

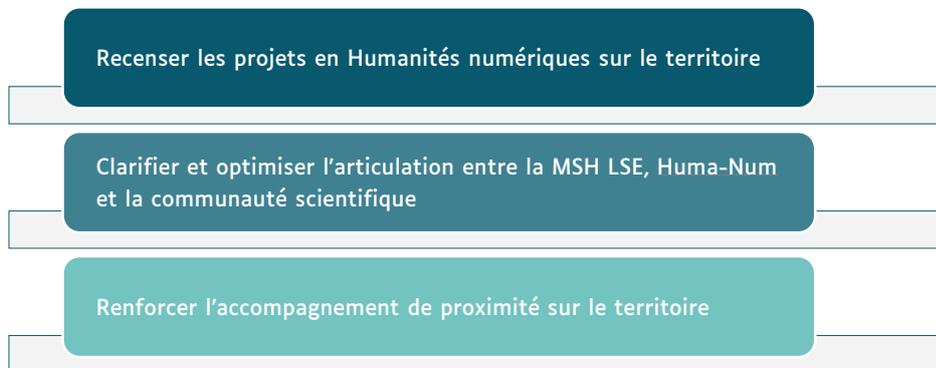
Synthèse : enquête sur les projets en Humanités numériques sur le site Lyon Saint-Étienne

Contexte

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, et sollicité plusieurs de ses agents : Orline Poulat, Sonia Guérin-Hamdi et Alexandra Dugué.

Objectifs de l'enquête

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. Son intention était triple :



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 à l'instar des ANR, ou locaux tels les appels à projets (AAP) de la MSH ou les projets financés par les établissements.

Nous nous focalisons sur les projets **en cours ou qui ont été clôturés récemment (entre 2022 et 2024)**. 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 et donc du site académique Lyon Saint-Etienne.

Quant à la cible de cette enquête, c'est-à-dire la population étudiée, elle vise les **personnels des 53 laboratoires associés à la MSH**, qu'ils soient chercheur-euse-s, enseignant-es-chercheur-es, (post)doctorant-es ou encore ingénieur-es.

De plus, nous avons pris en compte la manière dont cette population **se positionne vis-à-vis des Humanités numériques, des services de l'IR* Huma-Num et du pôle Humanités numériques de la MSH**, comme nous pouvons le synthétiser sur ce diagramme de Venn :

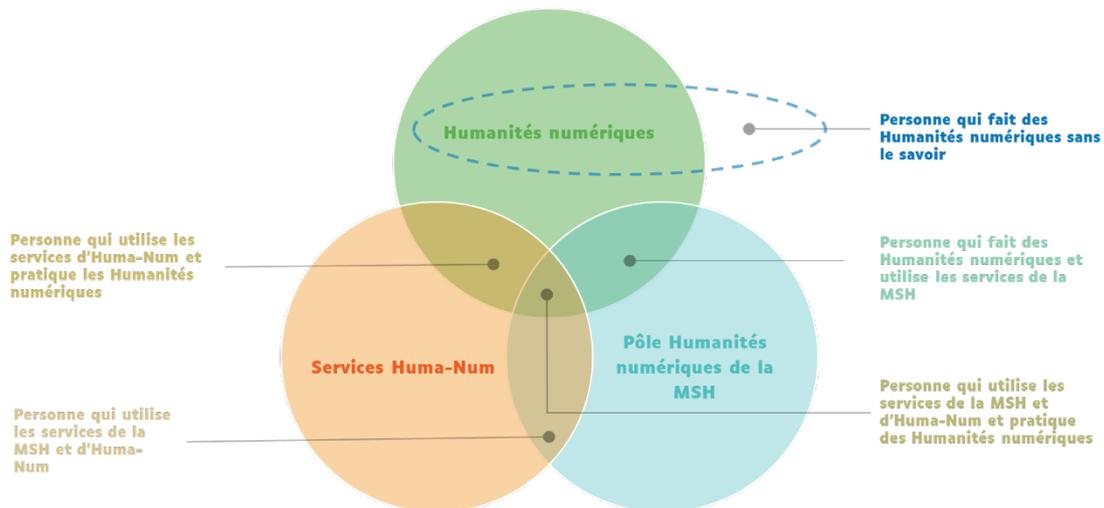


Figure 1 Diagramme de venn sur la population visée

Méthodologie de l'enquête

Cette enquête a été pensée en **quatre 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. Une vingtaine d'entretiens ont été effectués.

