

**casuHAL**

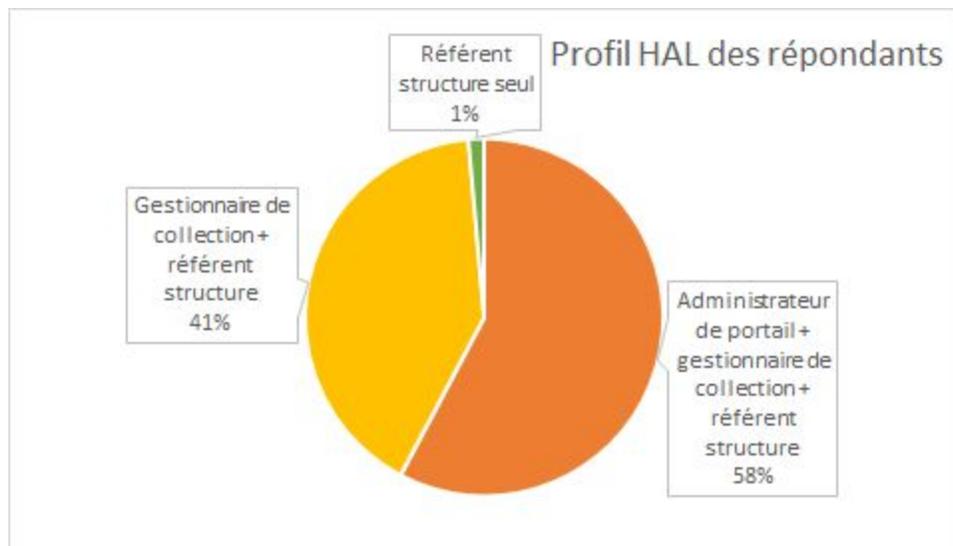
**Enquête sur  
l'organisation de la  
gestion des archives  
ouvertes sur HAL**

**GT 5 - Services aux administrateurs**

**Novembre 2016 - Février 2017**

Enquête conçue par le GT5 "service aux administrateurs" de CasuHal. Adressée aux administrateurs de portails, aux gestionnaires de collection et aux référents structures entre novembre 2016 et février 2017, elle a recueilli 71 réponses sur les 200+ membres de CasuHal.

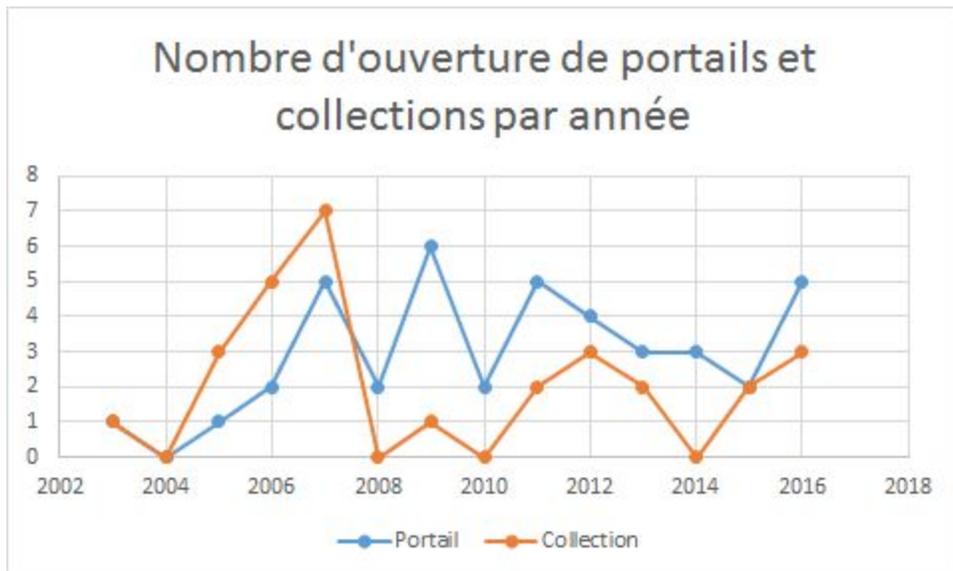
## 1/ Identification des répondants



HAL comptabilise actuellement au total 134 portails et 5207 collections. Nous avons obtenu 71 réponses, qui représentent 30 % des portails et 1,3 % des collections.

Nous avons différencié les réponses multiples (par défaut, un administrateur de portail est aussi gestionnaire de collection) pour distinguer les gestionnaires de collection exclusifs.

Le CCSD recense également 406 administrateurs, 2977 référents structures et 13447 gestionnaires de collections : ces données sont parfois éloignées de la réalité des établissements ; il est en effet fréquent que les droits d'administration (au sens large) de HAL ne soient pas supprimés lors du départ d'un agent ; il faudrait pouvoir repérer les détenteurs de droits sur HAL réellement actifs.

