



N°55 septembre 2018



## Édito

de Didier Torny,  
DAS InSHS

Au début de l'été 2018, le plan national pour la « science ouverte » a été présenté par notre ministère de tutelle. Cette expression, apparue il y a quelques années, peut sembler être un pléonasme [p2]

## NOUVELLES DE L'INSTITUT

Géographes, que valent vos articles ? Les bases de données commerciales en question

Les questions de la validité et de l'utilisation des bases de données bibliométriques pour caractériser les productions des chercheurs en SHS sont sans cesse relancées en France [p3]

## À PROPOS

L'ISC-PIF : un tiers-lieu de science pour mutualiser, innover et échanger autour des approches systèmes complexes

L'ISC-PIF est un tiers-lieu de sciences dédié au développement d'une recherche innovante sur les systèmes complexes, à la fois laboratoire, hôtel à projets, centre de ressources mutualisées et espace de co-working et de séminaires [p9]

## VALORISATION

La valorisation en sciences humaines et sociales : a never ending story

Depuis cinq ans environ, ce qui correspond à la date du premier salon *Innovatives SHS*, les communautés SHS se mobilisent de plus en plus, de façon interdisciplinaire, pour affirmer leurs activités de valorisation [p18]

## VIE DES RÉSEAUX

Le GIS « Études africaines en France » et les rencontres de Marseille

Du 9 au 12 juillet 2018 se sont tenues à Marseille, dans les locaux de l'Université Aix-Marseille (AMU), les 5<sup>e</sup> Rencontres des Études Africaines en France [p13]

## OUTILS DE LA RECHERCHE

Navigae : une plateforme de recherche et de valorisation des données géographiques scientifiques et patrimoniales

Navigae est un nouveau service à la disposition des communautés scientifiques pour rechercher et accéder selon des critères spatio-temporels à des données géographiques [p15]

## ZOOM SUR...

Les Humanités numériques 2.0

Alors qu'elles faisaient encore question et polémique il y a quelques années, les Humanités Numériques se sont imposées massivement dans tous les domaines des humanités et des SHS [p20]

## UN CARNET À LA UNE

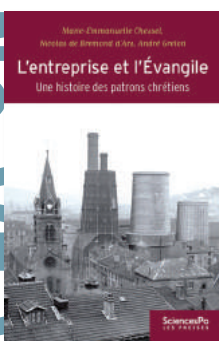
La 1<sup>ère</sup> Guerre vue de Paris

À travers ce carnet, l'historienne Odile Gaultier-Voitzier propose une édition critique des lettres envoyées par Étienne de Nalèche à son ami Pierre Lebaudy [p31]

## À SIGNALER

► Colloque *Tous mesureurs, tous mesurés. La science au cœur de la société*, 18-19 octobre, Paris  
Informations et inscription en ligne

## LIVRE

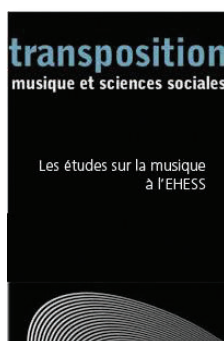


*L'Entreprise et l'Évangile. Une histoire des patrons chrétiens*, sous la direction de Marie-Emmanuelle Chesnel, Nicolas de Bremond d'Ars, André Grelon, Presses de Sciences Po, 2018  
Jusqu'ici peu étudié, le mouvement pa-

tronal chrétien a pourtant laissé une empreinte profonde sur l'éthique des affaires, y compris dans la lutte contre la corruption depuis les années 1990 [...]

voir toutes les publications

## REVUE

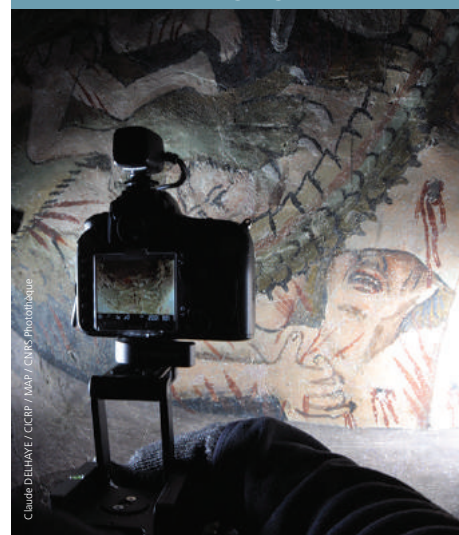


*Transposition. Musique et sciences sociales* est une revue scientifique interdisciplinaire, soutenue et coéditée par l'EHESS et la Cité de la musique-Philharmonie de Paris. Envisageant l'importance de la musique et des pra-

tiques musicales dans l'organisation des sociétés, *Transposition* entend questionner la manière dont celles-ci les pensent, les instituent et les mettent en scène [...]

voir toutes les revues

## PHOTO



Scientifique expérimentant un dispositif d'acquisition micro/macro photogramétrique sur un détail de la fresque du Leviathan, peinte sur l'un des murs de la chapelle Notre-Dame-des-Fontaines de la Brigue



# Édito

de Didier Torny  
DAS InSHS

Au début de l'été 2018, le [plan national pour la « science ouverte »](#) a été présenté par notre ministère de tutelle. Cette expression, apparue il y a quelques années, peut sembler être un pléonasme. En effet, les historiens nous ont montré que, depuis au moins quatre siècles, faire de la science, c'est écrire, lire, échanger, discuter, publier, dialoguer sous le regard des pairs et, souvent, de divers publics.

Pourtant l'électronisation de nos milieux et de nos sociétés, ces deux dernières décennies, a fait apparaître combien la diffusion de nos résultats était entravée par des barrières économiques — la nécessité d'un abonnement — technologiques — la limitation de circulation des écrits sur papier — ou professionnelles — la peur d'être plagié, l'excessive lenteur des procédures d'évaluation par les pairs, etc.

Trop souvent, nous nous sommes focalisés sur nos désirs de production de savoir, au cœur, cela va sans dire, de nos métiers de la recherche et de l'évaluation de nos carrières, au détriment de cette diffusion large, sous des formes adéquates et pérennes. Aussi, ce plan vise à remettre en première ligne cette diffusion, indispensable, des résultats mais aussi des données. Or, nos disciplines sont celles qui peuvent le plus bénéficier de la science ouverte, parce que nos résultats peuvent intéresser de vastes publics, tout en demeurant lisibles et appropriables, et parce que nos données, qu'elles soient textuelles, numériques, sonores ou visuelles, sont souvent très riches.

En pratique, le CNRS et, en son sein, l'InSHS contribuent déjà fortement à la science ouverte, au premier chef par la création et le soutien à ses infrastructures : HAL pour les archives ouvertes, Persée pour le patrimoine des revues, OpenEdition pour les publications contemporaines, Huma-Num pour les données de la recherche. Nous l'avons ensuite mise en œuvre dans notre politique de soutien aux revues : aujourd'hui, toutes les barrières mobiles des 109 revues sont égales ou inférieures à deux ans et près de la moitié d'entre elles a un contenu entièrement en libre accès. Enfin, nos collègues des métiers de la BAP F (Culture, communication, production et diffusion des savoirs) n'ont cessé d'informer et de former sur les pratiques de science ouverte, en soulignant l'intérêt de ces démarches pour les équipes de recherche elles-mêmes et en facilitant leur prise en main.

Le plan national conforte ces dispositifs en renforçant le financement des infrastructures, et favorise l'émergence de projets nouveaux en créant un fonds national. Il rend obligatoire la publication en accès libre des articles et ouvrages issus d'un financement public spécifique. Ce plan s'accompagne parallèlement d'un soutien à la bibliodiversité et à l'édition en accès ouvert, y compris dans des formats nouveaux (papiers de travail/*preprint*,

articles de données, évaluation ouverte...). Il s'agit également de dégager des moyens en diminuant les coûts très élevés de « l'édition fermée » détenue par l'oligopole des *big publishers*.

Dans le cadre de projets financés, les données produites devront, selon ce plan, également être rendues accessibles. Cela n'implique pas que tous nos matériaux devront être disponibles, alors qu'ils sont parfois soumis à d'autres contraintes comme des données sous protection intellectuelle ou des données personnelles et confidentielles, mais qu'un principe « aussi ouvert que possible, aussi fermé que nécessaire » sera mis en œuvre. Pour prendre un exemple trivial, cela signifie que, dans nos articles et ouvrages, nous publierons les tableaux, courbes et figures sous forme de données plutôt que de simples images. La politique de soutien de l'InSHS à la chaîne XML-TEI d'édition électronique facilite ce type d'édition, où la publication elle-même devient source de données pour ses lecteurs.

Enfin, si l'ensemble des mesures du plan s'inscrit dans une politique publique nationale, elle vise aussi à soutenir la science ouverte à l'échelon européen et international, notamment à travers les réseaux de la [Research Data Alliance](#) (RDA), les catalogues du [Directory of Open Access Journals](#) (DOAJ) et du [Directory of Open Access Books](#) (DOAB), les standards de données de type FAIR (*Findable, Accessible, Interoperable, Reusable*).

Pour les auteurs que nous sommes, faire circuler ses résultats dans un régime de science ouverte, c'est utiliser des licences libres afin de toujours associer des droits à ses publications et ses données, permettre de limiter une réappropriation par des intérêts privés (comme le pratiquent Academia et ResearchGate). Nous n'avons rien à craindre d'être lus partout dans le monde et par tous — collègues, associations et syndicats, décideurs publics, entreprises, citoyens — afin que nos recherches, longuement et soigneusement bâties, puissent éclairer, informer et questionner.

**Didier Torny**, DAS InSHS

## Géographes, que valent vos articles ? Les bases de données commerciales en question

Les questions de la validité et de l'utilisation des bases de données bibliométriques pour caractériser les productions des chercheurs en sciences humaines et sociales sont sans cesse relancées en France, non seulement par le Haut Comité à l'Évaluation de la Recherche et de l'Enseignement Supérieur (HCERES) qui établit actuellement ses critères d'évaluation, mais aussi par l'Observatoire des Sciences et Techniques (OST) qui a organisé en mai 2018 un séminaire regroupant des spécialistes de différents pays non anglophones. En nous appuyant sur les données objectives de publication produites par RIBAC et la base de données comparative JournalBase, il nous semble utile d'apporter ici une nouvelle contribution à cette réflexion.

En effet, nous avons déjà publié une étude<sup>1</sup> sur l'évaluation des principaux outils de recensement des revues en SHS, qui a abouti au développement de JournalBase. La vocation de JournalBase n'est pas bibliométrique, aucun indicateur ou facteur d'impact n'y est associé. Il permet principalement de faire des comparaisons de recensement des revues entre les outils bibliométriques privés et publics afin d'apporter des éléments fiables aux débats sur la possible utilisation de la bibliométrie en SHS. Les données sur les revues y sont enrichies avec notamment les langues d'utilisation, les pays d'édition, les ISSN, les URL vers les sites des revues, etc.

