

Open science et présence numérique en sciences humaines et sociales

Annaig Mahé (Dicen-IDF, URFIST de Paris – ENC)

Camille Prime-Claverie (Dicen-IDF, Université Paris Nanterre)

Contexte

- 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

Questions de recherche

Analyse de la visibilité en ligne de la recherche

- Mise en œuvre de l'interopérabilité dans l'univers de la science ouverte : technique, sémantique, organisationnelle
- Logiques d'appropriation des nouveaux modes de communication par les acteurs de la communauté scientifique (producteurs et utilisateurs de ressources)

Production scientifique : état des lieux en SHS

- Pas de recensement systématique de la production au niveau international ni en France
- Marché de l'édition globalement national
- Peu d'outils développés par acteurs commerciaux
- Des initiatives plus récentes et fragmentées qu'en STM

Objectifs de l'étude

- Evaluer les logiques communicationnelles des contributeurs
- Type de contenu, volume, discipline
- Comparaison de la présence numérique sur HAL-SHS et Hypotheses.org

Périmètre de l'étude

□ Hypotheses.org

- Plateforme de blogging scientifique créée en 2008
- Une des briques de la plateforme OpenEdition
- Actuellement, 2000 carnets de recherche en SHS

□ HAL-SHS

Périmètre de l'étude

□ Hypotheses.org

□ HAL-SHS

- Partie Sciences Humaines et sociales de HAL
- Spécificité française : une archive nationale
 - Hal, Centre pour la communication scientifique directe (CCSD, CNRS)

Périmètre de l'étude

□ Hypotheses.org

□ HAL-SHS

- Partie Sciences Humaines et sociales de HAL
- Spécificité française : une archive nationale
- **Motivations historiques de HAL**
 - **Modèle de la « communication scientifique directe » : ArXiv**

Périmètre de l'étude

□ Hypotheses.org

□ HAL-SHS

- Partie Sciences Humaines et sociales de HAL
- Spécificité française : une archive nationale
- Motivations historiques de HAL
- Développement de portails institutionnels

Périmètre de l'étude

□ Hypotheses.org

□ HAL-SHS

- Partie Sciences Humaines et sociales de HAL
- Spécificité française : une archive nationale
- Motivations historiques de HAL
- Développement de portails institutionnels

Aujourd'hui : dualité de HAL : archive thématique et des archives institutionnelles

Présentation des corpus

	HAL-SHS	Hypotheses.org	
Périmètre	Extraction complète de HAL-SHS	Extraction complète des billets de blogs publiés en 2015	Les cahiers de recherche de chercheurs et de thèses en langue française
Provenance	Extraction OAI-PMH	Extraction OAI-PMH	corpus CLEO
période	2002-2016	2015	2005-2016
type doc	Notices	Notices	Texte intégral
volume	336 160 notices	44 265 billets issus de 1059 blogs	16 196 billets issus de 256 blogs
format	XML-TEI	Oai-dc	XML
éléments	Très structurés en particulier pour les auteurs et institutions de recherche	Manque des information sur les blogs (type, langue, pays, thématique)	Titre ; auteur ; date de publication ; texte ; titre du blog ; thématiques blog

