	415
Participation	179 017 dépôts		16 196 billets issus de 256 blogs

Premiers résultats

Sujet HAL-SHS	% d'auteurs présents sur les deux plateformes	% d'auteurs sur HAL-SHS
History	26,4%	9%
Library and information sciences	13,2%	6,3%
Social Anthropology and ethnology	8,6%	3,8%
Geography	8,4%	9,1%
Literature	7,5%	6%
Sociology	7,2%	7,3%
Linguistics	6,4%	9,7%
Art and art history	5,7%	2,3%
Political science	5,5%	3,4%
History, Philosophy and Sociology of Sciences	4,3%	2,8%
Architecture, space management	3,0%	2,1%
Education	2,5%	6,6%
Philosophy	2,0%	2,4%
Economy and finances	2,0%	7,6%

Présence commune

- Chercheurs participant massivement au dépôt de texte intégral (pour plus de 91% d'entre eux)
- contribuant plutôt à la CSI (notices avec fichiers récents).
- 25 % entrés récemment sur HAL-SHS (2015/2016).
- Parmi les contributeurs plus anciens (entrés avant 2015), 70% sont des contributeurs réguliers, 20 % ponctuels et environ 10 % d'irréguliers.

Suite...

- ✓ Quelles évolutions des dépôts ?
 - ✓ Concurrence des réseaux sociaux ?
 - ✓ Evolutions plateforme HAL (procédures dépôt, données, logiciels) ?
 - ✓ Impact de la loi numérique ? Des politiques OA ?

- ✓ Présence numérique des chercheurs ?
 - ✓ Analyse sémantique, qualitative : contenus, motivations
 - ✓ Analyse socio-technique : quels espaces, quelles légitimités ?

- ✓ **Projet HYCAR** (GIS Réseau des Urfist) : présence numérique, des nouvelles formes de publications, des écritures ordinaires de la recherche, des pratiques de communication au sein de l'activité de recherche, de l'autorité numérique, de l'engagement, etc.

<https://metacarnet.hypotheses.org/36>

Camille Prime-Claverie et Mahé, Annaïg, « Sites de dépôt en libre accès et nouvelles formes de médiation : quelles évolutions ? », in *La Médiation Numérique: Renouveau et Diversification Des Pratiques ; Actes du colloque 'Document Numérique et Société'*, Zagreb, 2013, ed. by Joumana Boustany, Information & Stratégie (Bruxelles Paris: De Boeck, 2014), pp. 110–125, https://hal.archives-ouvertes.fr/sic_00866225v1

Annaïg Mahé et Camille Prime-Claverie, « 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* [En ligne], 11 | 2017, <http://journals.openedition.org/rfsic/3315> , DOI : [10.4000/rfsic.3315](https://doi.org/10.4000/rfsic.3315)

Camille Prime-Claverie et Annaïg Mahé, Le défi de l'interopérabilité entre plates-formes pour la construction de savoirs augmentés en sciences humaines et sociales, *Ecrilecture augmentée dans les communautés scientifiques*, ISTE éditions, 2017, https://hal.archives-ouvertes.fr/sic_01511618v1

Annaïg Mahé et Camille Prime-Claverie, Open Science et présence numérique des chercheurs en sciences humaines et sociales, *Document Numérique*, n°2-3/2017, 85^e Colloque de l'ACFAS, Analyser la science : les bibliothèques numériques comme objet de recherche, Montréal, 8-9 mai 2017, DOI : [10.3166/dn.2017.00010](https://doi.org/10.3166/dn.2017.00010)