La première étape a consisté à créer une **base de données** des projets en Humanités numériques correspondant aux critères que nous avons énoncés ci-dessus. Elle intègre différentes informations sur les projets tels que la description du projet, le type de financement et les laboratoires et institutions associés. Une soixantaine de projets ont été recensés. Cette base de données a pu être enrichie de diverses manières : par la conduite d'entretiens avec des directeur-trices de laboratoires, 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 sur les projets dans lesquels ils sont impliqués.

La seconde étape a consisté à 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 et la sollicitation ou non d'un accompagnement de la recherche. Nous avons effectué deux groupes de discussion, avec deux et trois personnes chacun.

Enfin, la dernière étape avait pour but d'élargir l'enquête et le panel de participant-es par l'envoi d'un **questionnaire anonyme à l'ensemble de la communauté des laboratoires associés à la MSH LSE**. Bien que les thématiques abordées soient similaires à celles de la

deuxième étape, le questionnaire ne concernait 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 des *focus groups* et d'avoir une approche plus quantitative. De plus, cela nous a permis de les comparer aux résultats obtenus auparavant.

Principaux résultats de l'enquête

Profils des répondants au questionnaire

Le questionnaire a recueilli 141 réponses complètes. Tout d'abord, nous avons posé une question relative aux statuts des personnes interrogées qui montre une prépondérance des EC, puis un nombre important d'ingénieurs, de chercheurs et de doctorants impliqués :



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

Trois laboratoires concentrent 30 % des réponses : IHRIM (14 %), EVS (8,8 %) et LARHRA (8 %). Nous pouvons souligner que sur les 53 laboratoires associés à la MSH, la majorité d'entre eux (34 laboratoires) ont répondu à l'enquête. 65 % des unités sont donc représentées dans les réponses obtenues. Un constat similaire est fait pour les disciplines les plus représentées : Histoire (13 %) et Littérature (12 %).

Nous pouvons également observer le profil des répondants sous l'angle des types de données exploitées :

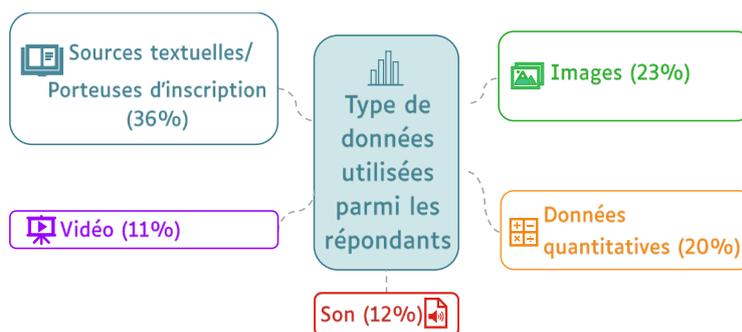


Figure 4 Profil des répondants selon le type de données

Comme nous pouvons le voir les types de données les plus utilisés parmi les répondants sont les sources textuelles (36%) et dans une moindre mesure les images (23%) et les données quantitatives (20%).

La connaissance du profil des répondants nous a permis de mieux cibler quelles sous-populations ont le plus besoin d'être davantage acculturées aux services proposés par Huma-Num ainsi que d'être mieux accompagnées par les pôles de compétences de la MSH.

Méconnaissance des services Huma-Num

Si nous regardons les résultats du questionnaire à la question «Connaissez-vous l'infrastructure de recherche "étoile" Huma-Num?», 37 % des répondants 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 et identifiée. Un **travail d'acculturation** à cette dernière semble en conséquence nécessaire.

D'autre part, si on analyse plus précisément les résultats de cette question, en s'intéressant aux statuts des répondants, on remarque que les disparités ne sont pas les mêmes. En effet, 80 % sont des ingénieur-es qui indiquent connaître et utiliser les services d'Huma-Num contre seulement 14 % des doctorant-es.