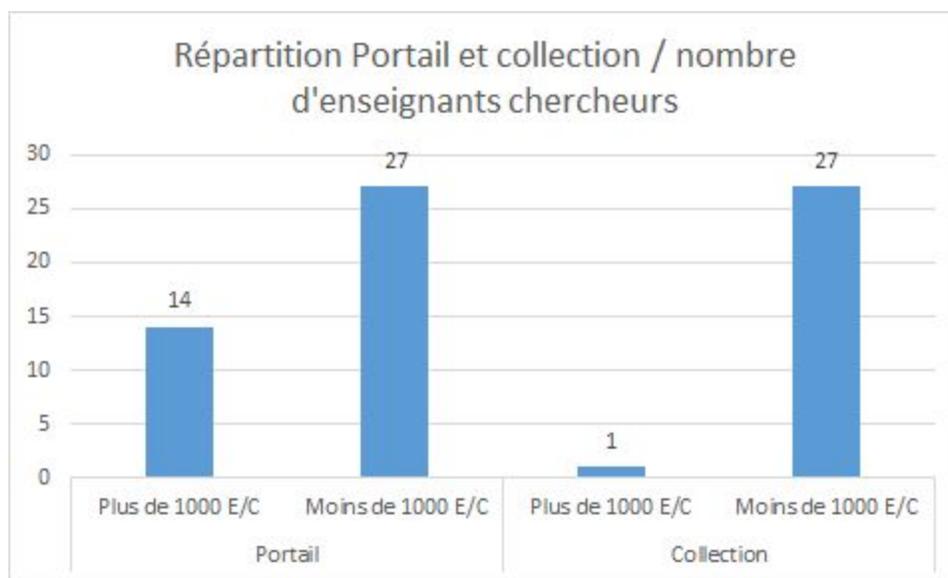


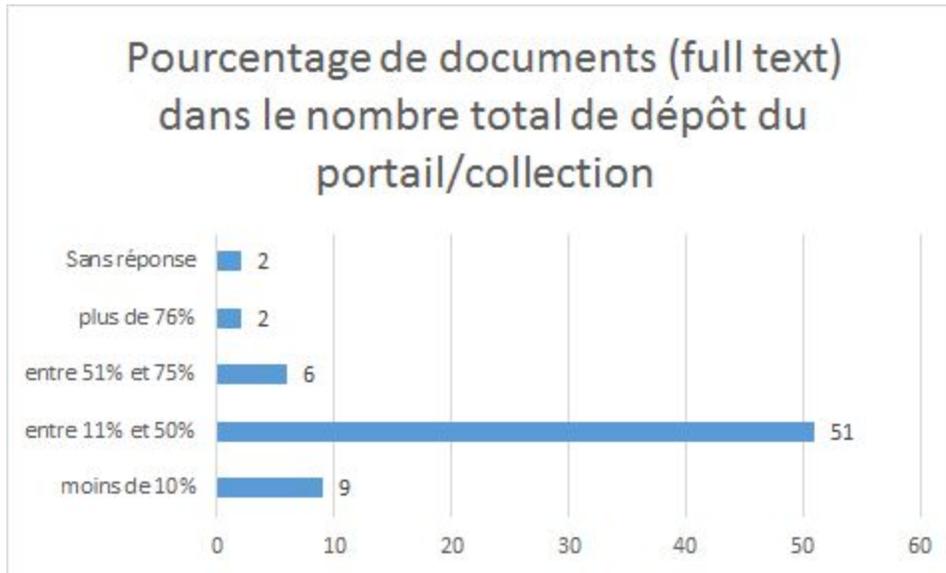
La possibilité de créer un portail sur HAL ayant été créée en 2003, la majorité des réponses antérieures concerne les collections. On observe que la création de portails se fait à un rythme régulier depuis 2003. 7 résultats non exploitables ne sont pas représentés sur ce graphique.

## 2/ Volumétrie

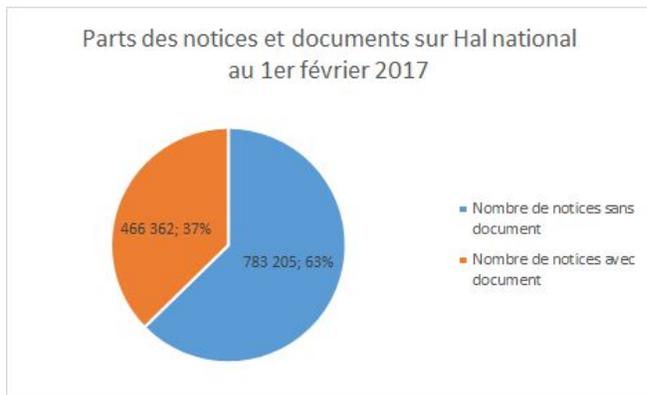
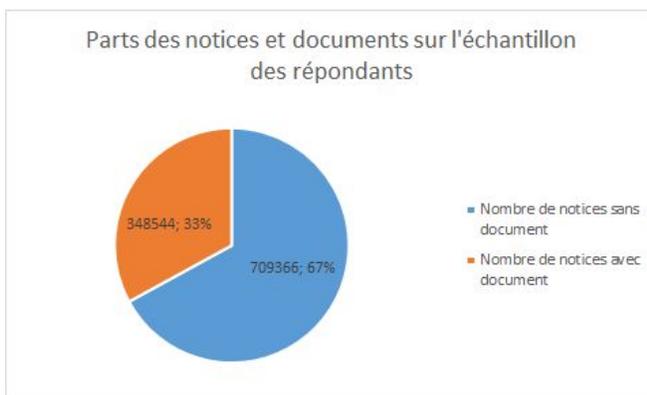
Sur 71 réponses, 41 concernent des établissements disposant d'un portail. Parmi ces établissements, 34 % sont des « gros » établissements (+ de 1000 EC), majoritairement des universités, mais aussi quelques organismes de recherche (INRIA, ONERA, Pasteur). Les 2 tiers restants (59%) se composent d'universités, d'organismes de recherche, de laboratoires, de grandes écoles, d'autres établissements.