Par ailleurs, le recensement exhaustif dans RIBAC, depuis 2011, de la production des chercheurs à l'InSHS, nous apporte maintenant de nouvelles données sur les publications. C'est pourquoi, en partant du référencement des articles dans les revues scientifiques saisis par les chercheurs dans l'observatoire RIBAC, nous nous proposons d'examiner ici dans quelles bases de données sont le mieux représentées les revues dans lesquelles publient les chercheurs. Sur la base de RIBAC, nous observerons dans un premier temps de manière comparative les données de Scopus (Elsevier), du WOS (Clarivate analytics) et des listes de l'HCERES ainsi que l'outil comparatif JournalBase qui réunit toutes ces sources (après élimination des doublons). Puis, dans un second temps, nous étudierons plus spécifiquement les résultats de Scopus/SCImago<sup>2</sup> confrontés aux résultats de RIBAC. SCImago est une application espagnole en accès ouvert basée exclusivement sur les données de Scopus et dont l'utilisation est actuellement débattue. Pour illustrer notre étude, nous prendrons l'exemple de la communauté des géographes du CNRS rattachée à la section 39 du comité national (Espaces, territoires et sociétés), qui regroupe en moyenne 150 chercheurs. Nous examinerons les données saisies sur trois années, de 2014 à 2016.

### Les articles de revues recensés dans RIBAC entre 2014 et 2016

Pour mémoire, RIBAC est une plateforme web sécurisée qui permet aux chercheurs CNRS-SHS d'enregistrer chaque année leurs productions scientifiques de différentes natures (articles,

### JournalBase : la plateforme d'interrogation des revues pour les sciences humaines et sociales

Construit à partir des données du *Web of Science-SHS* (WOS), de Scopus-SHS, des listes d'autorité ERIH et HCERES, JournalBase a été conçu et développé au CNRS pour fournir une ressource inédite sur les revues en SHS au service de la communauté scientifique. JournalBase recense 16 000 revues scientifiques éditées dans 111 pays, soit deux fois plus que Scopus (8 280 revues en SHS), près de quatre fois plus que le WOS (environ 4 300 en SHS) et presque deux fois plus que l'HCERES (8 777 en SHS).

La base de données est mise à jour annuellement depuis 2010. Accessible en version bilingue, c'est également un outil de comparaison du référencement des revues dans les ressources nationales et internationales citées ci-dessus.

En accès ouvert, le site web s'adresse aux :

- ▶ chercheurs, doctorants ;
- ▶ laboratoires, organismes scientifiques européens et internationaux ;
- ▶ producteurs de l'information scientifique, éditeurs ;
- ▶ experts en information scientifique et technique ;
- ▶ instances de politique scientifique.

Doté d'une recherche simple ou avancée, JournalBase est enrichi de toutes les informations indispensables sur chaque revue scientifique recensée, avec la possibilité de choix par discipline SHS.

Interactif, les internautes peuvent faire des suggestions d'ajouts ou de modifications. JournalBase enregistre chaque mois plus de 1 000 visiteurs et près de 15 000 pages sont consultées dont la moitié par des internautes américains.

[En savoir plus](#)

chapitres, sites web, media, bases de données, etc.). Depuis 2011, RIBAC regroupe 100 % des productions scientifiques des chercheurs CNRS de l'InSHS. Il n'existe à notre connaissance pas d'autres bases qui recensent la diversité des productions SHS décrites dans RIBAC. Nous nous intéressons ici en particulier aux articles scientifiques, seules données comparables aux outils « bibliométriques » Scopus/SCImago, WOS, HCERES.

Entre 2014 et 2016, les géographes ont publié, selon RIBAC, 629 articles scientifiques dans 354 revues différentes<sup>3</sup>. Parmi ces 354 revues, 17 concentrent 23 % des publications (soit 147 publications, voir Tableau 1). Pour les 337 revues restantes, on note un fort éparpillement puisque 243 revues (70 %) n'enregistrent qu'un seul article sur les trois années étudiées, les

1. Dassa M., Kosmopoulos C., Pumain D. 2010, « JournalBase. Comparer les bases de données scientifiques internationales en sciences humaines et sociales (SHS) », in *Cybergeo: European Journal of Geography*.

2. « Le portail comprend les revues et les indicateurs scientifiques nationaux développés à partir des informations contenues dans la base de données Scopus® (Elsevier B.V.) ».

3. Source Ribac 2014, 2015 et 2016.



autres se répartissant entre deux à quatre articles sur trois ans. Concernant le référencement des 354 revues dans lesquelles les géographes ont publié pendant cette période, 190 sont recensées dans JournalBase (soit 54 %) d'une part, et 154 dans les listes HCERES<sup>4</sup> (43 %), 102 dans Scopus (29 %) et 64 dans le WOS (18 %) d'autre part. Ainsi, en connectant les différentes sources HCERES, Scopus, WOS, on constate que JournalBase affiche une couverture plus représentative des revues où publient les géographes d'après les données extraites de RIBAC. En comparant les différentes sources individuellement, les listes de revues de l'HCERES (pour la plupart anciennement AERES) se placent devant Scopus et le WOS en nombre de revues SHS référencées par rapport aux données de RIBAC.

| Revues                                                                                  | nombre articles |
|-----------------------------------------------------------------------------------------|-----------------|
| Cybergeo: European Journal of Geography                                                 | 23              |
| L'Espace géographique                                                                   | 12              |
| Tous urbains                                                                            | 11              |
| Revue Internationale de Géomatique                                                      | 11              |
| Annales de Géographie                                                                   | 9               |
| Revue archéologique du Centre de la France                                              | 8               |
| Urban Studies                                                                           | 8               |
| Ecology and Society: a journal of integrative science for resilience and sustainability | 8               |
| Lecture Notes in Computer Science                                                       | 8               |
| L'Espace politique                                                                      | 8               |
| Environmental Modelling and Software                                                    | 7               |
| Les cahiers d'outre-mer                                                                 | 6               |
| Vertigo - La Revue électronique en Sciences de l'Environnement                          | 6               |
| L'Information Géographique                                                              | 6               |
| EchoGéo                                                                                 | 6               |
| Revue francophone sur la santé et les territoires                                       | 5               |
| PLOS One                                                                                | 5               |

Tableau 1 : Les dix-sept premières revues pour la section 39 en nombre d'articles publiés de 2014 à 2016 (source RIBAC)

Parmi les dix-sept revues majoritairement utilisées pour les publications en géographie entre 2014 et 2016, *Cybergeo : Revue électronique européenne de géographie*, se place en première position assez loin devant les autres revues en nombre d'articles publiés par les chercheurs CNRS (Tableau 1).

Sur ces dix-sept revues, douze sont répertoriées dans JournalBase (soit 70 %), douze également dans les listes de l'HCERES, quatre dans Scopus-SHS (24 %) et seulement une dans le WOS-SHS (SSCI et AHCI) (6 %) (voir Tableau 2). On constate que douze

des dix-sept revues sont éditées en France dont neuf sont effectivement recensées dans JournalBase et seulement cinq dans Scopus. On note également qu'au moins cinq revues françaises (en italiques dans le tableau 2) ne sont recensées dans aucune source bibliométrique, alors qu'elles sont très fréquentées par la communauté des géographes.

On remarquera que certaines revues, dans lesquelles publient majoritairement les chercheurs de la section 39, ne font pas partie du domaine *stricto sensu* de la discipline (par exemple, *Revue archéologique du Centre de la France*, *Lecture Notes in Computer Science*, *PLOS One*).

## Les revues de géographie référencées dans SCImago (Scopus)

Nous présentons ci-dessous une étude des données obtenues de SCImago sur les revues en géographie éditées en France, afin de les comparer avec les informations effectivement déclarées par les chercheurs dans RIBAC. SCImago est une application en ligne en accès ouvert développée par des universitaires espagnols spécialisés dans l'analyse, la représentation et la recherche d'informations scientifiques au moyen de techniques de visualisation et associés au Conseil supérieur de la recherche scientifique (CSIC) en Espagne. L'outil s'appuie exclusivement sur les revues indexées dans la base Scopus. À noter que SCImago classe d'emblée les revues selon l'indicateur de Scopus, le *Science Journal Rank* (SJR). C'est donc à travers ces classements que nous allons nous rendre compte de la visibilité des publications des géographes enregistrées dans RIBAC.

Pour accéder aux revues en géographie sur le site, il faut sélectionner les revues (type *Journals*) par grands domaines, puis par sous-domaines. Une fois le grand domaine « *Social Sciences* » sélectionné, nous avons étudié les trois sous-domaines qui nous semblent recouvrir la discipline Géographie : « *Urban studies* », « *Geography, planning and development* » et « *Transportation* ». Les requêtes ont été effectuées en juillet 2018.

### « *Urban studies* »

143 revues sont recensées dans le sous-domaine « *Urban studies* ». Le classement par indicateur SCImago SJR décroissant montre que la majorité des revues sont publiées aux États-Unis, au Royaume-Uni ou aux Pays-Bas (siège social d'Elsevier, producteur de Scopus). La première revue non éditée dans ces trois pays est tchèque (*Critical Housing Analysis*). Elle a été ajoutée en 2018 et est classée trente-neuvième. La première revue française est *Urban Morphology*, classée soixante-quatrième. Elle est aussi répertoriée dans JournalBase mais aucun chercheur CNRS n'a publié dans cette revue durant les trois années étudiées (source RIBAC). La deuxième revue française est *Articulo-Journal of urban research* qui est classée quatre-vingt-huitième. Elle est répertoriée dans JournalBase et un chercheur CNRS a publié deux articles dans cette revue durant les trois années étudiées (source RIBAC).

Les autres revues françaises sont *Méditerranée*, *Projections*, *Sapiens*, *Territoire en Mouvement* et *Histoire urbaine* classées respectivement quatre-vingt-onzième, quatre-vingt-douzième, quatre-vingt-dix-huitième, cent seizième et cent trente huitième. Parmi ces revues, on recense seulement deux articles dans la

4. Listes originellement produites par l'Agence d'évaluation de la recherche et de l'enseignement supérieur (AERES).

| Reuves                                                                                          | nombre<br>articles<br>publiés entre<br>2014 et 2016 | JournalBase | Listes<br>HCERES | Scopus-SHS | WOS-SHS | Pays        |
|-------------------------------------------------------------------------------------------------|-----------------------------------------------------|-------------|------------------|------------|---------|-------------|
| Cybergeog: European Journal of Geography                                                        | 23                                                  | oui         | oui              | oui        |         | France      |
| L'Espace géographique                                                                           | 12                                                  | oui         | oui              | oui        |         | France      |
| Tous urbains                                                                                    | 11                                                  |             |                  |            |         | France      |
| Revue Internationale de Géomatique                                                              | 11                                                  | oui         | oui              |            |         | France      |
| Annales de Géographie                                                                           | 9                                                   | oui         | oui              | oui        |         | France      |
| Revue archéologique du Centre de la France                                                      | 8                                                   | oui         | oui              |            |         | France      |
| Urban Studies                                                                                   | 8                                                   | oui         | oui              | oui        | oui     | Royaume-Uni |
| <i>Ecology and society : a journal of integrative science for resilience and sustainability</i> | 8                                                   |             |                  |            |         | Canada      |
| <i>Lecture Notes in Computer Science</i>                                                        | 8                                                   |             |                  |            |         | Suisse      |
| L'Espace politique                                                                              | 8                                                   | oui         | oui              |            |         | France      |
| Environmental Modelling and Software                                                            | 7                                                   | oui         | oui              |            |         | Royaume-Uni |
| Les Cahiers d'Outre-mer                                                                         | 6                                                   | oui         | oui              |            |         | France      |
| VertigO - La Revue électronique en Sciences de l'Environnement                                  | 6                                                   | oui         | oui              |            |         | France      |
| <i>L'Information Géographique</i>                                                               | 6                                                   |             |                  |            |         | France      |
| EchoGéo                                                                                         | 6                                                   | oui         | oui              |            |         | France      |
| <i>Revue francophone sur la santé et les territoires</i>                                        | 5                                                   |             |                  |            |         | France      |
| PLOS One                                                                                        | 5                                                   | oui         | oui              |            |         | USA         |

