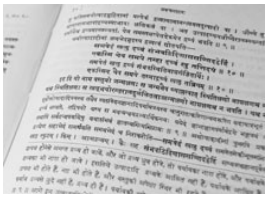
BULAC

[הפריה] [大学] [γλώσσες] [ཤེས་རིག]

Bibliothèque universitaire
des langues et civilisations

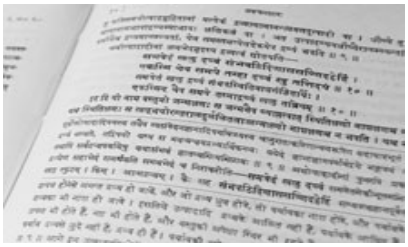


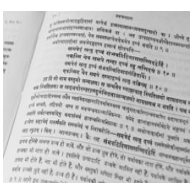
ENQUÊTE DE BESOIN

Besoins de formation des enseignants-chercheurs en études aréales

Pauline Belvèze

Version : Août 2023





GIP BULAC
65 rue des Grands Moulins
F-75013 PARIS

T +33 1 81 69 18 00
contact@bulac.fr
www.bulac.fr

E

TABLE DES MATIÈRES

CONTEXTE 3

1. LA FORMATION DES PUBLICS À LA BULAC....3

2. LA CONNAISSANCE DES BESOINS DES PUBLICS : PRATIQUES D'ENQUÊTE....3

ENQUÊTE 2023 SUR LES BESOINS DE FORMATION DES CHERCHEURS EN ÉTUDES ARÉALES 5

3. MÉTHODOLOGIE....5
- 3.1. Public ciblé....5
- 3.2. Construction de l'enquête et diffusion....5

4. PROFIL DES RÉPONDANTS....6

5. LES PRATIQUES DE LA RECHERCHE ET LES DIFFICULTÉS RENCONTRÉES....7
- 5.1. La collecte des matériaux de recherche....7
- 5.2. L'organisation et le traitement des matériaux de la recherche....9
- 5.3. Valorisation et publication des résultats de la recherche....10
- 5.4. Éthique et intégrité scientifique....11

6. LES BESOINS DE FORMATION EXPRIMÉS....11

7. BILAN ET PROSPECTIVE....13

CONTEXTE

1. LA FORMATION DES PUBLICS À LA BULAC

Depuis son ouverture en 2011, la BULAC déploie une offre de formation destinée à un large public qui comprend les enseignants-chercheurs et les étudiants des établissements membres du groupement d'intérêt public BULAC[1], des étudiants et des chercheurs issus d'autres établissements universitaires ainsi que des lecteurs sans rattachement académique qui utilisent les espaces et les collections. En découle une grande diversité de profils pour lesquels la bibliothèque veille à proposer des formations adaptées. Si l'accompagnement méthodologique des étudiants y occupe une place importante, la formation des chercheurs ne cesse également d'être enrichie.

L'activité de formation est organisée et mise en œuvre par une équipe de formateurs qui intervient auprès des publics de niveau études et recherche. Certaines séances mobilisent en outre les compétences des chargés de collections, de conservation et de l'informatique documentaire.

Pour le niveau recherche, les formations sont déclinées en deux volets complémentaires : les formations doctorales et celles intéressant l'ensemble des chercheurs. Les premières, conçues dans l'esprit des textes encadrant le diplôme du doctorat[2], doivent permettre l'appropriation des outils et des procédures nécessaires à la réalisation de la thèse et à la préparation de l'après-doctorat. Les secondes, également ouvertes aux doctorants, combinent des présentations synthétiques centrées sur les différentes étapes du cycle de la recherche et des ateliers destinés à certains outils.

La particularité de l'offre proposée réside dans la nature des domaines et des thèmes d'études des publics de la bibliothèque. L'approche pluridisciplinaires des terrains extra-occidentaux, qui caractérise les études aréales, nécessite une adaptation du contenu des formations et le recours à des exemples adaptés. La connaissance des besoins de cette communauté doit donc être sans cesse actualisée et approfondie.

2. LA CONNAISSANCE DES BESOINS DES PUBLICS : PRATIQUES D'ENQUÊTE

Le déploiement des ressources et des services de la BULAC repose sur un sondage périodique des besoins de ses publics. Cette démarche revêt une importance particulière du fait du souci d'amélioration continue des services rendus par la bibliothèque. La spécificité de ses collections et de ses publics invite à accorder une attention particulière aux besoins documentaires rencontrés et aux pratiques informationelles développées.