Contribution sur 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-SHS

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-SHS

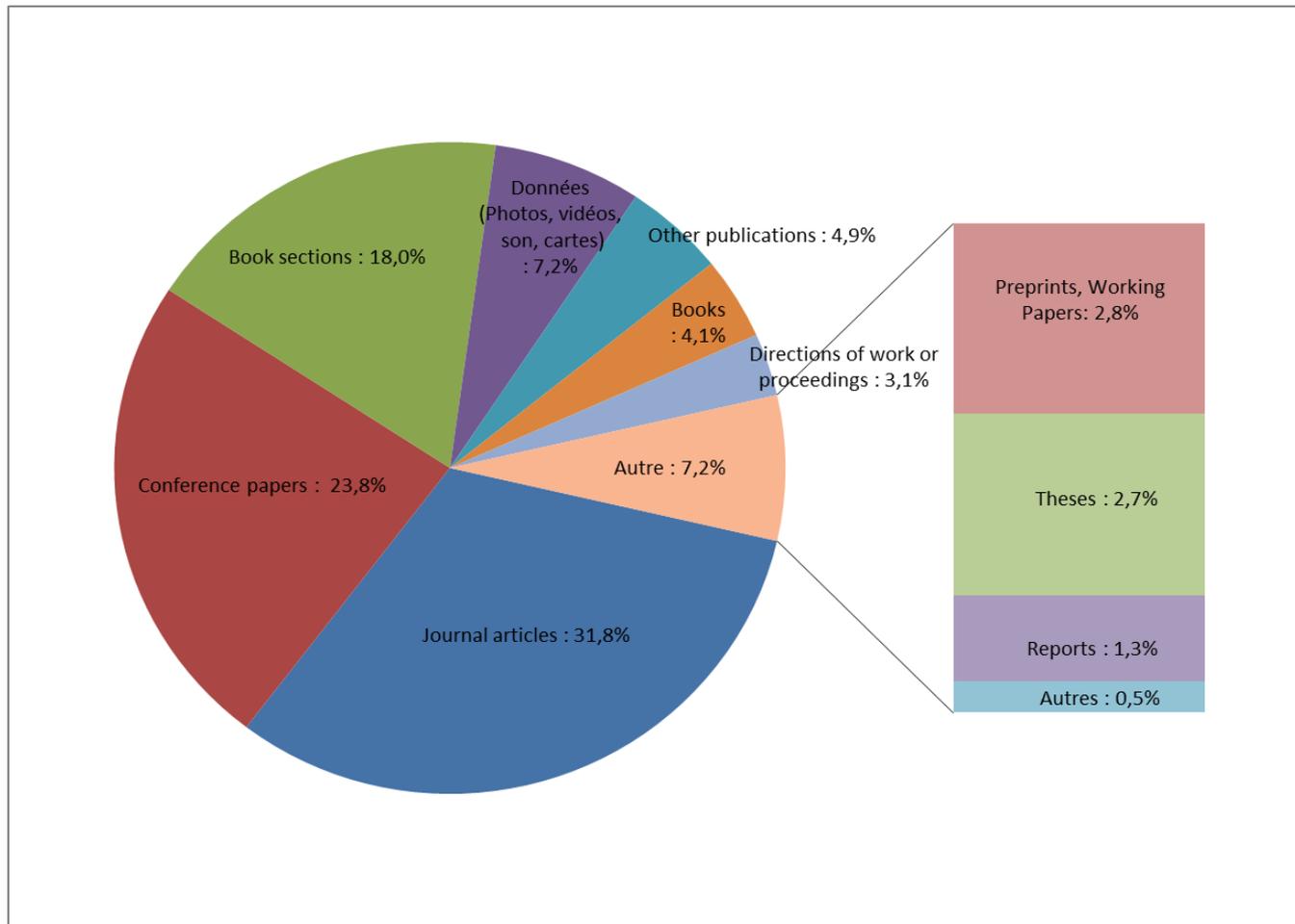
	Dépôt récent	Dépôt ancien
Dépôt avec fichier	Communication Scientifique immédiate	Archivage
Dépôt sans fichier	Visibilité de front de recherche (référencement)	Visibilité des recherches antérieures (recensement)