Nous avons également constaté que les réponses à la question relative aux principales fonctions de l'IR* sont d'une grande diversité comme l'illustrent les verbatims suivants :

«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). [...];
Archiver – diffuser ;
Intelligence artificielle de retranscription, océrisation, etc.»

Bien qu'une méconnaissance des services de l'IR* Huma-Num soit présente, il semble également que la présence de l'IR* est jugée indispensable par les personnes qui l'utilisent, malgré certaines difficultés rencontrées par certain-es. C'est ce que nous allons voir dorénavant.

Huma-Num : des services centralisés, 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. Les différents échanges durant les entretiens et les *focus groups* ont fait apparaître 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**. 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 d'accompagnement lors de la première utilisation et son ergonomie** rendent parfois l'utilisation de l'outil laborieuse.

Cependant, de nombreux éléments indiquent la nécessité d'avoir une infrastructure « étoile » telle qu'Huma-Num :

«Souvent quand il y a une DFG 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 ? », **90 % des répondants qui utilisent les services Huma-Num citent cet outil**. Une majorité de personnes indique également utiliser Nakala (51 %), Isidore (51 %) et Gitlab (54 %).

Nous avons également demandé si des besoins d'accompagnement paraissent nécessaires sur les trois services principaux : ShareDocs, Nakala, Isidore. 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** d'une part et de dispenser d'autre part une **formation sur Isidore**, réalisée à la MSH en septembre dernier, avec 19 participants.

Ces besoins d'accompagnement sur les services de l'IR* peuvent se recouper avec les pôles de compétences de la MSH. A ce titre, nous avons demandé dans notre questionnaire la manière dont les répondants perçoivent la MSH dans son rôle de relai des méthodes et des outils en Humanités numériques. C'est ce que nous allons voir maintenant

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 en tant que spécialiste des méthodes et des outils en Humanités numériques. Nous pouvons remarquer que pour les différents types d'accompagnement, une majorité des répondants considère la MSH comme un bon relai des méthodes et outils en Humanités numériques. C'est d'autant plus le cas pour son rôle de conseil où 68% des personnes interrogées indiquent que la **fonction de conseil est bien assurée par la MSH**. A l'inverse, pour l'aspect du partage, moins de la moitié des répondants indiquent que la MSH assure ce rôle. Cela montre l'importance pour la MSH d'apporter son **soutien sur le dépôt dans les entrepôts de données** tels que Nakala.

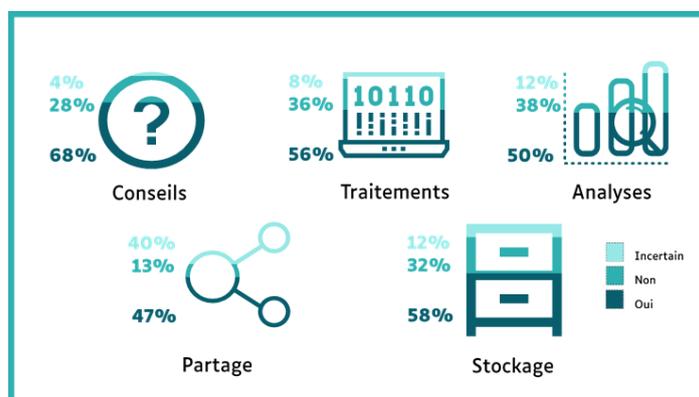


Figure 5 La MSH LSE en tant que relai des méthodes et outils des Humanités numériques exprimé en pourcentage

L'aspect de partage est d'autant plus important qu'il fait écho à un paradoxe que nous avons rencontré durant notre enquête entre la volonté d'ouverture des données et la réalité de cette ouverture dans les projets. Cela fait l'objet de la prochaine partie.

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

À la question « Souhaitez-vous mettre à disposition vos propres jeux de données de recherche dans le cadre de nouvelles recherches ? », **67 % des personnes ont indiqué vouloir partager leurs données.** » 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.

À la question relative à l'utilisation des jeux de données mis à disposition, 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 également observer qu'**une part non négligeable de personnes ne souhaitent pas partager leurs jeux de données** (34 %). On peut évoquer 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.