L'activité de formation s'accompagne à la BULAC de dispositifs d'évaluation déclinés pour le niveau Licence, Master et Doctorat. Sont étudiés le degré d'adéquation des formations aux attentes des publics et les propositions d'amélioration ou d'enrichissement de l'offre.

Conduite de manière systématique pour la formation doctorale, cette évaluation demeurait plus ponctuelle pour les autres formations de niveau recherche. L'enquête de 2023 visait donc à affiner la connaissance des besoins de ces publics en sondant leurs pratiques, les difficultés rencontrées et le degré de maîtrise des stratégies et des outils pour la recherche en études aréales.

1 Créé en novembre 2003, le GIP BULAC rassemble, en septembre 2023, le CNRS, l'École française d'Extrême-Orient (EFEO), l'École pratique des hautes études (EPHE), l'École des hautes études en sciences sociales (EHESS), L'Institut national des langues orientales (Inalco), l'Université Panthéon-Sorbonne, l'Université Sorbonne Nouvelle, Sorbonne Université et l'Université de Paris Cité.

2 Voir l'Arrêté du 26 août 2022 modifiant l'arrêté du 25 mai 2016 fixant le cadre national de la formation et les modalités conduisant à la délivrance du diplôme national de doctorat, disponible à cette adresse : https://www.legifrance.gouv.fr/eli/arrete/2022/8/26/ESRS2220637A/jo/texte [consulté le 23/08/2023].

E

Concernant le public des chercheurs, la BULAC avait mené en 2015 une vaste enquête ayant recueilli plus de 300 réponses[3]. Il s'agissait de sonder plus particulièrement les profils documentaires et informationnels des chercheurs en études aréales. Cette enquête révélait des usages encore peu développés des ressources électroniques et des outils de signalement ainsi que des pratiques limitées en matière de science ouverte.

Apparaissaient en creux des besoins de formation que les évolutions technologiques et leur impact sur la recherche n'ont fait qu'accroître. Il convenait donc de sonder à nouveau les besoins de formation en matière d'accès aux ressources documentaires et de prise en main d'outils au service de la recherche.

3 Les principaux résultats de l'enquête sont consultables dans le compte-rendu de Soline Lau-Suchet, "Enquête auprès des chercheurs : un premier retour…," dans *Le Carreau de la BULAC*, 9 octobre 2015, https://bulac.hypotheses.org/3347 [consulté le 23/08/2023].

ENQUÊTE 2023 SUR LES BESOINS DE FORMATION DES CHERCHEURS EN ÉTUDES ARÉALES

3. MÉTHODOLOGIE

3.1. PUBLIC CIBLÉ

L'enquête s'adressait en priorité aux chercheurs et aux enseignants, rattachés aux disciplines des sciences humaines et sociales, qui travaillent sur des objets et des terrains de recherche extra-occidentaux. Ceux-ci recouvrent l'ensemble des territoires et des aires culturelles situés hors de l'Europe occidentale et de la sphère occidentale anglophone pour lesquels la BULAC met à disposition des collections et des ressources adaptées.

Sans être exclus de l'enquête, en raison de leur implication dans la recherche, les doctorants ne constituaient pas la cible prioritaire.

Portée par un groupement d'intérêt public rassemblant 9 établissements partenaires[4], la BULAC est assidûment fréquentée par les chercheurs en études aréales relevant de ces institutions. Cela étant, ses ressources et ses services sont accessibles à l'ensemble de la communauté des chercheurs aréalistes oeuvrant dans le périmètre de la recherche française. L'enquête ayant pour but de sonder les besoins de formation de cette communauté, l'enquête a donc été ouverte à l'ensemble de ses membres, quel que soit leur rattachement institutionnel.

3.2. CONSTRUCTION DE L'ENQUÊTE ET DIFFUSION

L'enquête a été construite autour des quatre piliers de l'offre actuelle de formation à la BULAC :

- la collecte des matériaux pour la recherche ;
- le traitement et l'organisation de ces matériaux ;
- la valorisation et la publication des résultats de la recherche ;
- le travail dans le respect de l'éthique et de l'intégrité scientifique.

Une rubrique préliminaire permettait de cerner le domaine de recherche et le rattachement institutionnel des chercheurs répondants.

Lancée officiellement à la mi-janvier 2023, l'enquête a été close le 14 avril de la même année. Plusieurs canaux de diffusion ont été mobilisés :

- les services de formation des bibliothèques des établissements membres du GIP BULAC ;
- les listes de diffusion des unités de recherche rattachées à ces établissements du GIP ;
- la lettre d'information du consortium Huma-Num DISTAM, qui réunit les GIS aréaux[5] ;
- la lettre d'information et les autres canaux de communication de la BULAC (site Internet, réseaux sociaux, lettre institutionnelle).