A noter : 9 établissements à moins de 200 EC disposent d'un portail, et le regroupement d'établissement qui dénombre le plus d'EC (10300), la COMUE USPC, a fait le choix d'une collection.





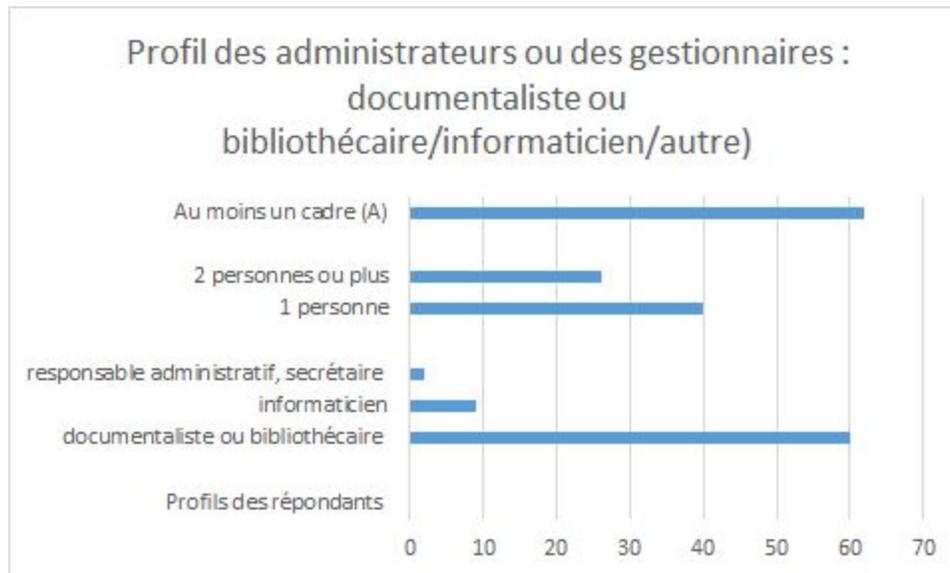
Sur la totalité de notre échantillon, le texte intégral représente un peu plus du tiers des dépôts. Notre échantillon se rapproche de la répartition en vigueur dans Hal générique, qui couvre 37% de texte intégral en février 2017. Noter qu'un établissement a fait le choix du texte intégral unique, il est donc à 100% de full text.



### 3/ Ressources humaines

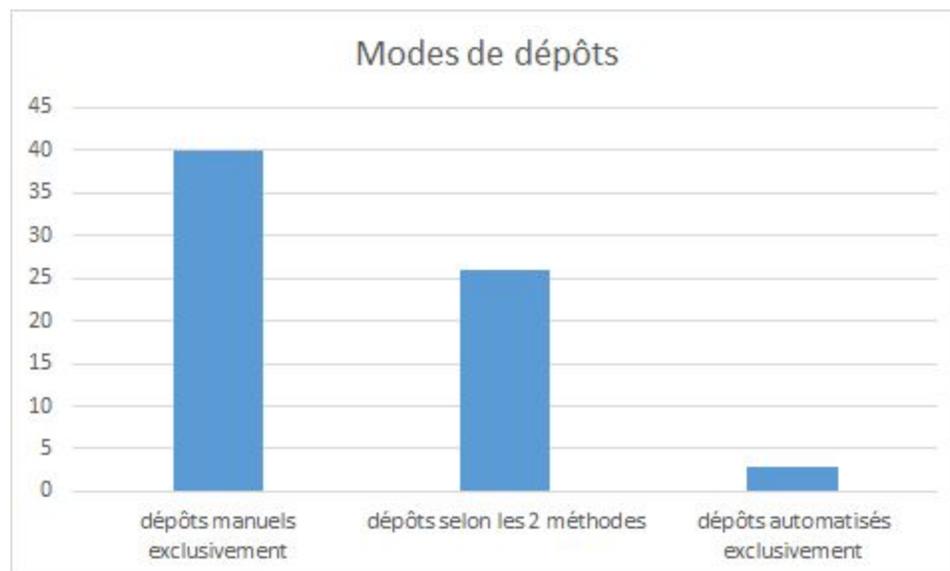
Nous avons voulu avoir une idée du profil professionnel des personnes intervenant dans HAL dans les établissements ; nous avons donc posé des questions concernant la catégorie, le statut (titulaire ou contractuel), le temps de travail consacré à HAL (nombre d'ETP). Nous avons obtenu des réponses extrêmement variées et incomplètes, qu'il nous a été impossible d'exploiter. Nous avons tout de même pu établir que les établissements investissent majoritairement sur des agents de catégorie A, parfois renforcés par des agents de catégorie B dans les structures les plus importantes.