Tableau 2 : Présence des dix-sept revues les plus publiantes (2014-2016) dans les différentes sources et liste de revues répertoriées sur la plateforme JournalBase (sources RIBAC et JournalBase)

| Rank SCImago | Titre de la revue                                           | Nombres d'articles publiés de 2014 à 2016 par les<br>chercheurs CNRS |
|--------------|-------------------------------------------------------------|----------------------------------------------------------------------|
| 386          | Géographie Economie Société                                 | 3                                                                    |
| 425          | Flux                                                        | 1                                                                    |
| 444          | <i>Annales de Géographie</i>                                | 9                                                                    |
| 462          | Revue de Géographie Alpine                                  | 3                                                                    |
| 488          | Méditerranée                                                | 0                                                                    |
| 494          | Projections                                                 | 0                                                                    |
| 506          | <i>Espace Géographique</i>                                  | 12                                                                   |
| 510          | Sapiens                                                     | 0                                                                    |
| 512          | Hérodote                                                    | 3                                                                    |
| 553          | Cahiers d'Etudes Africaines                                 | 0                                                                    |
| 575          | Field Actions Science Report                                | 0                                                                    |
| 580          | Etudes Rurales                                              | 1                                                                    |
| 583          | Bulletin de L'Association de Géographes Français            | 1                                                                    |
| 590          | Géographie et Cultures                                      | 0                                                                    |
| 601          | Territoire en Mouvement                                     | 2                                                                    |
| 618          | Sud-Ouest Européen                                          | 1                                                                    |
| 623          | Espace-Populations-Sociétés                                 | 2                                                                    |
| 635          | Revue du Marche Commun et de L'Union Européenne             | 0                                                                    |
| 662          | Histoire et Sociétés Rurales                                | 0                                                                    |
| 669          | Bulletin de liaison des membres de la Société de géographie | 0                                                                    |
| 679          | Histoire Urbaine                                            | 0                                                                    |
| 680          | OECD Observer                                               | 0                                                                    |
| Total        |                                                             | 38                                                                   |

Tableau 3 : Liste des vingt-deux revues françaises répertoriées dans le domaine « *Geography, planning and development* » et leur classement selon le SJR croissant (source SCImago, requête 20 juillet 2018)

revue *Territoire en Mouvement* et donc une seule revue sur les douze revues françaises les plus publiantes (source RIBAC). Au total, six revues éditées en France se trouvent référencées dans ce sous-domaine.

#### « Geography, planning and development »

634 revues sont recensées dans ce domaine, c'est donc de loin le plus gros domaine de Scopus/SCImago pour la géographie (pour information, JournalBase recense 946 revues en géographie). Quelle est la couverture des revues de ce domaine par rapport aux données de RIBAC ? Le classement par SJR décroissant montre que la majorité des revues est publiée là encore aux États-Unis, Royaume-Uni ou Pays-Bas. La première revue non éditée dans ces trois pays est suédoise (*Ambio*), classée vingt-neuvième. Il ressort de ces résultats que les géographes chercheurs du CNRS ont publié au total trente-huit articles de 2014 à 2016, dans onze revues françaises parmi les vingt-deux revues françaises recensées dans SCImago dans le domaine « *Geography, planning and development* » (voir Tableau 3) qui se répartissent dans les revues suivantes : *Géographie Economie Société* (trois articles), *Flux* (un article), *Annales de Géographie* (neuf articles), *Revue de Géographie Alpine* (trois articles), *L'Espace Géographique* (douze articles), *Herodote* (trois articles), *Etudes Rurales* (un article), *Bulletin de L'Association de Géographes Français* (un article), *Territoire en Mouvement* (deux articles), *Sud-Ouest Européen* (un article), *Espace-Populations-Sociétés* (deux articles). On constate aussi que dans ce gros domaine, seules deux revues (en italique dans le tableau 3) font partie des douze revues françaises les plus publiantes recensées dans Ribac. On notera également qu'on ne trouve aucune trace de *Cybergeo* dans ces résultats, nous reviendrons sur ce sujet au moment de la discussion.

#### « Transportation »

Quatre-vingt onze revues sont recensées dans le sous-domaine « *Transportation* ». Le classement par indicateur Scopus SJR décroissant montre que 70 % des revues de cette catégorie sont publiées aux États-Unis, au Royaume-Uni ou aux Pays-Bas. La première revue non éditée dans ces trois pays est allemande mais là encore anglophone (*Public Transport*), classée quarante-deuxième. Il y a seulement deux revues françaises recensées dans cette catégorie : *Recherche - Transports - Sécurité* et *Cahiers Scientifiques du Transport*, respectivement classées soixante-dix-neuvième et quatre-vingt-cinquième. Ces deux revues sont peu utilisées par les chercheurs puisqu'on recense un seul article publié de 2014 à 2016 dans *Recherche - Transports - Sécurité*.

Ces trois sous-domaines confondus regroupent donc quarante-deux articles publiés entre 2014 et 2016 dans treize revues différentes et recensent au total trente revues de géographie éditées en France.

### Comparaison avec les résultats dans Ribac et JournalBase

JournalBase ne classe pas les revues contrairement à Scopus/SCImago et ne fournit aucun indicateur, mais permet d'avoir des informations précises (langue, url, éditeur, pays éditeur, langue, sources dans lesquelles la revue est référencée, etc.) sur les revues par disciplines. Ribac enregistre de façon exhaustive toutes les publications des chercheurs CNRS.

Si l'on s'en tient aux articles publiés dans les revues éditées en France entre 2014 et 2016 en géographie, on constate que les

| Revues                                                                   | Nombre d'articles |
|--------------------------------------------------------------------------|-------------------|
| Cybergeo: European Journal of Geography                                  | 23                |
| L'Espace géographique                                                    | 12                |
| Revue Internationale de Géomatique                                       | 11                |
| Annales de Géographie                                                    | 9                 |
| L'Espace politique                                                       | 8                 |
| Cahiers d'Outre-Mer                                                      | 6                 |
| EchoGéo                                                                  | 6                 |
| AutrePart. Revue de Sciences Sociales au sud                             | 4                 |
| Natures Sciences Sociétés                                                | 4                 |
| Revue d'Histoire des Sciences Humaines                                   | 4                 |
| EspacesTemps.net                                                         | 3                 |
| Géographie Economie Société                                              | 3                 |
| Hérodote                                                                 | 3                 |
| Revue de Géographie Alpine / Journal of Alpine Research                  | 3                 |
| NETCOM - Networks and Communication Studies                              | 3                 |
| Espaces et Sociétés                                                      | 2                 |
| Espace, Populations, Sociétés                                            | 2                 |
| M@ppeMonde                                                               | 2                 |
| Noroiis – Revue géographique de l'ouest et des pays de l'Atlantique nord | 2                 |
| Revue Européenne des Migrations Internationales                          | 2                 |
| Territoire en mouvement                                                  | 2                 |
| Cahiers Agricultures                                                     | 2                 |
| Bulletin de l'Association de géographes français                         | 1                 |
| Développement Durable & Territoires                                      | 1                 |
| Etudes rurales                                                           | 1                 |
| Flux                                                                     | 1                 |
| Géomorphologie : relief, processus, environnement                        | 1                 |
| Maghreb Machrek                                                          | 1                 |
| Problèmes d'Amérique latine                                              | 1                 |
| Sud Ouest Européen                                                       | 1                 |
| Tiers Monde                                                              | 1                 |
| Pollution atmosphérique                                                  | 1                 |
| Recherche-Transport-Sécurité                                             | 1                 |
| Projets de paysage                                                       | 1                 |
| Climatologie                                                             | 1                 |

Tableau 4 : liste des trente-cinq revues françaises répertoriées dans le domaine « géographie » dans JournalBase dans lesquelles ont publié les chercheurs CNRS de 2014 à 2016 (sources Ribac et JournalBase)

chercheurs CNRS ont publié 129 articles (voir Tableau 4) dans trente-cinq revues (source Ribac) parmi les soixante-cinq revues françaises recensées dans JournalBase dans la discipline, alors que SCImago ne recense que quarante-deux articles dans treize revues sur trente revues françaises recensées au total dans SCImago/Scopus dans les trois sous-domaines étudiés ci-dessus.

## Discussion

D'après l'échantillon prélevé par Ribac sur trois années de publication, les bases de données bibliographiques donnent des représentations inégalement partielles de l'univers des revues utilisées par les chercheurs. Parmi les 354 revues dans lesquelles les chercheurs de la section 39 ont publié entre 2014 et 2016 (source Ribac), 190 sont recensées dans JournalBase (54 %) dont 154 dans la liste Géographie de l'HCERES (43 %), 102 dans Scopus/SCImago (29 %) et 64 dans le WOS (18 %). Si l'on se concentre sur les revues éditées en France, on retrouve trente-cinq revues dans lesquelles les géographes ont publié 129 articles (source Ribac) alors que Scopus/SCImago n'en recense qu'un tiers. Si l'on zoome encore plus précisément sur les douze revues françaises les plus publiantes de la discipline d'après Ribac (voir Tableau 1) pour la période étudiée, on constate là un déficit encore plus fort en terme de représentativité dans les sources commerciales puisqu'on retrouve dans les sous-domaines *Géographie* de Scopus/SCImago seulement trois revues françaises sur les douze revues, alors que dix sont recensées dans JournalBase, sous la rubrique *Géographie*. JournalBase affiche donc une couverture des revues beaucoup plus proche de la réalité observée dans Ribac<sup>5</sup>, alors que les données de SCImago sont très éloignées de la réalité des pratiques de publication des chercheurs CNRS.

Comme nous l'avons vu plus haut, les données de Ribac indiquent que *Cybergeo* est la revue en géographie la plus utilisée pour diffuser les travaux de recherche des chercheurs CNRS de la discipline. Il s'agit d'une [revue en accès ouvert](#), disponible sur le site d'OpenEdition. Or, *Cybergeo*, bien que référencée dans Scopus, est invisible lorsqu'on la recherche dans les revues en géographie de SCImago/Scopus. La raison en est que *Cybergeo* y est identifiée non comme une revue des domaines géographiques de SCImago/Scopus, mais comme une revue mixte à tendance géographie physique<sup>6</sup>, ce qui est à l'opposé de la réalité. Ce n'est sans doute pas le seul cas puisque seulement trois revues sur les dix-sept plus publiantes déclarées dans Ribac (voir Tableau 2) ont été retrouvées dans les rubriques de géographie de SCImago. On notera par exemple aussi que la revue *Géomorphologie : relief, processus, environnement*, éditée par le Groupe Français de Géomorphologie, est classée non en géographie, mais dans le domaine *Earth and Planetary Sciences / Earth-Surface Processes*. De même, la revue *L'Information géographique* est recensée dans le domaine *Medicine*.

Nous avons déjà pointé dans notre publication de 2010 la qualité déplorable des données de Scopus et les problèmes récurrents de référencement pour les sciences humaines et so-

ciales qui avaient amené Scopus à supprimer à l'époque tous les classements par mots-clés de leurs quelque 20 000 revues pour adopter un référencement plus généraliste (comme décrit plus haut dans SCImago). Il n'en reste pas moins que, dans sa hâte de concurrencer le WOS en intégrant beaucoup plus de revues que celui-ci, Scopus a bien multiplié le nombre de revues en SHS dans sa base pour arriver aujourd'hui à près de 8 300 revues toutes disciplines confondues. Cela peut apparaître comme un avantage sur le WOS qui en comptabilise environ 4300, mais le problème de la qualité des données reste entier.