4 Voir note 1.

5 Sur le consortium DISTAM (UAR 2999) et les groupements d'intérêt scientifique qui le portent, voir la page de présentation de cette unité d'appui à la recherche sur le site du CNRS : https://etudes-areales.cnrs.fr/distam/ [consulté le 23/08/2023].

4. PROFIL DES RÉPONDANTS

Conformément à la cible choisie pour l'enquête, la quasi-totalité des répondants exerce une activité de recherche à temps plein ou partagée avec l'activité d'enseignement.

Statut des répondants	Nombre
Enseignants-chercheurs ou chercheur avec mission d'enseignement	46
Chercheurs sans mission d'enseignement	11
ATER	3
Post-doctorants	2
Doctorants	7
PRAG	1
Total	70

Malgré l'ouverture du champ de l'enquête et le relais au-delà du GIP, la majorité des répondants est rattachée à des laboratoires d'institutions membres de ce groupement. Seuls 8 chercheurs relèvent d'établissements hors de ce périmètre.

Périmètre	Nom du laboratoire	chercheurs
Avec au moins un établissement membre du GIP BULAC	PLIDAM - Pluralité des langues et des identités en didactique : acquisition, médiations	6
	CCJ - Chine Corée Japon (UMR 8173)	6
	CREE - Centre de recherche Europes-Eurasie	5
	CASE - Centre d'Asie du Sud-Est (UMR 8170)	5
	IFRAE - Institut français de recherche sur l'Asie de l'Est (UMR 8043)	5
	CRCAO - Centre de recherche sur les civilisations de l'Asie orientale	5
	LLACAN - Langues, langages et cultures d'Afrique	4
	IMAF - Institut des mondes africains	4
	CERMOM - Centre d'études et de recherche Moyen-Orient, Méditerranée	3
	CESSMA - Centre d'études en sciences sociales sur les mondes africains, américains et asiatiques	3
	CEPED - Centre Populations et Développement	3
	EFEO – 2 équipe de recherche	3
	LACITO - Langues et civilisations à tradition orale	2
	LACNAD - Langage, langues et cultures d'Afrique Noire	2
	SEDYL - Structure et dynamique des langues	2
	CRLAO - Centre de recherches linguistiques sur l'Asie orientale	1
	CETOBaC – Centre d'études turques, ottomanes, balkaniques et centrasiatiques	1
	GSRL - Groupe sociétés, religions, laïcités	1

E

	UMR 8215 Trajectoires - Panthéon Sorbonne, CNRS, INRAP	1
UMR hors Île-de-France & UR sans membre du GIP BULAC	UMR 7306 IrAsia – Institut de recherches asiatiques - Aix-Marseille Université, CNRS	2
	UMR 9022 Héritages - Université Cergy-Pontoise, CNRS + Ministère (MCC)	1
	UMR 7220 CERI – Centre de recherches internationales - Sciences Po Paris, CNRS	1
	UMR 7220 ISP – Institut des sciences sociales du politique – Nanterre Universite, CNRS, ENS Paris-Saclay	1
	UM5 5062 IAO – Institut d'Asie orientale – ENS Lyon, CNRS, Lyon 2, Sciences Po Lyon	1
	UR changes – Aix-Marseille Université	1
	UMRU 24140 Dynamiques, interactions, nterculturalité asiatiques (D2IA) – Université Bordeaux Montaigne, La Rochelle Université	1
		70

La répartition des répondants par aire géolinguistique laisse apparaître des disparités qui reflètent un degré de pénétration inégal de l'enquête selon les communautés de chercheurs.

Domaine	Nombre de répondants
Afrique	14
Moyen-Orient, Monde musulman	8
Europe Balkanique, centrale et orientale - Russie	10
Asie orientale	25
Asie du Sud et du Sud-Est	10
Océanie & Extrême-Nord	1 + 1
Aires combinées	1
Total	70

Parmi les répondants, 25 avaient déjà suivi une formation au cours des cinq années précédentes, soit 35% . Ces formations avaient trait aux quatre grands domaines ciblés par l'enquête ; elles étaient principalement dispensées par :

- le laboratoire de rattachement (13 personnes) ;
- une structure documentaire, dont la BULAC (16 personnes) ;
- le CNRS (7 personnes) ;
- les URFIST (Unités régionales de formation à l'information scientifique et technique) et les Maisons des sciences de l'homme (respectivement 2 personnes).

