

Huma-Num : l'humain en majesté numérique

La TGIR Huma-Num développe des réponses aux besoins nouveaux, apparus avec l'irruption du numérique, des différents acteurs de la recherche en sciences humaines et sociales en France. Elle offre un appui aux dispositifs français et, s'appuyant sur l'expertise réunie dans les consortiums, coordonne la production de corpus de sources. Survol de sa stratégie et de son fonctionnement.

Dans le domaine des sciences humaines et sociales (SHS), le tournant numérique de la recherche a provoqué une évolution considérable des pratiques de recherche. Aujourd'hui, les chercheurs produisent et utilisent collectivement d'importants volumes de données, qui nécessitent des outils spécialisés pour y accéder, les manipuler et les visualiser et d'importants moyens technologiques. La maîtrise de ce nouvel environnement, composé de sources hétérogènes – des bibliothèques numériques aux données numériques récentes, exclusivement traitées et diffusées en ligne – appelle des compétences en constante évolution, aussi réflexives que les changements sont rapides. Un tel écosystème ne peut plus être réalisé avec les moyens dont disposent les chercheurs, ni avec ceux d'un laboratoire, ni même avec ceux d'un programme de recherche collectif. Une véritable infrastructure, au sens élargi du terme, peut fournir et mutualiser ces différents services. La très grande infrastructure de recherche nationale (TGIR) Huma-Num est tournée vers les services pour la production et la réutilisation de données numériques, mais aussi vers l'accompagnement d'outils émergents et de bonnes pratiques issues des communautés scientifiques des SHS elles-mêmes. Elle est associée étroitement à la construction d'infrastructures européennes et internationales.

DES OUTILS POUR TOUT LE CYCLE DE VIE DES DONNÉES

Les services d'Huma-Num doivent fournir des outils tout au long du cycle de vie des données : de la production de données brutes jusqu'à leur préservation après traitement. Pour proposer des ressources bien dimensionnées et adaptées, Huma-Num n'a pas créé son propre centre de données, mais s'appuie sur les moyens de centres de calcul nationaux (centre de calcul de l'IN2P3-CNRS, CINES).

Un objectif sous-jacent consiste à amener les chercheurs à s'approprier les méthodes et services numériques au cœur des processus de la recherche. Les technologies d'interopérabilité, en particulier celles proposées par le Web sémantique, sont au centre de cette chaîne de traitement, permettant des échanges fluides entre les différentes « briques » qui la composent.



Huma-Num développe à la fois des services propres (Nakala, Nakalona, Isidore), tous interopérables, munis d'API et utilisant le Web sémantique, et met aussi à disposition des services forgés par les communautés SHS, principalement issus des consortiums et Maisons des sciences de l'homme. Ainsi, Nakala (qui veut dire « copie conforme » en langue swahili) permet de déposer des données, de les documenter à l'aide de métadonnées. Nakalona (qui veut dire « voir Nakala » en swahili) utilise le logiciel *open source* Omeka pour permettre à un chercheur d'éditorialiser les corpus de données sous la forme de bibliothèque numérique. L'ensemble de ces données, issues des corpus en SHS, qu'elles utilisent nos services ou qu'elles viennent d'ailleurs, peuvent être signalées dans notre moteur de recherche sémantique Isidore.

Un point important dans le fonctionnement de cet écosystème complet de services pour les données numériques en SHS est la sensibilisation à la pérenni-

↖ Cologne, grand pont sur le Rhin [fin du XIX^e siècle]. Photographie extraite du fonds du lycée Colbert (Paris), mise en ligne sur le site du labex Ehne, site hébergé par Huma-Num.

[1] European Research Infrastructure Consortium.

[2] Common Language Resources and Technology Infrastructure, www.clarin.eu

[3] Digital Research Infrastructure for the Arts and Humanities, www.dariah.eu

sation, au partage des données et à la normalisation par une concertation (disciplinaire et interdisciplinaire) entre chercheurs. Cette sensibilisation devrait être systématique lors d'une demande d'utilisation.

UNE ACTION EN LIEN AVEC LES CONSORTIUMS

Les consortiums réunissent des acteurs issus de différents projets de recherche autour de thématiques et d'objets communs coédités. Ils favorisent l'appropriation des dispositifs numériques par ces communautés et créent ainsi des synergies en lien avec les services d'Huma-Num. Leurs activités vont de la création d'applications web à la définition de schémas de métadonnées cogérés, en passant par la production de guides de bonnes pratiques et des formations spécifiques.