Nos hypothèses

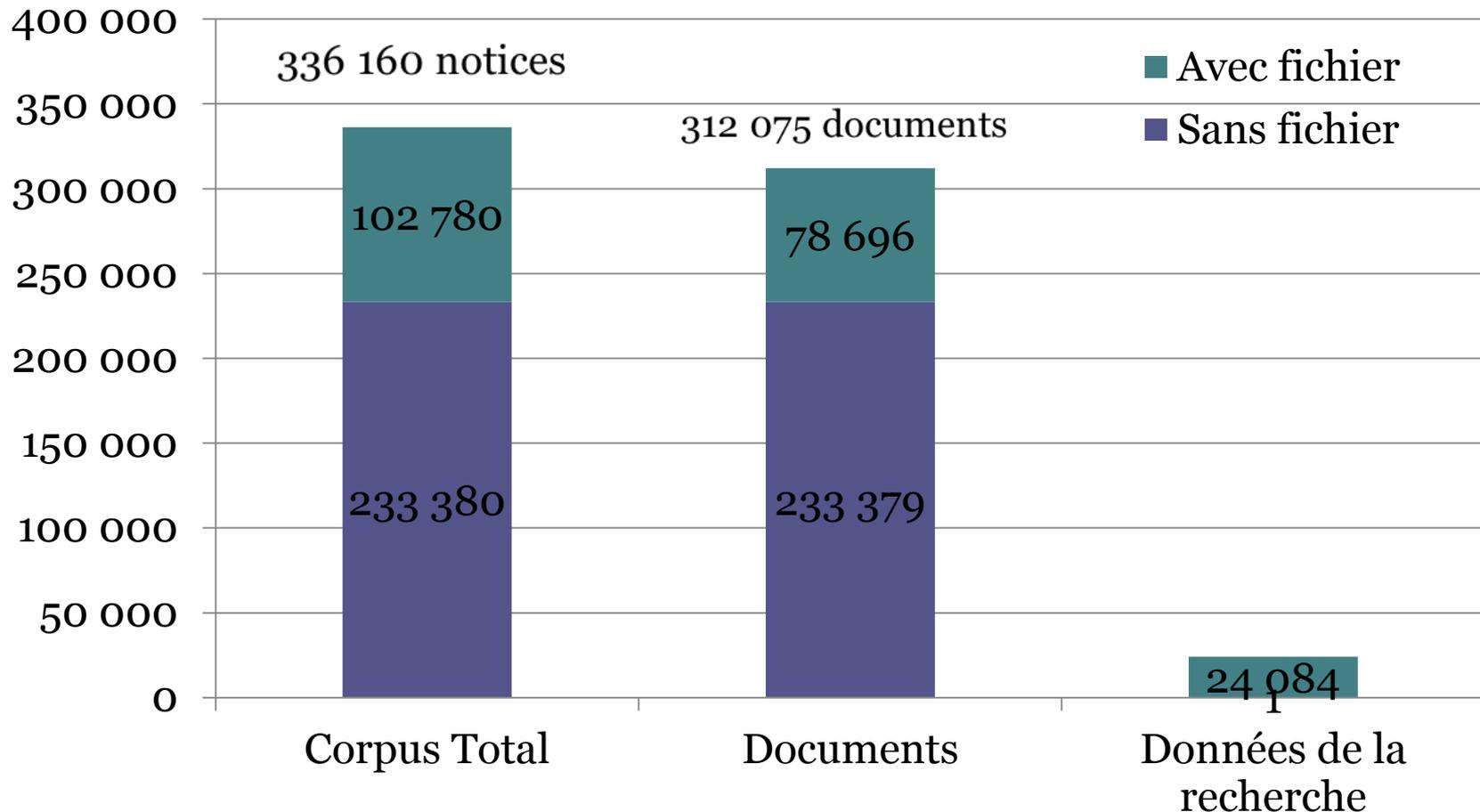
		Auteur	chercheur non auteur du document	Intermédiaire (non chercheur)	Transfert entre plateformes
Notice avec texte intégral	Document récent	CSD immédiate			Diffusion du texte intégral
	Document ancien				
Notice seule	Document récent			Référencement / recensement	
	Document ancien				