Nous avons eu plus de succès dans les réponses concernant le type de profil (bibliothécaire/documentaliste ou informaticien) : il s'avère que l'administration de HAL est majoritairement une activité confiée à des professionnels de l'IST.



En marge des questions sur les ressources humaines, nous nous sommes interrogés sur les modes de dépôt privilégiés par les administrateurs et gestionnaires. Nous avons distingué le mode manuel (la saisie des notices se fait à la pièce), du mode automatique (l'administrateur ou le gestionnaire procède à des imports de notices par lot, généralement avec l'outil Bib2Hal).

On note que la majorité des répondants ne fait que du dépôt manuel, et que le dépôt automatisé n'est pas encore bien entré dans les usages. Il serait intéressant de creuser cette question ultérieurement par une enquête spécifique, pour voir notamment quels peuvent être les freins au dépôt automatique, quels sont les outils utilisés, dans quel cas une méthode est privilégiée par rapport à l'autre (à partir de combien de notices les imports sont-ils "rentables"), etc.



## 4/ Données bibliographiques

Nous nous sommes également intéressés au processus de collecte des données dans HAL : les répondants participent davantage à la collecte des données bibliographiques (70% interviennent régulièrement ou occasionnellement sur les métadonnées) qu'à la collecte des documents en texte intégral - dans ce cas, c'est plutôt sous une forme incitative (communication, accompagnement, formations dans un peu moins de 50% des cas).

Ce travail sur les métadonnées se compose de 3 activités principales :

- La vérification de la qualité des métadonnées (pour 60 répondants, dont 46 régulièrement)
- La collecte de données par la communication avec les chercheurs (pour 66 répondants, dont 42 régulièrement)
- La collecte de données par la mise en place d'une veille sur internet pour la communauté est moins répandue (40 personnes en font régulièrement ou occasionnellement) ; par ailleurs, c'est une activité qui semble cliver les répondants en 2 groupes (28 en font régulièrement, 21 n'en font jamais).

Parmi les activités plus rares, car vraisemblablement plus ponctuelles, on note la reprise de listes bibliographiques et/ou de rapports quinquennaux, ou encore la reprise de bases bibliographiques existantes. La reprise de listes bibliographiques semble une activité légèrement plus courante que la reprise de bases bibliographiques.

A noter qu'un tiers des répondants (reprises de listes) et un peu plus de la moitié des répondants (reprise de bases bibliographiques) n'ont jamais eu à intervenir de la sorte.

La plupart des commentaires concernant les éventuelles méthodes alternatives pour récupérer des données permettant d'alimenter HAL relèvent des différentes méthodes de veille documentaire (sur les réseaux sociaux, dans les lettres d'infos, dans les sommaires des revues et les tables de matières des ouvrages, etc), et mettent également en avant les activités de formation et de soutien au dépôt par les professionnels de l'IST (on peut noter que l'implication de secrétaires de laboratoires dans la saisie des références n'est mentionnée que par un seul répondant).

Au coté de ce type d'activité, on notera d'autres méthodes originales et/ou moins répandues, comme le suivi des déplacements donnant lieu à communication dans des conférences (travail en lien avec le service financier), ou l'envoi de messages invitant le chercheur à déposer son article ou à envoyer son manuscrit au service documentaire qui se chargera du dépôt.

Une grande partie du travail des administrateurs et des référents structure consiste à mettre à jour les données dans AuréHAL, la base des référentiels de HAL : les répondants ont donc été invité à indiquer sur quel(s) référentiel(s) portaient leurs efforts, et, dans la mesure du possible, ce que cela représentait en terme de temps de travail.