5. LES PRATIQUES DE LA RECHERCHE ET LES DIFFICULTÉS RENCONTRÉES

5.1. LA COLLECTE DES MATÉRIAUX DE RECHERCHE

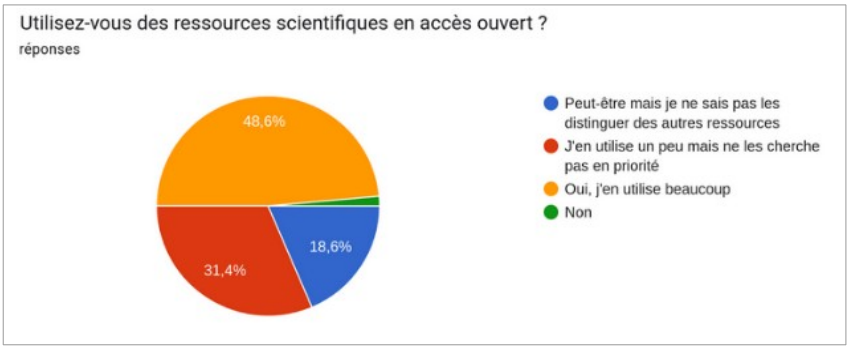
La diversité des matériaux de recherche laisse toutefois ressortir une prépondérance des corpus textuels, qu'ils soient imprimés, manuscrits ou en format numérique. Plus de la moitié

E

des sondés travaillent sur des données issues du Web. Les corpus audiovisuels sont analysés par une part plus réduite de répondants.

	Travaillent assez souvent ou très régulièrement sur	Poucentages (%)
Corpus imprimés	57	81,4
Corpus numériques ou numérisés	63	90
Données du Web	41	58,6
Corpus sonores	15	21,4
Corpus audiovisuels	17	24,3
Données chiffrées	17	24,3

La documentation en accès ouvert occupe une place non négligeable dans le travail de collecte des sources et de constitution des corpus.



Toutefois, on observe une utilisation très inégale des plate-formes de ressources documentaires parmi les plus connues dans le monde de la recherche francophone.

Les plate-forme les plus utilisées sont OpenEdition Journals (75 %), Persée (77 %) et HAL (60 %). La bibliothèque numérique Gallica et le site Internet Archive complètent cet usage. Sont en revanche très peu utilisées les plate-formes Zenodo, OpenAIRE ainsi qu'ISTEX (entre 0 et 1,5 % chacune).

À l'inverse, près de 86 % des répondants se procurent fréquemment des publications sur les réseaux sociaux académiques ou sur des sites pirates tels SciHub.

Les réponses libres révèlent un usage complémentaire et régulier de certaines bases de données muséales françaises et étrangères ainsi que diverses plate-formes des pays d'Asie de l'Est, du monde arabe ou bien d'Europe centrale, balkanique et orientale.

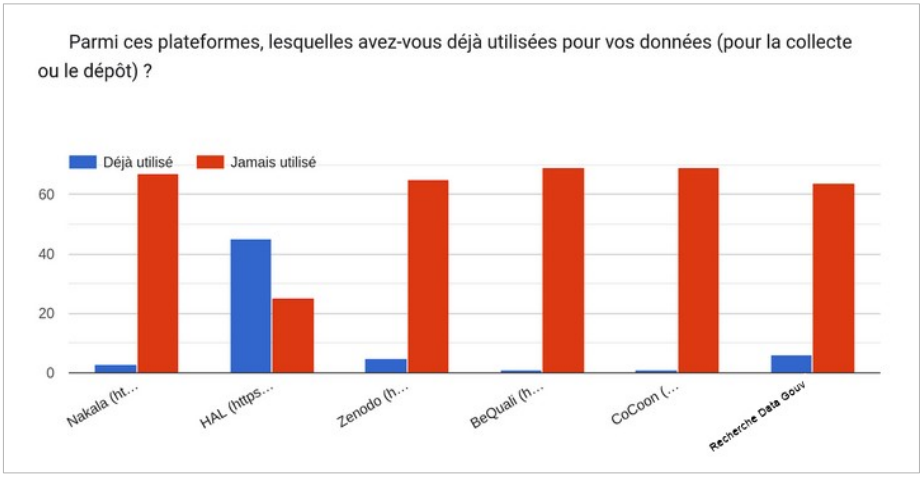
Les répondants ne distinguent toutefois pas toujours les bases offrant un accès entièrement ouvert de celles dont les contenus demeurent en partie payants.

Par ailleurs, les moteurs de recherche moissonnant les plate-formes de ressources en accès ouvert (OpenAIRE, BASE, Jurn ou Paperity) sont assez peu utilisés, soit qu'ils ne sont pas connus, soit qu'ils ne sont pas jugés utiles. Seul le moteur Isidore semble un peu mieux connu (21 %).