Concrètement, sur les bases de notre étude comparative, la mauvaise qualité informationnelle de la base Scopus a pour effet de produire non seulement des classements de revues françaises qui ne correspondent pas, loin s'en faut, à ce que nous observons dans Ribac, mais aussi des répartitions disciplinaires erronées, des analyses de contenus bâclées non exhaustives et, en fin de compte, des indicateurs dont la valeur scientifique ne saurait être démontrée. Cette sous-représentation a déjà été mise en évidence dans nombre d'études et dans d'autres domaines des sciences humaines et sociales, car ces outils commerciaux recensent une majorité de revues anglo-américaines<sup>7</sup>.

De par son ancienneté dans la pratique du recensement, le WOS avance plus prudemment dans l'intégration de revues non anglophones et non anglo-saxonnes avec pour résultat d'avoir quelques revues européennes, pour l'instant très peu représentatives de la production scientifique dans nos disciplines, si l'on s'appuie là encore sur les données de Ribac. Il reste donc à recommander comme sources de références les listes sélectives de l'HCERES ou les revues recensées dans JournalBase qui recouvrent presque les deux tiers de la production scientifique officiellement recensée dans Ribac. Elles présentent l'avantage d'être évolutives et leurs informations sont contrôlées par la communauté scientifique. Ces listes de revues représentant un périmètre de scientificité avaient d'ailleurs été préférées aux classements de revues, il y a quelques années en France, lors de la mise en place des listes AERES. Il est donc tout à fait possible de définir un ensemble de revues sélectionnées par les communautés scientifiques sans utiliser les classements qui font appel à des considérations plus complexes à établir et qui sont moins consensuels.

D'un autre côté, Ribac a précisément été conçu pour rendre compte de la grande diversité des activités et des supports de publication en SHS. Cet outil a l'avantage de recenser selon les disciplines la variété des productions des chercheurs et de générer, par conséquent, des données réelles sur les activités de la recherche. Ces données apportent ainsi des arguments fiables, vérifiables, qui nous permettent désormais d'analyser la validité des résultats obtenus avec les bases de données commerciales comme Scopus et WOS. Accusé parfois à tort d'être un outil d'évaluation, Ribac a été élaboré au plus près des pratiques des chercheurs, avec leur participation et dans leur intérêt, face à l'hégémonie des entreprises commerciales de bibliométrie. Son recensement exhaustif et inédit permet d'afficher tant les spéci-

5. On notera toutefois que 164 revues enregistrées dans Ribac dans lesquelles publient les géographes ne sont présentes dans aucune des trois sources (HCERES, Scopus/SCImago, WOS) et n'apparaissent donc pas dans JournalBase.

6. *Sciences sociales* (divers) - *Cybergeo* y est décrit de la manière suivante : « Les publications en ligne regroupées sous l'égide de « Geoscience e-Journals » (un ancien réseau Web converti en un véritable portail) intègrent les revues électroniques à comité de lecture consacrées à la publication en libre accès dans le domaine des sciences de la Terre (géochimie, géophysique, paléontologie, géologie régionale, stratigraphie, géographie, etc.), c'est-à-dire les revues électroniques qui permettent un accès en ligne gratuit à l'ensemble ou au moins à une partie importante de leur édition et/ou de leurs archives actuelles » (page de SCImago consultée le 23/7/2018).

7. Mongeon P. et Paul-Hus A. 2016, "The journal coverage of Web of Science and Scopus: a comparative analysis", in *Scientometrics*, n°106 : 213-228.

ficités que la variété et la complexité des pratiques scientifiques au sein des différentes disciplines, lesquelles ne sauraient se réduire à des modèles de pratiques uniformes, ni à des indicateurs quantitatifs générés par des bases de données sur lesquelles la communauté scientifique n'a d'ailleurs aucun contrôle et pour qui les revues ne correspondent pas à la réalité, comme nous le montrons dans cet article. Dans le même esprit, d'ailleurs, le [manifeste de Leiden](#) pour la mesure de la recherche préconise dans son sixième principe<sup>8</sup> de tenir compte des différences entre disciplines en matière de publication.

## Conclusion

Les bases de données Scopus/SCImago et WOS offrent une représentation partielle et tronquée des publications d'articles scientifiques des chercheurs CNRS géographes au regard des données extraites de Ribac pour cette discipline. Conscients des limites et de la faiblesse du nombre de revues européennes de leurs bases de données en SHS, les producteurs de ces outils (*Elsevier* et *Clarivate Analytics*) ont cherché à améliorer leur couverture. Scopus ajoute régulièrement de nouvelles revues, Thomson Reuters racheté par *Clarivate Analytics* a lancé, en 2015, une nouvelle base de données, *Emerging Sources Citation Index* (ESCI). Bien qu'il y ait eu des améliorations, le problème de la couverture majoritairement anglo-américaine persiste, pour les bases SSCI et AHCI du WOS et pour Scopus. En 2017, on recense 60 % de revues anglo-américaines dans Scopus-SHS, 61 % dans la base Sciences Humaines du WOS (AHCI) et 76 % dans la base Sciences Sociales du WOS (SSCI) contre seulement 48 % dans JournalBase.

À cela vient s'ajouter la question de la qualité des données informées dans ces bases avec les effets de classements erronés. L'utilisation d'indicateurs bibliométriques comme l'indicateur SJR (Scopus/SCImago) ou le facteur d'impact (WOS) pour classer les revues et évaluer la production scientifique des chercheurs et des unités se trouve donc de fait biaisée si on les rapporte à la diffusion des travaux scientifiques observée dans Ribac. D'ailleurs, le CNRS a indiqué sur son [compte Twitter](#) le 14 juillet

2018 qu'il est désormais signataire de la [Déclaration de San Francisco](#) (DORA) qui vise à améliorer l'évaluation de la recherche : « Ce texte remet en cause l'utilisation abusive de l'indicateur de notoriété *Journal Impact Factor* ».

En conclusion, l'examen des revues en géographie montre que les indicateurs générés par Scopus/SCImago ne sont pas représentatifs de la réalité qu'ils sont censés mesurer et qu'ils ne s'adaptent pas non plus aux spécificités et aux pratiques de chaque discipline. Il est donc essentiel d'être conscient des risques que l'utilisation de ces indicateurs et de ces bases commerciales comporte, si l'on en venait à s'appuyer sur les résultats qu'elles produisent, tant dans le référencement des revues que dans leur impact, comme critères d'évaluation scientifique. Le développement de JournalBase et sa mise à disposition pour la communauté scientifique SHS en accès ouvert apportent des données concrètes pour comparer les différentes sources bibliométriques et contextualiser les productions scientifiques des chercheurs. Sans être exhaustif, JournalBase reste à ce jour la source documentaire la plus représentative des revues couramment utilisées en SHS, par ailleurs sans classement ou indicateur bibliométrique.

► Sources :  
[JournalBase](#)  
[SCImago](#)

### contact&info

► Michèle Dassa,  
InSHS-CNRS  
[Michele.DASSA@cnrs-dir.fr](mailto:Michele.DASSA@cnrs-dir.fr)  
► Christine Kosmopoulos,  
Géographie-cités  
[christine.kosmopoulos@parisgeo.cnrs.fr](mailto:christine.kosmopoulos@parisgeo.cnrs.fr)

8. Hicks D., Wouters P., Waltman L., De Rijcke S., Rafols I. 2015, "Bibliometrics: The Leiden Manifesto for research metrics", in *Nature*, Vol.520, N°7548 : 429-431.



# À PROPOS

## L'ISC-PIF : un tiers-lieu de science pour mutualiser, innover et échanger autour des approches systèmes complexes



© ISC-PIF

L'Institut des Systèmes Complexes de Paris Île-de-France (ISC-PIF, UPS3611, CNRS) est un tiers-lieu de sciences dédié au développement d'une recherche innovante sur les systèmes complexes, à la fois laboratoire, hôtel à projets, centre de ressources mutualisées et espace de *co-working* et de séminaires. Il catalyse depuis 2005 l'émergence de pratiques communes et interdisciplinaires et joue un rôle décisif dans la structuration et l'animation de la communauté systèmes complexes en Région Île-de-France.

### Les approches « systèmes complexes » au cœur des sciences contemporaines

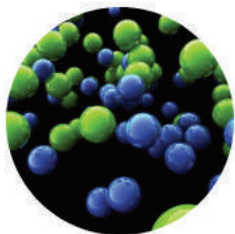
On qualifie de complexe un système composé d'une multitude d'entités dont les interactions locales font émerger des propriétés globales difficilement prédictibles par la seule connaissance des propriétés de ces entités. Une nuée d'oiseaux, un réseau social, une fourmilière, un flux de piétons ou encore un réseau de neurones peuvent être qualifiés de systèmes complexes.

L'interconnexion et l'interdépendance croissantes des réseaux technologiques, économiques, sociaux et médiatiques des sociétés contemporaines rendent de plus en plus prégnants leurs

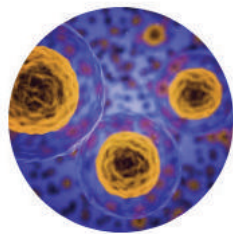
comportements en tant que systèmes complexes. Ils ont été illustrés autant par les récentes crises financières que par les révolutions en cascade qu'ont vécu les pays arabes ou par les défaillances pan-continentales qu'ont pu connaître certains réseaux technologiques comme les réseaux électriques.

L'étude des systèmes complexes dynamiques, qu'ils soient naturels (écosystèmes, organismes vivants, climat, etc.) ou artificiels — construits ou influencés par l'homme — (systèmes sociaux, économiques, villes, Web, etc.) couvre des questions portant sur leur fonctionnalité, leur organisation et désorganisation en tant que système multi-échelles, leur dynamique de développement, leur résilience et viabilité, leur capacité de réparation et réactivité, leur vieillissement et points de rupture. Ces questions sont transversales à de nombreux systèmes complexes.

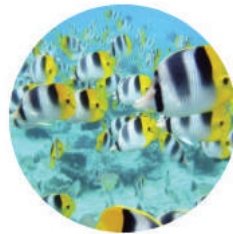
L'étude de la désorganisation des systèmes complexes concerne de nombreux enjeux contemporains — climatiques, écologiques, sociétaux — auxquels l'homme, au cœur de ces interactions multi-échelles, est confronté. La compréhension de ces phénomènes est indispensable à une bonne gouvernance.



Molécules



Cellules



Animaux



Humains & Tech

## Relever des défis nouveaux par une approche interdisciplinaire

La maîtrise des dynamiques des systèmes complexes est un facteur majeur d'innovations sociétales, un grand nombre de technologies futures et émergentes faisant le pari de la décentralisation et de l'auto-organisation (*blockchains*, flotte de voitures autonomes, robotique collective, réseaux électriques intelligents ou *smart grids*, technologies pair-à-pair, dispositifs pour les villes intelligentes, etc.).

L'étude de ces systèmes est interdisciplinaire à double titre. D'une part, l'articulation de différents niveaux d'organisation implique de mobiliser plusieurs disciplines. D'autre part, des méthodologies peuvent être développées de manière transversale pour leur étude (par exemple, la théorie des réseaux complexes dynamiques, la théorie de la percolation, les techniques de simulations agents, etc.).