HAL-SHS : Répartition des dépôts

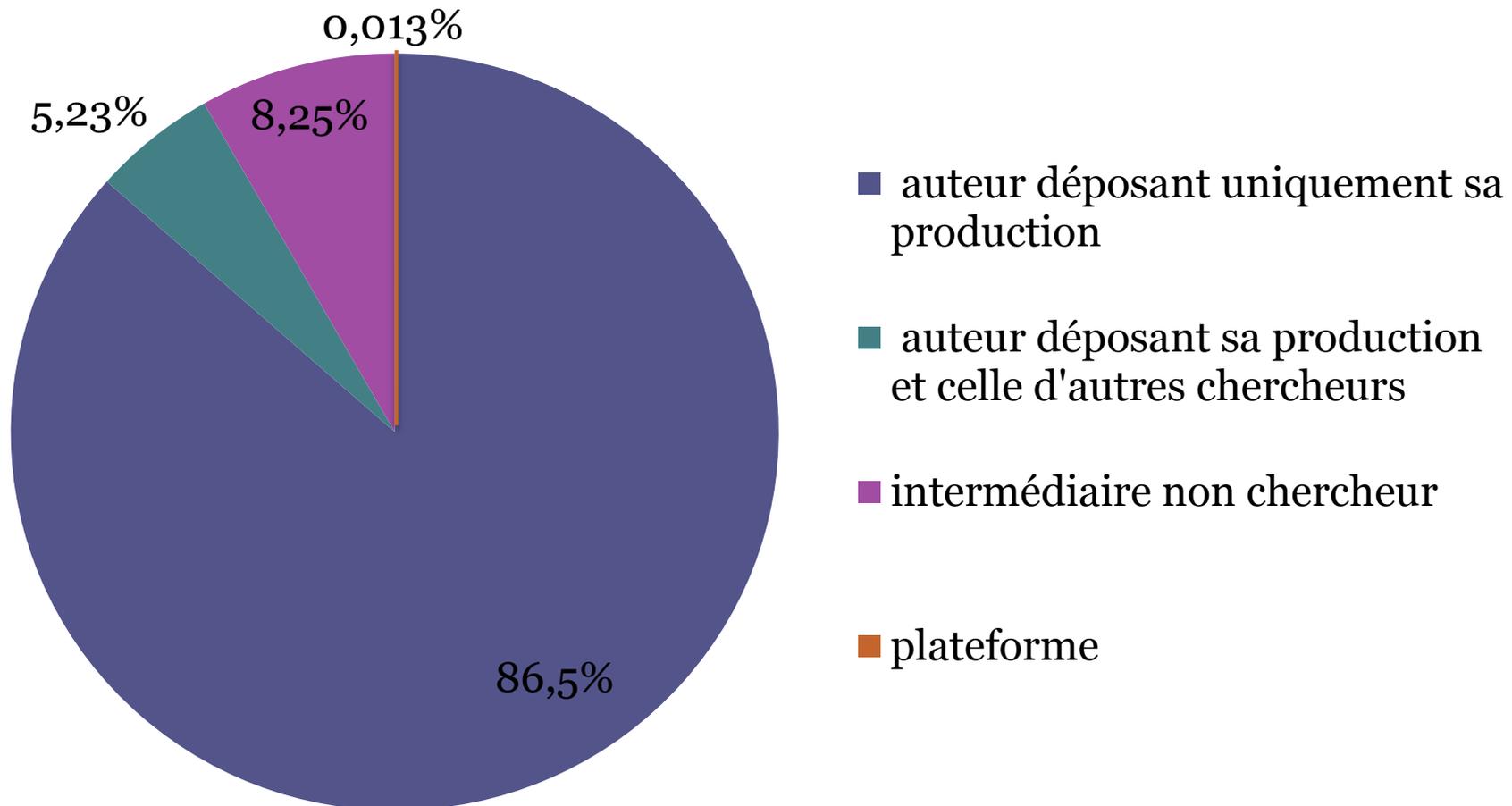
Type de document



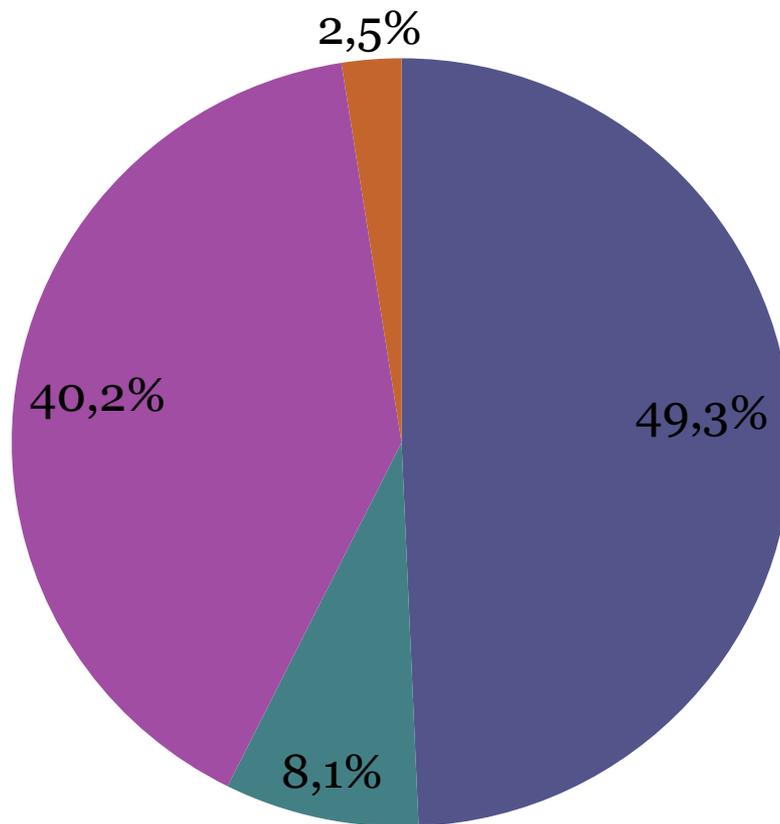
HAL-SHS : Répartition des dépôts



HAL-SHS : Répartition des déposants

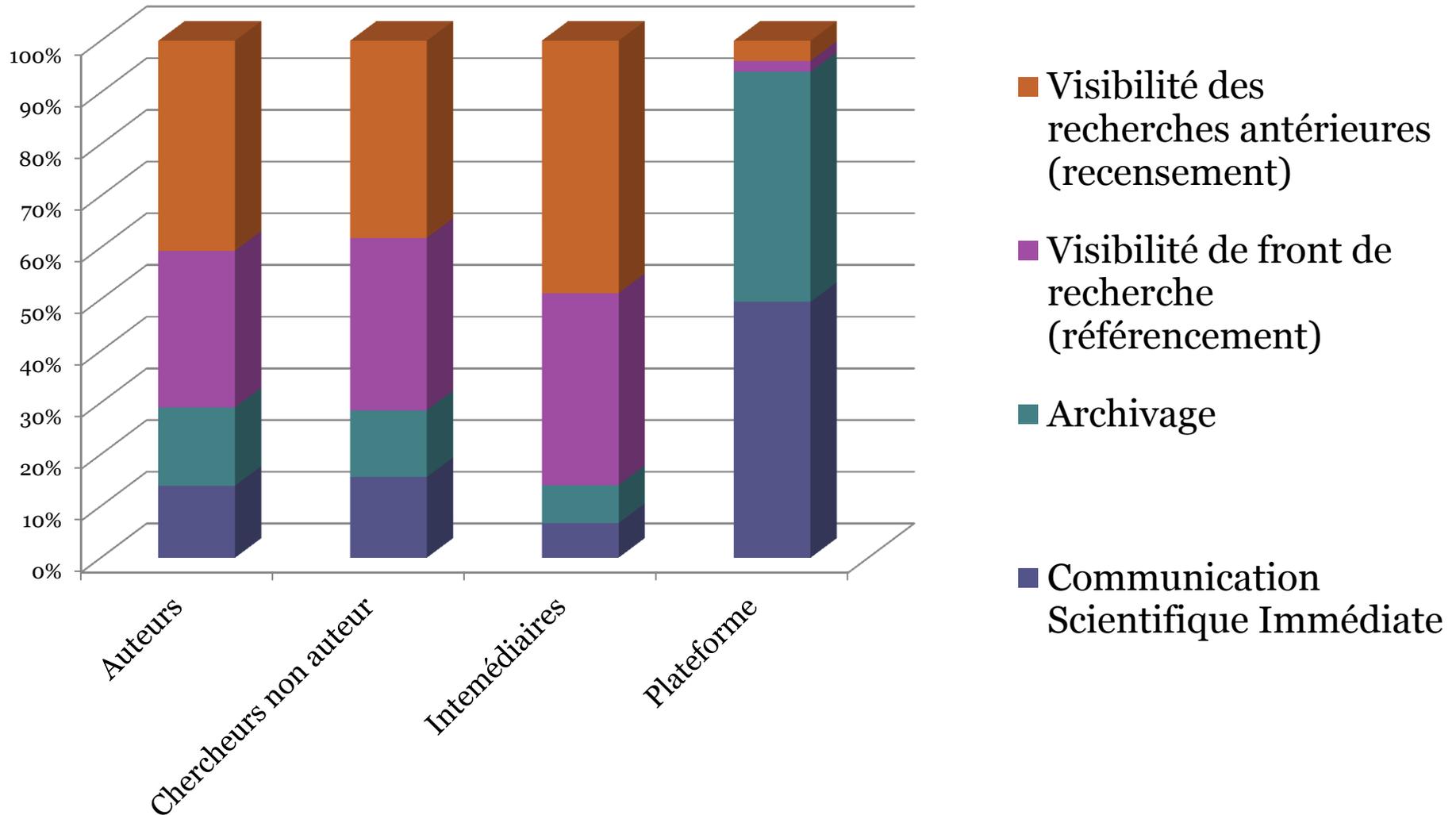


HAL-SHS : Répartition des dépôts en fonction du type déposant



- Niveau 1 : dépôt effectué par un des auteurs
- Niveau 2 : dépôt effectué par un autre chercheur
- Niveau 3 : dépôt effectué par un intermédiaire non chercheur
- Niveau 4 : dépôt automatique (plateformes)

HAL-SHS : Logiques de dépôts



HAL-SHS : des différences disciplinaires

Discipline	CSI	Archivage	Référence ment	Recensement
History	8,50%	13,30%	32,50%	45,60%
Economies and finances	21,00%	13,30%	35,60%	30,10%
Geography	14,70%	16,80%	32,70%	35,80%
Sociology	11,80%	15,30%	34,60%	38,30%