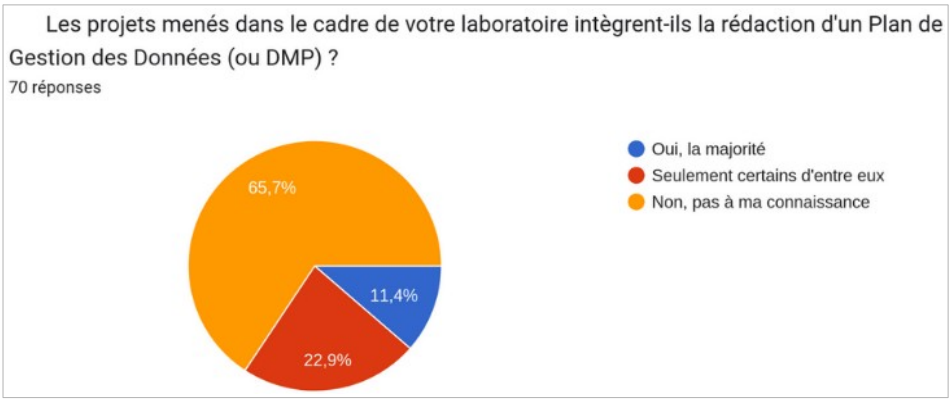
E

5.2. L'ORGANISATION ET LE TRAITEMENT DES MATÉRIAUX DE LA RECHERCHE

Les réponses aux questions laissent transparaître une certaine confusion entre les résultats publiés et les données collectées ou produites. Cette confusion est entretenue par la polyvalence de certaines plate-formes, telles Zenodo ou HAL, sur lesquelles les deux types de matériaux sont présents.



Concernant la gestion globale des données, moins de 12 % disent être confrontés à la nécessité de produire un plan de gestion des données (PGD).



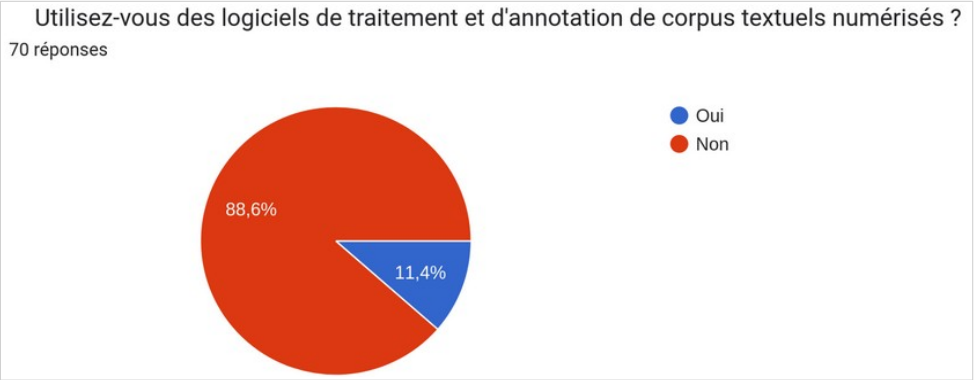
Si près de 66 % indiquent participer à des projets pour lesquels la rédaction d'un PGD ne semble pas exigée, près de 60 % expriment malgré tout le besoin d'être accompagnés dans la gestion globale de leurs données. L'utilisation et le degré de maîtrise des outils de gestion de ces données diffèrent toutefois grandement selon les profils et les types de données traités.

Pour les données bibliographiques, près de 59 % des répondants utilisent un outil de gestion dédié (Zotero pour 50 %). Parmi les motifs de non utilisation évoqués figure l'absence de prise en charge des références en multi-écriture. L'outil Jurism, variante de Zotero conçue pour satisfaire ce besoin, n'étant pas mentionné dans les commentaires libres, son existence semble largement méconnue.

Concernant le traitement des données iconographiques, on note une utilisation très restreinte du logiciel Tropy corrélée, chez 50 % des répondants, au souhait d'être formé à son usage.

