



HAL
open science

Enquête sur les projets en Humanités numériques

Sylvain Besson

► **To cite this version:**

Sylvain Besson. Enquête sur les projets en Humanités numériques. MSH Lyon - Saint-Étienne; Huma-Num. 2024. hal-04886620

HAL Id: hal-04886620

<https://hal.science/hal-04886620v1>

Submitted on 21 Feb 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Enquête sur les projets en Humanités numériques

Sylvain Besson

Ingénieur d'étude en Humanités numériques, chargé
d'enquête, MSH Lyon Saint-Étienne



Novembre 2024

Table des matières

Introduction	3
Contexte général	3
Objectifs de l'enquête.....	6
Cible de l'enquête.....	7
Définition des Humanités numériques	8
Méthodologie de l'enquête	9
Limites de l'enquête	12
Résultats de l'enquête	14
Profils des répondants au questionnaire	14
Méconnaissance des services Huma-Num et des pôles des compétences de la MSH Lyon Saint-Étienne.....	17
Huma-Num : des services centralisés, mais jugés indispensables.....	20
La MSH Lyon Saint-Étienne comme relai des méthodes et des outils en Humanités numériques	23
Paradoxe entre la volonté d'ouverture des données et la réalité des projets.....	25
Des projets hétéroclites dans le périmètre de la MSH Lyon Saint-Étienne.....	27
Collaboration entre les laboratoires associés	28
Pistes d'amélioration proposées par la MSH LSE	29
Territorialisation des services Huma-num : des services de proximité.....	29
Des accompagnements intensifiés : les doctorant-es au cœur de la démarche	31
Renforcer l'interconnaissance entre les laboratoires.....	32
Accompagnement sur l'ouverture des données.....	33
Conclusion : une enquête montrant les besoins d'accompagnement.....	35
Annexes.....	37
Services de l'IR* Huma-Num	37
Services tiers hébergés.....	38
Outils.....	38
Financements	39

Maisons des Sciences de l'Homme.....	39
Laboratoires.....	40
Calendrier	44
Graphiques	45
Graphe de réseaux.....	58
Typologie.....	59
Questionnaire.....	60

Introduction

Contexte général

L'enquête s'est déroulée sur une durée de 12 mois et a été menée par un ingénieur d'étude en Humanités numériques et chargé d'enquête, Sylvain Besson. Elle a été pilotée par la **MSH Lyon Saint-Étienne** et financée par l'**IR* Huma-Num**. Elle a également fait appel à l'**axe Sociétés et Humanités numériques** de la MSH LSE, et sollicité plusieurs de ses agents : Orline Poulat, Sonia Guérin-Hamdi, et Alexandra Dugué.

Qu'est-ce que la **MSH Lyon Saint-Étienne** ?



La **Maison des Sciences de l'Homme** (MSH) Lyon Saint-Étienne (LSE) fait partie des 21 Maisons du Réseau national des MSH (Infrastructure Nationale de Recherche). C'est une **Unité d'Appui et de Recherche** (UAR), constituée de 20 agents et à laquelle sont rattachées les 53 unités SHS du site (3800 pers.). Ses tutelles principales sont les universités Lyon 2, Lyon 3, l'université de Saint-Étienne et le CNRS, et ses tutelles secondaires sont Lyon 1, l'ENS et l'IEP. Elle compte également trois établissements partenaires : ENSSIB, ENTPE et INSA.

Dispositif transversal et mutualisé sur le site, sa mission est triple, dans une logique générale d'**interdisciplinarité**, de **science ouverte** et d'**internationalisation des SHS**. Elle participe tout d'abord à la structuration

de la recherche, en SHS et à ses interfaces, par le biais de ses **quatre axes scientifiques**, de ses **Appels à Projets** et de son **hôtel à projets**. Elle a ensuite une fonction d'accompagnement de la recherche à travers ses **huit pôles d'expertise**, ses services et plateformes ainsi que ses liens avec les IR* Progedo et Huma-Num. Enfin, elle propose des opérations de formation et de médiation scientifique.

Cette enquête fait suite à plusieurs analyses menées entre 2019 et 2023. Une enquête préliminaire a tout d'abord été entreprise par l'IR* Huma-Num *via* trois MSH : la MSH LSE, la MMSH¹ et la MSH Ange-Guépin². Elle a été effectuée dans un contexte de volonté de fusion entre les services **ShareDocs** (pour le stockage chaud/tiède) et **Huma-Num Box** (stockage froid)³. Sur le site Lyon Saint-Étienne, elle a ensuite impulsé le déploiement de la préenquête sur l'identification des besoins des chercheurs et des laboratoires, l'amélioration des services et de l'accompagnement. Cette préenquête a été envoyée aux 52 laboratoires alors associés à la MSH LSE (elle en compte actuellement 53) mais n'a pas donné de résultats concluants, le taux de réponse ayant été faible (7 laboratoires sur 52).

Concomitamment à ce travail de préenquête, un projet de cartographie des projets en Humanités numériques (HN) avait été amorcé à l'initiative de l'axe Sociétés et Humanités numériques de la MSH LSE. Une stagiaire, Ambrine Soltani, avait débuté le travail cartographique. Ce travail intégrait également

¹ Voir les **Maisons des Sciences de l'Homme** à la page 28.

² *Idem*.

³ Voir les **services de l'IR* Huma-Num** présentés à la page 26.

des associations *via* la boutique des sciences et était circonscrit au site lyonnais. Ce projet a abouti à la création d'un gabarit de cartographie avec une identification des informations à recueillir pour chaque projet. Ce travail de réflexion a également permis de mieux définir la notion d'Humanités numériques, ce que nous avons pu affiner lors de la présente enquête.

Qu'est-ce que l'IR* Huma-Num ?

La **principale mission de l'IR*** est de construire, avec les communautés et à partir d'un **pilotage scientifique**, une infrastructure numérique de niveau international pour les SHS.



L'IR* structure, par l'intermédiaire de **consortiums** regroupant des acteurs des communautés scientifiques et d'un réseau de référent-es dans les MSH, l'accompagnement des communautés scientifiques SHS en matière d'infrastructure numérique pour les données de la recherche.

Elle met en œuvre une infrastructure numérique au moyen de différents services permettant aux communautés SHS de développer, de réaliser et de

préservé sur le long terme les programmes de recherche – leurs données et outils – dans un contexte de science ouverte et de partage des données.

L'ensemble de l'infrastructure s'inscrit dans le cadre des principes dits **FAIR** (**F**acile à trouver, **A**ccessible, **I**nteropérable, **R**éutilisable) qui favorisent, outre l'ouverture des données, leur mise à disposition avec un triple objectif de qualité des données et des métadonnées, d'inscription dans un cycle de vie maîtrisé par les scientifiques et enfin de pérennité des données sur le long terme (accès, intégrité, contextualisation de la production des données).

Huma-Num IR* est une infrastructure de recherche « étoile », inscrite sur la feuille de route des infras du ministère de l'Enseignement supérieur et de la Recherche, mise en œuvre par le CNRS avec le Campus Condorcet et Aix-Marseille Université.

Objectifs de l'enquête

À la suite de ces deux études, il est apparu nécessaire de réaliser une enquête plus approfondie sur les projets en Humanités Numériques (HN) sur le site stéphano-lyonnais. Cette enquête a comme objectifs de recenser les usages, les besoins, les questionnements et les freins concernant les projets intégrant une dimension HN à l'échelle de la MSH LSE. Au-delà du recensement, l'intention était double :

- Clarifier et optimiser l'articulation entre la MSH LSE, Huma-Num et la communauté scientifique concernant le recours aux services proposés par l'IR*, par la prise de contact, les services dédiés et proposés, etc. ;

- Renforcer l'accompagnement de proximité sur le territoire.

Cible de l'enquête

Cette enquête s'intéresse, plus particulièrement, aux **projets ayant bénéficié d'un financement**. Les financeurs sont variés : financements européens tels que les ERC, nationaux tels que les ANR, ou locaux tels les appels à projet (APP) de la MSH LSE. Nous avons fait ce choix afin de circonscrire le nombre de projets et pour nous focaliser sur des projets ayant une date de début et une date limite déterminée (même si certains peuvent faire l'objet de plusieurs financements).

Nous nous focalisons sur les projets **en cours ou qui ont été clôturés récemment**. Cela nous permet d'avoir une photographie suffisamment récente des pratiques sur le site au moment de l'enquête. Nous avons également fait le choix de nous limiter au périmètre de la MSH LSE.

Quant à la cible de cette enquête, c'est-à-dire la population étudiée, elle vise d'une part les **personnels des laboratoires associés à la MSH LSE**, qu'ils soient chercheur·euse·s, enseignant·es-chercheur·es, post-doctorant·es ou encore ingénieur·es. D'autre part elle vise les **personnels avec ou sans compétences informatiques**. Et enfin, elle tient compte de la manière dont la cible **se positionne vis-à-vis des Humanités numériques, des services de l'IR* Huma-Num et du pôle Humanités numériques de la MSH LSE**. Nous pouvons synthétiser ces objectifs par le diagramme de Venn suivant :

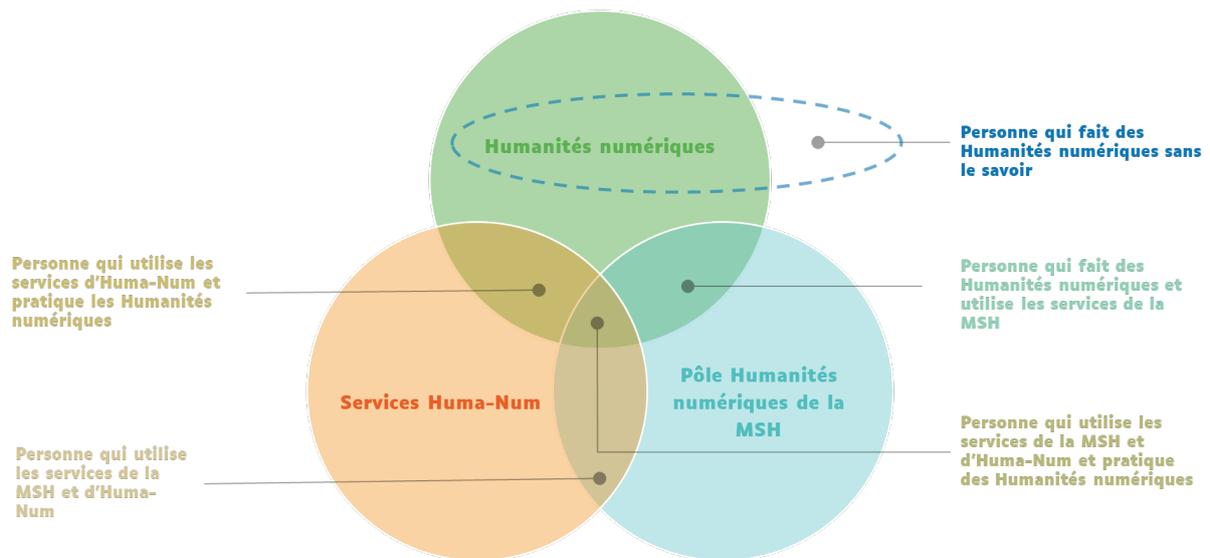


Diagramme de Venn sur les personnes ciblées

Figure 1 Diagramme de Venn sur la population étudiée

Nous pouvons alors constater que la population globale concernée et les usages sont assez divers. Il est donc essentiel de chercher à conserver cette cible protéiforme tout au long de notre enquête.

Définition des Humanités numériques

Il semble important de définir ce que nous entendons par Humanités numériques. Sa définition dans la littérature a pu fluctuer voire se métamorphoser. Bien qu'on puisse rapprocher le terme Humanités numériques de celui, anglo-saxon, de *Digital Humanities*, les deux termes ne se valent pas. Le premier a pour sens premier l'application de méthodes numériques aux disciplines des Humanités (Lettres et Littérature). Le second désigne une acception beaucoup plus large des différentes méthodes computationnelles appliquées aux SHS⁴. C'est la raison pour laquelle, à l'origine, l'édition numérique était la principale manière de faire des Humanités numériques. Aujourd'hui, il semblerait que la vision française et la

⁴ Michael Piotrowski et Aris Xanthos, « Décomposer les humanités numériques », Humanités numériques [En ligne], 1 | 2020, mis en ligne le 01 janvier 2020 ; DOI : <https://doi.org/10.4000/revuehn.381>.

vision anglophone se soient beaucoup rapprochées. Nous parlons ainsi d'Humanités numériques dès lors qu'une discipline des SHS pratique des méthodes computationnelles. Cela amène souvent à penser qu'il n'y a désormais plus aucune barrière entre les disciplines SHS d'un côté et les Humanités numériques de l'autre.

Pour cette raison, nous avons dû dès le début du projet nous approprier cette notion pour mieux échanger avec les personnes sollicitées. Nous avons également pu construire cette définition à travers différents entretiens préliminaires avec différents personnels des unités associées. Selon notre définition, les projets en humanités numériques sont ceux qui ont une **interaction forte entre les Sciences Informatiques et au moins une discipline des Sciences Humaines et Sociales**. Cette interaction peut se faire sur l'ensemble ou une partie d'un projet. Cela se traduit par des **outils et méthodologies computationnels et numériques** qui peuvent être développés ou qui existent déjà, et qui vont permettre d'appuyer les SHS dans leurs projets de recherche. Cela peut être à l'inverse des outils qui n'existent pas, mais dont un **besoin est ressenti dans un champ des Sciences Humaines et Sociales**.

Il est possible que cela fasse intervenir une **collaboration entre un ou des laboratoires SHS et un ou plusieurs laboratoires d'informatique** avec *a minima* un laboratoire associé à la MSH LSE, mais ce n'est pas une condition obligatoire.

Méthodologie de l'enquête

Cette enquête a eu pour objectif de mettre en évidence les projets en Humanités numériques par le recensement des usages, des besoins, des questionnements et des freins. Elle a été pensée en **trois temps principaux**.

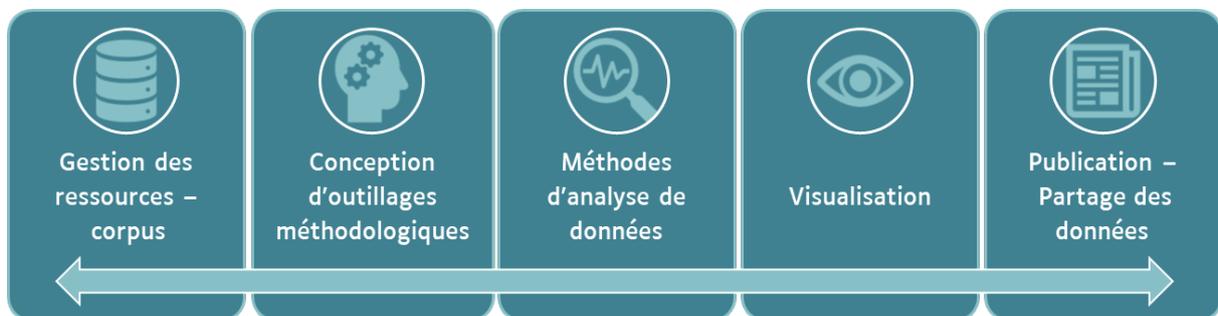


Figure 2 Calendrier de l'enquête

Préalablement à ces trois temps, une **phase préparatoire** a été nécessaire afin de préciser plus clairement les objectifs de l'enquête et ses attendus. Cette phase préparatoire a impliqué la réalisation d'une **série d'entretiens exploratoires** avec des directeur-trices de laboratoires, des chercheur-euses et des ingénieur-es du site. Ces derniers ont ainsi déjà pu faire remonter une série de besoins et de difficultés qu'ils pouvaient rencontrer dans leur projet et indiquer la façon dont ils percevaient les Humanités numériques. C'est également à partir de ces entretiens que nous avons pu définir la méthodologie du projet de cette enquête qui s'est déroulée en trois étapes.

La première consiste à créer une **base de données** des projets en Humanités numériques correspondants aux critères que nous avons énoncés ci-dessus. Elle intègre différentes informations sur les projets telles que la description du projet, le type de financement et les laboratoires et institutions associés. Cette base de données a pu être enrichie de diverses manières : par la conduite d'entretiens avec les directeur-trices de laboratoires et des chercheur-euses et ingénieur-es, par la consultation de pages des laboratoires dédiées aux projets, par l'interrogation des bases de données de l'ANR et de l'ERC et par la discussion avec d'autres membres de la MSH LSE sur les projets dans lesquels ils sont impliqués. L'enrichissement de la base de données, créée dès décembre, s'est poursuivi jusqu'en septembre 2024 et a vocation à se développer continuellement. Ainsi, des projets ont pu être ajoutés jusqu'au terme de la période d'analyse. Le nombre actuel de projets recensés est de 61.

La seconde étape a été de réaliser des **groupes de discussion (ou focus groups)** pour échanger et débattre de différentes thématiques : les outils utilisés, leur connaissance de l'IR* Huma-Num et le recours à ses services, ainsi que la connaissance des pôles de compétence de la MSH LSE et la sollicitation ou non d'accompagnement de la recherche. Nous avons également pu présenter aux participant-es une typologie des projets, que nous avons élaborée à la suite des entretiens préliminaires, afin d'éprouver auprès d'elles-eux cette dernière. Cette typologie permet de mieux cibler quel(s) aspect(s) des Humanités numériques est (sont) mis en œuvre au sein de chaque projet. Nous avons ainsi dressé cinq grandes catégories :



Chacune ayant elle-même des sous-catégories⁵. Il est important de souligner que chaque projet peut être présent dans plusieurs catégories.

Nous avons réalisé deux *focus groups*, l'un composé de trois personnes et le second de deux personnes.

Enfin, la dernière étape avait pour but d'élargir l'enquête et le panel de participant-es par l'envoi d'un **questionnaire anonyme⁶ à l'ensemble de la communauté associée à la MSH LSE**. Bien que les thématiques abordées soient similaires à celles de la deuxième étape, le questionnaire ne concerne plus seulement les projets financés actuels, mais toutes les personnes impliquées dans des projets en Humanités numériques. Cet élargissement a permis d'impliquer davantage de chercheur-euses, doctorant-es, ITA que lors

⁵ Voir la **typologie** complète en annexe à la page 45.

⁶ Voir le **questionnaire** complet en annexe à la page 47.

des *focus groups* et d'avoir une approche plus quantitative. De plus, cela nous a permis de les comparer aux résultats obtenus auparavant. Certains éléments ont pu confirmer des hypothèses formulées lors de la phase qualitative. Le questionnaire a été lancé le 12 juin pour une durée d'un mois. 141 personnes y ont répondu de façon complète.

À l'issue de ces trois étapes, une **phase de comparaison et d'analyse** des multiples données acquises *via* les entretiens, les *focus groups* et le questionnaire a pu être effectuée. Cette phase s'est déroulée entre juillet et septembre et a abouti à la rédaction du présent **rapport**, ainsi qu'à une **synthèse** permettant une diffusion plus large auprès de la communauté sollicitée. Enfin, un **site internet**[®], issu de la base de données Heurist et déjà évoqué précédemment, a été créé pour permettre d'avoir accès à des ressources et à des informations sur les différents projets que nous avons recensés⁷.

Limites de l'enquête

Comme précédemment évoqué, la population étudiée est très protéiforme. Elle est répartie entre les 7 établissements tutelles de la MSH LSE : CNRS, Lyon 2, UJM, Lyon 3, Lyon 1, l'ENS de Lyon et Sciences Po Lyon. Cela recouvre potentiellement 53 laboratoires associés à la MSH LSE. La présentation de l'enquête dans certains événements comme la semaine Data-SHS, les conseils des directeur-ices d'unité ou des assemblées générales de laboratoires associés à la MSH LES, ainsi que des sessions de réseaux professionnels comme celui de Maté-SHS, a grandement facilité la prise de contact avec certaines unités.