La science des systèmes complexes s'intéresse de manière privilégiée aux interactions entre niveaux d'organisation micro/meso/macro et investit les méthodes de reconstruction et d'analyse capables de prendre en considération de multiples échelles de temps et d'espace. Elle mobilise pour cela les masses de données produites par ces systèmes — qui participent pleinement du phénomène *Big Data* — et le calcul haute performance pour modéliser leur comportement.

Ce parti pris scientifique amène les chercheurs confrontés aux systèmes complexes à relever des défis radicalement nouveaux, aussi bien du point de vue théorique et méthodologique que technologique.

## L'ISC-PIF, au service de la recherche Systèmes Complexes

L'ISC-PIF a pour mission de développer des méthodes et outils transversaux pour mieux comprendre et maîtriser les systèmes

complexes. Il structure et anime une communauté interdisciplinaire et inter-organismes de recherche à travers trois axes stratégiques.

### Hôtel à projets

L'ISC-PIF propose des résidences pour les chercheurs et groupes de recherches aux institutions partenaires qui souhaitent développer leur projet dans un contexte dynamique. Les résidents ont accès à tous les services et sont invités à contribuer à la vie de l'institut, en participant à des séminaires, des formations, etc.

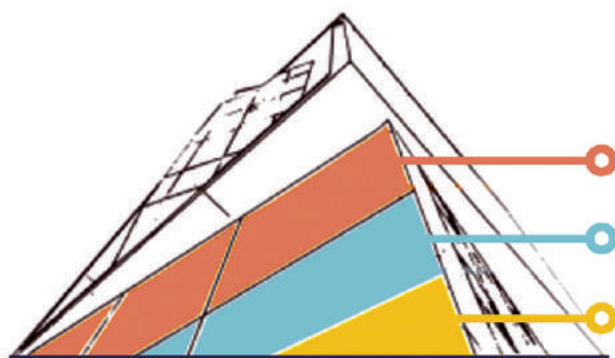
### Plateformes technologiques

Parce que la compréhension des systèmes complexes requiert des ressources spécifiques pour leur modélisation, pour leur simulation ainsi que pour l'analyse des masses de données, les chercheurs qui étudient les systèmes complexes sont confrontés chaque jour à de nombreux défis méthodologiques et technologiques. L'ISC-PIF fait le pari que ces défis pourront être relevés en intensifiant les collaborations entre chercheurs. C'est pourquoi l'ISC-PIF propose la mutualisation des moyens techniques et des infrastructures à l'échelle de la communauté scientifique et offre aux chercheurs plusieurs dispositifs permettant d'abaisser ces barrières technologiques :

- MULTIVAC est une plateforme de traitement et d'analyse en temps réel des masses de données qui offre de nombreuses fonctionnalités : apprentissage automatique, apprentissage profond, cluster Hadoop/Spark en nuage...
- GARGANTEXT (logiciel libre) est un service web connecté à plusieurs grands entrepôts d'articles scientifiques qui propose des outils innovants d'analyse textuelle, d'analyse de réseaux et de visualisation interactive.
- OPENMOLE (logiciel libre) offre des outils pour exécuter, explorer, diagnostiquer et optimiser votre modèle numérique, en tirant parti des environnements informatiques distribués (grille, *cluster*, *cloud*).

### Dissémination

L'une des missions de l'ISC-PIF est de favoriser de nouvelles formes de travail collaboratif et de dialogue entre sphères académique, citoyenne, politique et industrielle. Tout au long de l'année, l'institut organise des événements, séminaires, ateliers et formations ouverts autour des systèmes complexes. Il soutient par ailleurs des projets de manifestation scientifique par la mise à disposition de lieux (salles de réunion et de séminaires, auditorium) et par une aide logistique et en communication.



**3e étage :**  
Hôtel à projets

**2e étage :**  
Plateformes technologiques

**1e étage et espace TOTEM :**  
séminaires et co-work



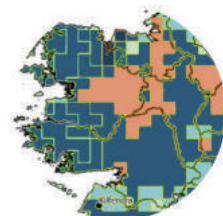
**Marchés financiers**



**Internet, Web**



**Réseaux sociaux**



**Villes et Populations**

## Un lieu ouvert tourné vers la société

Les défis auxquels font face nos sociétés nécessitent davantage d'interactions science-société afin de fournir aux citoyens les outils adaptés aux enjeux à venir. Dans cette optique, l'ISC-PIF développe plusieurs initiatives.

En novembre 2017, l'ISC-PIF a inauguré avec la Ville de Paris, la RIVP, la Mairie du XIII<sup>e</sup> arrondissement de Paris et l'association Musaiques un **nouvel espace dédié à la culture scientifique au sens large** : 700 m<sup>2</sup> comprenant un auditorium, un espace de *co-working* et un espace d'exposition. L'ouverture de ces nouveaux espaces constitue pour l'ISC-PIF et ses partenaires une véritable opportunité pour développer pleinement ces synergies science/citoyens, en bénéficiant de modalités d'échange originales et proactives avec la société civile, stimulées par les interactions entre arts et sciences. Ces espaces vont également renforcer leur capacité à favoriser le brassage des idées et la fertilisation croisée entre pôles de compétences territoriaux, leur donnant ainsi une plus large visibilité et de nouvelles occasions d'interagir avec les sphères citoyennes, économiques et politiques.

L'ISC-PIF a également été choisi par la Direction de la Communication du CNRS comme partenaire pour la coordination de **l'incubateur de projets de médiation scientifique sur les systèmes complexes**, en partenariat avec l'Amcsti, le réseau national des professionnels des cultures scientifique, technique et industrielle. À partir de son lancement en octobre prochain et pendant neuf mois, cet incubateur a vocation à accompagner les chercheurs et médiateurs dans la conception de projets de médiation sur les systèmes complexes.

## La recherche à l'ISC-PIF, une rencontre entre innovation et SHS

En provoquant la migration d'un grand nombre de contextes de production de connaissances vers des supports numériques — depuis les discussions informelles jusqu'à la production de théories scientifiques en passant par le débat politique —, l'essor des technologies de l'information et de la communication (TIC) a révolutionné notre rapport aux productions culturelles de nos sociétés. L'entrée des TIC dans l'ère du web 2.0 a ouvert, entre autres, de nombreuses opportunités pour une recherche qui se situe à la croisée entre l'innovation et les sciences humaines et sociales. Grâce aux outils du *Big Data* et des méthodes issues de l'analyse et de la modélisation des systèmes complexes, il est aujourd'hui possible d'extraire simultanément et à grande échelle les contenus sémantiques et les réseaux d'acteurs présents dans les différentes sphères des TIC.

L'ISC-PIF est spécialisé dans le **développement de ces outils** qui permettent l'analyse de systèmes complexes tels que les réseaux socio-sémantiques et proposent de nouveaux modes d'interactions avec les productions scientifiques et culturelles des sociétés. Ces outils sont appelés des « **macroscopes sociaux** », en référence à cet outil virtuel décrit par Joël de Rosnay qui permet d'observer un système complexe en adoptant un point de vue « extérieur ». Ces macroscopes ambitionnent d'accompagner les citoyens, responsables politiques, industriels et scientifiques dans la connaissance de nos sociétés comme ont pu le faire le microscope ou le télescope pour la connaissance du vivant ou de l'Univers.



L'Auditorium, Le Shakerspace © ISC-PIF





Blockchain



Voitures autonomes



Smart cities



Robotique

## Politoscope

Parce que le débat d'idées et la remise en question de nos convictions est un élément très important de la démocratie et de la vie citoyenne, les chercheurs et ingénieurs de l'ISC-PIF ont développé une application mobile ainsi qu'une plateforme permettant aux utilisateurs de naviguer facilement dans les masses de données générées par le débat politique français sur Twitter et ses commentaires. Le **Politoscope** propose ainsi des vues chronologiques et thématiques qui s'appuient sur les toutes dernières innovations du CNRS en terme d'analyse du *Big Data*. En plus d'un fil d'actualité politique en continu « à la une », l'application Politoscope propose de découvrir les groupes de *hashtags* qui définissent le mieux l'actualité de ces derniers jours, ainsi que de suivre la diversité des préoccupations des politiques à partir d'une détection, *via* une analyse automatique de l'ensemble des contenus des tweets, des sujets qui font débats. Tout ceci, en temps réel.

► Pour s'abonner à la newsletter, cliquer [ici](#)

► Retrouver l'ISC-PIF sur 

### contact&info

- David Chavalarias,  
ISC-PIF  
[contactus@iscpif.fr](mailto:contactus@iscpif.fr)
- Pour en savoir plus  
<https://iscpif.fr/>



Vues chronologiques et thématiques sur l'actualité politique française à découvrir avec l'application Politoscope © ISC-PIF



# VIE DES RÉSEAUX

## Le GIS « Études africaines en France » et les rencontres de Marseille (REAF)



Du 9 au 12 juillet 2018 se sont tenues à Marseille, dans les locaux de l'Université Aix-Marseille (AMU), les 5<sup>e</sup> Rencontres des Études Africaines en France (REAF 2018). Ces quatre jours ont une nouvelle fois donné l'occasion à plusieurs centaines de chercheurs consacrant leurs travaux aux Afriques de se réunir, de se rencontrer, de s'écouter et de partager leurs avancées scientifiques. Qui plus est, ces rencontres ont été associées pour la première fois à celles des JCEA (Jeunes Chercheur.e.s en Études africaines), dont c'était la 4<sup>e</sup> édition. Étudiants et chercheurs confirmés ont ainsi pu participer à des dizaines d'ateliers et débattre ensemble sur les thèmes « Afriques enchantées, Afriques en chantiers » et « Les Afriques en débat ». Moments indispensables de réflexion commune, mais aussi de convivialité, une riche programmation culturelle (concerts, expositions, projections, présentation d'ouvrages) a complété un programme bien rempli.

Ces rencontres sont les cinquièmes organisées par le Groupe-ment d'intérêt scientifique (GIS) « Études africaines en France ». Elles s'étaient déjà tenues en 2014 à Bordeaux et en 2016 à Paris, pour les dernières éditions. À chaque fois, il a été demandé à des unités de recherche locales de prendre en charge l'organisation sous l'égide du GIS. En 2014, le laboratoire *Les Afriques dans le monde* (LAM)<sup>1</sup> et l'*Institut des mondes africains* (Imaf)<sup>2</sup> ont mis en place la manifestation ; en 2016, ce fut un consortium de cinq unités, le Centre d'études en sciences sociales sur les mondes africains, américains et asiatiques (CESSMA), l'Unité de Recherche Migrations et Société (URMIS), le laboratoire Langage, Langues et Cultures d'Afrique Noire (LLACAN), le Pôle de recherche pour l'organisation et la diffusion de l'information géographique (PRODIG) et le laboratoire Développement et société (D&S)<sup>3</sup> qui s'en chargea. Cette année, ce sont le Centre Interdisciplinaire d'étude des littératures d'Aix-Marseille (CIELAM, EA4235), le Centre Nor-



1. Les Afriques dans le monde (LAM, UMR5115, CNRS / Sciences Po Bordeaux).

2. Institut des mondes africains (Imaf, UMR8171, CNRS / IRD / EHESS / EPHE / Université Paris 1 Panthéon-Sorbonne / AMU).