E

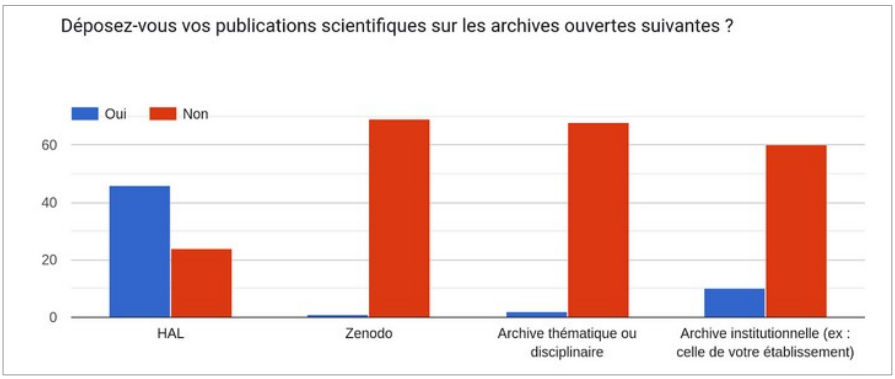
La question de la gestion des données se pose aussi dans le cadre des recherches mobilisant des outils de traitement et d'annotation de corpus numériques. Au vu du profil des sondés, les corpus textuels numérisés sont ici particulièrement concernés. Pour autant, l'enquête fait ressortir un usage encore peu répandu de ces outils de traitement. Seuls 11,4 % disent les utiliser.



Les réponses libres laissent apparaître un usage plus fréquent des outils de traitement et d'analyse de corpus par les chercheurs travaillant sur des corpus audiovisuels et ceux recourant aux méthodes de traitement automatique des langues.

5.3. VALORISATION ET PUBLICATION DES RÉSULTATS DE LA RECHERCHE

L'auto-archivage de publications est pratiqué par environ deux tiers des répondants. Parmi les plate-formes dédiées, HAL est de loin la plus utilisée, suivie des archives ouvertes institutionnelles.



Les formations à l'auto-archivage dans HAL ont été suivies par 64 % des répondants. Toutefois, 27 % expriment encore le souhait d'être formés et accompagnés dans cette tâche.

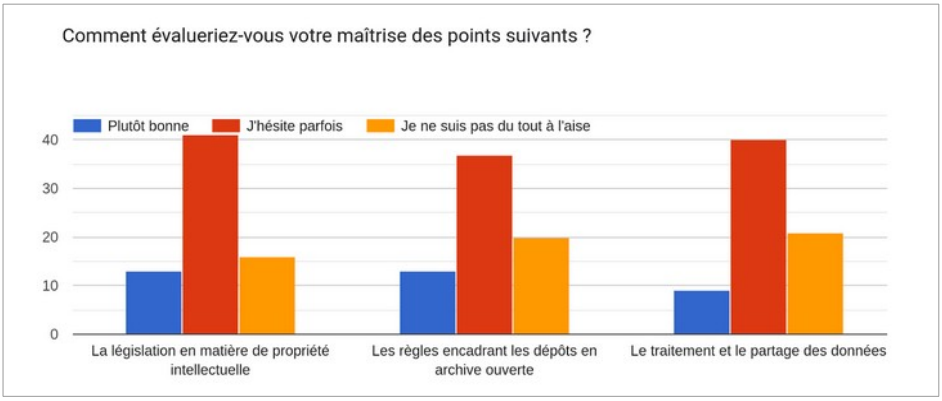
Le souci de visibilité des travaux et des activités sur le Web est bien présent chez les répondants. Si 70 % disent posséder un identifiant sur la plateforme HAL (idHAL) et 57,1 % un identifiant ORCID, 57 % estiment ne pas être très à l'aise dans la gestion de leur identité numérique.

De même, l'activité sur le Web par le biais d'une contribution à un carnet de recherche ou à l'encyclopédie collaborative Wikipédia est assez inégale et demeure, dans l'ensemble, assez limitée. 77 % des répondants disent ne contribuer à aucun blog ni carnet de recherche en ligne. En outre, seuls 11 % contribuent à l'encyclopédie en ligne Wikipédia.

E

5.4. ÉTHIQUE ET INTÉGRITÉ SCIENTIFIQUE

La maîtrise des règles encadrant la propriété intellectuelle, la publication scientifique et le traitement des données est inégale selon les répondants mais révèle dans l'ensemble des hésitations.



Les termes des contrats signés par ceux qui ont déjà publié semblent assez bien compris. Toutefois 27 % mentionnent des difficultés à saisir le contenu de toutes les clauses mentionnées.

6. LES BESOINS DE FORMATION EXPRIMÉS

Qu'elle concerne la collecte ou le traitement des matériaux de recherche, les stratégies de publication ou l'application des règles juridiques, l'activité courante des chercheurs interrogés laisse transparaître des besoins de formation plus ou moins marqués.