Par ailleurs, la notion d'Humanités numériques, par sa nature même et son évolution sémantique, peut laisser une liberté interprétative. Cette dimension est particulièrement ressortie dans les entretiens préliminaires. Pour cette

⁷ Voir **calendrier** complet en annexe à la page 31.

raison, on peut faire l'hypothèse que certaines personnes n'ont pas répondu à notre enquête, car elles se sentaient étrangères à la notion d'Humanités numériques et/ou à la définition que nous en avons donnée. Il a été également difficile d'évaluer *a priori* le nombre de personnes qui correspondent à notre population cible, au vu de son hétérogénéité.

Une autre difficulté que nous avons pu rencontrer est celle de la constitution des *focus groups*. Afin de pouvoir échanger plus facilement, nous avons fait le choix de les organiser uniquement en présentiel. L'indisponibilité des personnes sollicitées et la coordination de leurs agendas contraints ont rendu l'organisation des *focus groups* difficiles. Pour cette raison, nous avons fait le choix de réduire le nombre de *focus groups* initialement prévu, de 3 à 2, avec un nombre moins important de participant-es par *focus group*. C'est l'une des raisons pour lesquelles il nous a semblé nécessaire d'adopter une approche quantitative à la suite de cette phase. Néanmoins, le profil des participant-es étant varié (chercheur-euse, enseignant-e-chercheur-euse, ingénieur-e) et leurs disciplines diverses (histoire, littérature, linguistique, informatique), les discussions furent denses, contradictoires et représentatives d'un panel éclectique. L'analyse qui en découle reste donc, selon nous, pertinente et éclairante.

Ce présent rapport est composé de deux parties. La première a pour but de présenter les différents **résultats** et analyses que nous avons pu mettre en évidence durant les trois étapes du projet. La seconde partie, en s'appuyant sur les analyses de la première partie, permet quant à elle d'apporter des **propositions de perspectives** possibles, et des actions à mettre en place, en répondant aux objectifs de l'enquête.

Résultats de l'enquête

Profils des répondants au questionnaire

Le questionnaire a recueilli 141 réponses complètes⁸. Tout d'abord, nous avons posé une question relative aux statuts des personnes interrogées⁹. Ce sont majoritairement des enseignant-es chercheur-euses qui ont répondu, car ils+ représentent 47 % de l'ensemble des répondants. Ce sont ensuite les ingénieur-es et les doctorant-es qui sont les plus représenté-es avec respectivement 19 % et 17 %.

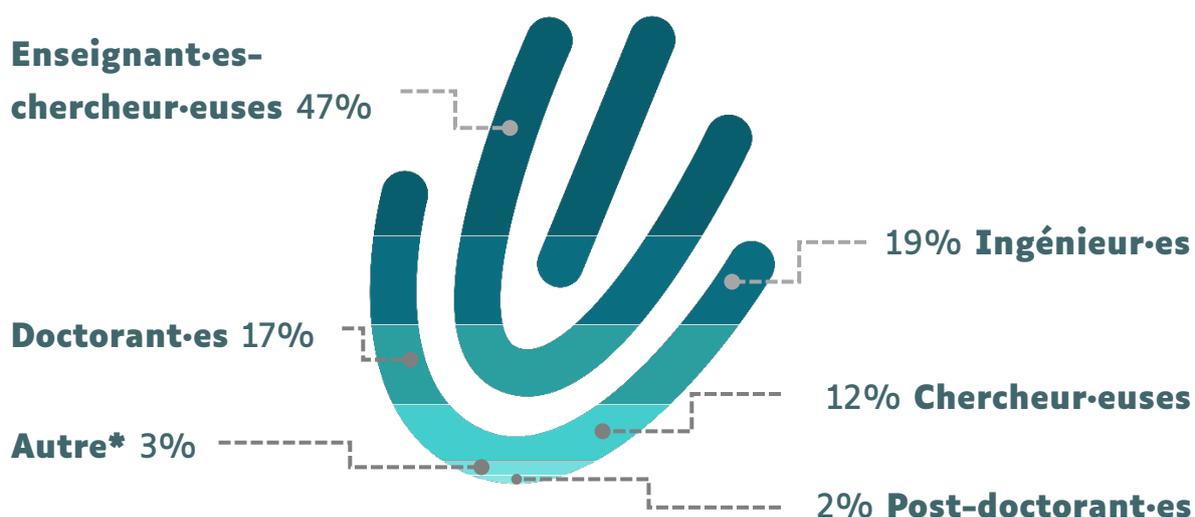


Figure 3 Répartition des répondants en fonction de leur statut exprimé en pourcentage

Comme nous n'avons pas connaissance de la répartition des personnels au sein des unités par statut, la part des enseignant-es-chercheur-euses dans la population étudiée peut ainsi sembler surreprésentée. Nous avons donc

⁸ Nous avons mis en avant les graphiques par effectif. L'ensemble des graphiques sont disponibles en annexe, à la fois exprimés en effectif et en pourcentage. De plus, des extraits de questions posées dans le questionnaire sont en **questionnaire**. Nous avons aussi fait le choix de mettre davantage en avant des graphiques en barres, plus juste statistiquement, au détriment d'autres représentations davantage esthétiques, mais plus compliquées à interpréter.

⁹ Le **graphique** montrant le statut des personnes ayant répondu à l'enquête exprimés en effectif complet est disponible en annexe à la page 32. * La modalité « Autre » correspond à des émérites et à un « praticien hospitalier ».

confronté cette proportion à celle de quatre laboratoires de taille significative (DDL, CMW, LARHRA, GATE¹⁰). Ainsi les chercheur-euses et enseignant-es-chercheur-euses représentent 26 % dans ces laboratoires contre 6 % pour les ITRF. À l'inverse, les doctorant-es représentent également 26 % des effectifs dans ces laboratoires. Il semble donc que la catégorie des doctorant-es soit sous-représentée dans notre enquête. Dans ces quatre unités, les chercheur-euses associée-s représentent 23 %, cependant nous ne les avons pas distingué-es spécifiquement dans les statuts proposés dans notre questionnaire. Le reste des effectifs dans ces laboratoires concernent les émérites (3 %) et des CDD/Post-doc (2 %).

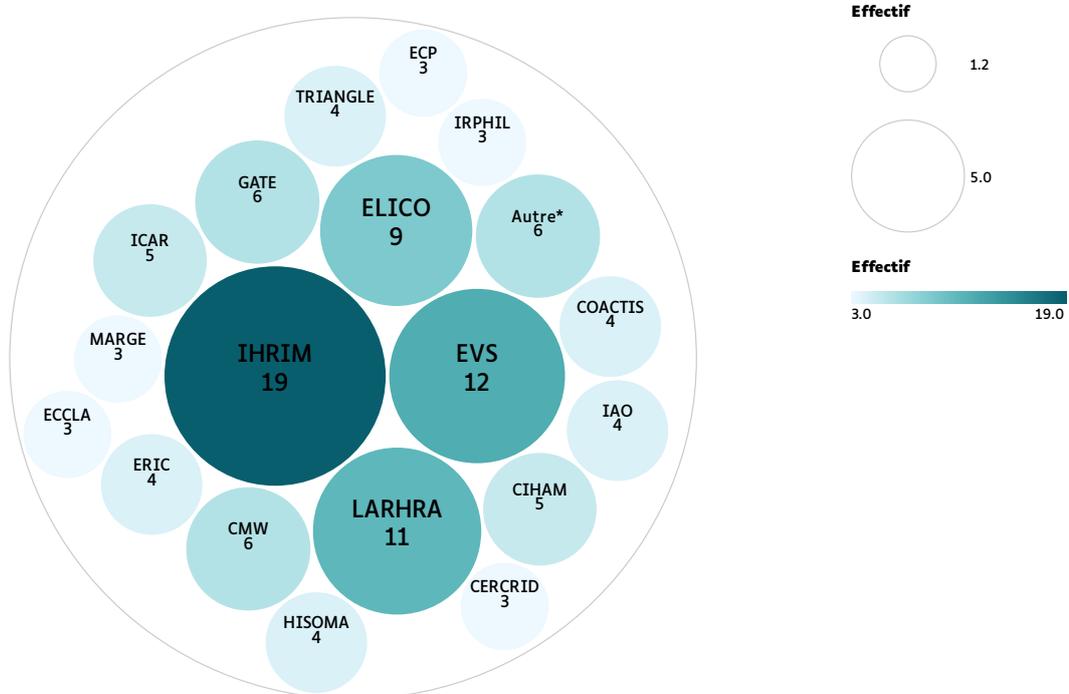


Figure 4 Répartition par effectif des répondants au questionnaire selon leur laboratoire de rattachement

Comme nous pouvons le voir, trois laboratoires sont les principaux répondants : IHRIM (14 %), EVS (8,8 %) et LARHRA (8 %). Il est à noter que nous avons fait le choix de ne pas afficher sur le graphique les effectifs

¹⁰ Voir les intitulés complets des laboratoires en annexe à la page 28.

inférieurs à trois¹¹. L'IHRIM étant un laboratoire à coloration littéraire et très porté sur les Humanités numériques¹², nous nous attendions à ce qu'il soit très représenté. De même pour le LARHRA qui comprend également un axe de recherche en Histoire numérique. Nous pouvons souligner que sur les 53 laboratoires associés à la MSH LSE, la majorité d'entre eux (34 laboratoires) ont répondu à l'enquête¹³. 65 % des unités sont donc représentées dans les réponses obtenues.

Le même constat est fait pour les disciplines les plus représentées : Histoire (13 %) et Littérature (12 %) ¹⁴.

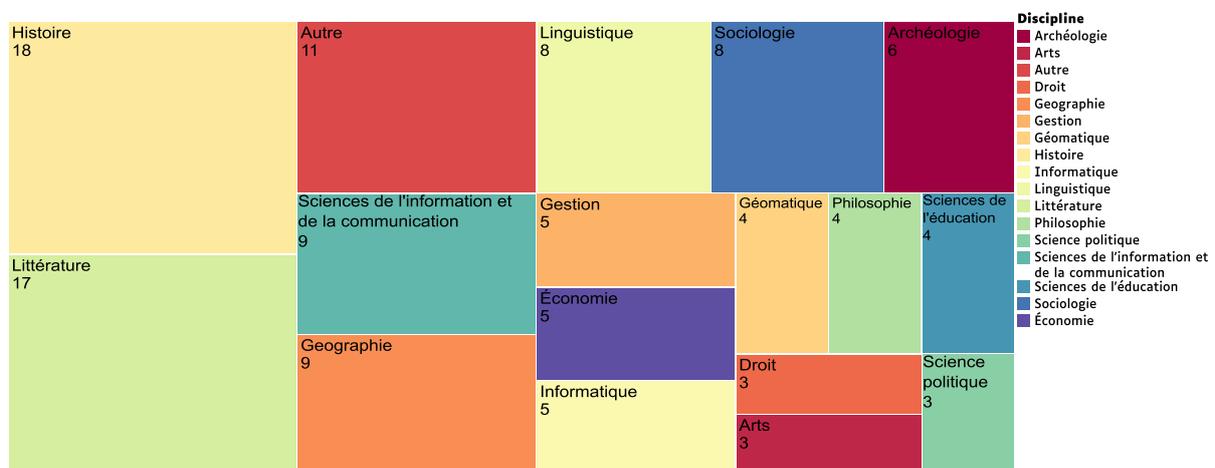


Figure 5 Répartition par effectif des répondants au questionnaire selon leur discipline principale

Nous pouvons également observer le profil des répondants sous l'angle des types de données exploitées. Nous pouvons constater que ce sont davantage

¹¹ Le **graphique** est dans son entièreté disponible en annexe à la page 33.

¹² Il existe en sein de l'ENS un **atelier des Humanités numériques (AHN)** porté en partie par des personnels de l'IHRIM.

¹³ La modalité « autre » correspond essentiellement à des personnes de la MSH LSE et à une personne ayant indiqué n'être plus en stage.

¹⁴ Le **graphique** indiquant la discipline principale étudiée par les répondants exprimés en effectif et en pourcentage complet est en annexe à la page 34. La modalité « autre » correspond surtout à des personnes ayant indiqué la modalité « Humanités numériques » mais aussi des personnes ayant indiqué les modalités « Philologie », « Informatique linguistique » et « Sportive ».

les sources textuelles/porteuses d'inscription (36 %) qui sont manipulées par les répondants, puis des images (23 %) et enfin des données quantitatives (20 %) ¹⁵. Néanmoins, le son (12 %) et la vidéo (11 %) sont aussi des données utilisées.

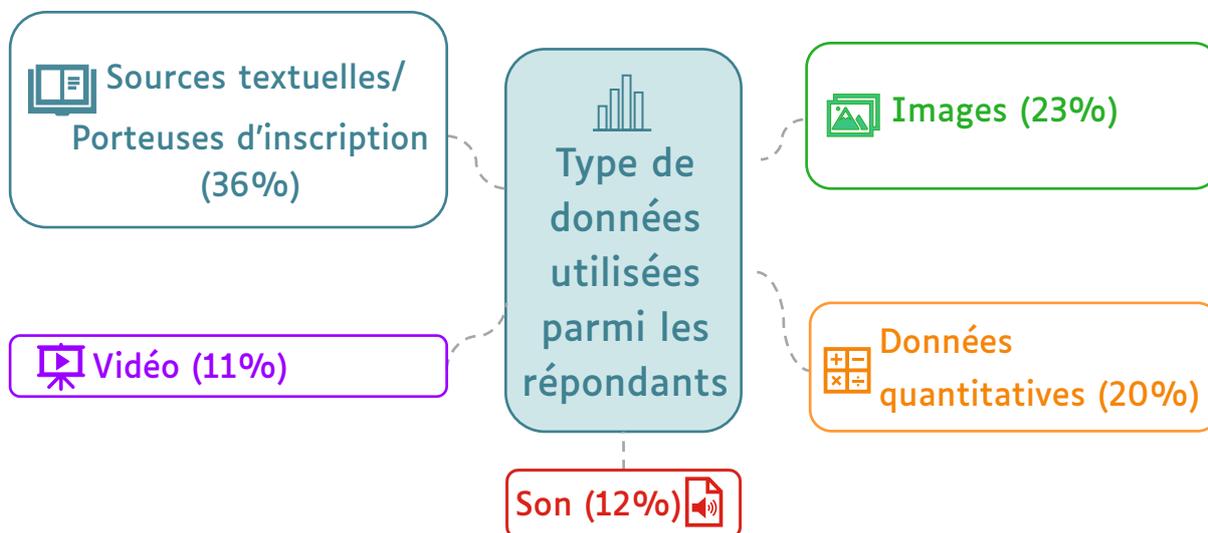


Figure 6 Profil des répondants selon le type de données

Enfin, nous pouvons également analyser le profil des répondants suivant leur site géographique de rattachement entre Lyon et Saint-Étienne ¹⁶. Ils sont 81 % à être situés sur Lyon contre 17 % sur Saint-Étienne, le restant ayant répondu « autre ». Cette proportion étant cohérente avec la répartition de la communauté SHS associée à la MSH LSE.

Méconnaissance des services Huma-Num et des pôles des compétences de la MSH Lyon Saint-Étienne

Dès les entretiens préliminaires, la **méconnaissance de l'IR* Huma-Num** et de ses services nous a été rapportée. Ce constat a été confirmé lors des *focus groups* et encore davantage par le questionnaire. La présentation de cette enquête, lors d'instances variées telles des AG d'unités, puis son déploiement ont d'ailleurs permis d'effectuer un travail d'acculturation auprès de la communauté associée à la MSH LSE. Une confusion entre les Humanités

¹⁵ Le **graphique** complet est en annexe à la page 35.

¹⁶ Le **graphique** est en annexe à la page 32.

numériques et l'IR* Huma-Num est également remontée dans les entretiens préliminaires. Cette confusion peut s'expliquer par la proximité lexicale et graphique entre les deux termes. Cela a rendu certains échanges autour des enjeux de l'enquête plus compliqués. Une confusion entre les **périmètres de l'IR* et de la MSH LSE** a également émergé. En effet, lors des entretiens, certains membres de la communauté ont parfois exprimé, des besoins de services pas encore proposés par l'IR*. Cette enquête étant portée par l'IR* et la MSH LSE, le chargé d'enquête a souvent été perçu comme étant un agent territorialisé Huma-Num et cela a été perçu par la communauté comme une opportunité de faire des remontées sur leurs besoins.

Par ailleurs, nous avons pu constater que la connaissance de l'IR* est très variable. Si nous regardons les résultats du questionnaire à la question « Connaissez-vous l'infrastructure de recherche "étoile" Huma-Num ? », nous avons 37 % des répondants qui indiquent « connaître et utiliser », 13 % « connaître mais ne pas utiliser » et 50 % « ne pas connaître ». Force est de constater que la **moitié des répondants indiquent ne pas connaître l'IR***. Cependant, ne pas connaître l'IR* ne signifie pas ne pas connaître les services qui lui sont associés. En effet, les services Nakala ou Isidore peuvent être connus, sans que l'infrastructure soit quant à elle connue. Un **travail d'acculturation** de cette dernière semble en conséquence nécessaire.

D'autre part, si on analyse plus précisément les résultats à cette question, en s'intéressant aux statuts des répondants, on remarque que les disparités ne sont pas les mêmes¹⁷. En effet, 80 % sont des ingénieur-es qui indiquent connaître et utiliser les services d'Huma-Num mais seulement 14 % sont des doctorant-es. La prépondérance des ingénieur-es semble cohérente puisque ce sont les principaux demandeurs auprès de l'IR*. Concernant les doctorant-es, cela met en exergue leur méconnaissance de l'écosystème de la recherche dans son intégralité. Cela a d'ailleurs été évoqué par l'une des

¹⁷ Voir les **graphique** en annexe à la page 37.

participantes des *focus groups* : elle n'avait elle-même pas eu connaissance des services durant sa thèse et n'en a eu vent que par la suite, par le « bouche-à-oreille ».

Il est à souligner qu'un écart significatif existe entre les enseignant-es-chercheur-euses (26 %) et les chercheur-euses (64 %). Cet écart peut sans doute s'expliquer par la différence de temps à consacrer à la recherche. D'ailleurs la proportion des enseignant-es-chercheur-euses « connaissant, mais n'utilisant pas les services » est la plus importante (18 %). Il semblerait pertinent d'affiner l'analyse de ces résultats pour en connaître les raisons : un manque de temps ? des outils qui ne correspondent pas à leur besoin ? des modalités demandant trop de contraintes ? ...

Il est également intéressant de constater que les réponses à la question relative aux principales fonctions de l'IR* sont d'une grande diversité :

«Huma-Num est une "infrastructure" qui offre des services gratuits aux projets de recherche SHS de deux ordres : mettre en place des services et outils mutualisés cloud (hébergement web, [...]) et opérer le soutien à des réseaux thématiques (consortiums). [...];

Proposer des outils numériques divers à disposition de la recherche SHS ;

Archiver – diffuser ;

Intelligence artificielle de retranscription, océrisation, etc. ;

J'utilise Humanum uniquement pour Whisper, je ne connais pas les autres fonctionnalités ;

Proposer une infrastructure pour la gestion des données de la recherche aux différentes étapes du cycle de vie des données»

Cet échantillon de réponses montre la **méconnaissance des principales fonctions** de l'IR*, telle que dans la réponse « Intelligence artificielle de retranscription, océrisation, etc. ». Mais par ailleurs, cela fait ressortir la **diversité des services proposés** par l'IR* et force est de constater que certains identifient parfaitement les principales fonctions de l'IR*, comme dans la réponse « Proposer une infrastructure pour la gestion des données de la recherche aux différentes étapes du cycle de vie des données ». On peut

donc légitimement émettre l'hypothèse que cette méconnaissance des principales fonctions peut avoir pour conséquence une utilisation partielle des services.