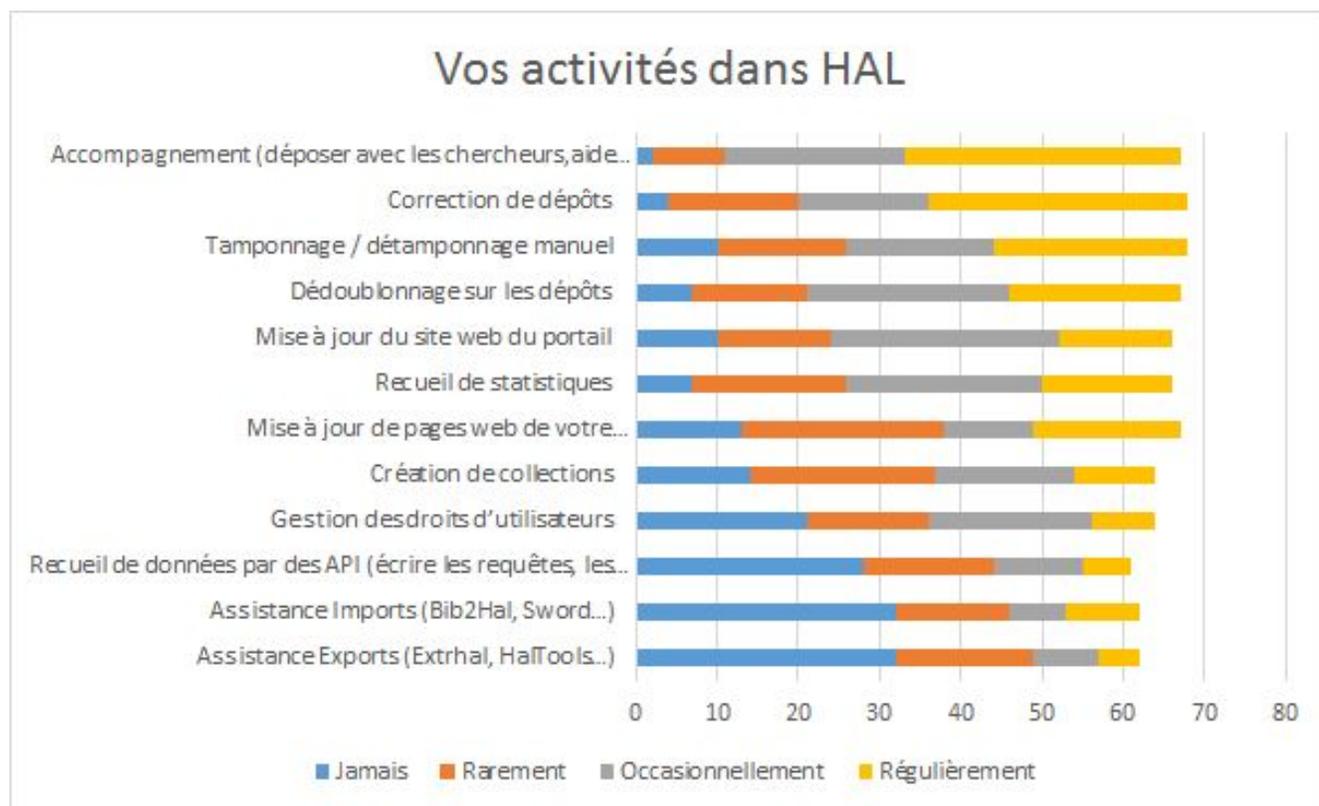
Les référentiel "Auteurs" et "Structures" sont ceux sur lesquels les répondants interviennent le plus régulièrement/occasionnellement (62%), suivi du référentiel "Revue" (28%). Il est à noter qu'une majorité des répondants ne travaillent jamais sur les référentiels "Projets ANR" ou "Projets européens" (pour mémoire, le référentiel "Domaines" n'est modifiable que par le CCSD).

La question du temps de travail consacré aux référentiels a donné des réponses trop hétérogènes pour

pouvoir être exploitées, les répondants soulignant la difficulté, voire l'impossibilité de quantifier cette activité. Certains répondant ont déploré le fait de ne pas pouvoir intervenir sur les référentiels en fonction de leurs droits, limités à leur propre structure.

## 5 / Administration du portail

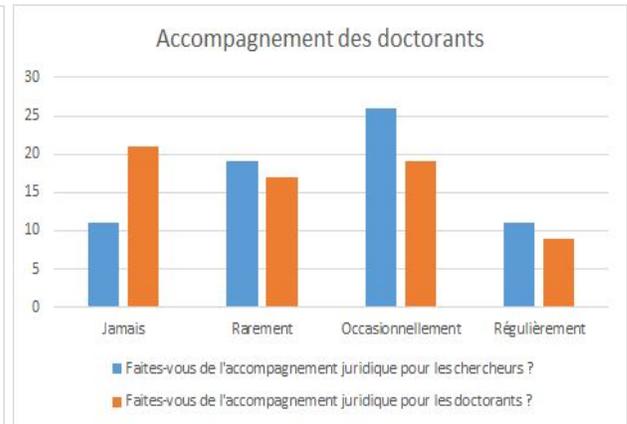
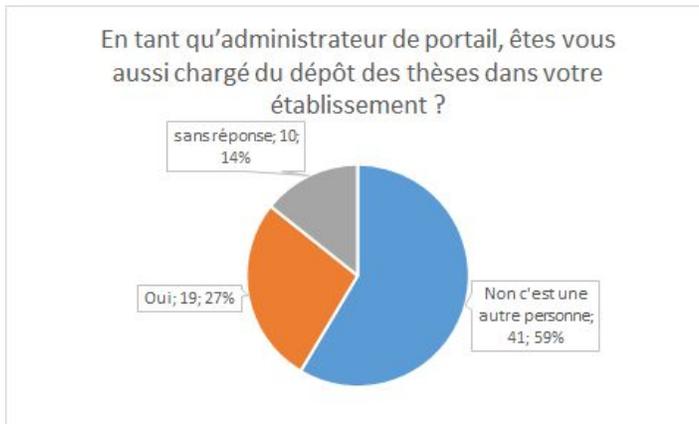
L'administration d'un portail dans HAL se décline en une gamme variée d'activités, couvrant aussi bien des tâches plutôt "techniques" (imports, exports...) que "documentaires" (dédoublonnage, tamponnage...) ou de sensibilisation. L'accompagnement, la formation, l'assistance au dépôt sont les activités principales des administrateurs/trices de portail, suivies du travail sur les dépôts (corrections et dédoublonnage), les aspects plus techniques (API, imports, exports) étant les moins cités par les répondants.