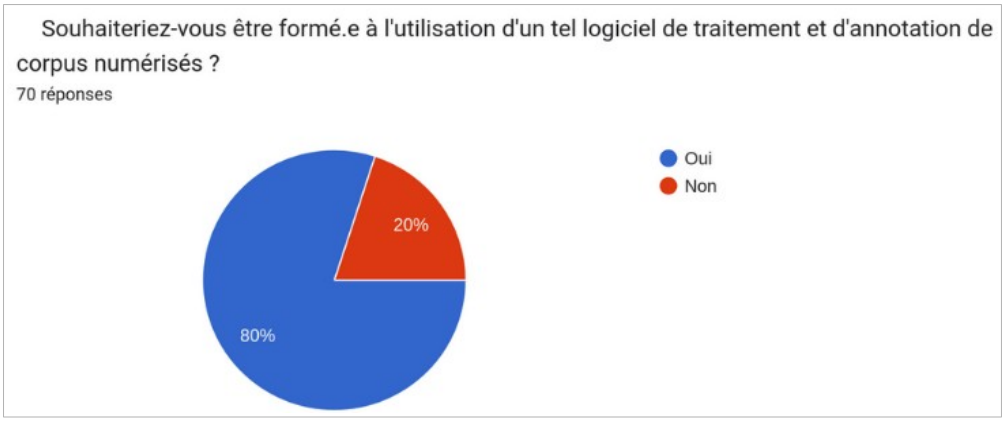
Utilisateurs de vastes corpus imprimés ou numériques, 80 % des chercheurs interrogés souhaitent être formés à une utilisation avancée des bases documentaires et des outils de signalement en ligne. Ce besoin recouvre la connaissance des outils et des plate-formes disponibles à la fois dans et hors du périmètre occidental ainsi que des méthodes d'interrogation dans les langues cibles. L'identification des bases en accès ouvert arrive en tête des besoins exprimés en la matière.

De façon complémentaire, une montée en compétences dans la collecte distante des données sur les pays étudiées est exprimée. Près de 43 % souhaiteraient notamment mieux maîtriser les techniques de collecte et d'archivage des données du web vivant.

Concernant le traitement des données, les besoins portent sur les données visuelles aussi bien que textuelles.

E



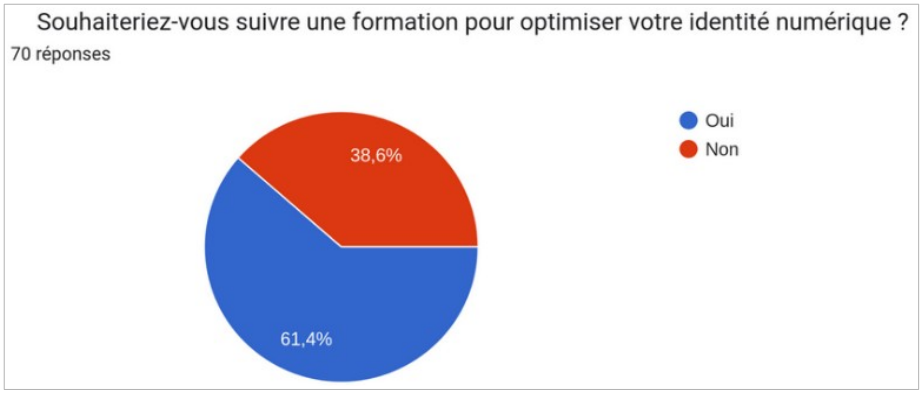


Si la moitié des répondants se dit intéressée par une formation à la prise en main de Tropy pour annoter les données iconographiques, ce sont 80 % qui expriment le souhait d'être formés à l'utilisation d'un logiciel de traitement et d'annotation de corpus textuels numérisés.

En complément, les réponses libres expriment à plusieurs reprises le souhait d'être formé à des outils de visualisation de données sous forme de réseaux ou de graphes ou à des outils de cartographie thématique à partir de données quantitatives.

De façon générale, 59 % indiquent avoir besoin d'un accompagnement dans la gestion globale de leurs données, quelle que soit l'étape de leur traitement.

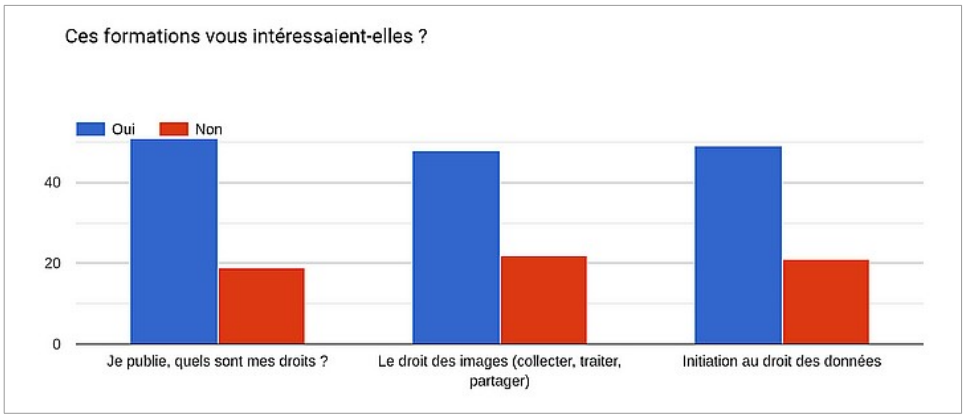
La valorisation des activités de recherche constitue aussi un enjeu important.