Huma-Num : des services centralisés, mais jugés indispensables

La place de l'IR* dans l'ESR est attestée et jugée primordiale. Néanmoins, elle évolue dans un écosystème complexe pour les acteurs de la recherche. À plusieurs reprises, une **confusion entre l'IR* Huma-Num et l'IR* Progedo**¹⁸ a été évoquée, de même qu'entre les IR* et la MSH LSE. Lors d'un des groupes de discussion, une personne a demandé :

« Est-ce qu'il est possible d'avoir une explication pour les gens comme moi d'exactly comprendre l'exact différence entre Huma-Num et Progedo sur les données ? »

Les différents échanges ont fait resurgir le besoin, pour la communauté de la recherche, **de clarification entre les différents consortiums**. Cela semble essentiel même pour des personnes directement impliquées dans ces derniers.

Sur la base des entretiens préliminaires et des *focus groups*, des difficultés dans la **prise de contact** avec l'IR* et dans l'**accès aux services** sont apparues. Certain-es participant-es ont indiqué que cela a pu être un **frein à l'utilisation des services**. D'autres ont déclaré privilégier le contact direct avec certains membres de l'IR* plutôt que passer par les adresses de contact afin d'accélérer et de fluidifier les échanges. Cette complexité est également ressortie concernant les modalités d'utilisation des services eux-mêmes, en particulier, dans l'utilisation de ShareDocs. En effet, le **manque**

¹⁸ **Progedo** (pour PROduction et Gestion des DONnées) est une IR* visant à développer la culture des données en sciences humaines et sociales. Son champ est centré sur les données d'enquête. Elle soutient la création et la diffusion de grandes enquêtes internationales et l'usage des statistiques publiques par les chercheur·euses.

d'accompagnement lors de la première utilisation et son ergonomie rendent parfois laborieuse l'utilisation de l'outil.

Cependant, de nombreux éléments indiquent la nécessité d'avoir une infrastructure « étoile » telle qu'Huma-Num. Tout d'abord, **le besoin et l'attachement aux services Huma-Num sont prégnants**. Lors d'un *focus group*, une personne a réagi en indiquant :

«Souvent quand il y a une DFG¹⁹ par exemple France-Allemagne, les serveurs c'est systématiquement ceux d'Huma-Num parce que l'Allemagne n'est pas capable de fournir ça. Donc il ne faut pas qu'ils baissent les bras en fait. On se plaint beaucoup mais...»

En particulier pour le service ShareDocs, à la question « Quel(s) service(s) d'Huma-Num utilisez-vous ? »²⁰, **90 % citent cet outil**. Une majorité de personnes indique également utiliser Nakala (51 %), Isidore (51 %) et Gitlab (54 %)²¹. Contrairement à Mattermost (31 %) et Kanboard (19 %) qui sont moins utilisés. Il est à noter que concernant Gitlab, il n'était pas spécifié s'il s'agissait de l'instance d'Huma-Num. Ces résultats démontrent le rôle essentiel de ces services pour les acteurs de la recherche qui y recourent.

Si nous faisons un focus sur ShareDocs, **90 % des personnes l'utilisent pour du stockage**, fonction initiale et principale²² de l'outil. Mais nous remarquons qu'il est également utilisé pour ses services complémentaires à l'intérieur, de **Speech-to-Text** (34 %) et d'**OCR** (34 %)²³, et à la marge de **conversion** (17 %). L'outil de compression est en revanche extrêmement peu utilisé (2 %). Cela montre l'intérêt de ses services complémentaires et leur

¹⁹ Voir le glossaire des **financements** en annexe à la page 27.

²⁰ Voir le **graphique** en annexe à la page 38.

²¹ Voir le glossaire des **services de l'IR* Huma-Num** à la page 26.

²² Voir le **graphique** en annexe à la page 39.

²³ Voir le glossaire des **outils** à la page 27.

utilisation au sein des projets. Seuls quelques cas montrent une utilisation de ShareDocs pour les services complémentaires uniquement.

À la marge, d'autres services et logiciels ont été énumérés par les répondants tels que Huma-Num Box, Stylo²⁴, des licences d'outils, de l'hébergement web et des machines virtuelles. Une nuance est à apporter sur ces services, car nous avons conçu cette interrogation sous la forme d'une question ouverte. Cela signifie que certaines personnes utilisant ces services ne les ont pas forcément indiqués.

Nous avons également demandé si les personnes utilisent les **outils de traitement** proposés par Huma-Num (licences d'outils qualitatifs/quantitatifs, machines virtuelles, etc.). Ils sont 29 % à avoir répondu favorablement. Cette proportion est d'autant plus importante que ces services sont peu mis en avant et ne disposent pas de « produits phares ». Parmi les répondants, 55 % utilisent d'autres outils de traitements que ceux proposés par Huma-Num, les deux questions ne s'excluant pas l'une l'autre.

Nous avons également demandé si des besoins d'accompagnement seraient nécessaires sur les trois services principaux : ShareDocs, Nakala, Isidore²⁵. On peut remarquer que la majorité des personnes souhaiterait un plus grand accompagnement sur ces services. Mais ce besoin est d'autant plus marqué pour Nakala (61 %) et Isidore (48 %). Cela confirme la pertinence pour Huma-Num d'avoir instauré un système de modération de proximité pour les demandes **Nakala**²⁶ d'une part et de dispenser d'autre part une **formation**

²⁴ Voir le glossaire des **services tiers hébergés** à la page 27.

²⁵ Voir le glossaire des **services de l'IR* Huma-Num** en annexe à la page 26.

²⁶ Une nouvelle fonctionnalité permet à un gestionnaire de données de demander la modération du dépôt afin d'en évaluer la qualité documentaire. Si le dépôt est conforme aux critères qualité définis, il pourra lui être attribué le statut « Modéré ». La modération est faite à la demande par un réseau de modérateurs identifiés par Huma-Num. La MSH LSE a fait partie des sites pilotes pour cette nouvelle fonctionnalité. Source : Rédaction (24 juin 2024). Évaluer la qualité documentaire des données : le projet de modération dans NAKALA.

sur **Isidore**, réalisée à la MSH LSE en septembre dernier. Par ailleurs, bien que le besoin de formation sur ShareDocs soit moindre (44 %), il reste tout de même important. Ce service a également fait l'objet d'une formation organisée par la MSH LSE en septembre 2023 (avec l'intervention de deux membres de l'IR*), et dans laquelle Mattermost et Kanboard ont également été évoqués.

Ces besoins d'accompagnement sur les services Huma-Num peuvent se recouper avec avec les pôles de compétences de la MSH. C'est pourquoi nous avons interrogé dans notre questionnaire la perception que peuvent avoir les répondant-es sur la MSH dans son rôle de relai des méthodes et des outils en Humanités numériques. C'est l'objet de la prochaine sous-partie.

La MSH Lyon Saint-Étienne comme relai des méthodes et des outils en Humanités numériques

Afin de poursuivre la réflexion sur l'accompagnement, le questionnaire comportait également des questions relatives à la perception de la MSH LSE en tant que spécialiste des méthodes et outils en Humanités numériques²⁷. Nous pouvons remarquer que pour les différents types d'accompagnement, une majorité des répondants considère la MSH LSE comme un relai des méthodes et outils en Humanités numériques. Nous pouvons néanmoins nuancer cela en soulignant qu'environ un tiers des réponses évoquent une incertitude sur les types d'accompagnement proposés par la MSH LSE. Cela révèle que les services proposés par la MSH LSE demeurent assez méconnus par un grand nombre d'agents au sein des laboratoires.

Si on analyse ces résultats en détails, c'est sur sa **fonction de conseil que la MSH LSE est perçue comme un relai** (68 %). C'est également sur ce type d'accompagnement où les personnes ont le moins répondu

Le blog d'Huma-Num et des Consortiums-HN. Consulté le 25 octobre 2024 à l'adresse <https://doi.org/10.58079/11v7o>.

²⁷ Voir le **graphique** en annexe à la page 41.

défavorablement (4 %). À l'inverse, l'accompagnement sur l'aspect du partage est le seul à être inférieur à la moitié des répondants (47 %). Bien que la notion de partage dépasse les entrepôts de données, cela renforce l'idée d'**accompagner davantage sur les entrepôts de données** tels que Nakala. Cette notion recouvre également le RGPD²⁸, le dépôt de logiciels²⁹ et le dépôt de publications.

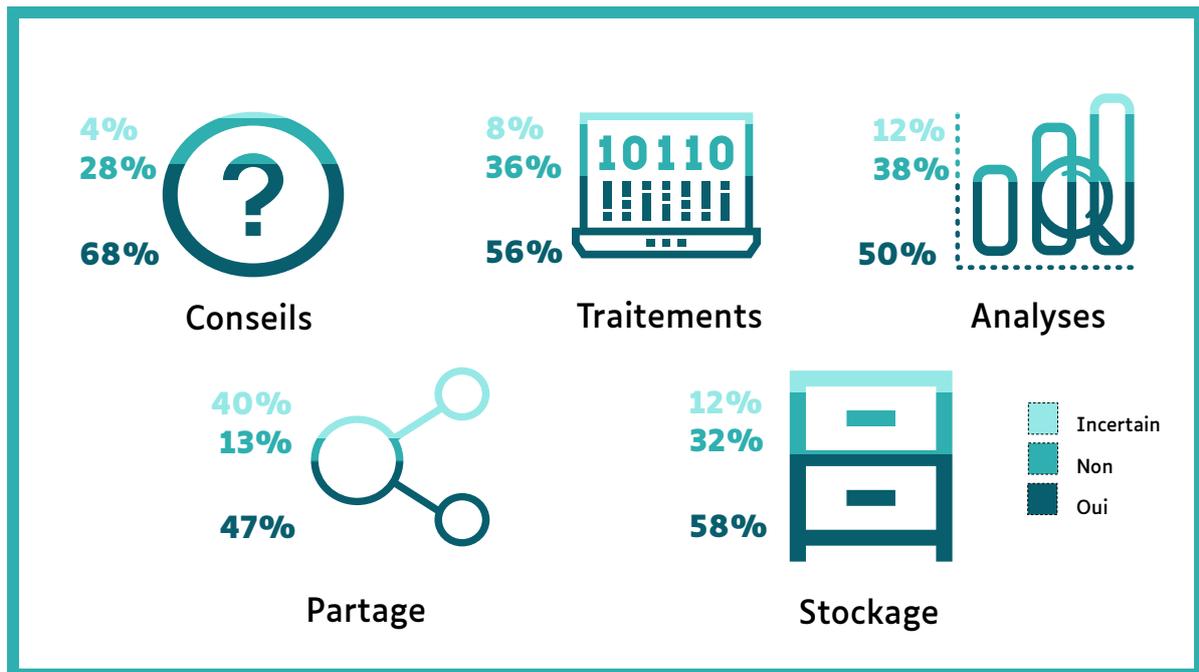


Figure 7 La MSH LSE en tant que relai des méthodes et outils des Humanités numériques exprimé en pourcentage

²⁸ Le règlement général sur la protection des données, relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE (règlement général sur la protection des données). Source : « Le règlement général sur la protection des données - RGPD », sur CNIL, sans date (en ligne : <https://www.cnil.fr/fr/reglement-europeen-protection-donnees>).

²⁹ Dans le contexte du second Plan national pour la science ouverte, il a été établi qu'il était essentiel, à l'instar des données et des publications, que les logiciels fassent l'objet d'une ouverture et d'un archivage. Software Heritage (développé par l'INRIA) a été présenté comme une solution à cet archivage des logiciels notamment par la création d'un SWHID complément logiciel du DOI pour les publications et les jeux de données. Source : https://www.ouvrirlascience.fr/wp-content/uploads/2021/06/Deuxieme-Plan-National-Science-Ouverte_2021-2024.pdf.

Paradoxe entre la volonté d'ouverture des données et la réalité des projets

Au cours du questionnaire, nous avons évoqué plus précisément l'ouverture des données de la recherche. À la question « Souhaitez-vous mettre à disposition vos propres jeux de données de recherche dans le cadre de nouvelles recherches ? »³⁰, **67 % des personnes ont indiqué vouloir partager leurs données**. C'est un résultat important si nous reprenons les résultats de l'enquête sur la gestion et ouverture des données de recherche en SHS où « 75 % des répondants ne s'y sont jamais livrés et où seuls 7 % disent publier leurs données en accès ouvert régulièrement »³¹. Les principales raisons évoquées pour ouvrir les données sont :

*“Les principes des Sciences Ouvertes ;
Répondre aux principes FAIR ;
La difficulté à les obtenir donc l'envie de les partager ;
Une obligation de l'ANR ;
Ce sont des projets publics ou financés
par de l'argent public”*

Néanmoins, ce chiffre important représente un paradoxe au regard de la réalité de l'ouverture des données dans les projets. Dans le recensement des projets, **seuls 10 % de ceux-ci ont ainsi une partie ou la totalité de leurs données ouvertes**. Ce chiffre est très faible en comparaison de la propension des personnes à vouloir le faire. On peut tout d'abord nuancer nos propos, car les projets recensés dans la présente enquête sont des projets en cours ou qui viennent de se terminer. Ce sont également parfois des projets qui ont bénéficié de plusieurs financements successifs. Par ailleurs, le partage des données est souvent réalisé en dernière phase du projet et non en amont du projet.

³⁰ Voir le **graphique** en annexe à la page 42.

³¹ Marie-Emilia Herbet, Aurore Montébran, Alexandra Dugué, Raphaëlle Billy, Mariana Maire, et al., *Gestion et ouverture des données de recherche en SHS : une enquête auprès des communautés de Lyon 2, Lyon 3 et la MSH Lyon St-Étienne*. Université Jean Moulin Lyon 3 ; Université Lumière Lyon 2 ; MSH Lyon - Saint-Étienne. 2023, p. 23.

À la question relative à l'utilisation des jeux de données mis à disposition³², les personnes n'ont majoritairement jamais réutilisé des jeux de données existants (63 %). Nous pouvons en conclure que les personnes sont davantage prêtes à partager leurs propres données plutôt qu'à recourir à celles produites par des tiers. On peut émettre plusieurs hypothèses face à ce constat. Il est tout d'abord possible que **l'injonction à partager des données** soit présente, de même que la sensation que ce partage soit utile dans un contexte de Science Ouverte. Mais *a contrario* cette volonté de partage ne coïncide pas nécessairement avec un besoin d'obtenir des données produites par d'autres. On peut se demander en conséquence si une **démarche de recherche sur les données déjà produites** est réalisée en début de projet.

Enfin, on peut également observer qu'**une part non négligeable de personnes ne souhaitent pas partager leurs jeux de données** (34 %). Parmi elles, quelques personnes ont déjà eu recours à des jeux de données mis à disposition (six personnes), avec quelques freins :

«*Confidentialité, données sensibles, propriété intellectuelle ;
Risque de plagiat ;
Trop long à anonymiser ;
Problématique particulière pour que
leur partage ait un sens.*»

On peut noter deux raisons principales au non-partage de ses données : le **temps de travail** que cela peut demander pour rendre les **données partageables** et **anonymisées** mais également la crainte d'être dépossédé de son travail de recherche. Concernant le temps de travail que cela implique, celui-ci pourrait être réduit si cela était anticipé dans un **plan de gestion de données** (PGD) en début de projet. Les pôles de compétence de la MSH LSE ont déjà initié ce rôle d'accompagnement et doivent continuer à le développer, que ce soit pour le PGD ou pour l'anonymisation des données.

³² Tel que BnF Data, Progedo, Recherche Data Gouv.

Des projets hétéroclites dans le périmètre de la MSH Lyon Saint-Étienne

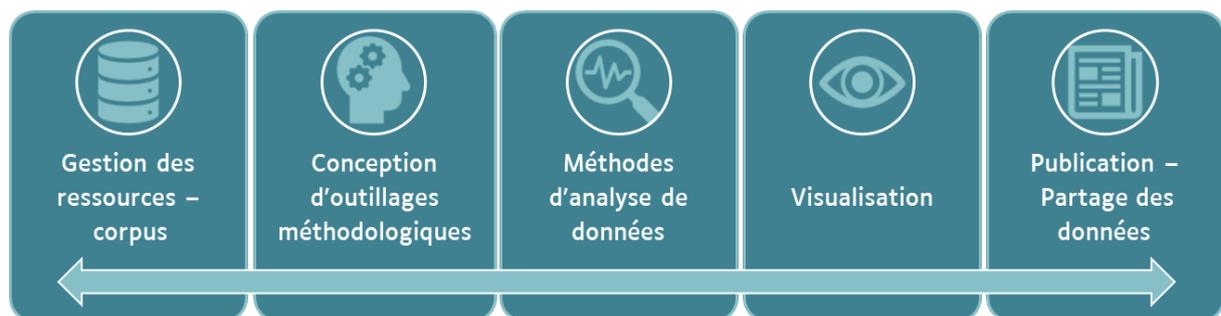
Comme évoqué précédemment, plus d'une soixantaine de projets – ayant pour critère être interdisciplinaire entre les Sciences Informatiques et une des disciplines des SHS, d'avoir été financés et d'être toujours en cours ou terminés récemment – ont pu être recensés et ajoutés à la base de données. Nous dénombrons 29 disciplines représentées parmi lesquelles nous retrouvons davantage des projets en Histoire (14 %), en Linguistique (8 %), en Géographie (5 %) et en Sciences de l'éducation (5 %). L'Informatique est la plus représentée, car elle fait partie intégrante des critères de sélection des projets. De la même façon, le LIRIS (10 %), laboratoire d'Informatique, est le plus représenté, suivi du LARHRA (8 %), laboratoire d'Histoire. 24 laboratoires associés à la MSH LSE sur les 53 sont représentés. Cette représentation de « seulement » 45 % des laboratoires associées à la MSH LSE s'explique par deux facteurs. D'une part, tous les laboratoires n'ont pas une dimension en Humanités numériques de recherche. D'autre part les projets recensés sont uniquement ceux bénéficiant d'un financement. Donc les projets ayant une dimension Humanités numériques mais sans financement n'ont pas été recensés, ce qui constitue un biais sur les projets que nous avons répertoriés et la représentation des unités.

Concernant les types de financement de ces projets, ce sont incontestablement les projets ANR (44 %) qui sont les plus présents. Les appels à projets de la MSH LSE arrivent seulement 5^e (6 %) après les LABEX (13 %), les APPI (10 %) et les ERC (8 %). Bien que cela ne soit pas l'unique indicateur de la connaissance des projets par la MSH LSE, cela montre que la majorité des projets ne sont pas connus de cette dernière.

Afin de pouvoir disposer d'une cartographie plus globale des projets existants, nous avons mis au point avec Marie Deprès-Lonnet³³ une **typologie des**

³³ Marie Desprès-Lonnet était professeure des universités en information et communication au laboratoire Elico. Elle est également responsable de l'axe Sociétés et Humanités numériques. Elle s'intéresse au rôle et à la place des dispositifs d'information et de médiation

projets³⁴. Cette typologie a pour fonction de mieux appréhender les forces du site, au regard des Humanités numériques. Nous avons pu l'éprouver notamment lors des deux *focus groups* que nous avons effectués. Néanmoins, cette typologie nécessiterait d'être davantage expérimentée auprès des porteur-euses pour s'assurer de sa concordance avec la réalité de leurs pratiques. Comme précédemment indiqué, nous avons ainsi dressé cinq grandes catégories :



Chacune de ces catégories ayant elles-mêmes des sous-catégories. Chaque projet d'ailleurs peut être mentionné dans plusieurs catégories. Il aurait été intéressant d'associer à ces catégories les outils, eux-mêmes associés aux projets. Mais le nombre d'outils que nous avons pu recenser n'a pas été suffisant pour réaliser un travail pertinent.