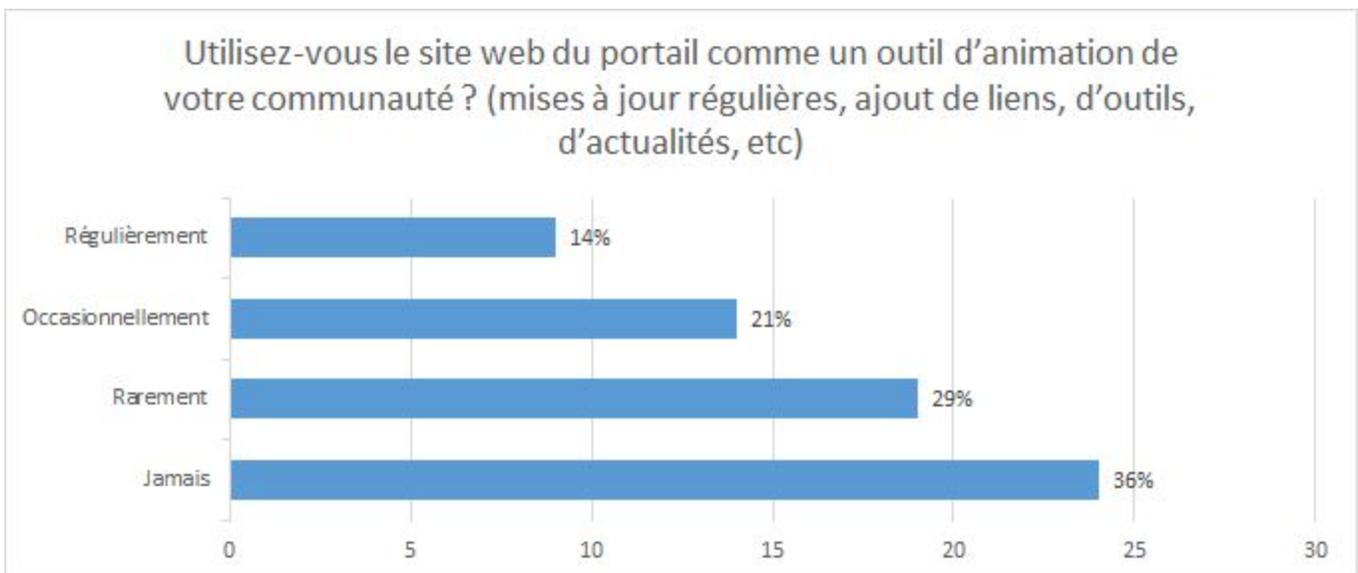
## 6 / Autres missions

D'autres missions peuvent être dévolues aux administrateurs/trices de portail et aux gestionnaires de collections, en rapport avec avec HAL et/ou avec l'open access :

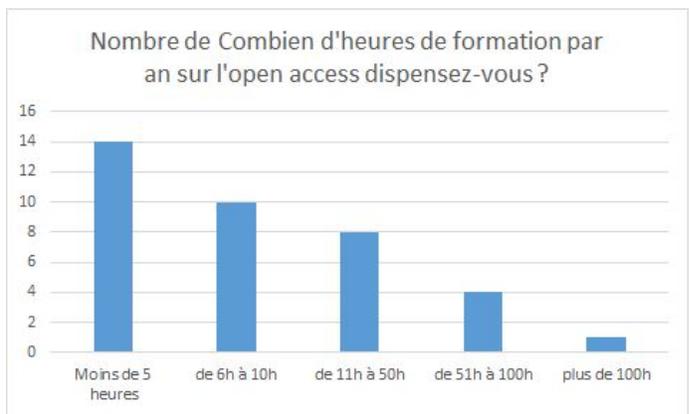
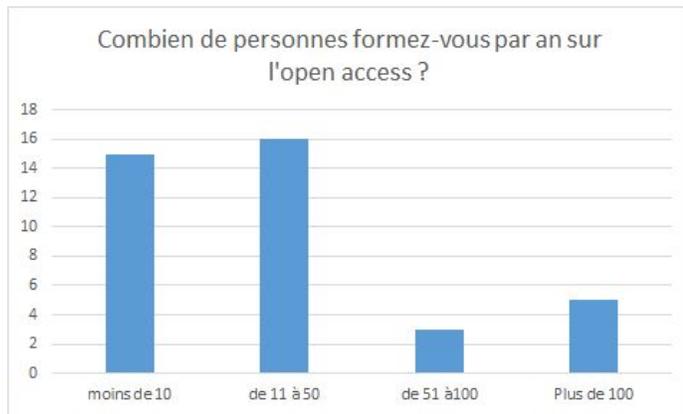
- Missions liées aux doctorants et aux thèses : un tiers des répondants déclare avoir des responsabilités dans le dépôt des thèses pour son établissement. Dans ce cadre, ils sont amenés à faire occasionnellement ou régulièrement de l'accompagnement juridique pour les doctorants.