Si plus de 60 % des sondés souhaitent être formés à une gestion efficace de leur identité numérique, seul un tiers indique souhaiter apprendre à créer ou à enrichir un carnet de recherche en ligne. Certains (27 %) préfèrent approfondir leur maîtrise de l'auto-archivage, notamment sur HAL.

En matière de partage du savoir, la contribution à Wikipédia par le biais d'ateliers mobilisant les étudiants suscite l'intérêt d'un tiers des répondants.

Enfin, les mises en garde contre les inconduites scientifiques semblent avoir renforcé le souhait de mieux maîtriser le cadre réglementaire et légal des activités de recherche. Les besoins de formation concernent la législation encadrant la collecte et l'usage des différents types de données ainsi que les règles s'appliquant à la publication scientifique et à l'auto-archivage.



De manière générale, 70 % des répondants se disent intéressés par une formation dédiée à l'intégrité scientifique et à ses zones grises.

7. BILAN ET PROSPECTIVE

L'enquête a fait ressortir des besoins ressortissant aux quatre domaines suivants :

- la collecte de sources et de données ciblées sur les aires géolinguistiques étudiées
- l'utilisation d'outils de gestion des données
- la valorisation du profil, de l'activité et des publications scientifiques
- la clarification des cadres réglementaires et juridiques s'appliquant à la recherche dans et hors de l'espace communautaire européen

Bien que souvent rattachés à des UMR, et bénéficiant à ce titre de formations dispensées par le CNRS ou l'IRD, la plupart des répondants manifestent des difficultés ou des hésitations persistantes qui influent sur la conduite de leurs travaux.

En matière de collecte et de traitement des matériaux de recherche, une confusion demeure entre les données et les résultats de la recherche. Elle se double d'un degré inégal de maîtrise des bases documentaires spécialisées et *a fortiori* des outils de traitement et de visualisation des données. De façon générale, un certain nombre de processus et d'outils mobilisés dans le champ des humanités numériques semblent encore peu connus ou sous-utilisés par une part non négligeable des chercheurs en études aréales.

À titre d'exemple, les services déployés par l'infrastructure Huma-Num pour la gestion des données demeurent sous-employés. À l'inverse, l'usage de la plate-forme HAL pour l'auto-archivage des publications progresse, même si des souhaits d'accompagnement persistent en la matière.

Concernant la maîtrise du cadre juridique de la recherche, des difficultés résultent du périmètre d'activité des chercheurs, partagés entre l'espace européen et les terrains extra-communautaires. Or, si la connaissance des volets français et européens de la législation en matière de recherche peut progresser grâce au suivi de formations adaptées, celle des périmètres extra-européens relève souvent de l'accompagnement au cas par cas qu'il est plus difficile de mettre en œuvre.

Les résultats de l'enquête offrent malgré tout matière à un enrichissement des formations actuellement proposées par la BULAC aux chercheurs. Parmi les pistes d'enrichissement à l'étude pour le court et le moyen termes figurent entre autres :

- le recours systématique à des exemples et des cas pratiques empruntés aux études aréales pour l'ensemble des formations ;

- le renforcement des formations aux bases documentaires en langues extra-européennes et aux stratégies d'interrogation dans la langue cible ;

- la conception d'un module de prise en main de l'outil de gestion bibliographique Jurism, variante de Zotero qui gère les données bibliographiques en multi-écritures ;

- l'approfondissement des formations dédiées aux services et aux outils de l'infrastructure Huma-Num ainsi que du logiciel libre Tropy ;

- la conception d'ateliers de cartographie à partir de logiciels libres dédiés ;

- la conception de formations dédiées aux différentes formes de publication scientifique et aux stratégies de publication pour les chercheurs en SHS travaillant sur les terrains extra-occidentaux.

Déployé dans le temps, ce processus s'inscrira dans le cadre général du développement de l'offre documentaire et de la politique d'établissement en faveur de la science ouverte. Il sera alimenté par l'évaluation régulière des formations proposées et l'affinement progressif des réponses aux besoins exprimés.