Collaboration entre les laboratoires associés

À travers le recensement des projets, nous avons aussi mis en évidence certaines collaborations entre des laboratoires associés à la MSH LSE. En effet, nous avons remarqué, lors d'un *focus group*, des personnes provenant de laboratoires et de disciplines différentes, mais travaillant dans des projets communs. Nous avons ainsi pu réaliser un **graphe de réseaux** des laboratoires afin d'identifier quels sont ceux impliqués et centraux dans

dans la production, la circulation et l'appropriation des savoirs et de la culture, principalement dans les « environnements numériques » mais également dans différents lieux de pratique comme les bibliothèques ou les musées. <https://elico-recherche.msh-lse.fr/membres/marie-despres-lonnet>.

³⁴ Voir **typologie** complète en annexe à la page 45.

plusieurs projets³⁵. Le LIRIS (laboratoire informatique associé à la MSH LSE via son pan SHS) est ainsi le plus en lien avec d'autres laboratoires. Mais le laboratoire ERIC, spécialisé dans la science des données et l'informatique décisionnelle, est aussi lié à d'autres laboratoires associés à la MSH LSE à travers plusieurs projets. On peut aussi noter que certains laboratoires du site collaborent avec des laboratoires extérieurs, hors site. C'est en particulier le cas pour des ANR et pour des ERC où des collaborations nationales ou européennes sont attendues.

La base de données ainsi constituée pourrait créer un **terrain fertile de collaborations** entre les différents laboratoires du site en mettant en valeur les compétences et les thématiques de ces derniers. C'est la raison pour laquelle nous avons souhaité donner aux membres des laboratoires la possibilité de collaborer au développement de la base de données *via* un **formulaire**[®] disponible sur le site web que nous avons créé³⁶.

Nous avons présenté sous différents aspects et à différentes étapes de la vie de la donnée les résultats de l'enquête que nous avons mené. Suite à ces résultats, un certain nombre de piste d'amélioration peuvent être énoncé en vu d'améliorer l'accompagnement sur le territoire et de faire évoluer les liens avec l'IR* Huma-Num. C'est que nous allons voir dans la partie suivante.

Pistes d'amélioration proposées par la MSH LSE

Territorialisation des services Huma-num : des services de proximité

Le premier constat que nous avons pu faire sur les services d'Huma-Num, c'est que leur **accès** est parfois compliqué. Cela s'accroît lorsqu'on est isolé dans un laboratoire, par l'absence d'ingénieur-es notamment. De même, l'infrastructure ayant une portée nationale, elle peut être perçue comme inaccessible. C'est la raison pour laquelle nous pensons que les MSH – dans le

³⁵ Voir le **graphe de réseaux** en annexe à la page 44

³⁶ Voir **site web des projets en Humanités numériques**[®] (lien : https://heurist.huma-num.fr/heurist/?db=sylvain_enquete_HumaNum&website&id=396&lang=FRE)

cas présent la MSH LSE – peuvent assurer le **rôle d'intermédiaire** entre les chercheur·euses des laboratoires et l'IR*. C'est déjà la vocation des **référent·es Huma-Num**, mais leur rôle pourrait être étendu. Cela pourrait se traduire par une **décentralisation partielle des demandes** faites sur le site pour une **évaluation et une prise en charge par les référent·es**, au moins pour un premier niveau de demande. Cela permettrait à la MSH LSE d'accroître sa connaissance des besoins et à Huma-Num de répondre plus efficacement aux différentes demandes. Pour cela, une piste serait de disposer d'une visibilité sur les demandes d'accès aux services de l'IR* dans le périmètre de la MSH LSE. Concrètement, l'indication de l'unité ou de l'institution d'appartenance par le demandeur lors d'une demande d'ouverture d'un service permettrait de notifier ces demandes en local. Cette redirection pourrait se faire au moyen de l'envoi d'un mail automatique aux référent·es ou par la création d'un outil dédié à la gestion des demandes. Le traitement de la demande serait alors encadré par l'IR*, *via* des formations ou des guides, afin de sécuriser une gestion de la demande en totale conformité avec les procédures et attendus d'Huma-Num. Ce transfert de la demande devra alors est explicité à l'utilisateur·ice et faire l'objet d'un traitement conforme au RGPD. La place qu'occupe la MSH LSE au sein du site, par son caractère transversal, lui permet d'avoir **une connaissance plus fine des acteurs et de leurs besoins**. Cette connaissance permet d'avoir un contact plus direct avec les acteurs et d'améliorer la définition, la détermination de leurs besoins. Ainsi ce premier niveau de traitement par les MSH permettrait, si nécessaire, de les rediriger vers les services de l'IR* les plus adaptés ou vers des solutions plus adéquates, permettant ainsi aux services centralisés d'alléger leur charge et de se concentrer sur les demandes nécessitant un niveau d'expertise plus pointu.

La territorialisation des services Huma-Num ne peut pas se faire sans une intensification des accompagnements. C'est la raison pour laquelle vous verrons dans la partie suivante la proposition de renforcer l'accompagnement en particulier à l'égard des doctorant·es par la mise en place de temps de formation.

Des accompagnements intensifiés : les doctorant·es au cœur de la démarche

Compte tenu du besoin d'accompagnement sur les services d'Huma-Num³⁷, il est nécessaire de **renforcer les accompagnements** sur des initiatives déjà existantes. En premier lieu, il semble nécessaire d'accroître le **temps de formation** sur les différents services proposés par Huma-Num. Ce temps de formation pourrait être assuré en proximité par les agents de la MSH LSE, en particulier les référentes Huma-Num, et par des membres d'Huma-Num. Il semble prioritaire que dans un premier temps, ces formations traitent des services principaux d'Huma-Num : Nakala, ShareDocs et Isidore. Dans un second temps, à l'issue d'une consultation de la communauté, d'autres services proposés par l'IR* pourraient être l'objet d'une nouvelle vague de formations.

Au vu du nombre très important de **doctorant·es** qui ne connaissent pas l'IR* Huma-Num, il semble essentiel de réaliser des **temps d'échanges et de formations en ciblant ce public spécifiquement**. De ce fait, la création de nouveaux partenariats avec les écoles doctorales (ED) serait un bon moyen d'atteindre cet objectif, car cela permettrait aux doctorant·es de pouvoir valoriser ce temps de formation dans leur cursus. Par ailleurs, des collaborations avec des écoles doctorales ont déjà vu le jour avec le pôle audio-visio de la MSH LSE³⁸. Nous pourrions donc étendre ces **collaborations entre le pôle Humanités numériques de la MSH LSE et les différentes ED**. Cette possibilité avait été évoquée lors d'un *focus group*. L'intérêt est de pouvoir offrir des outils supplémentaires pour des personnes qui débutent dans la recherche et qui seraient amenées à les employer tout au long de leur activité de recherche. De la même manière, cela pourrait amener des échanges

³⁷ Voir le **graphique** des accompagnements sur les services Huma-Num à la page 40.

³⁸ À titre d'illustration, le PI2A a organisé une formation en 2017 avec l'ED483 SCIENCES SOCIALES et CMW autour de « L'image dans les recherches en sciences humaines et sociales » et en 2023 une formation sur la « Médiation avec vidéo et audio ».

avec leurs encadrants et ainsi augmenter l'acculturation autour de la question.

Bien qu'il soit essentiel d'intensifier les accompagnements, il nous semble également très important de pouvoir positionner la MSH en tant que pont entre les laboratoires avec leurs disciplines et compétences respectifs. Nous allons voir cela dans la partie suivante.

Renforcer l'interconnaissance entre les laboratoires

Au cours de notre enquête, nous avons pu voir, à travers le très grand nombre de laboratoires associés, des compétences et des expertises très différentes. Des collaborations existent déjà entre certains laboratoires du site. Nous pensons que la MSH LSE, par sa **position transversale au niveau disciplinaire et institutionnel**, peut favoriser la mise en réseaux des acteurs de la recherche par la connaissance des autres laboratoires et de leurs compétences propres. Ce renforcement de l'interconnaissance s'illustre en particulier dans le cadre des Humanités numériques, à la fois par la mise en lien de laboratoires en informatique (tel que le LIRIS ou ERIC) avec des laboratoires SHS, de même qu'au sein même des laboratoires en SHS ayant des compétences dans les Humanités numériques. À titre d'exemple, le LARHRA développe un outil, Geovistory, permettant de faire du web sémantique et d'avoir des données nativement FAIR, et le CIHAM regroupe plusieurs personnes expertes en reconnaissance de l'écriture manuscrite (HTR). Ces exemples illustrent la capacité de nombreux laboratoires du site à développer des outils dont le champ d'application et l'utilité s'étendent au-delà de leur seule structure. Bien que ces outils soient pensés pour des usages disciplinaires, ils pourraient être généralisés à d'autres disciplines. L'HTR, qui permet de faire de la reconnaissance de caractères manuscrits est à ce jour en majorité utilisé par les historiens et les littéraires. Mais cette technologie peut être appliquée à l'ensemble des corpus écrits, anciens comme contemporains. Néanmoins, elle doit faire l'objet d'une mise en lien des spécialistes de cette dernière vers des personnes susceptibles de l'utiliser.

Cette mise en lien pourrait se faire de deux façons. Tout d'abord, pérenniser la **base de données sur les projets en Humanités numériques** et s'assurer de la modification et de la mise à jour des contenus par une personne référente. Cela passe également par la promotion de l'outil et son enrichissement au moyen d'un espace de collaboration. De plus, la typologie que nous avons créée, si elle est approuvée par un ensemble de personnes concernées, peut permettre de mieux identifier des besoins et des méthodes nécessaires à la réalisation d'un projet en Humanités numériques. La base de données peut jouer un rôle de vecteur entre les personnels du site quand certaines compétences ne sont pas présentes dans leur laboratoire de rattachement. Pour la MSH LSE, un tel outil permet de connaître les projets du site afin de pouvoir mieux identifier les accompagnements possibles et/ou d'assurer la mise en lien entre les laboratoires du site. De même, pour Huma-Num, et à plus forte raison si un tel outil est proposé à une échelle nationale, la base de données assurerait une fonction d'observatoire des projets pour mieux en comprendre les besoins et orienter les choix stratégiques à déployer.

Une seconde façon de renforcer cette mise en lien est la constitution d'une ou plusieurs **journées annuelles sur la présentation de projets en Humanités numériques**. Ces journées permettraient la rencontre de personnes qui font des Humanités numériques et qui peuvent avoir des méthodes et des besoins communs. Ces journées pourraient être portées par la plateforme PAGES de la MSH LSE, dans le cadre de sa labellisation nouvelle. Elles pourraient également être co-portées par l'axe Sociétés et Humanités numériques de la MSH LSE.

L'interconnaissance entre les différences laboratoires peut se jouer à un autre niveau, celui de l'ouverture des données, permettant leur réutilisation par d'autres. Pour cette raison, l'accompagnement sur l'ouverture des données nous semble un enjeu majeur. Nous le verrons dans la partie suivante.

Accompagnement sur l'ouverture des données

Comme nous avons pu le voir précédemment, malgré la réelle volonté d'ouvrir les données de recherche, la réalité dans les projets montre une faible

ouverture de ces dernières. Il semble impératif d'œuvrer à l'ouverture des données de la recherche dans un contexte où les incitations à la science ouverte sont de plus en plus présentes. De même, la création de Recherche Data Gouv³⁹, incitant les entrepôts de données à être pourvus de standards pour garantir la qualité de leurs métadonnées, induit la nécessité de rendre accessible des jeux de données pour qu'ils puissent être réutilisés.

De par ces enjeux, il semble impératif de garantir la qualité et donc **l'accompagnement sur l'ouverture des données**. Cela passe déjà par la constitution de PGD qui permettent d'anticiper ces enjeux (rendus obligatoires par des organismes comme l'ANR), mais également par l'accompagnement au dépôt sur des entrepôts de données adaptés. Pour cette raison, l'initiative, lancée par Huma-Num, de modération et de création d'un réseau de référent-es Nakala, semble être une solution parfaitement adaptée pour répondre à ces enjeux. Cette procédure a été expérimentée, plusieurs rendez-vous individuels entre des membres de la MSH LSE et des déposants du site s'étant déjà tenus. Cependant, il est nécessaire d'aller plus loin sur ces **incitations à déposer des jeux de données** en donnant la visibilité sur l'ensemble de l'écosystème des entrepôts de données qui pourraient être plus appropriés à un projet, comme CoCoON⁴⁰ ou Ortolang⁴¹. De plus, il semble essentiel d'apporter un soutien supplémentaire sur la constitution de jeux de données et d'utiliser les bonnes métadonnées en vue de leur réutilisation. C'est en particulier le cas pour Nakala qui utilise le DublinCore comme vocabulaire contrôlé pour la description de ses métadonnées mais dont un

³⁹ **Recherche Data Gouv** est un écosystème au service du partage et de l'ouverture des données de la recherche créé en 2022 à la suite du deuxième plan pour la science ouverte. Il donne à la fois des outils pour mieux accompagner et centralise des entrepôts de données de qualité par du moissonnage.

⁴⁰ **CoCoON** pour « COLlections de CORpus Oraux Numériques » est une plateforme technique qui accompagne les producteurs de ressources orales, à créer, structurer et archiver leurs corpus ; un corpus pouvant se composer d'enregistrements (en général audio) accompagnés éventuellement d'annotations de ces enregistrements.

⁴¹ **Ortolang** est un équipement dans le domaine des ressources et outils de traitement de la langue. Il a notamment pour but d'archiver et enrichir des données dans le domaine la langue.

coût d'entrée est nécessaire si l'on veut remplir les métadonnées de façon adéquate.

Conclusion : une enquête montrant les besoins d'accompagnement

Ce travail d'enquête a permis de mettre en évidence des constats déjà perceptibles sur le site, en y apportant de l'objectivité *via* l'utilisation de méthodes qualitatives et quantitatives. Il a aussi donné lieu à la mise en place de nouveaux outils tels que la création d'une base de données sur les projets⁴².

L'enquête a fait ressortir des éléments saillants sur les besoins et les problématiques que les personnes peuvent rencontrer dans le cadre des projets en Humanités numériques et souligné l'importance de l'IR* Huma-Num dans un site comme celui de Lyon Saint-Étienne. Néanmoins, cela a aussi montré que des actions doivent être menées pour mieux accompagner sur les services de l'IR* Huma-Num et les faire connaître. Ces actions doivent être réalisés à destination de la communauté dans sa globalité, avec une attention particulière pour les doctorant-es.

L'ouverture, le partage des données, ainsi que la sensibilisation sont à poursuivre afin d'amoinrir l'écart entre la volonté de partager les données et le déploiement effectif de cette ouverture.

Ce travail d'accompagnement ne peut être fait qu'avec l'appui de l'IR* Huma-Num, en particulier sur une **territorialisation partielle de certains services**, pour lesquels la MSH LSE jouerait le rôle de médiateur, au niveau territorial, entre les agents dans leur laboratoire et l'IR* Huma-Num située à un niveau national.

Afin de jouer ce rôle de médiateur entre les laboratoires eux-mêmes, la MSH LSE peut aussi, par sa position transverse disciplinaire et institutionnelle,

⁴² Voir [site web des projets en Humanités numériques](https://heurist.huma-num.fr/heurist/?db=sylvain_enquete_HumaNum&website&id=396&lang=FRE) (lien : https://heurist.huma-num.fr/heurist/?db=sylvain_enquete_HumaNum&website&id=396&lang=FRE).

incarner un intermédiaire, un relais entre ces derniers, dont les modalités seraient à définir avec Huma-Num.

Enfin, bien que la MSH LSE– au vu du nombre et de la diversité de laboratoires qu'elle accompagne – puisse paraître représentative d'une conjoncture nationale, il serait opportun de mener ce **type d'enquête dans d'autres MSH et sur d'autres sites académiques** avec des configurations et des dynamiques scientifiques et institutionnelles pouvant être différentes. De la même manière, certaines données de l'enquête n'ont pas été intégrées à ce présent rapport, en particulier sur la question de la préservation des données à long terme dans les laboratoires. Ces dernières pourraient faire l'objet de réflexions futures.

Annexes

Services de l'IR* Huma-Num



ShareDocs : ShareDocs est un gestionnaire de fichiers mis en œuvre par l'IR* Huma-Num, sur ses propres serveurs, pouvant être utilisé via un navigateur web, un client WebDAV ou un logiciel de synchronisation de fichiers. Il s'agit d'un stockage déporté en ligne sur le réseau de recherche français (semblable à des outils comme DropBox hébergés quant à eux, sur le cloud public).

Nakala : Nakala est un entrepôt de données dont le but est de préserver et de disséminer les données produites par les productions des projets de recherche français en Sciences Humaines et Sociales dans le respect des principes FAIR. Nakala est destiné en premier lieu aux projets de recherche des institutions ayant une affiliation dépendante du MESR (ministère de l'Enseignement supérieur et de la Recherche).



Isidore : Isidore est un moteur de recherche et un assistant de recherche permettant de découvrir et de trouver des publications, des données numériques et profils de chercheurs et chercheuses en SHS venant du monde entier.

GitLab : L'instance GitLab d'Huma-Num permet l'hébergement sécurisé et le partage maîtrisé de fichiers, dont des fichiers de code (pour tous les langages) selon le protocole de versionnage git.



KB Kanboard : Kanboard est un service de gestion de projet basé sur la méthode organisationnelle « Kanban ». Ce service est réservé aux porteur-euses et participant-es de projets de recherche en SHS et mis à disposition dans la limite des ressources dont dispose Huma-Num.

Mattermost : Mattermost permet de partager des messages et des fichiers sur différentes plateformes tels que les PC, les téléphones et les tablettes.



Huma-Num Box : Le service Huma-Num Box est un service de stockage sécurisé pour la préservation de volumes importants de données (plusieurs téra-octets), par réplication et distribution sur plusieurs sites en France. C'est un service de stockage qui s'utilise comme un disque réseau ou via le protocole SFTP (avec un client comme Filezilla). Il est accessible depuis tout Internet, indépendamment de la présence ou non d'un serveur du dispositif sur son site géographique. Il est destiné aux données dites froides ou tièdes, qui ne sont pas fréquemment accédées.

Services tiers hébergés

Heurist : Heurist est un outil de création et publication de base de données en ligne, Open Source et à destination des SHS. Il a notamment été développé par Ian Johnson et est, hébergé par Huma-Num, en France pour l'ESR.

Stylo : Stylo est un éditeur de texte simplifiant la rédaction et l'édition d'articles scientifiques en Sciences Humaines et Sociales. Il est disponible en France via Huma-Num.

Outils

OCR : La reconnaissance optique de caractères (OCR pour l'anglais Optical Character Recognition), ou « océrisation », désigne les procédés informatiques pour la traduction d'images de textes imprimés ou dactylographiés en fichiers de texte.

Speech-to-Text : La reconnaissance automatique de la parole (souvent improprement appelée reconnaissance vocale) est une technique informatique qui permet d'analyser la voix humaine captée au moyen d'un microphone pour la transcrire sous la forme d'un texte exploitable par une machine. Sur le micro-service inclus dans ShareDocs, il s'agit d'un modèle de reconnaissance automatique de la parole multilingue intitulé Whisper et développé par OpenAI.

Financements

ANR : L'Agence nationale de la recherche (ANR) est une agence de moyens créée le 7 février 2005, qui finance la recherche publique et la recherche partenariale en France.

APPI : L'Appel à projets pluridisciplinaires interne (APPI) permet de favoriser l'émergence de projets de recherche originaux et/ou innovants qui, à terme, pourront se préparer à répondre à des appels à projets d'autres financeurs nationaux ou internationaux (Région, ANR, Europe, etc.).