3. Centre d'études en sciences sociales sur les mondes africains, américains et asiatiques (CESSMA, UMR 245, Université Paris Diderot / Inalco / IRD), Unité de Recherche Migrations et Société (URMIS, UMR8245, CNRS / IRD / Université Paris Diderot / Université Nice Sophia Antipolis), laboratoire Langage, Langues et Cultures d'Afrique Noire (LLACAN, UMR8135, CNRS / Inalco), Pôle de recherche pour l'organisation et la diffusion de l'information géographique (PRODIG, UMR8586, CNRS / IRD / AgroParisTech / Université Paris 1 Panthéon-Sorbonne / Université Paris Diderot), laboratoire Développement et société (D&S, UMR201, IRD / Université Paris 1 Panthéon-Sorbonne).

bert Elias (CNE), le laboratoire [Interactions, Corpus, Apprentissages, Représentations](#) (Icar), le Laboratoire Population Environnement Développement (LPED)<sup>4</sup>, l'IMAF et l'URMIS qui ont pris la responsabilité de l'organisation.

Ces rencontres nationales proposent des panels consacrés à une thématique particulière ou à des recherches émergentes. Occasions irremplaçables de contacts et d'une meilleure connaissance des thèmes de recherche des uns ou des autres, elles ont eu un fort effet fédérateur pour nos études. Elles ont aussi été l'occasion d'inviter des chercheurs africains ou venant d'autres continents et ont largement contribué à une meilleure visibilité des recherches sur les Afriques. Enfin, elles ont à chaque fois été le lieu d'une réflexion collective sur l'avenir de nos recherches et sur les actions à poursuivre ou à entreprendre par le GIS.

La fondation de ce dernier est issue d'une histoire relativement récente, à partir du constat, fait à plusieurs reprises dès les années 2000, d'un manque de structuration et de visibilité des études africaines en France. Ayant vécu de près cet émiettement, les dissensions entre chapelles, les discordances entre individus, bien au-delà d'une émulation qui aurait pu être saine, nous avons pris langue avec l'équipe de direction du CNRS pour obtenir un outil susceptible d'aider à fédérer les études africaines. Dès 2004, le département des sciences de l'homme et de la société du CNRS donna son accord pour la création d'un Réseau thématique pluridisciplinaire (RTP) dédié aux études africaines. La première grande rencontre nationale du RTP « Études africaines » eut lieu en novembre 2006, au siège du CNRS, autour du thème « État des lieux et des savoirs en France ». Il s'agissait déjà de faire se rencontrer acteurs et unités de recherche s'intéressant aux Afriques dans les diverses disciplines des sciences humaines et sociales et de comprendre quels étaient les thèmes récurrents ou émergents dans cette spécialité. L'intérêt fut important et quarante-sept ateliers furent présentés durant cette manifestation. Les interventions sont d'ailleurs [en ligne sur le site actuel du GIS](#). À l'issue de ces premières rencontres, le département SHS proposa la création d'un GIS destiné à remplacer et améliorer le dispositif RTP. Néanmoins, une [deuxième rencontre des études africaines](#) fut organisée en 2010 à Bordeaux sans que ce groupement fût créé. Il fallut attendre 2012 pour que l'équipe de direction du nouvel Institut des sciences humaines et sociales, consciente des enjeux inhérents aux études aréales et aux études africaines en particulier, crée le Groupement d'intérêt scientifique « Etudes africaines en France ».

La convention de création précise les buts du GIS et son organisation. Dans le préambule, une phrase synthétise la vocation du GIS : « fédérer les différents laboratoires, mettre en réseau les équipes et les chercheurs, favoriser la diffusion de la recherche francophone et en améliorer la visibilité ». L'article 1 précise les missions que se donne le groupement : coordonner et promouvoir les activités des chercheurs et des unités de recherche en sciences humaines et sociales travaillant sur l'Afrique et ses diasporas ; organiser et animer le réseau articulé, en particulier, autour d'un site web (portail, annuaire en ligne, agenda des événements, veille scientifique) et d'une rencontre nationale

biennale ; renforcer les relations scientifiques internationales ; susciter des réponses communes aux grands appels d'offres nationaux et internationaux. Les missions du GIS relèvent aussi de la valorisation des études africaines auprès des médias et du grand public, de la sensibilisation auprès des pouvoirs publics sur l'Afrique et les études africaines et du renforcement de la présence des études africaines dans les réseaux de recherche européens et internationaux.

Pour ce faire, le GIS a pour membres vingt-trois tutelles (établissements de recherche et universités), représentées dans le comité directeur. Trente-trois unités de recherche dépendantes de ces tutelles sont engagées dans le réseau<sup>5</sup>, consacrant tout ou partie de leurs recherches à l'Afrique et ses diasporas. Le groupement a à sa tête un directeur (Pierre Boilley a été remplacé au cours de l'année 2018 par Hervé Pennec) et un comité scientifique composé d'une vingtaine de personnalités de la recherche.

Les actions du GIS ont été multiples ces dernières années. Outre les REAF déjà citées, une importante réalisation a été la rédaction et la publication de l'ouvrage *Les études africaines en France, un état des lieux*, plus communément appelé *Livre blanc des études africaines*. Publié sous format papier et [disponible en ligne au format PDF](#), cet ouvrage a eu pour but, pour la première fois au sein des études africaines, de tenter de comprendre les forces et faiblesses de notre spécialité et ses évolutions depuis sa fondation autour des années 1960. Ce bilan donne des raisons d'être en partie optimiste, mais pose de nombreuses interrogations. Les fusions de petites unités ont permis de créer des laboratoires de taille conséquente, les études africaines se sont structurées autour du GIS et de ces gros laboratoires et la visibilité de nos recherches a été améliorée, notamment auprès des pouvoirs publics et des médias. Mais s'il est indéniable que dans la recherche le vocabulaire a changé, ainsi que le regard posé sur les Afriques, s'il est vrai que de nouveaux thèmes tels que l'étude du genre ou de l'esclavage ont émergé, que les terrains peu labourés des Afriques anglophones commencent à l'être, que de nouveaux paradigmes théoriques se sont ancrés, beaucoup reste à faire pour vulgariser et rendre davantage visibles les connaissances et les savoirs produits par les chercheurs, de manière à casser les stéréotypes et les clichés qui restent bien présents dans la vision française des Afriques.

► [Visionner le film réalisé pour l'événement](#)

#### contact&info

► Pierre Boilley,

Études africaines en France  
[pierre.boilley@univ-paris1.fr](mailto:pierre.boilley@univ-paris1.fr)

► Pour en savoir plus  
<http://etudes-africaines.cnrs.fr>

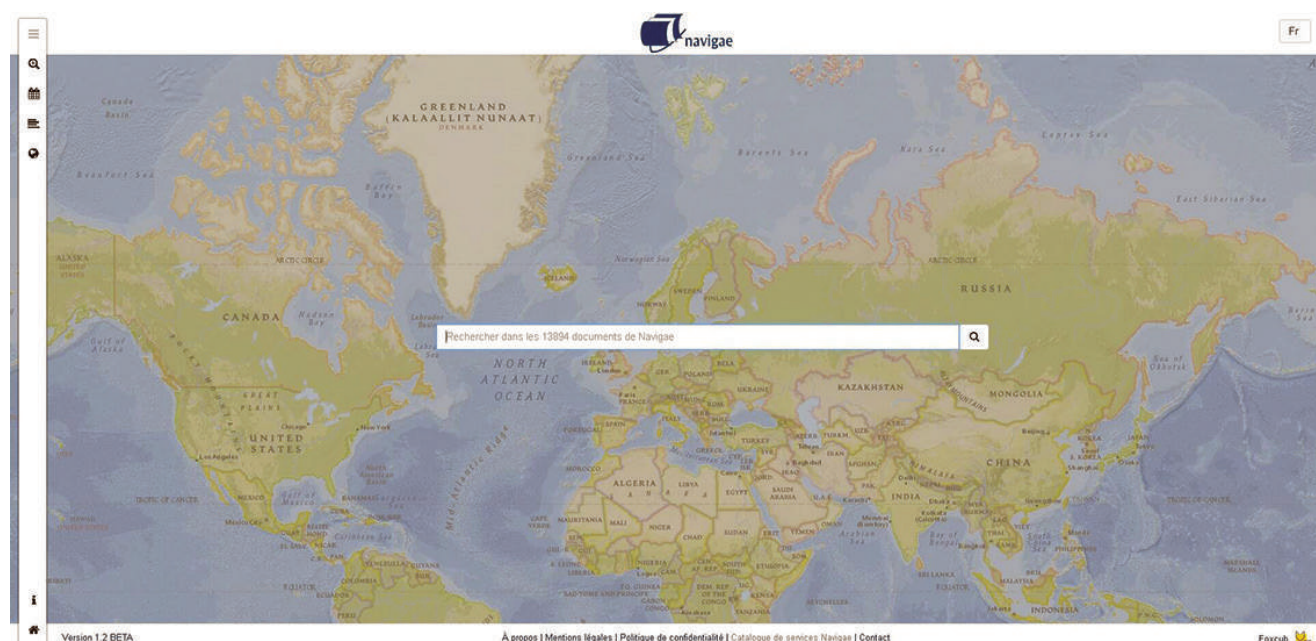
4. Centre Interdisciplinaire d'étude des littératures d'Aix-Marseille (CIELAM, EA4235), Centre Norbert Elias (CNE, UMR8562, CNRS / EHESS / AMU / Université d'Avignon et des pays de Vaucluse), laboratoire Interactions, Corpus, Apprentissages, Représentations (Icar, UMR5191, CNRS / Université Lumière Lyon 2, ENS de Lyon), Laboratoire Population Environnement Développement (LPED, UMR151, AMU / IRD).

5. Unités mixtes de recherche (UMR), Unités mixtes des instituts français de recherche à l'étranger (UMIFRE), Unités mixtes internationales (UMI), Unités de service et de recherche (USR), Équipe d'accueil (EA).



# OUTILS DE LA RECHERCHE

## Navigae : une plateforme de recherche et de valorisation des données géographiques scientifiques et patrimoniales



Page d'accueil de la plateforme Navigae

Navigae est un nouveau service à la disposition des communautés scientifiques pour rechercher et accéder selon des critères spatio-temporels à des données géographiques.

Réalisation du [Centre d'Information Scientifique et Technique Regards](#) de l'unité [Passages](#) (UMR5319, CNRS / Université de Bordeaux / Université Bordeaux Montaigne / Université de Pau et des Pays de l'Adour / Ensap Bordeaux) dans le cadre du [consortium ImaGEO](#) (labellisé et financé par la Très Grande Infrastructure de Recherche des Sciences Humaines et Sociales [Huma-Num](#)), cette plateforme a pour objectif d'améliorer la diffusion et la réutilisation des données géographiques issues de fonds de laboratoires de recherche ou de bibliothèques.

Ces fonds, fort précieux par leur caractère patrimonial, par la diversité des sujets représentés ainsi que par l'étendue des zones géographiques couvertes, sont encore peu utilisés.

Ces données iconographiques (cartes, photographies de terrain) restent peu exploitées pour diverses raisons : elles ne sont pas toujours disponibles en format numérique et les lieux de consultation ne sont pas toujours identifiés en raison du caractère épars de ces fonds.

En outre, une part importante de ces données géographiques n'est paradoxalement pas géolocalisée ce qui rend difficile leur réutilisation.

Navigae est donc né de la volonté commune de l'ensemble des membres du consortium ImaGEO de diffuser ces fonds auprès de chercheurs en géographie et plus largement en sciences humaines et sociales tout en permettant une réutilisation optimale de ces données. Il s'agit de favoriser une mise à disposition ouverte.

### Genèse du projet Navigae

Depuis sa création fin 2012, les différents membres du consortium ImaGEO se sont consacrés à la numérisation, au traitement et à la diffusion numérique de leurs collections de cartes anciennes ou de photographies de terrain.