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

Plus d'une soixantaine de projets – ayant pour critère le fait d'être interdisciplinaires 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 le processus de sélection de tous les projets implique une dimension informatique. 24 laboratoires associés à la MSH sur les 53 sont représentés dans les projets recensés.

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 arrivent seulement 5^{ème} (6 %) après les LABEX (13 %), les APPI (10 %) et les ERC (8 %). Bien que les appels à projets de la MSH ne soient pas l'unique indicateur de la connaissance des projets par la MSH, cela montre qu'une grande part des projets ne sont pas connus de cette dernière.

À travers le recensement des projets, nous avons aussi mis en évidence certaines collaborations entre des laboratoires associés à la MSH, comme nous allons le voir dorénavant.

Collaboration entre les laboratoires associés

Nous avons identifié, 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 les plus centraux dans plusieurs projets. Le LIRIS (laboratoire informatique associé à la MSH *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 à travers plusieurs projets.

Les différents constats qui ont été évoqués précédemment nous amènent à réfléchir sur des

pistes d'amélioration à un niveau local et à un niveau national. C'est ce que nous allons voir maintenant.

Pistes d'amélioration proposées par la MSH

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 lorsque la personne est isolée dans un laboratoire, du fait de l'absence d'ingénieur·es d'études ou de recherche 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 cas présent la MSH Lyon Saint-Étienne - peuvent assurer le **rôle d'intermédiaire** entre les chercheur·euses des laboratoires et l'IR*. C'est déjà la vocation des **réfèrent·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 réfèrent·es**, au moins pour un premier niveau de demande. Cela permettrait à la MSH 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.

Cette territorialisation ne peut se faire sans une intensification des accompagnements sur les services proposés par l'IR*. C'est ce que nous verrons par la suite.

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

Compte tenu du besoin d'accompagnement sur les services d'Huma-Num, 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, en particulier les réfèrent·e.s 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* comme Mattermost et Kanboard comme pourraient faire l'objet d'une nouvelle vague de formations.

Au-delà de l'accompagnement qui peut être fait au niveau de la MSH, c'est aussi par l'interconnaissance entre les laboratoires associés à la MSH que la diffusion et le partage des compétences peut se faire. C'est ce que nous allons voir maintenant.

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, 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.

Cette mise en lien pourrait se faire de deux façons. Tout d'abord, en pérennisant la **base de données sur les projets en Humanités numériques** et en s'assurant de la modification et de la mise à jour des contenus par une personne référente. Cela peut également passer par la promotion de l'outil et son enrichissement au moyen d'un espace de collaboration. La base de données doit jouer un rôle de vecteur/médiateur entre les personnels du site quand certaines compétences ne sont pas présentes dans leur laboratoire de rattachement. Pour la MSH, un tel outil permettrait 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 pourrait être l'organisation 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, dans le cadre de sa labellisation nouvelle. Elles pourraient également être co-portées par l'axe scientifique Sociétés et Humanités numériques de la MSH.

Un autre levier pour accroître l'expertise de chacun-e est l'ouverture des données qui permet la réutilisation et la reproductibilité des méthodes. Nous verrons dans la partie suivante comment nous pouvons davantage accompagner sur l'ouverture des données.

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 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, 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.

Pour aller plus loin :

Sylvain Besson. *Enquête sur les projets en Humanités numériques*. MSH Lyon Saint-Étienne ; Huma-Num. 2024. ([hal-04886620](#))

Sylvain Besson. *Base de données sur les projets en Humanités numériques* [En ligne]. MSH Lyon Saint-Étienne ; 2024, (Huma-Num, disponible : https://heurist.huma-num.fr/heurist/?db=sylvain_enquete_HumaNum&website&id=396)

Sylvain Besson. *Script pour l'analyse du questionnaire lié à l'enquête sur les projets en Humanités numériques sur le site Lyon Saint-Étienne*. MSH Lyon Saint-Étienne ; Huma-Num. 2024. ([hal-04935698](#))

Sylvain Besson. *Jeux de données issus d'une enquête sur les projets en Humanités numériques*. NAKALA - <https://nakala.fr> (Huma-Num - CNRS), 2025. ([10.34847/NKL.1BEDV5UJ](#)).