DFG : Le Deutsche Forschungsgemeinschaft (DFG) est un service allemand de la science dans toutes ses ramifications qui apporte un soutien financier aux projets de recherche et favorise la coopération nationale et internationale entre chercheurs.

ERC : Le Conseil européen de la recherche (en anglais : European Research Council, ERC) est un organe de l'Union européenne chargé de coordonner les efforts de la recherche entre les États membres de l'UE et la première agence de financement pan-européenne pour une « recherche à la frontière de la connaissance ».

LABEX : LabEx (mot-valise fabriqué à partir de la contraction de Laboratoire d'excellence) est un des instruments du programme d'investissements d'avenir, destiné à soutenir la recherche d'ensembles d'équipes sur une thématique scientifique donnée.

Maisons des Sciences de l'Homme

MMSH : Maison méditerranéenne des Sciences de l'homme ; située à Marseille

MSH Ange-Guepin : Maison des Sciences de l'Homme Ange Guépin ; implantée à Angers, Nantes et Le Mans

RnMSH : Réseau national des Maisons des Sciences de l'Homme

Laboratoires

ARAR : Archéologie et Archéométrie

ARCHEORIENT : Environnements et sociétés de l'Orient ancien

CEL : Centre d'Etudes Linguistiques

CERGIC : Center for economic research on governance, inequality and conflict

CERCOR : Centre européen de recherche sur les communautés, congrégations et ordres religieux

CERCRID : Centre de Recherches CRItiques sur le Droit

CERGIC : Center for economic research on governance, inequality and conflict

CERLA : Centre de recherche en linguistique appliquée

CGN : Centre Gabriel Naudé

CIHAM : Histoire Archéologie Littératures des mondes chrétiens et musulmans médiévaux

CLHDPP : Centre Lyonnais d'Histoire du Droit et de la Pensée Politique

CMW : Centre Max Weber

COACTIS : COncption de l'ACTIon en Situation

CRDMS : Centre de recherche en droit et management des services de santé

CRPPC : Centre de Recherche en Psychopathologie et Psychologie Clinique

DDL : Dynamique Du Langage

DIPHE : Développement, Individu, Processus, Handicap Éducation

ECLLA : Etudes du Contemporain en Littératures, Langues et Arts

ECP : Éducation, Cultures, Politiques

EDIEC : Équipe de Droit International Européen et Comparé Équipe de droit public de Lyon

EDPL : Équipe de droit public de Lyon

ELICO : Équipe de recherche de Lyon en sciences de l'Information et de la Communication

ELJ : Équipe Louis Josserand

EMC : Étude des Mécanismes Cognitifs

ERIC : Entrepôts, Représentation et Ingénierie des Connaissances

EVS: Environnement, ville, société

GATE : Groupe d'Analyse et de Théorie Economique Lyon Saint-Etienne

GREPS : Groupe de Recherche En Psychologie Sociale

HISOMA : Histoire et Sources des Mondes Antiques

IAO : Institut d'Asie Orientale

ICAR : Interactions, Corpus, Apprentissage, Représentation

IETT : Institut d'Etudes Transtextuelles et Transculturelles

IHRIM : Institut d'Histoire des Représentations et des Idées dans les Modernités

IRAA : Institut de Recherche sur l'Architecture Antique

IRPHIL : Institut de Recherches Philosophiques de Lyon

ISERL : Institut Supérieur d'Étude des Religions et de la Laïcité

LADEC : Laboratoire d'anthropologie des enjeux contemporains

LAET : Laboratoire Aménagement Economie Transports

LCE : Lettres et civilisations étrangères

LER : Laboratoire d'Etudes Rurales

LARHRA : Laboratoire de Recherche Historique Rhône-Alpes

LIRIS : Laboratoire d'InfoRmatique en Image et Systèmes d'information

MAELYSE : Management Économie Lyon St-Etienne

MOM : Maison de l'Orient et de la Méditerranée - Jean Pouilloux

P2S : Parcours, santé, systémique

PASSAGES (XX-XXI) : Passages, arts et littératures (XX-XXI)

RADIATIONS : Défense, santé, environnement

RESHAPE : Recherche en santé publique/Research on healthcare performance

S2HEP : Sciences, Société, Historicité, Éducation, Pratiques

SAF : Sciences actuarielle et financière

TRANSVERSALES : Unité de recherche en droit

TRIANGLE : Action, discours, pensée politique et économique

Calendrier

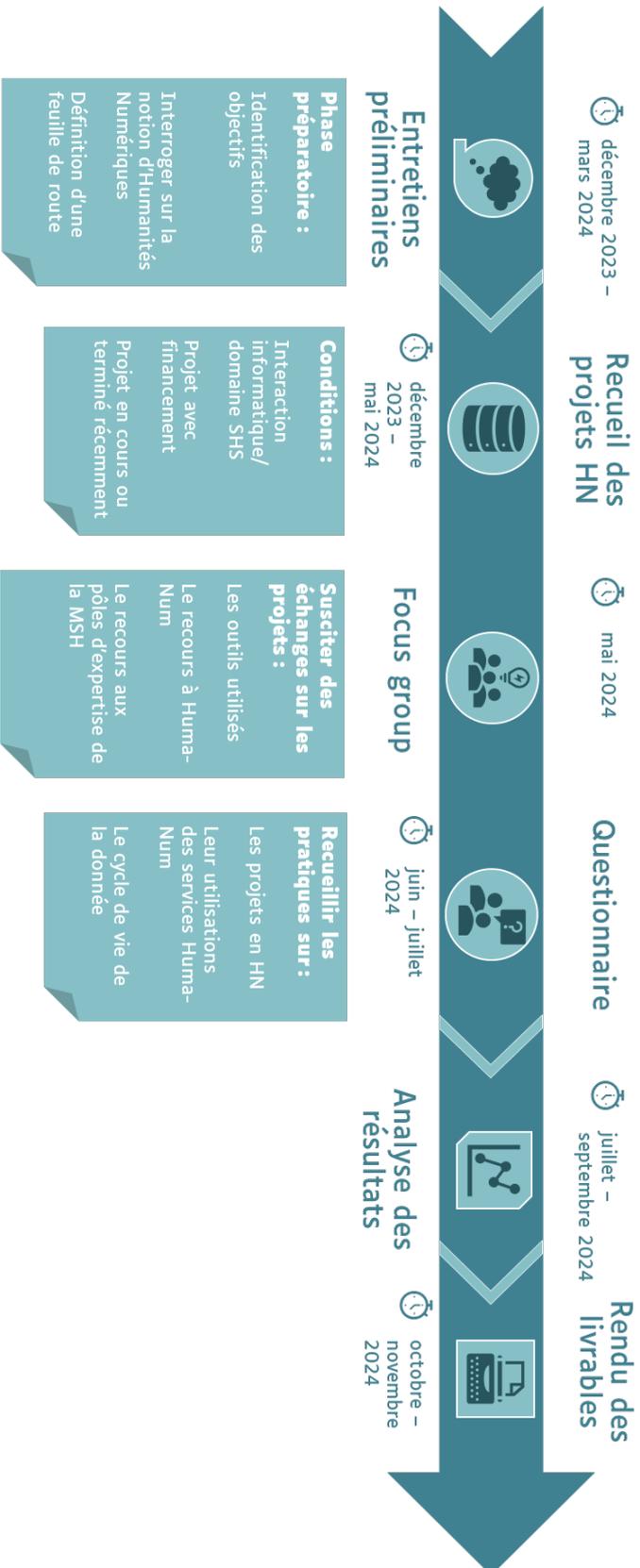
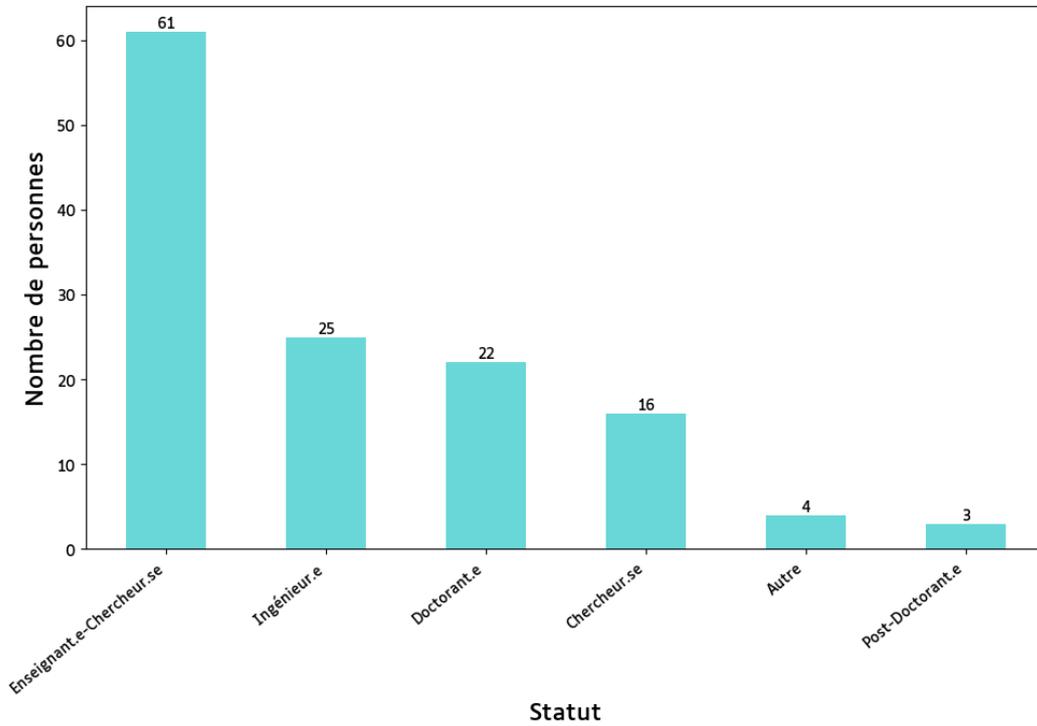


Figure 8 Frise des différentes phases du projet

Graphiques

Graphique montrant le statut des personnes
qui ont répondu à l'enquête en effectif (n=131)



Répartition des statuts des personnes
ayant répondu à l'enquête en pourcentage (n=131)

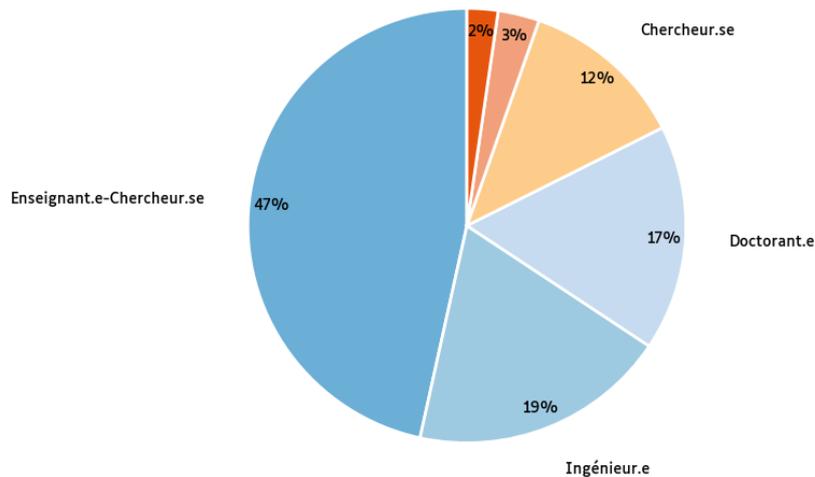
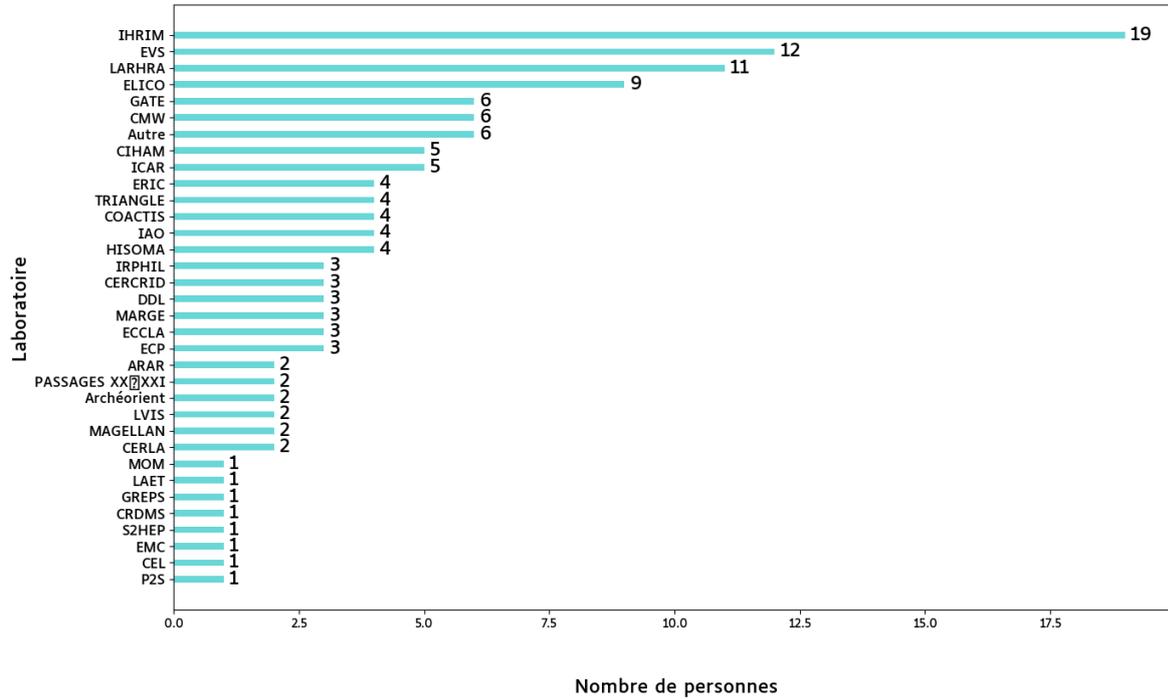


Figure 9 Graphiques montrant le statut des personnes ayant répondu à l'enquête exprimés en effectif et en pourcentage

Graphique montrant les laboratoires au sein desquels
les personnes ont répondu à l'enquête en effectif (n=137)



Graphique montrant les laboratoires au sein desquels
les personnes ont répondu à l'enquête en pourcentage (n=137)

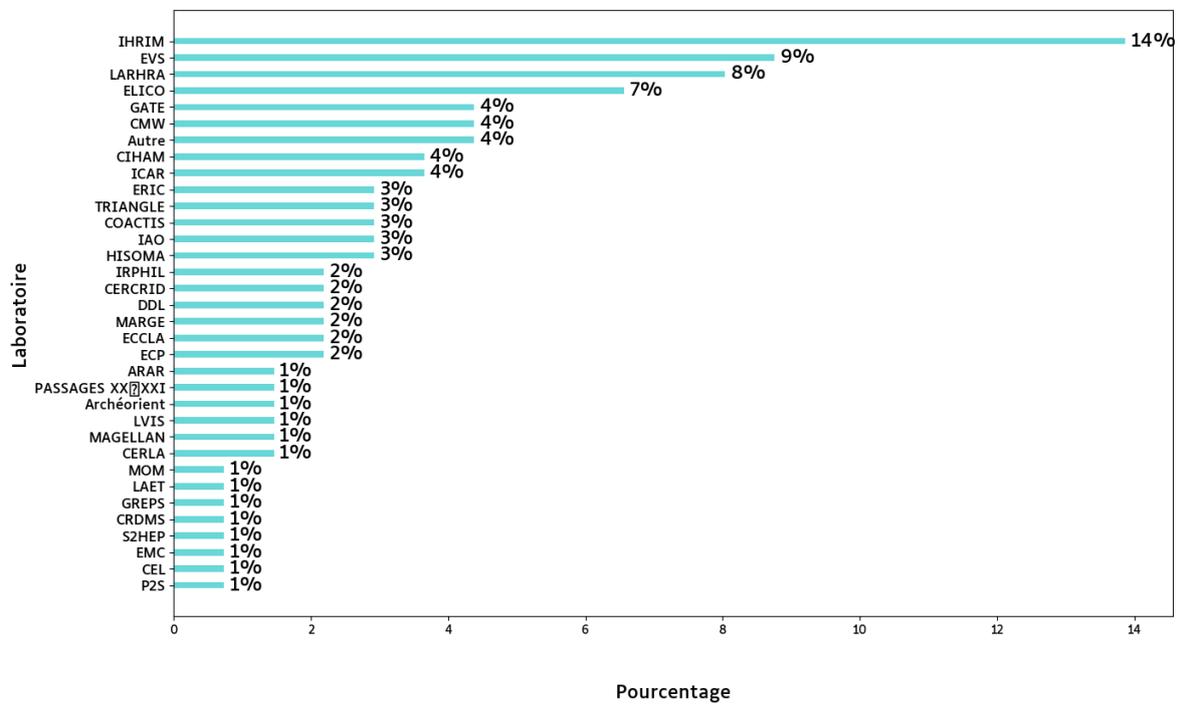
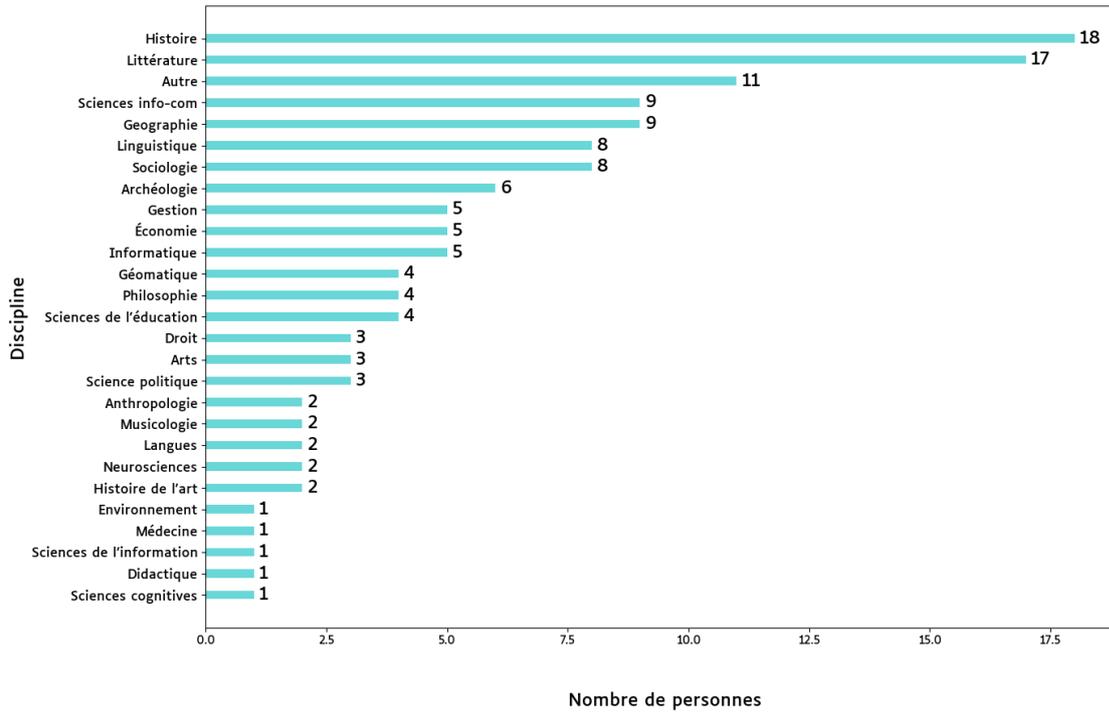