Une méthodologie commune sur la numérisation et le traitement documentaire des images a ainsi permis la mise en ligne de plus de 15000 documents sur l'archive ouverte [MédiHAL](#) ou dans des bibliothèques numériques.

Après ce long travail de mise à disposition des collections des différents partenaires du consortium, il est apparu nécessaire de disposer d'un outil commun permettant à la fois d'offrir une vision unifiée et harmonisée de l'ensemble des fonds d'archives du consortium et ainsi de favoriser leur rapprochement, mais aussi de bénéficier d'une solution innovante capable de fournir une interface conçue et adaptée aux données géographiques scientifiques et patrimoniales (ce qui ne peut être le cas avec des plateformes de recherche généralistes) et de valoriser et compléter les données géospatiales grâce à des enrichissements obtenus à partir de sources de données *Open Data*.

### Une plateforme composée de différents processus

La plateforme Navigae devait être en mesure de couvrir différents besoins exprimés par les membres du consortium ImaGEO.

Elle devait notamment permettre de récupérer les données de l'ensemble des membres du consortium disséminées dans différents entrepôts, d'apporter des informations



complémentaires en matière de géolocalisation si l'information originelle est manquante et de proposer une interface de recherche et de visualisation de ces données.

Navigae est donc un ensemble de différents composants dont voici les principaux :

- ▶ un processus permettant la récupération des données et des métadonnées associées (titre, auteur, coordonnées géographiques, etc.) ;
- ▶ une chaîne de normalisation et d'enrichissement des données, utilisée notamment pour récupérer les coordonnées géographiques quand la donnée ne présente pas cette information ;
- ▶ une [interface web](#) pour rechercher et visualiser l'ensemble des données.

Le consortium a choisi de proposer l'accès à ces documents non seulement à l'ensemble de la communauté scientifique mais aussi à tout type de public. Ainsi, l'interface web de Navigae est ouverte à tous depuis juin 2018 et est disponible en trois langues : français, anglais et espagnol.

Navigae permet actuellement la consultation des fonds des membres du consortium ImaGEO, mais il est ouvert à l'ensemble des laboratoires de recherche et des bibliothèques et peut intégrer d'autres fonds.

## Une ergonomie adaptée aux données géographiques

Navigae a été conçu pour pouvoir interroger de manière cartographique des données géolocalisées ou géoréférencées à l'aide de logiciels SIG.

L'interface web permet donc d'accéder aux données *via* un fond de carte dynamique. Plusieurs fonds de cartes *open source* sont disponibles et il est possible de changer au fur et à mesure de sa consultation le fond de carte sélectionné.

Depuis la page d'accueil, plusieurs modalités de recherche sont proposées : recherche par facettes, par date, par géolocalisation (système décimal et sexagésimal) ou par mots clés — dont les toponymes — depuis la barre de recherche principale.

Après avoir effectué une première recherche, il est possible d'affiner les résultats en sélectionnant une emprise sur le planisphère correspondant à la zone géographique souhaitée. Les

données comprises, recouvrant ou à l'intersection de l'emprise, sont alors affichées. L'utilisateur peut également choisir de combiner plusieurs filtres afin de restreindre sa recherche.

Les résultats sur le planisphère sont représentés par un système de *clusters* qui permet de différencier les types d'objets (cartes, photographies). Chaque résultat peut être déployé afin d'afficher le document et ses métadonnées.

Un système de visionneuse permet de zoomer très finement sur l'image tout en pouvant la comparer dynamiquement avec le fond de carte. Il est également possible *via* la visionneuse d'afficher le document sélectionné en plein écran.

L'équipe a privilégié une interface épurée et intuitive, permettant à la fois aux institutions qui diffusent leurs données dans Navigae d'exposer leurs fonds, tout en garantissant le confort de l'utilisateur qui, grâce à la recherche et visualisation cartographique, peut voir et trier rapidement les données qui l'intéressent.

## Navigae : un outil précieux pour l'analyse diachronique

Les données numériques présentes dans Navigae ont permis de concevoir des techniques d'analyse spatiale dédiées à ces objets. En effet, les cartes ou les photographies ne sont plus de simples illustrations mais deviennent des objets manipulables, facilitant ainsi la mise en parallèle de documents anciens avec des sources de données cartographiques récentes.

Les géographes ont depuis longtemps recours à l'analyse diachronique qui consiste à étudier et comparer différents documents en les mettant côte à côte ou en les superposant.

Néanmoins, l'utilisation de données numériques facilite grandement la manipulation et la visualisation simultanée de plusieurs données.

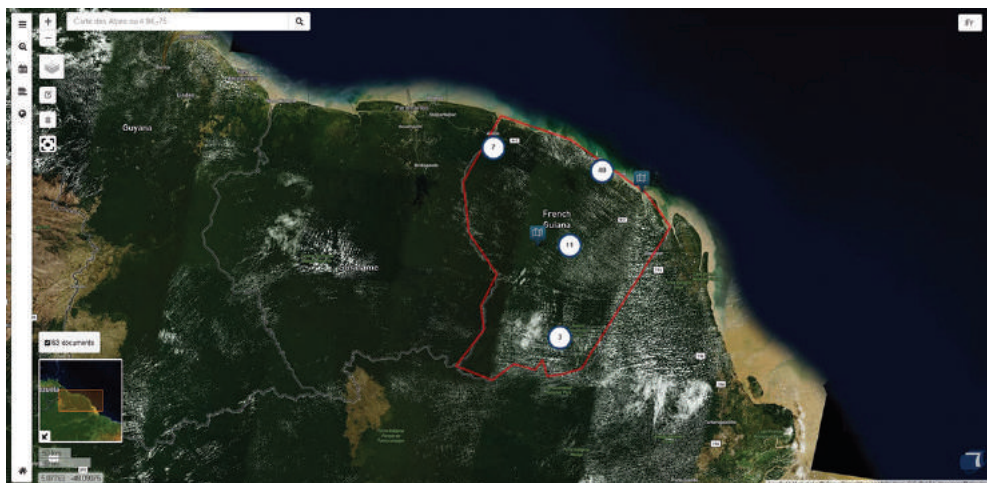
Avec Navigae, l'équipe a souhaité profiter largement de cette opportunité et a donc mis en place un système de comparaison permettant l'observation des transformations d'une zone géographique donnée en confrontant des sources anciennes aux données contemporaines.

Cette possibilité était fondamentale dès les premiers développements de Navigae, car c'est par cette fonctionnalité que peut s'opérer une véritable réutilisation des données par la communauté scientifique.

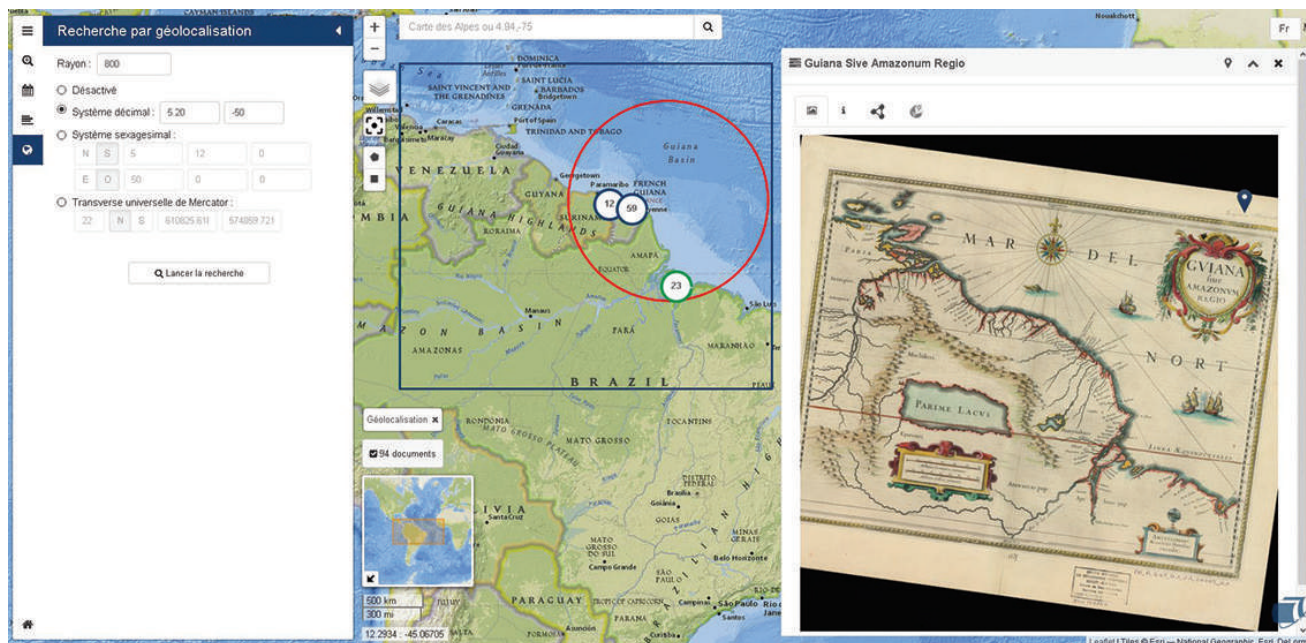
Le vis-à-vis d'une carte ancienne avec des données cartographiques présentes permet d'analyser des dynamiques territoriales ou de constater des évolutions de paysages.

Les cartes anciennes donnent des indications précieuses sur les aspects d'un lieu à une époque donnée qu'il est ainsi possible de mettre en perspective avec des enjeux actuels.

Navigae offre un formidable outil d'appropriation de documents géographiques patrimoniaux utiles pour



Recherche par emprise géographique dans Navigae



Recherche par géolocalisation dans Navigae. Carte *Guiana Sive Amazonum Regi*. Blaeuw, Guiljemus. Domaine public. Source Centre IST Regards

comprendre les modalités de construction de l'information territoriale d'un lieu.

La mise à disposition de ces nombreux documents anciens permet également de questionner la manière dont les connaissances géographiques s'établissent et d'analyser les changements environnementaux et territoriaux sur une longue période.

## Navigae : perspectives d'évolutions

Navigae est actuellement en version bêta et la totalité des développements prévus ne sont pas encore mis en production sur l'interface web.

Il est prévu d'étendre le système de comparaison des documents avec les fonds de cartes contemporains, aux documents eux-mêmes. Il est également envisagé de pouvoir comparer, par proximité géographique et par intervalle de dates, différents documents du corpus entre eux.

Enfin, différents services d'exposition des données de Navigae permettant ainsi leur réutilisation seront proposés.

Le consortium étudie actuellement la possibilité d'utiliser la norme GeoSPARQL définie par l'*Open Geospatial Consortium* (OGC) qui permet d'interroger des données géographiques selon les principes du Web sémantique.

Navigae a donc été réalisé de manière à répondre à deux enjeux majeurs : rendre visibles des documents géographiques patrimoniaux méconnus tout en garantissant les conditions de leur réutilisation grâce à l'application de normes permettant l'interopérabilité des données et spécifiques aux données géographiques.