Linguistics	14,70%	15,40%	27,40%	42,50%
Archaeology and Prehistory	5,20%	10,60%	30,50%	53,80%
Law	4,70%	4,00%	44,90%	46,40%
Business administration	15,50%	13,40%	36,10%	35,00%
Literature	7,90%	11,60%	26,10%	54,50%
Political science	14,80%	17,40%	42,20%	25,60%
Education	25,90%	24,50%	23,20%	26,40%
Library and information sciences	26,30%	19,70%	22,70%	31,30%

Niveaux de contribution sur Hypotheses.org

Qui participe ?	Personne appartenant à la communauté scientifique : Chercheur, enseignant-chercheur, ingénieur, doctorant, etc.	
Niveau de contribution	Rédacteur(s) en chef (Responsable du blog) Niveau 1	Auteur(s) de Billets de blog Niveau 2
Limites	<i>Information non disponible dans nos corpus</i>	<i>Information parfois mal structurée</i>

Logiques de contribution sur Hypotheses.org

Diffuser les résultats de sa recherche

Discuter, commenter des résultats de recherche récents

**Partager des connaissances entre chercheurs
(outils, méthodes, etc.)**

Mettre en place une veille informationnelle sur un sujet

**Informar, mettre en visibilité ses activités (manifestations,
colloques), sa production scientifique, etc.**

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

Connexion entre les plateformes

- Des difficultés techniques
 - Sur [Hypotheses.org](https://www.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

Premiers résultats

	HAL-SHS	Présence commune	Hypotheses.org (carnets de recherche)
Contributeurs-chercheurs	14 200	107	415
Participation	336 160 notices		16 196 billets issus de 256 blogs
Auteurs	100 000 environ	174	415

Perspectives