Figure 10 Graphiques indiquant les laboratoires de rattachement des personnes ayant répondu à l'enquête exprimés en effectif et en pourcentage

Graphique montrant les disciplines qui ont répondu à l'enquête en effectif (n=137)



Graphique montrant les disciplines qui ont répondu à l'enquête en pourcentage (n=137)

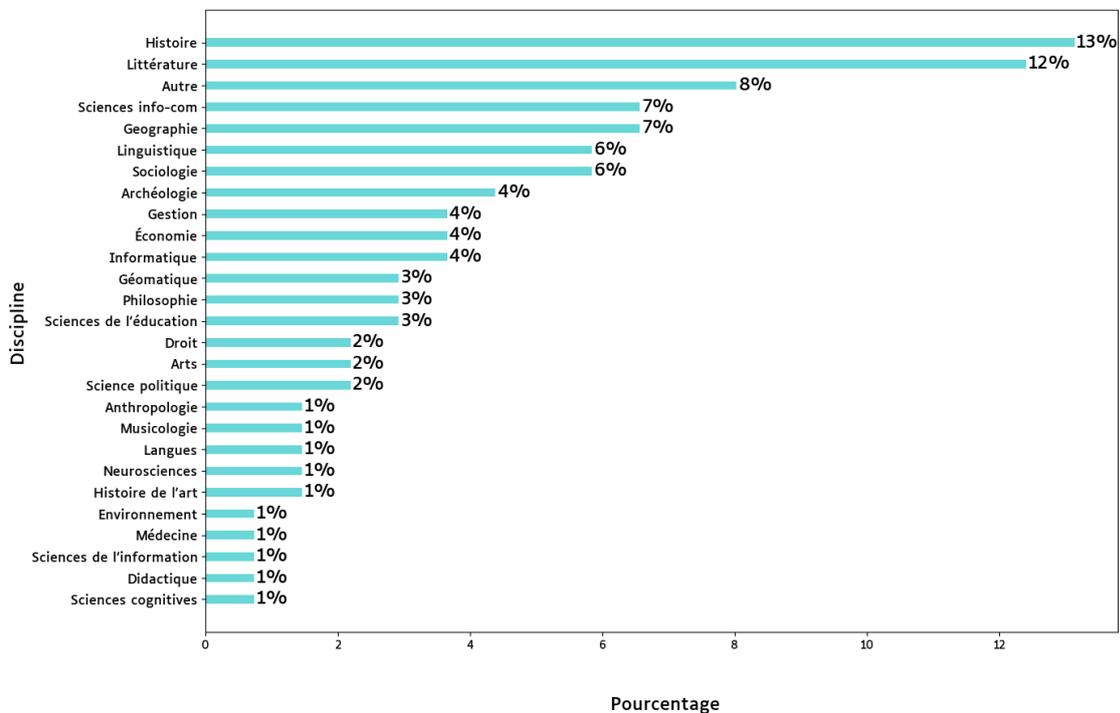
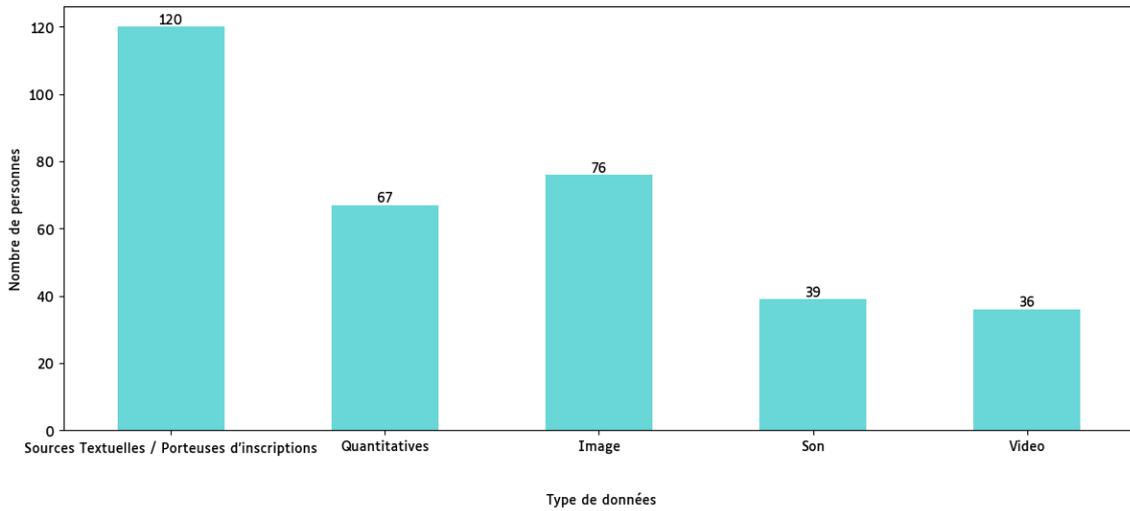


Figure 11 Graphiques indiquant la discipline principale étudiée par les répondants exprimés en effectif et en pourcentage

Graphique indiquant quels sont les types de données manipulées parmi les interrogés en effectif (n=141)



Répartition des types de données manipulées parmi les interrogés en pourcentage (n=141)

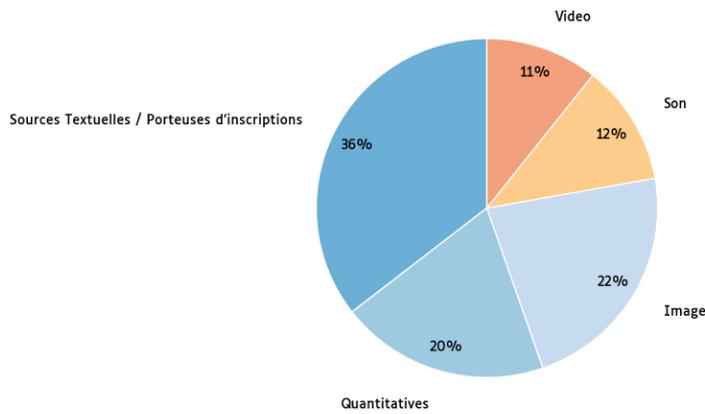
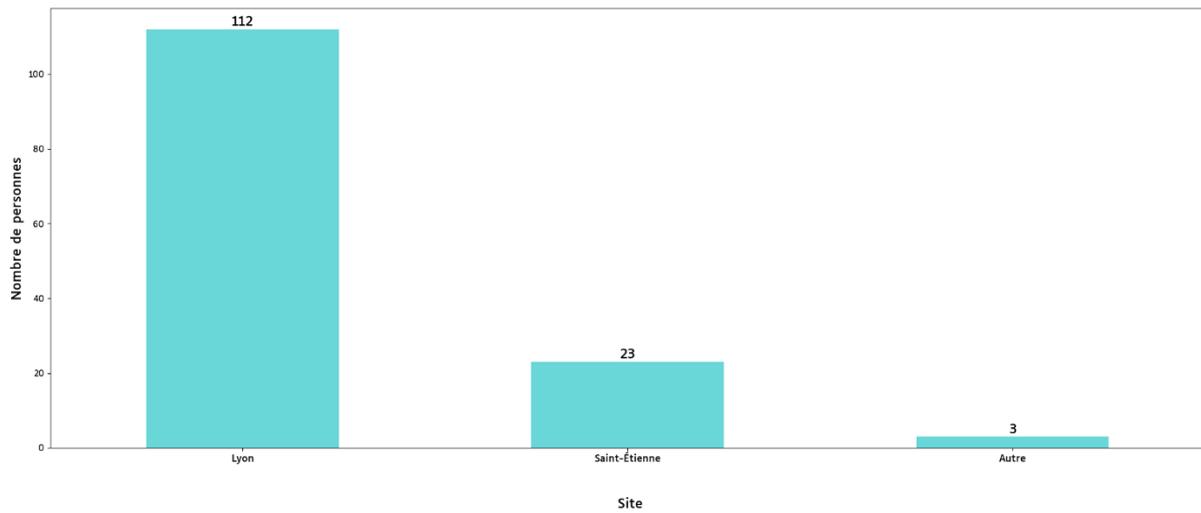


Figure 12 Graphique montrant les types de données manipulées par les interrogés en effectifs et en pourcentage

Graphique montrant sur quel site les personnes ayant répondu à l'enquête sont situées en effectif (n=138)



Graphique montrant sur quel site les personnes ayant répondu à l'enquête sont situées en pourcentage (n=138)

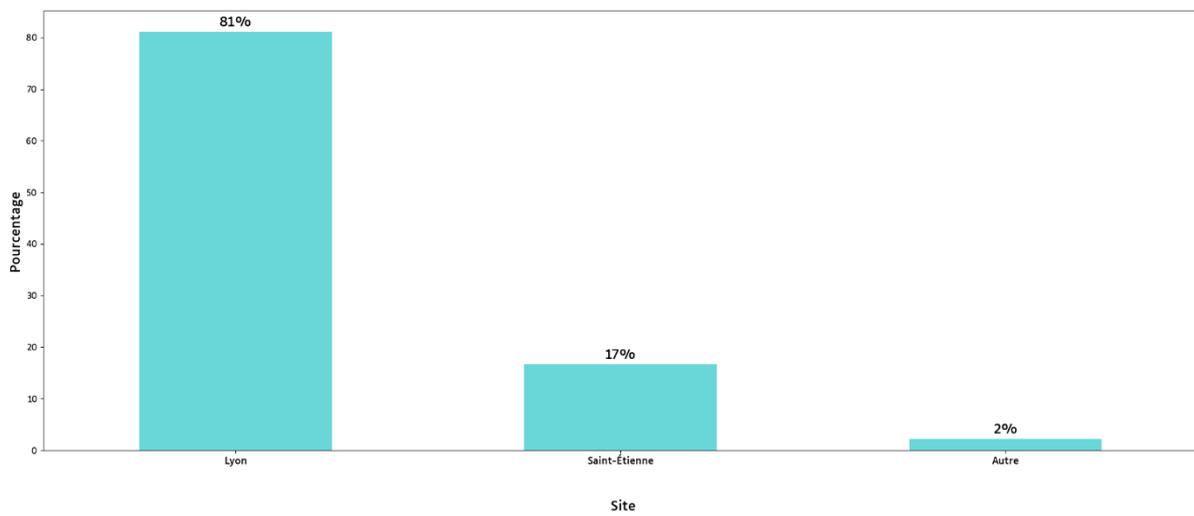
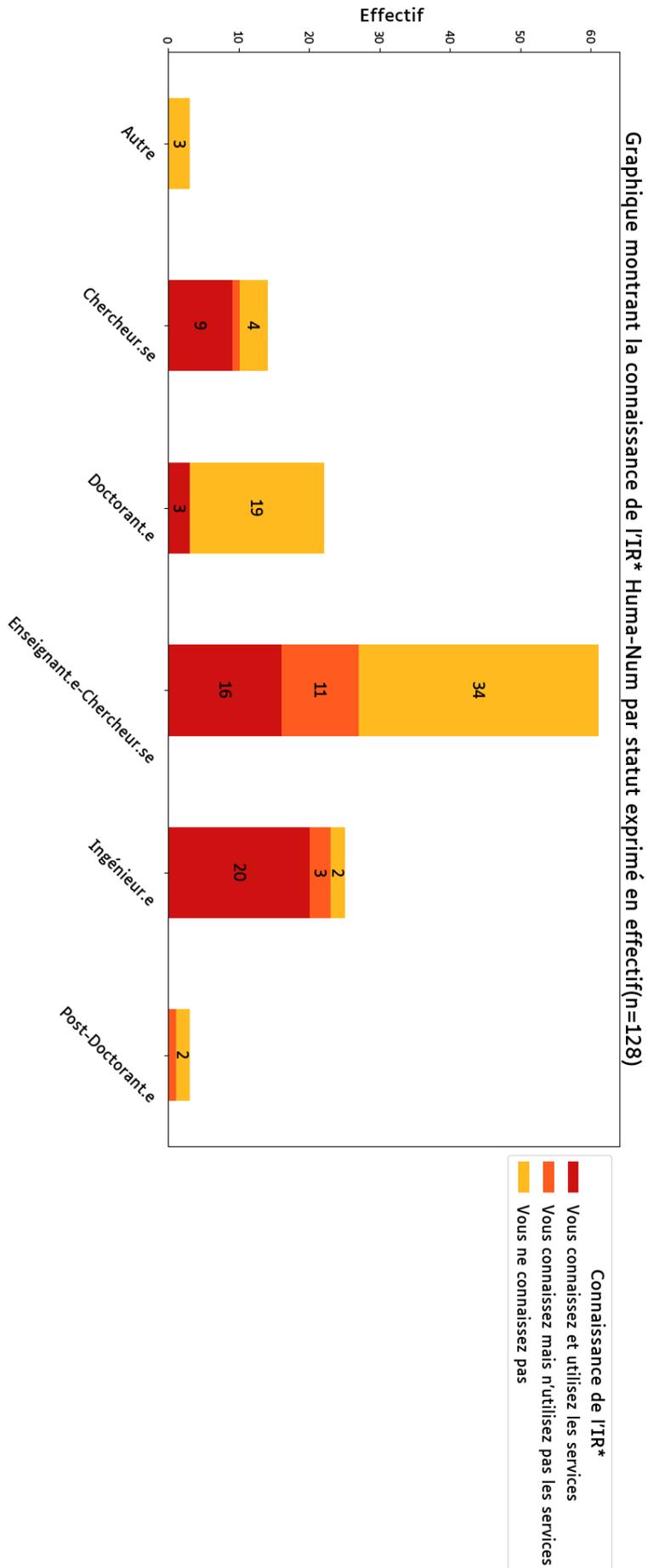


Figure 13 Graphiques montrant sur quel site les personnes ayant répondu à l'enquête sont situées exprimés en effectif et en pourcentage

Figure 14 Graphique montrant la connaissance de l'IR Huma-Num, selon le statut, exprimés en effectif*



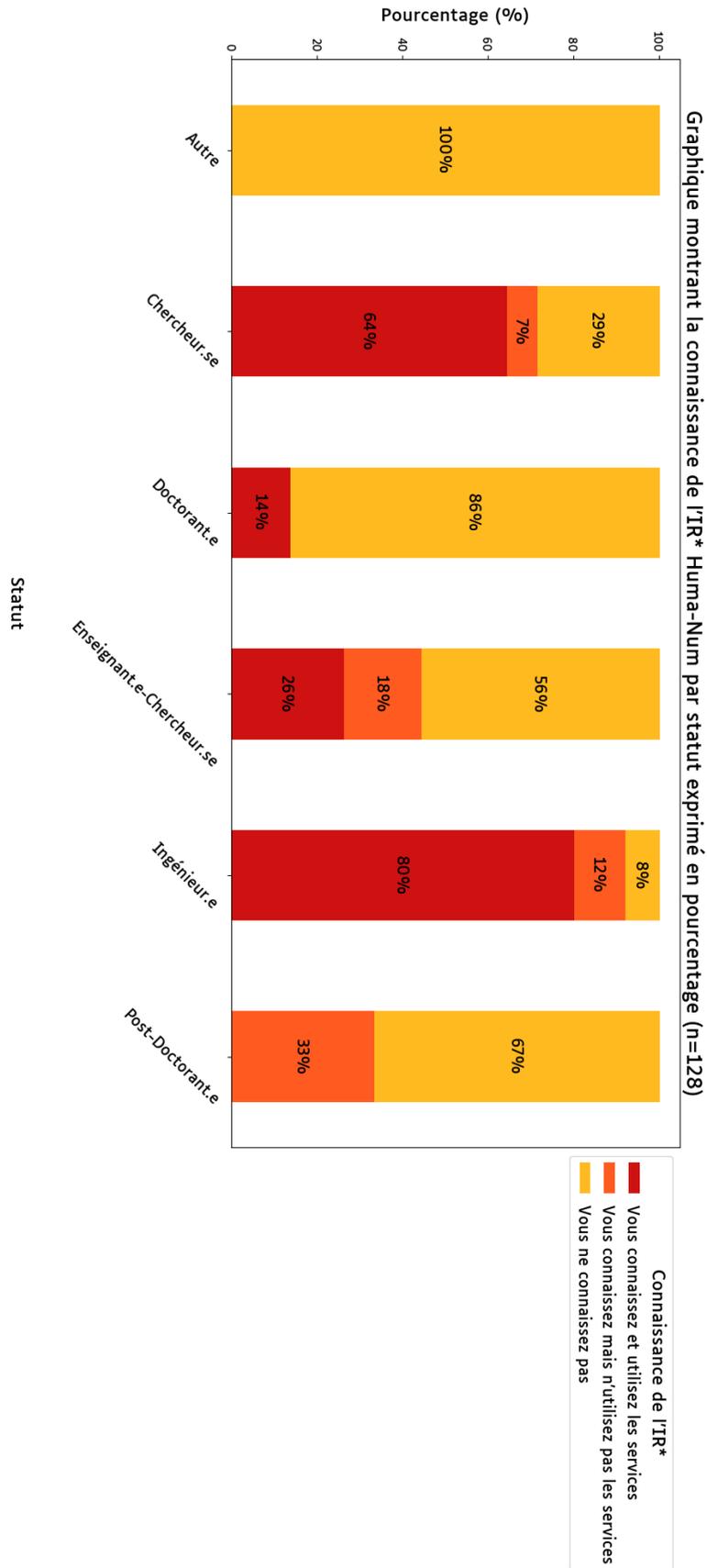


Figure 15 Graphique montrant la connaissance de l'IR* Huma-Num, selon le statut, exprimés en pourcentage

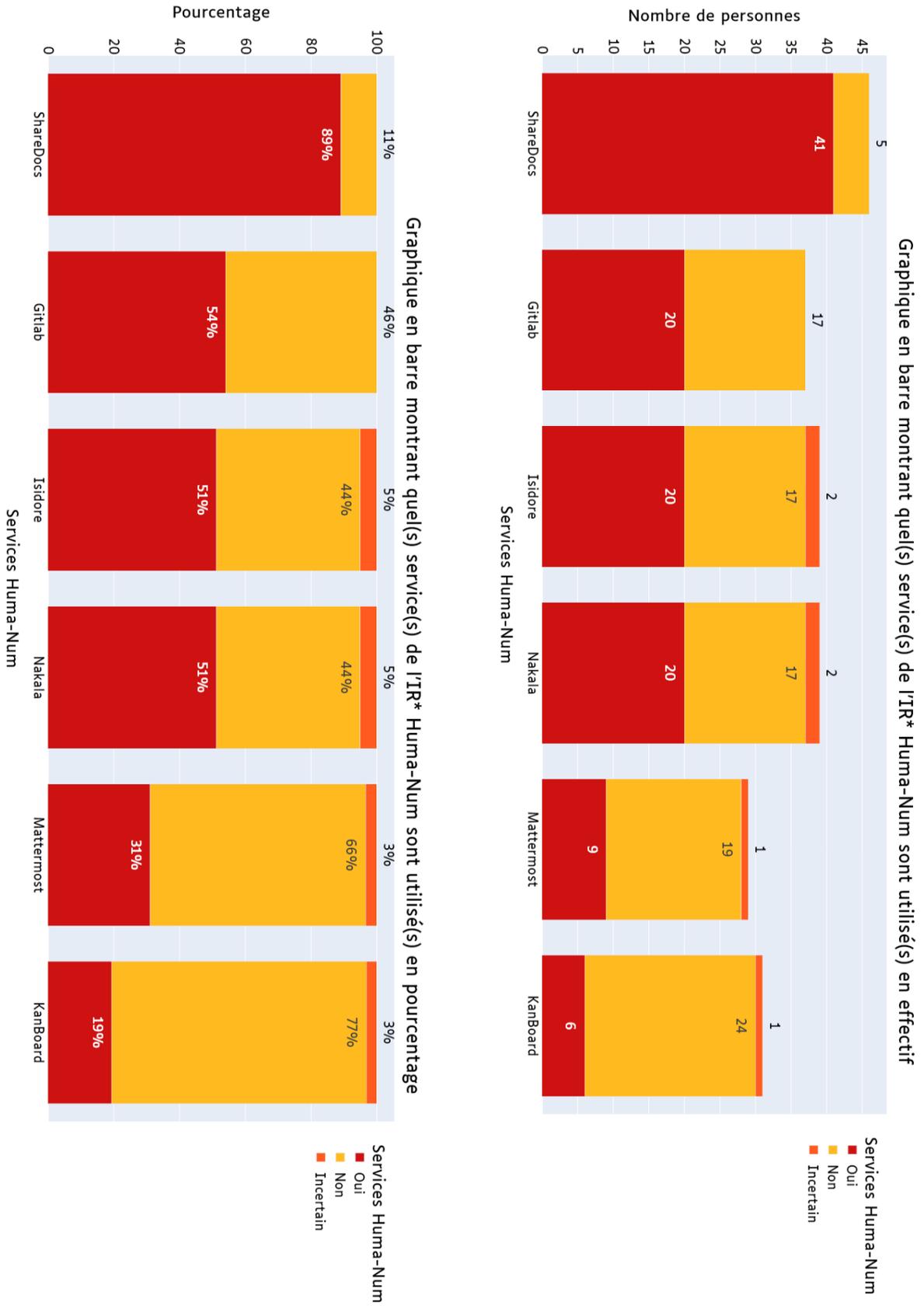


Figure 16 Graphique montrant quel(s) service(s) de l'IR* Huma-Num sont utilisé(s) en effectif et en pourcentage