Actuellement, Huma-Num soutient les activités de dix consortiums, principalement disciplinaires. Il s'agit de : **Cahier** (Corpus d'auteurs pour les humanités : vise à fédérer les différentes initiatives françaises autour de corpus d'auteurs, qu'ils relèvent de la littérature, de la philosophie ou d'une thématique liée à une école ou à une pratique); **Archives des ethnologues** (conservation et pérennisation des données collectées sur le terrain ainsi que de tous documents pouvant aider à la contextualisation de ces données ou à une meilleure compréhension de l'élaboration des travaux anthropologiques); **Musica** (qui coordonne la réflexion épistémologique des chercheurs, ainsi que le développement de nouveaux outils de représentation et d'analyse de la musique); **Corli** (Corpus, langues et interaction, né du rapprochement de deux consortiums de linguistique, Corpus écrits et Corpus oraux et multimodaux); **3D** (qui se propose de fédérer les équipes de recherche travaillant sur l'usage des méthodologies et technologies 3D dans l'étude des sociétés humaines); **Masa** (Mémoires des archéologues et des sites archéologiques, qui réunit Musée d'archéologie nationale des acteurs nationaux déjà rassemblés dans le réseau des Maisons des sciences de l'homme traitant d'archéologie); **Cosme** (Consortium Sources médiévales, qui réunit des scientifiques de toutes les disciplines autour de

la recherche sur le passé médiéval); **Imageo** (qui numérise, archive et valorise les cartes et photographies par différents laboratoires du CNRS et la cartothèque de Bordeaux-Montaigne); **Archipolis** (Archives des sciences sociales du politique); **ArcMC** (Archives des mondes contemporains, se consacrant à la problématique de la fragilité et dispersion des archives des XIX^e et XX^e siècles – microsillons, photographies, disquettes...).

VERS UN ENSEMBLE COHÉRENT, NATIONAL ET EUROPÉEN

Les activités internationales d'Huma-Num sont de natures différentes. Elles visent essentiellement à valoriser les dispositifs français d'infrastructures et de recherche et de participer aux différents réseaux des humanités numériques de par le monde. Concrètement et sur le long terme, Huma-Num porte la participation française à des Eric¹, disciplinaires, comme **Clarin**², ou incluant le domaine culturel, comme **Dariah**³. Les Eric ne financent pas les projets qui les composent; ceux-ci et l'Eric sont financés, pour toute leur durée, par chacun des États membres qui y participent. L'originalité du budget de ces structures consiste dans le fait qu'il est en majeure partie constitué de contributions « en nature », afin de ne pas dupliquer à l'échelle européenne des services qui existeraient déjà aux niveaux nationaux. Le rôle d'Huma-Num est donc d'identifier les services et expertises nationaux – en particulier ceux en provenance des consortiums – et d'en faire un ensemble cohérent, qui intégrera en la complétant l'offre européenne. Ce rôle englobe la participation d'Huma-Num, de façon plus directe, à des projets à plus court terme, qui soutiennent la construction de ces grandes infrastructures du numérique et qui sont, eux, financés par la Commission européenne par le programme-cadre Horizon 2020.

Enfin, Huma-Num maintient des liens importants avec les mondes francophones (le Québec) et hispanophones (l'Amérique du Sud). Par exemple, depuis 2016, Huma-Num travaille en coopération avec l'Université de Montréal et le consortium Érudit autour du développement de la cyberinfrastructure COSHS, en partageant de l'information sur la façon de développer des services, d',k,animer des communautés en humanités numériques, et en réalisant des démonstrateurs web avec Isidore, le moteur de recherche.

À travers ces différentes composantes, Huma-Num s'emploie à développer des réponses aux besoins nouveaux, provoqués par l'utilisation du numérique, des différents acteurs de la recherche. Le fil rouge qui les relie est de rendre possibles de nouvelles recherches dans le domaine des sciences humaines et sociales.

STÉPHANE POUYLLAU

Directeur technique d'Huma-Num
stephane.pouyllau@cnrs.fr

LES 3 AXES D'HUMA-NUM

- **les services technologiques** pour l'outillage des données de la recherche (traitement, conservation pérenne, accès et l'interopérabilité)
- **la création et le soutien de consortiums**, disciplinaires ou non, qui nourrissent Huma-Num en guides de bonnes pratiques, formations, plateformes de données ancrées dans les Maisons des sciences de l'homme
- **l'ouverture à l'international** : participation à des projets d'infrastructures européennes (en particulier d'expertise), en complément d'échanges bilatéraux plus ciblés portant sur le déploiement des services d'Huma-Num.