- Missions de communication avec HAL : il semble que les portails et les collections ne soient pas vraiment perçus comme des outils de communication voire d'animation des communautés, seul un tiers des répondants indiquent les utiliser à ces fins.



- Missions liées à la formation : on l'a vu plus haut, les activités de sensibilisation et d'accompagnement remplissent une partie importante du travail d'administration. Les répondants semblent cependant éprouver des difficultés à quantifier l'impact de leurs actions : on peut supposer que ce sont des missions qui émergent dans les établissements, et sont en cours de mise en place. Une autre explication peut venir de la difficulté à déterminer ce qui est de l'ordre de la formation, ou de la sensibilisation et de l'information des communautés ; l'assistance au dépôt présente aussi un caractère pédagogique... Le nombre d'heures déclarées est également caractéristique d'un service en émergence dans les établissements.



- Missions en lien avec l'open access : un tiers des répondants à l'enquête a indiqué avoir des missions plus larges sur l'open access. Parmi celles-ci reviennent le plus souvent des responsabilités sur l'accès aux données de la recherche, ainsi qu'en matière d'édition électronique et d'open access "en général". Sont également cités : les carnets de recherche, l'accès aux mémoires sur DUMAS, le suivi des APC, la bibliométrie, la création d'outils autour de HAL.

## 7/ Réseau

En tant que club utilisateur, nous ne pouvons pas ne pas nous interroger sur l'intérêt des répondants pour la participation à un ou des réseau(x). Plus de la moitié des répondants déclare adhérer à CasuHal, et un quart au groupe de travail sur l'open access du consortium Couperin (GTAO) ; parfois ils adhèrent également à un réseau local d'établissement (Toulouse, Montpellier, Grenoble sont cités).

Près d'un tiers des répondants ne participe à aucun réseau local ; ceux qui sont actifs dans les réseaux le sont à égalité en présentiel (réunions régulières) et/ou à distance via des listes de diffusion ; ils peuvent participer aux actions variées de formation et d'information mises en place dans le cadre de ces réseaux.

# ANNEXES

## Détails des réponses

### Profil des répondants

<b>Vous êtes</b>	
Administrateur de portail + gestionnaire de collection + référent structure	41
Gestionnaire de collection + référent structure	29
Référent structure seul	1

<b>Nombre d'ouverture de portails et collections par année</b>		
<b>Année</b>	<b>Portail</b>	<b>Collection</b>
2003	1	1
2004	0	0
2005	1	3
2006	2	5
2007	5	7
2008	2	0
2009	6	1
2010	2	0
2011	5	2
2012	4	3
2013	3	2
2014	3	0
2015	2	2
2016	5	3

## Volumétrie

	Nombre de documents en texte intégral du portail ou de la collection	Nombre de notices du portail ou de la collection	Somme total d'entrées dans le portail ou collection	pourcentage du nombre de texte intégral dans les notices totale pour chaque portail/collection
	3185	0	3185	100,00
	15030	4298	19328	77,76
	299	171	470	63,62
	2136	1437	3573	59,78
	2251	1567	3818	58,96
	429	307	736	58,29
	47489	35425	82914	57,28
	1582	1299	2881	54,91
	7774	7774	15548	50,00
	1998	2191	4189	47,70
	2326	2579	4905	47,42
	2233	2595	4828	46,25
	7294	8924	16218	44,97
	5090	6452	11542	44,10
	106	141	247	42,91
	748	1044	1792	41,74
	2346	3391	5737	40,89
	8999	13346	22345	40,27
	386	597	983	39,27
	69	110	179	38,55
	5532	8993	14525	38,09
	1590	2634	4224	37,64
	32190	55014	87204	36,91
	15	26	41	36,59
	11380	22024	33404	34,07
	6358	12416	18774	33,87
	5237	10544	15781	33,19
	5237	10544	15781	33,19
	25640	51673	77313	33,16
	4277	8815	13092	32,67
	1132	2371	3503	32,32
	15761	33288	49049	32,13