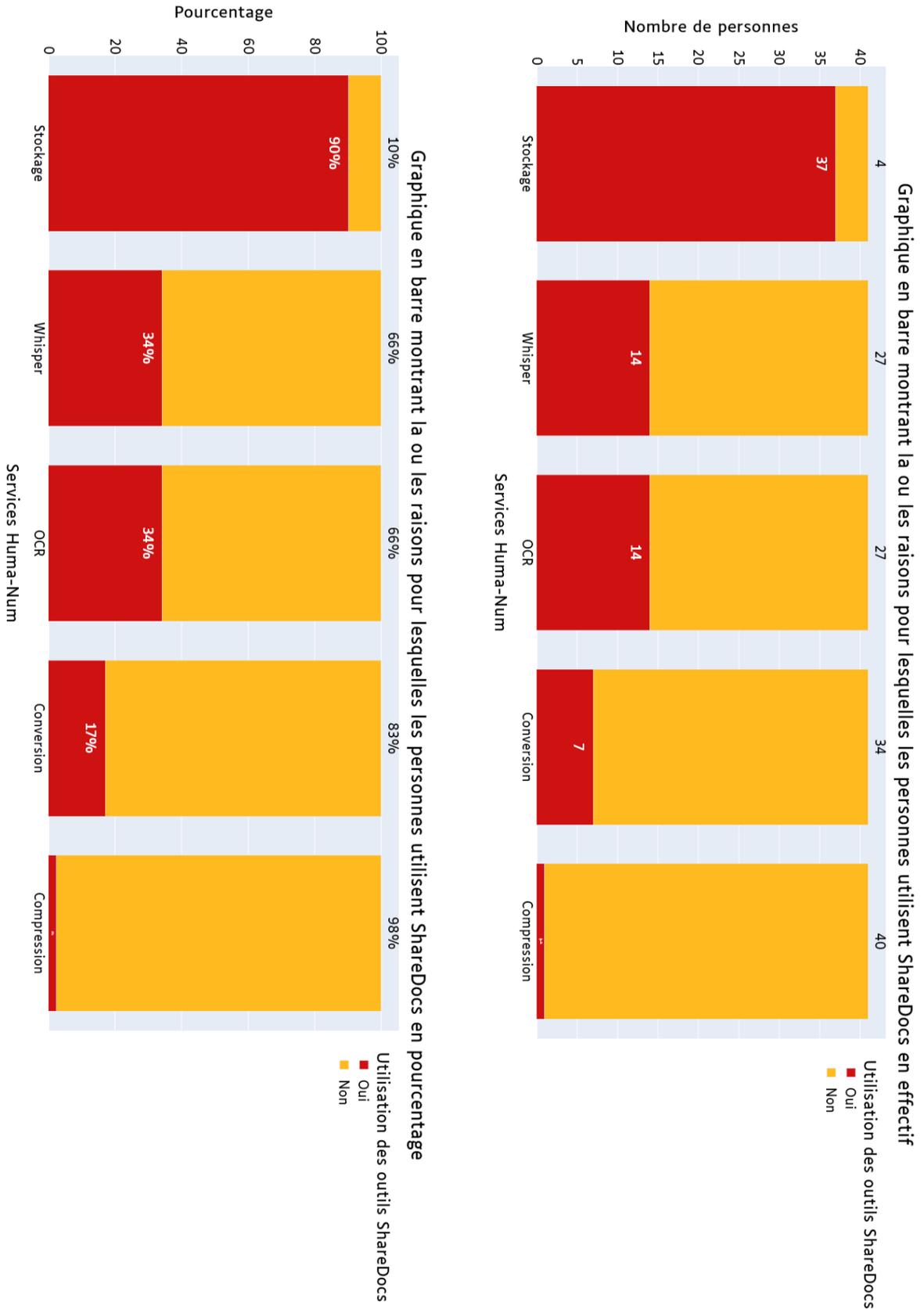


Figure 17 Graphique montrant la ou les raisons pour lesquelles les personnes utilisent ShareDocs en effectif et en pourcentage

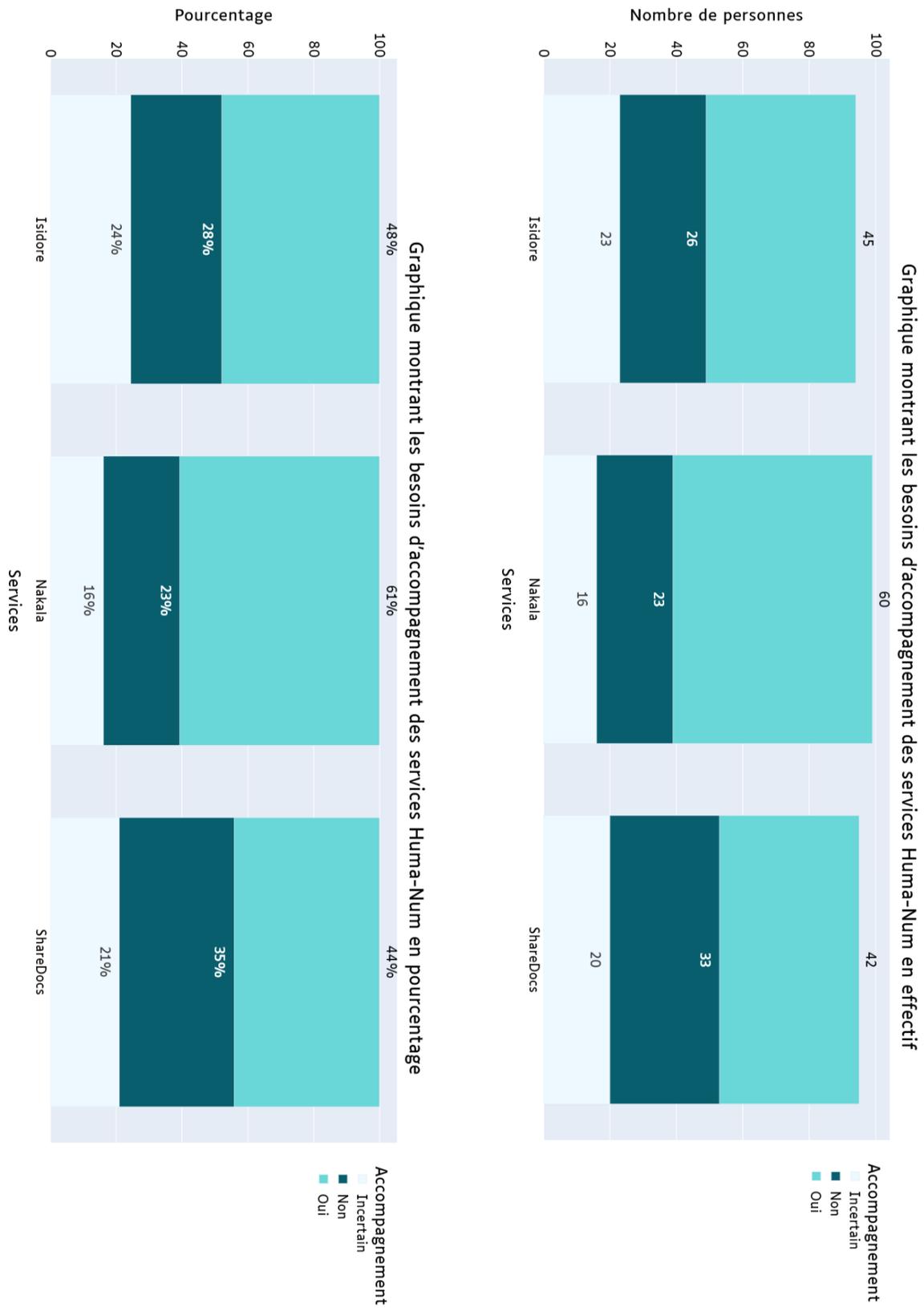


Figure 18 Graphiques montrant les besoins d'accompagnement des services exprimés en effectif et en pourcentage

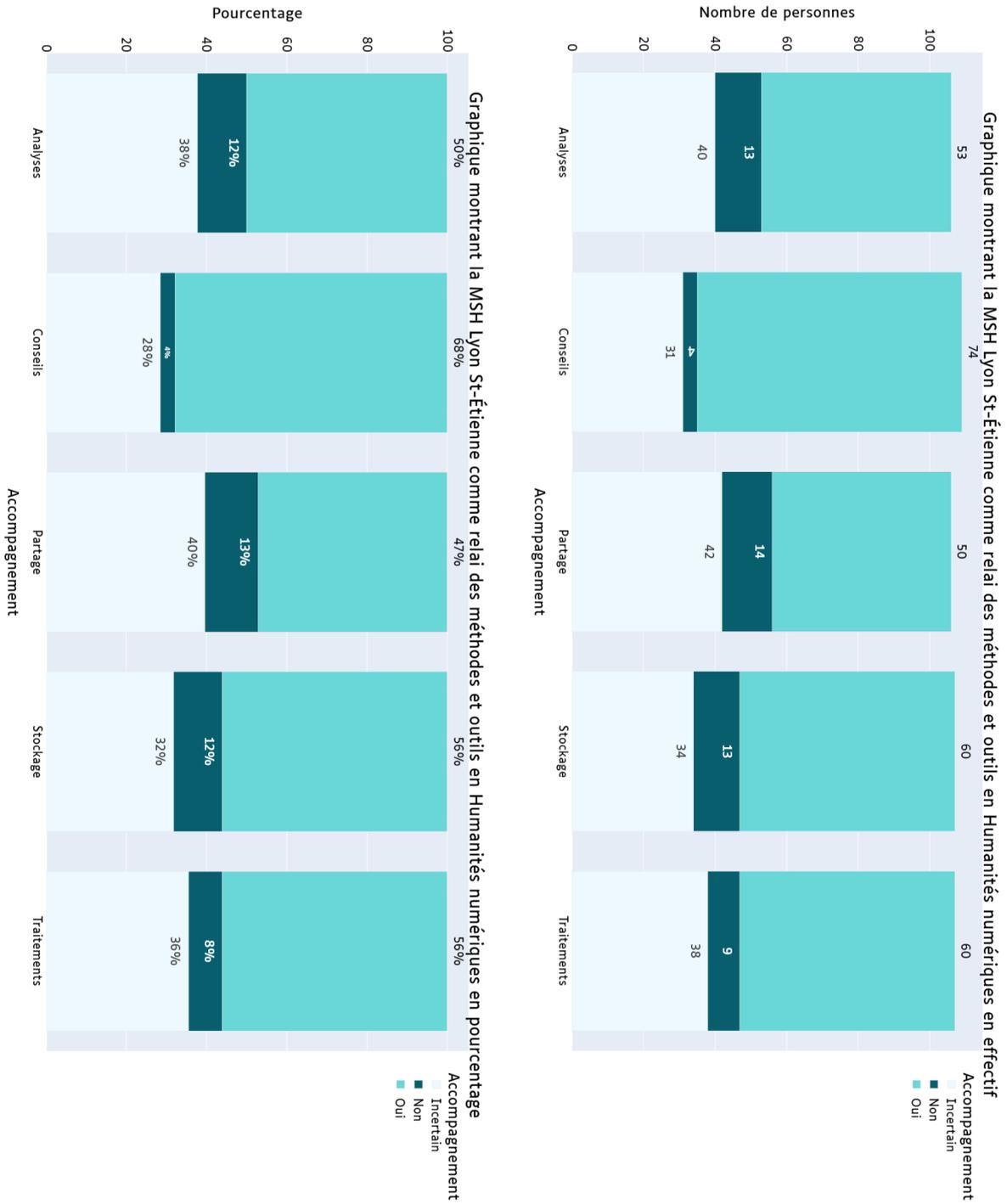
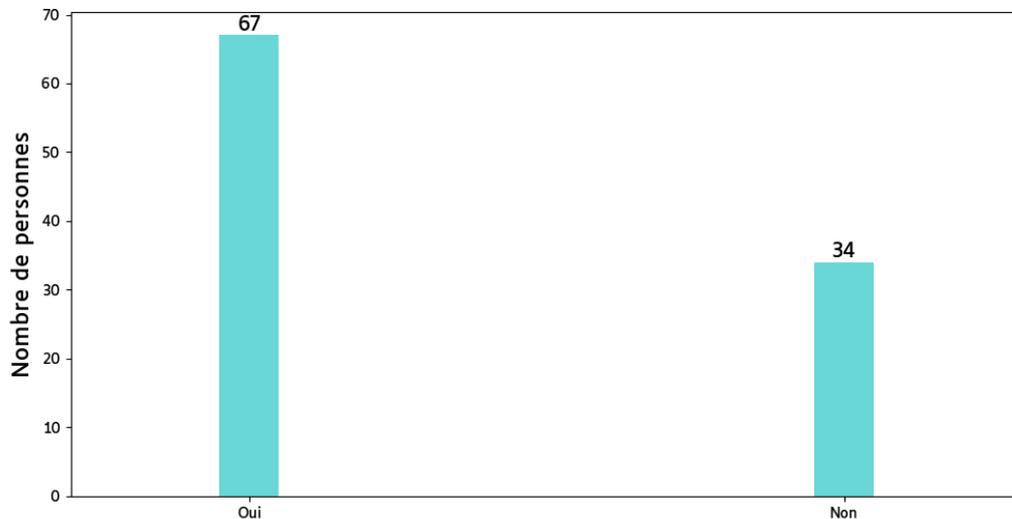


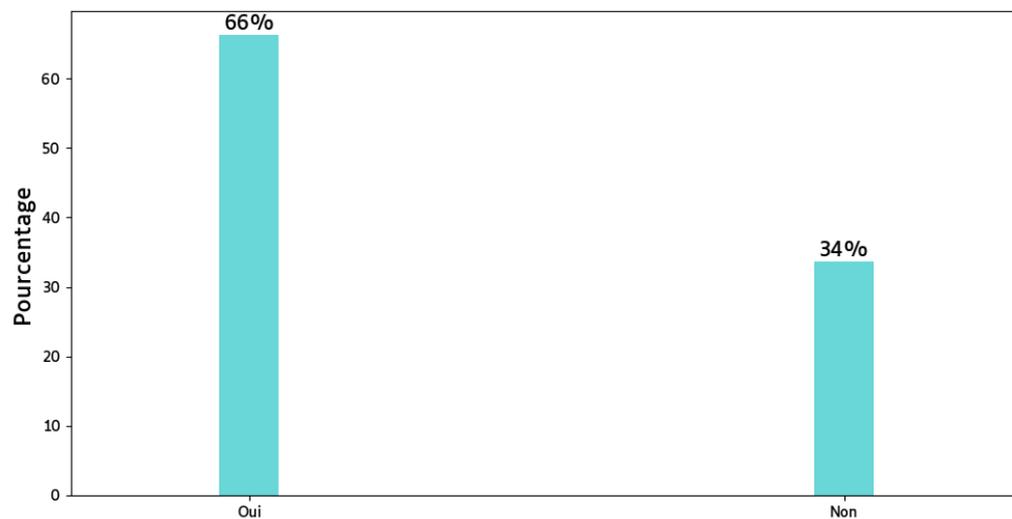
Figure 19 Graphiques montrant la perception de la MSH LSE comme relai des méthodes et des outils en Humanités numériques exprimés en effectif et en pourcentage

Souhaitez-vous mettre à disposition vos propres jeux de données de recherche
dans le cadre de nouvelles recherches ? En effectif (n=101)



Modalité

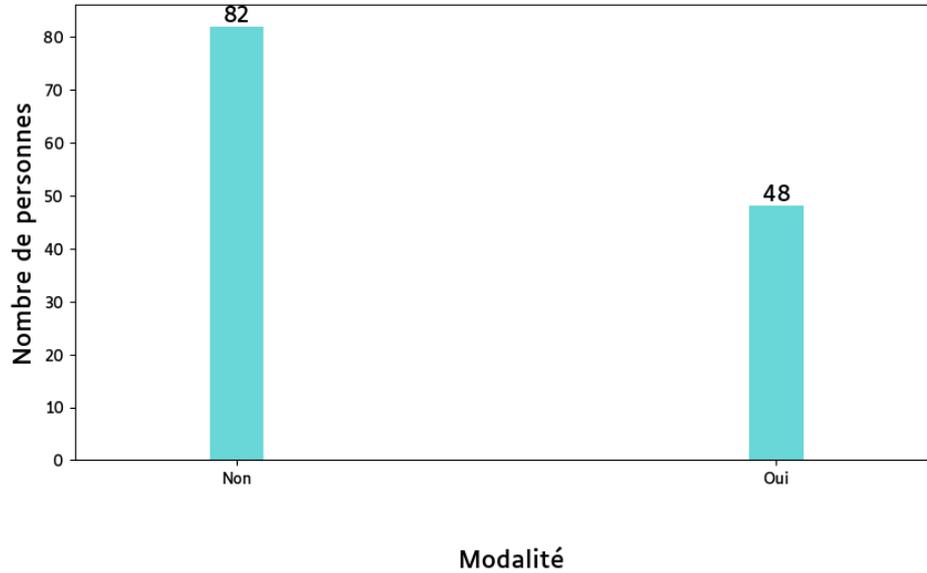
Souhaitez-vous mettre à disposition vos propres jeux de données de recherche
dans le cadre de nouvelles recherches ? En pourcentage (n=101)



Modalité

Figure 20 Graphiques révélant le souhait des enquêtés de mettre à disposition leurs propres jeux de données de recherche exprimés en effectif et en pourcentage

Graphique montrant si les personnes interrogées ont déjà utilisé des jeux de données mises à disposition en effectif (n=130)



Graphique montrant si les personnes interrogées ont déjà utilisé des jeux de données

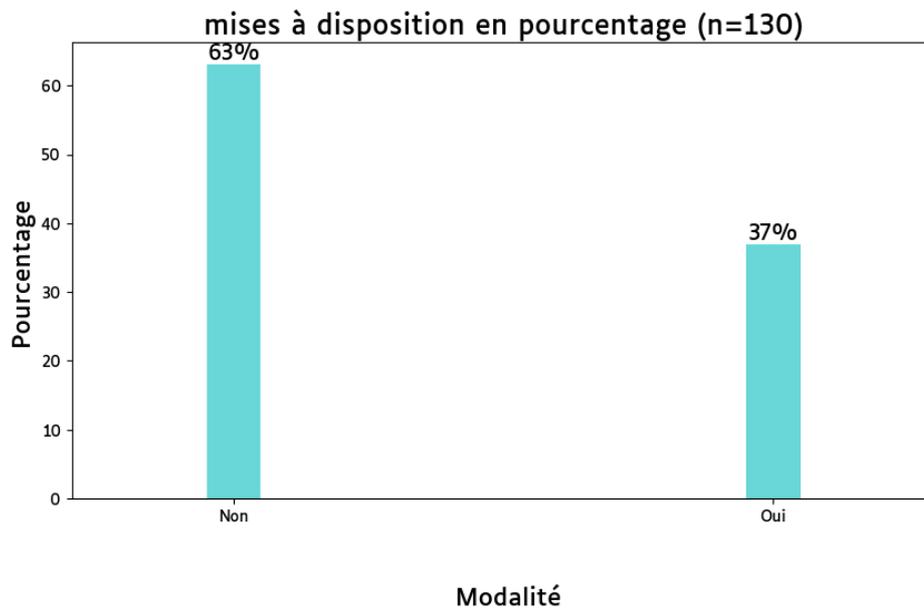


Figure 21 Graphiques représentant si les personnes interrogées ont déjà utilisé des jeux de données mises à disposition exprimés en effectif et en pourcentage

Graphe de réseaux

Légende:

- Projet financé en Humanités numériques
- Unité associée à la MSH LSE
- Unité pas associée à la MSH LSE

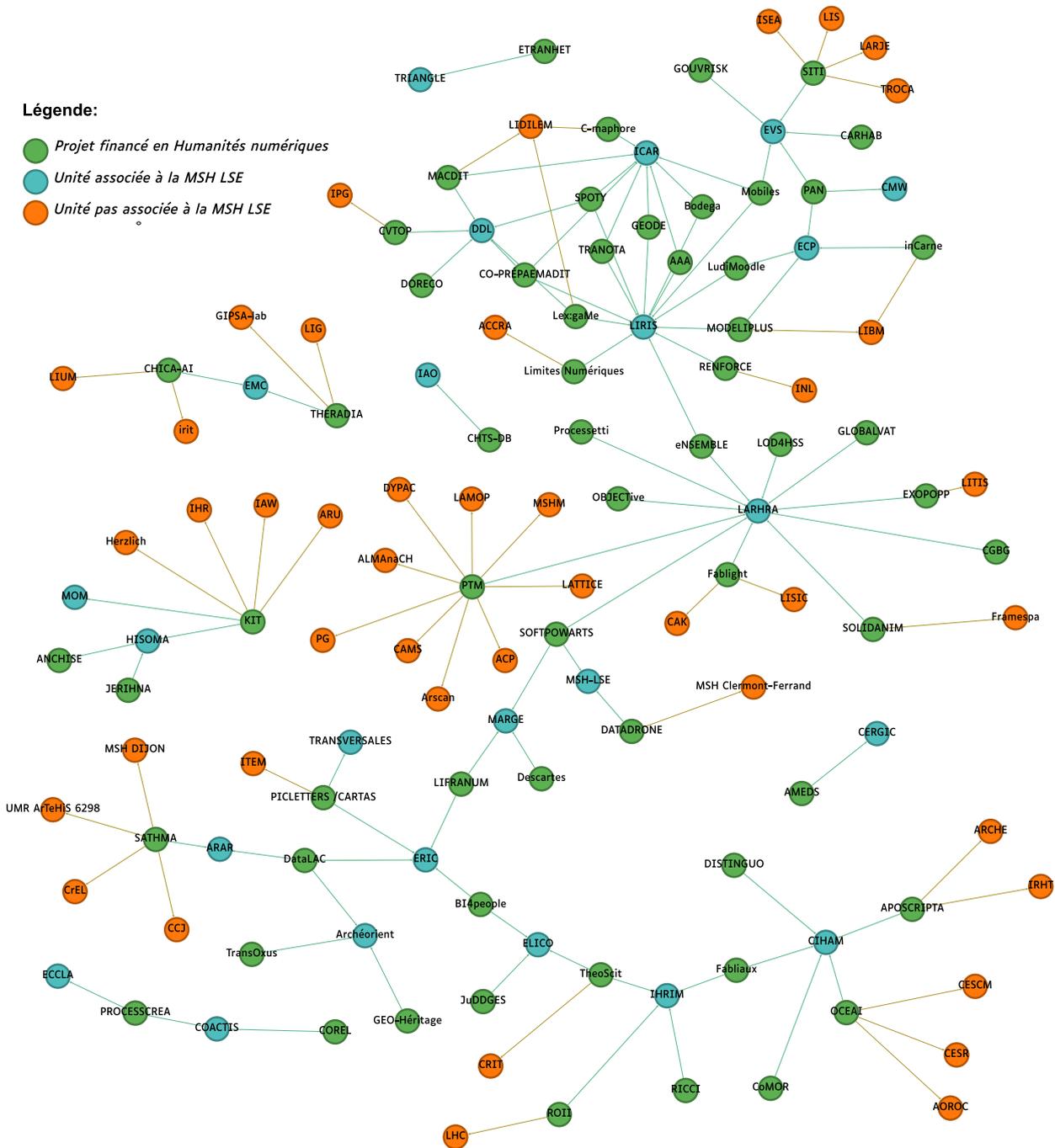


Figure 22 Graphe de réseaux des projets en Humanités numériques financés et leurs laboratoires associés

Typologie



Figure 23 Typologies des projets en Humanités numériques

Questionnaire

Nous avons essayé ici de retranscrire le plus fidèlement possible le questionnaire effectué sur LimeSurvey. Mais certains éléments tels que des menus déroulants ou des pop-ups n'ont pas été retranscrits. De même que certaines questions ne s'affichaient qu'à la condition qu'une question précédente ait reçu une réponse, ce que nous n'avons pas pu réaliser dans ce rapport. Les questions obligatoires sont indiquées par un « * ».

Questionnaire Humanités numériques ~ MSH Lyon Saint-Étienne ~ IR* Huma-Num

Ce questionnaire s'inscrit le cadre d'un projet financé par l'IR* Huma-Num. Son périmètre est celui de la MSH Lyon Saint-Étienne et des laboratoires qui lui sont associés.

Les enjeux de cette enquête sont triples :

- Mieux connaître les **pratiques**, les **besoins** et les **questionnements** dans les projets en Humanités numériques
- Améliorer et développer l'**interface** entre la MSH Lyon Saint-Étienne et l'IR* Huma-Num (contact, services proposés, etc.)
- Renforcer l'**accompagnement** de proximité/local sur le territoire

Cette enquête concerne les porteurs des projets en Humanités numériques mais également tous les personnels impliqués dans ces projets sur le territoire. Elle s'adresse aux **ingénieur·es**, aux **chercheurs·euses**, aux **enseignant·es-chercheurs·euses** ainsi qu'aux **doctorants·es**.

Le questionnaire est structuré de la façon suivante : votre **rapport à l'IR* Huma-num** et à la **MSH Lyon Saint-Étienne** dans vos projets / **collecte**

/archivage/réutilisation / accompagnement / questions socio-professionnelles

La durée moyenne de ce questionnaire est de **10 mn**.

Il y a 34 questions dans ce questionnaire.

Votre rapport à l'IR* Huma-Num et à la MSH Lyon Saint-Étienne dans vos projets

Cette première partie du questionnaire a pour but de connaître votre participation à des projets de recherche en Humanités numériques.

Et dans le même temps de réunir des informations sur votre rapport à l'IR* Huma-Num et ses services ainsi que ceux de la MSH Lyon Saint-Étienne au sein de vos projets.

Entre 2020 et maintenant, avez-vous participé à un ou des projets scientifiques en Humanités numériques ? *

Qu'est-ce qu'un projet scientifique en Humanités numériques ?

Selon notre définition dans le cadre de l'enquête, les projets de recherche en Humanités numériques sont ceux qui ont une interaction entre les Sciences informatiques et au moins un domaine des Sciences humaines et sociales. Cette interaction peut se faire sur l'ensemble du projet ou sur une partie.

Cela se traduit par des **outils** et **méthodologies** qui peuvent être développés ou qui existent déjà et qui vont permettre d'appuyer les SHS dans leurs thématiques de recherche. Cela peut être à l'inverse des outils qui n'existent pas mais dont il y a un besoin dans un champ des Sciences humaines et sociales.

Il est possible que cela fasse intervenir une collaboration entre un ou plusieurs laboratoire(s) en SHS et un ou plusieurs laboratoire(s) en

informatique, mais ce n'est pas une condition obligatoire. Néanmoins, il faut **au moins un laboratoire associé à la MSH Lyon Saint-Étienne.**

Veillez sélectionner une seule des propositions suivantes :

- Oui
- Non
- Vous ne savez pas

Dans combien de projets en Humanités numériques avez-vous été impliqué entre 2020 et maintenant ?

Veillez sélectionner une réponse ci-dessous.

Veillez sélectionner une seule des propositions suivantes :

- 1
- 2-3
- 4-5
- + de 5

Connaissez-vous l'infrastructure de recherche « étoile » Huma-Num ?

Veillez sélectionner une réponse ci-dessous.

Veillez sélectionner une seule des propositions suivantes :

- Vous connaissez et utilisez les services
- Vous connaissez mais n'utilisez pas les services
- Vous ne connaissez pas

Selon vous, quelles sont les principales fonctions/missions d'Huma-Num ?

Veillez écrire votre réponse ici :

Quel(s) service(s) d'Huma-Num utilisez-vous ?

Choisissez la réponse appropriée pour chaque élément :

	Oui	Incertain	Non
	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Autre(s) service(s) :	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Quelle(s) utilisation(s) de ShareDocs faites-vous ?

Cochez tout ce qui s'applique.

Veuillez choisir toutes les réponses qui conviennent :

- Stockage de documents
- OCR (reconnaissance optique de caractères)
- Whisper (transcription automatique d'audio)
- Compression des PDF
- Conversion de fichiers Audio / Vidéo

Quel(s) autre(s) service(s) hébergé(s) par Huma-Num utilisez-vous ?

Veuillez écrire votre réponse ici :

Vous pouvez les lister en les séparant par des virgules ou des points-virgules.

D'après vous, la MSH Lyon Saint-Étienne est-elle un relai des méthodes et outils en Humanités numériques ?

Choisissez la réponse appropriée pour chaque élément :

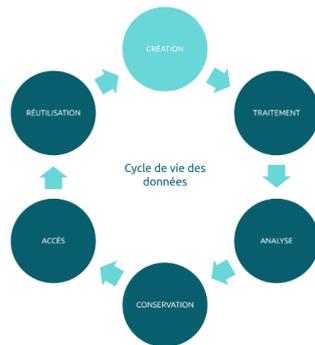
	Oui	Incertain	Non
Conseils méthodologiques	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Traitements des données	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Analyses des données	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Stockage des données	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Partage des données	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Vos pratiques individuelles relatives au cycle de vie des données (collecte, stockage, etc.) diffèrent-elles des pratiques édictées par votre laboratoire ?

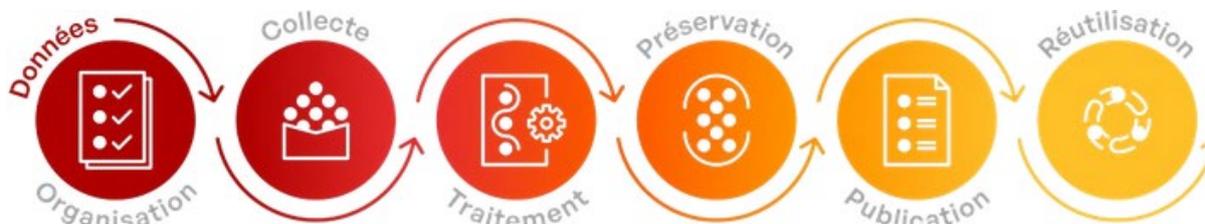
Qu'est-ce que le cycle de vie des données ?

Le cycle de vie des données est l'ensemble des étapes de gestion, de conservation et de diffusion des données de recherche, associées aux activités de recherche.

Il est le plus souvent représenté en 5 ou 6 étapes :



Ce cycle de vie des données est réemployé dans les étapes des services Huma-Num :



Sources :

Inist - CNRS, 2014, Le cycle de vie des données in *Une introduction à la gestion et au partage des données de la recherche*, https://www.inist.fr/wp-content/uploads/donnees/co/module_Donnees_recherche_7.html, consulté le 06 juin 2024.

Huma-Num, s.d., Les services par étapes in *Des services pour les données*, <https://www.huma-num.fr/les-services-par-etapes/>, consulté le 06 juin 2024.

Veillez sélectionner une réponse ci-dessous.

Veillez sélectionner une seule des propositions suivantes :

- Oui
- Non
- Vous ne savez pas

Si oui, précisez les différences entre vos pratiques et celle du laboratoire en quelques mots.

Veillez écrire votre réponse ici :

Collecte

L'étape de collecte des données regroupe la collecte de données existantes et la création de données nouvelles.

Quels types de données manipulez-vous ?

- Sources Textuelles/Porteuses d'inscriptions
- Quantitatives
- Image
- Son
- Vidéo
- Autre :

Pour structurer vos données, créez-vous des bases de données ?

Veillez sélectionner une seule des propositions suivantes :

- Oui
- Non

Si oui, quel(s) outil(s)/logiciel(s) utilisez-vous pour créer votre base de données ? *

Veillez écrire votre réponse ici :

Traitement

Les services de traitement répondent aux besoins de transformation et d'analyses de données : extraction et annotation de données sérielles ou textuelles, calcul, ou annotation des données multimédia, audio/vidéo, 3D, cartographie, visualisation, encodage, puissance de calcul, etc.

Pour le traitement de vos données, utilisez-vous les services Huma-Num ?

Veillez sélectionner une seule des propositions suivantes :

- Oui
- Non

Utilisez-vous également d'autres outils de traitement de données que ceux proposés par Huma-Num ?

Veillez sélectionner une seule des propositions suivantes :

- Oui
- Non

Si oui, quel(s) outil(s) de traitement utilisez-vous ?

Veillez écrire votre réponse ici :

Archivage

L'archivage des données numériques regroupe un ensemble de bonnes pratiques et d'outils complexes à articuler. On peut présenter cette complexité par les différentes dimensions qui entrent en œuvre dans la préservation des données.

Est-ce que vous-même ou votre laboratoire avez des pratiques de conservation des données sur le long terme ?

Veillez sélectionner une seule des propositions suivantes :

- Oui
- Non

Si oui, comment ?

Veillez écrire votre réponse ici :

Réutilisation des données

Les données peuvent être réutilisées pour la validation du modèle ou de l'expérience.

Elles peuvent également servir à d'autres travaux scientifiques permettant de faire avancer ou tester de nouvelles hypothèses.

Source : Nathalie Lac, 2023, Le cycle de vie de la données, <https://www.universite-paris-saclay.fr/recherche/science-ouverte/le-cycle-de-vie-des-donnees>, consulté le 06 juin 2024.

Souhaitez-vous mettre à disposition vos propres jeux de données de recherche dans le cadre de nouvelles recherches ?

Veillez sélectionner une seule des propositions suivantes :

- Oui
- Non

Si oui, quelles sont vos motivations ?

Veillez écrire votre réponse ici :

Si non, quels sont vos freins pour le faire ?

Veillez écrire votre réponse ici :

Avez-vous déjà utilisé des jeux de données mises à disposition ?

Tel que BnF Data, Progedo, Recherche Data Gouv

- Oui
- Non

Combien de données Nakala avez-vous mises en ligne ?

Le terme « donnée » dans Nakala désigne l'association d'un ou plusieurs fichiers et d'un ensemble d'informations les décrivant (les métadonnées).

Veillez sélectionner une seule des propositions suivantes :

- 1
- 2
- 3
- 4
- 5 et +

Avez-vous déjà réutilisé ou cité des données déposées sur Nakala ?

Veillez sélectionner une seule des propositions suivantes :

- Oui
- Non

Avez-vous déjà déposé sur un autre entrepôt de données que Nakala ?

Exemple : Ortholang

Veillez sélectionner une seule des propositions suivantes :

- Oui
- Non

Quel(s) entrepôt(s), à l'exception de Nakala, utilisez-vous ?

Veillez écrire votre réponse ici :

Accompagnement

Avec cette question, nous cherchons à améliorer nos pratiques d'accompagnement des services de l'IR* Huma-Num en identifiant vos besoins.

Est-ce que vous auriez besoin d'un accompagnement supplémentaire sur les services Huma-Num ?

	Oui	Incertain	Non
	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Autre(s) service(s) :	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Pour quel(s) autre(s) service(s) Huma-Num souhaiteriez-vous avoir un plus grand accompagnement ?

Exemple : Mattermost, Kanboard, Gitlab, Heurist, etc.

Veillez écrire votre réponse ici :

Questions socio-professionnelles

Quelle discipline étudiez-vous principalement ? *

Cela peut par exemple être la discipline dans laquelle vous avez fait votre thèse ou votre HDR.

Veillez sélectionner une seule des propositions suivantes :

- Anthropologie
- Archéologie
- Archivistique
- Arts
- Biologie
- Business intelligence
- Didactique
- Doctrine
- Droit
- Économie
- Environnement
- Géographie
- Géomatique
- Gestion
- Histoire
- Histoire de l'art
- Histoire du livre
- Informatique
- Langues
- Linguistique

- Littérature
- Médecine
- Musicologie
- Neurosciences
- Philosophie
- Psychologie
- Science politique
- Sciences cognitives
- Sciences de l'éducation
- Sciences de l'information
- Sciences de l'information et de la communication
- Sociolinguistique
- Sociologie
- Urbanisme
- Autre :

Quel est votre laboratoire d'accueil ou de rattachement actuel ? *

Veillez sélectionner une seule des propositions suivantes :

- ARAR
- Archéorient
- CEL
- CERCOR
- CERCRID
- CERGIC
- CERLA
- CGN
- CIHAM
- CLHDPP
- CMW
- COACTIS
- CRDMS
- CRPPC
- DDL
- DIPHE
- ECCLA

- ECP
- EDIEC
- EDPL
- ELICO
- ELJ
- EMC
- ERIC
- EVS
- GATE
- GREPS
- HISOMA
- IAO
- ICAR
- IETT
- IHRIM
- IRAA
- IRPHIL
- ISERL
- LADEC
- LAET
- LARHRA
- LCE
- LER
- LIRIS
- LVIS
- MAELYSE
- MAGELLAN
- MARGE
- MOM
- P2S
- PASSAGES XX-XXI
- RADIATIONS
- RESHAPE
- S2HEP
- SAF

- TRANSVERSALES
- TRIANGLE
- Autre :

Quel est votre statut ?

Veillez sélectionner une seule des propositions suivantes :

Ingénieur·e

Chercheur·se

Enseignant·e-Chercheur·se

Post-Doctorant·e

Doctorant·e

Autre :

Sur quel site êtes-vous situé·e ? *

Veillez sélectionner une seule des propositions suivantes :

Lyon

Saint-Étienne

Autre

Avez-vous des remarques ou des commentaires à nous faire remonter ?

Veillez écrire votre réponse ici :

Merci d'avoir répondu à ce questionnaire. Si vous voulez des informations supplémentaires ou discuter davantage sur l'enquête, n'hésitez pas à me contacter à l'adresse suivante :

sylvain.besson@MSH.fr

15/07/2024 – 08:33

Envoyer votre questionnaire.

Merci d'avoir complété ce questionnaire



Novembre 2024