	455	1015	1470	30,95
	34345	77256	111601	30,77
	4558	10563	15121	30,14
	3456	8318	11774	29,35
	695	1929	2624	26,49
	140	391	531	26,37
	1823	5208	7031	25,93
	15373	44701	60074	25,59
	3056	9076	12132	25,19
	4669	14164	18833	24,79
	639	1964	2603	24,55
	500	1600	2100	23,81
	4150	13356	17506	23,71
	4881	16274	21155	23,07
	5963	20299	26262	22,71
	4407	15654	20061	21,97
	1036	3957	4993	20,75
	6239	23839	30078	20,74
	753	2886	3639	20,69
	13978	54500	68478	20,41
	4720	19018	23738	19,88
	1596	6838	8434	18,92
	583	2675	3258	17,89
	387	1878	2265	17,09
	284	1568	1852	15,33
	318	2018	2336	13,61
	888	6800	7688	11,55
	183	1577	1760	10,40
	139	1260	1399	9,94
	269	2764	3033	8,87
	187	1973	2160	8,66
	60	658	718	8,36
	70	805	875	8,00
	23	272	295	7,80
	1503	17907	19410	7,74
	99	2415	2514	3,94
Total	348544	709366	1057910	32,95

## Données bibliographiques

	Régulièrement	Occasionnellement	Rarement	Jamais
En vue d'alimenter HAL, intervenez-vous directement dans le processus de récolte des données bibliographiques ? (imports...)	25	18	14	14
En vue d'alimenter HAL, intervenez-vous directement dans le processus de récolte des articles eux-mêmes ?	19	16	19	17

Comment intervenez-vous dans ces processus ?	Régulièrement	Occasionnellement	Rarement	Jamais	Vide
[Veille internet/BDD sur les publications de votre communauté ?]	23	12	8	22	6
[Communication régulière avec les chercheurs]	37	23	5	1	5
[Reprise de rapports quinquennaux et listes bibliographiques diverses]	8	22	13	20	8
[Reprise d'anciennes bases de données biblio]	5	13	9	34	10
[Vérification de la qualité des données]	42	13	9	2	5
Autre	7	1	1	9	53

## Administration du portail

Vos activités dans HAL				
	Jamais	Occasionnelle ment	Rarem ent	Régulière ment
Recueil de données par des API (écrire les requêtes, les exploiter)	28	11	16	6
Tamponnage / détamponnage manuel	10	18	16	24
Dédoublonnage sur les dépôts	7	25	14	21
Correction de dépôts	4	16	16	32
Gestion des droits d'utilisateurs	21	20	15	8
Création de collections	14	17	23	10
Assistance Imports (Bib2Hal, Sword...)	32	7	14	9
Assistance Exports (Extrhal, HalTools...)	32	8	17	5
Mise à jour du site web du portail	10	28	14	14
Mise à jour de pages web de votre établissement/laboratoire avec des données provenant de HAL	13	11	25	18

Accompagnement (déposer avec les chercheurs, aide pour créer son idhal, son CV, conseiller sur la méthodologie pour une personne qui dépose pour les autres...)	2	22	9	34
Recueil de statistiques	7	24	19	16

## Autres missions

	Non	Oui					
<b>En tant qu'administrateur de portail, êtes vous aussi chargé du dépôt des thèses dans votre établissement ?</b>	41	18					
	Jamais	Occasionnellement	Rarement	Régulièrement			
<b>Faites-vous de l'accompagnement juridique pour les chercheurs ?</b>	11	26	19	11			
	Jamais	Occasionnellement	Rarement	Régulièrement			
<b>Faites-vous de l'accompagnement juridique pour les doctorants ?</b>	21	19	17	9			
	Je ne sais pas	Correspondant Star/Thèses	Ecole doctorale, encadrant	Services juridiques	SCD/Service documentaire	Pas d'accompagnement juridique	Pas de doctorant
<b>Sinon qui fait de l'accompagnement juridique pour les doctorants ?</b>	4	14	2	6	2	11	1
	Jamais	Occasionnellement	Rarement	Régulièrement			
<b>Utilisez-vous le site web du portail comme un outil d'animation de votre communauté ? (mises à jour régulières, ajout de liens, d'outils, d'actualités, etc)</b>	24	14	19	9			
	0	moins de 10	de 11 à 50	de 51 à 100	Plus de 100		
<b>Combien de personnes formez-vous par an sur l'open access ?</b>	3	12	16	3	5		
	0	moins de 5	de 6 à 10	de 11 à 50	de 51 à 100	plus de 100	

Combien d'heures de formation par an sur l'open access dispensez-vous ?	7	7	10	8	4	1	
	<b>Liste de discussion entre contributeurs</b>	<b>réunions régulières</b>	<b>actions du réseau</b>	<b>no réseau</b>	<b>Autre</b>		
En tant que gestionnaire de collections, participez-vous à un réseau local ? Si oui, quel est son fonctionnement ?	17	16	14	20	7		
	<b>GTAO Couperin</b>	<b>CasuHal</b>	<b>Autre</b>				
Etes-vous membre d'un réseau sur l'Open access ?	20	51	10				