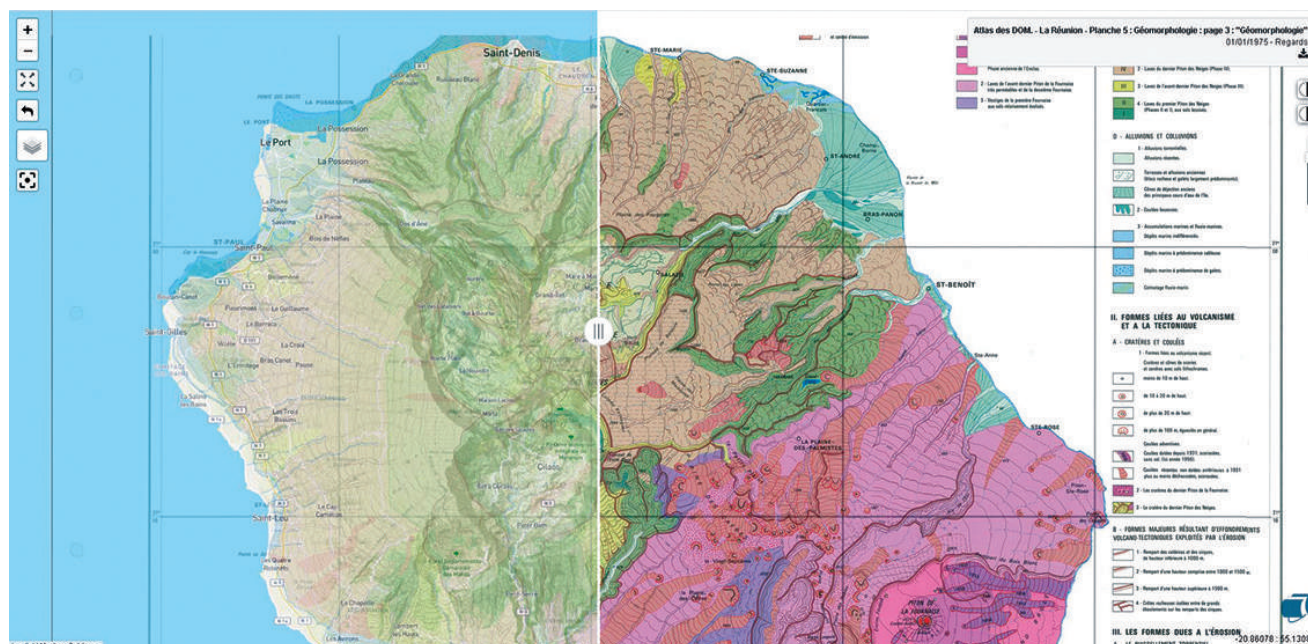
► Pour diffuser vos données dans Navigae : [contact@navigae.fr](mailto:contact@navigae.fr)

### contact&info

► Shadia Kilouchi,  
Passages

[shadia.kilouchi@cnrs.fr](mailto:shadia.kilouchi@cnrs.fr)

► Pour en savoir plus  
<https://www.navigae.fr>



Système de comparaison d'une carte géoréférencée dans Navigae. Carte *Atlas des Départements d'Outre-Mer, La Réunion planche 5 Géomorphologie*. Chaussalet, Jeanine ; Fontaine, Jean-Paul ; Nicolas, Françoise. Centre d'Études de Géographie tropicale (CNRS) ; Institut Géographique National. Source Centre IST Regards



# VALORISATION

## La valorisation en sciences humaines et sociales : a never ending story

Depuis cinq ans environ, ce qui correspond à la date du premier salon *Innovatives SHS*, les communautés SHS se mobilisent de plus en plus, de façon interdisciplinaire, pour affirmer leurs activités de valorisation. La valorisation est bien présente dans les recherches en SHS, qu'il s'agisse d'une valorisation par le produit ou du résultat, par l'innovation sociale ou par le transfert d'une méthode, d'un savoir-faire. D'autant que la valorisation nourrit aussi, en retour, la dynamique scientifique.

Un chercheur en SHS s'inscrit toujours dans une démarche de valorisation de ses recherches, quand il publie, crée des logiciels, organise des expositions, intervient dans les médias, participe à des réunions publiques. En voici quelques exemples.

L'Infrastructure de recherche nationale *OpenEdition*, labellisée Investissements d'avenir en 2012 et développée par OpenEdition Center depuis 2016, est une unité de service et de recherche (USR 2004) du CNRS, de l'université d'Aix-Marseille, de l'EHESS et de l'Université d'Avignon et des Pays de Vaucluse. Initiative publique à but non lucratif soutenue par des grandes institutions de recherche, sa principale mission est la promotion de l'édition électronique en accès ouvert. OpenEdition inscrit également son action dans le cadre de la *Bibliothèque Scientifique Numérique*. Les principales missions d'OpenEdition sont le développement de l'édition électronique en accès ouvert, la diffusion des usages et compétences liées à l'édition électronique, la recherche et l'innovation autour des méthodes de valorisation et de recherche d'information induites par le numérique.

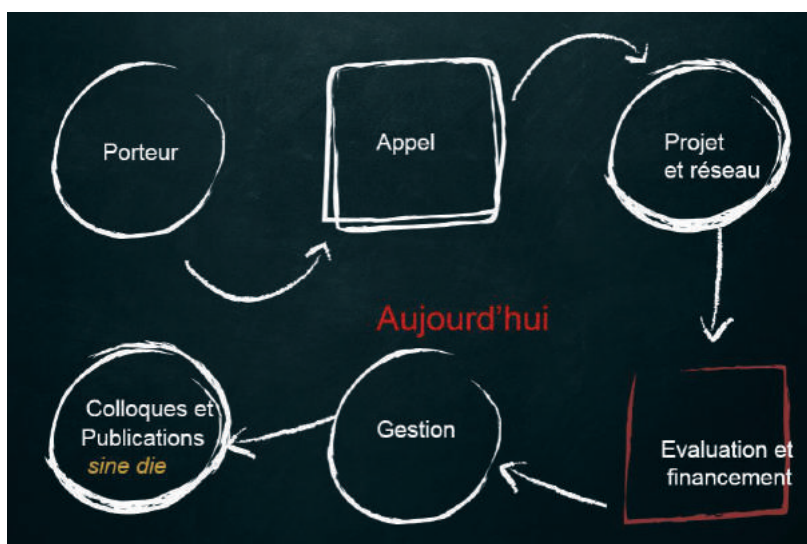
Le projet LABΩ, porté par le directeur du *Centre Gilles-Gaston Granger* (UMR7304, CNRS / AMU) développe une nouvelle idée de plateforme de recherche en ligne pour la réalisation de projets thématiques collaboratifs dans le domaine SHS. LABΩ répond à un constat et à un besoin. La recherche en SHS sous-emploie en effet les possibilités offertes par le numérique en réseau, car ce dernier est encore pensé, la plupart du temps, simplement comme la version dématérialisée des revues (*HAL*, *OpenEdition Journals*, etc.), des bibliothèques (*Classiques des sciences sociales*, *Gallica*, etc.), ou encore des encyclopédies (*Stanford Encyclopedia of Philosophy*, *Internet Encyclopedia of Philosophy*, etc.). Ce sont là des ressources et des bases de données indispensables, mais ce ne sont pas encore des outils structurels de constitution du savoir.

Pour constituer le savoir, il faudrait pouvoir commencer et mener à bien une recherche collaborative sans avoir besoin de financements préalables. C'est le sens du projet LABΩ qui offre à chaque chercheur la possibilité de coordonner et de synchroniser une équipe thématique internationale dans un laboratoire dématérialisé. Des objectifs très élevés peuvent donc être atteints avec des moyens qui sont déjà accessibles (serveurs, logiciels, algorithmes d'intelligence artificielle, ressources juridiques, *cloud computing*, etc.) et reviennent pour l'essentiel à créer une plateforme en ligne inspirée des *serious games*. Il s'agit bien d'une valorisation par le transfert car le projet réunit un ensemble de technologies qui existent déjà, pour une finalité qui n'a encore jamais été visée : un nouveau modèle pour l'Enseignement supérieur et la Recherche (ESR) qui tire pleinement parti de la dématérialisation numérique.

L'enjeu pour les SHS est de faire connaître ces différentes formes de valorisation. Il faut distinguer une valorisation qui a un impact sur les usages ou sur la société, de celle qui a un impact économique. Ces deux formes sont complémentaires mais il faut souvent mettre en avant des indicateurs spécifiques pour mesurer leur impact, notamment pour la première : une valorisation liée à l'innovation sociale ne rapporte pas comme un brevet, mais cela ne doit pas lui enlever sa valeur et son impact est mesurable.

Un exemple peut bien être représenté par la plateforme *VIA Inno*, rattaché au *Groupe de Recherche en Économie Théorique et Appliquée* (GREThA, UMR5113, CNRS / Université de Bordeaux). La plateforme, qui prend appui sur des outils et logiciels déjà existants mais payants, promeut le développement de méthodes d'intelligence technologique originales pour comprendre à la fois les dynamiques de l'innovation au sein des industries et l'environnement concurrentiel des entreprises de haute technologie. Les conditions requises pour la qualification de la démarche engagée sont connues : VIA Inno a développé un savoir-faire dans la construction et la mise à disposition de méthodes et d'outils de pilotage d'activités stratégiques. La large implication de VIA Inno dans la réalisation du Schéma Régional de l'Enseignement Supérieur, de la Recherche et de l'Innovation, adopté à la fin de l'année 2012 par le Conseil Régional d'Aquitaine, illustre la reconnaissance rapide dont la plateforme a bénéficié comme la mise en place, par la suite, de laboratoires communs avec de grands groupes industriels (PSA, Michelin...) en appui de leur stratégie de développement et du renforcement de leur positionnement technologique.

L'impact sociétal est également à la base du projet *Collective Science* porté par Fabienne Cazalis du *Centre d'analyses et de mathématiques sociales* (CAMS, UMR 8557, CNRS / EHESS). Dans son programme de recherche MACA (*Mapping Autistic Cognitive Abilities*), la chercheuse tente de répondre à des préoccupations tant scientifiques que sociales sur l'hétérogénéité de l'autisme. Elle se base sur l'usage du jeu vidéo, la recherche des pics d'habileté cognitive, l'étude de vastes cohortes et la protection des données personnelles dans le cadre d'expérimentations scientifiques sur la personne humaine, via un dispositif qui permet la création de tâches expérimentales ludifiées (*gamification*, *serious game*), la collecte de données massives, le stockage et le traitement des données en conformité avec le Règlement général sur la protection des données (RGPD). Le dispositif permet égale-



Configuration de Lab Omega © Pascal Taranto



ment la mutualisation des tâches expérimentales ainsi que la réplique des études. Le module RGPD de protection des données garantit l'authentification des participants et des chercheurs, la sécurisation du stockage des données personnelles et résultats expérimentaux, la gestion des consentements, la gestion des autorisations d'accès, le journal des opérations effectuées... Il est conçu en architecture mixte : décentralisation partielle *blockchain* combinée à des méthodologies d'authentification classiques.

La valorisation économique classique qui débouche sur une licence ou un brevet n'est pas exclue en SHS, bien au contraire. Ainsi en témoigne le projet qu'a développé Eloi Gattet, ingénieur de recherche au laboratoire *Modèles et simulations pour l'Architecture et le Patrimoine* (MAP, UMR 3495, CNRS / Ministère de la Culture), qui a débouché sur un prototype et est désormais une start-up. Le laboratoire MAP développe depuis près de dix ans des méthodes d'acquisition numérique 3D par photogrammétrie d'objets et de sites patrimoniaux. Cette technologie, à l'instar des scanners 3D laser, a l'avantage d'être peu coûteuse (le prix d'un appareil photo est moins élevé que l'achat d'un scanner qui se compte en milliers d'euros) et plus flexible. Elle permet en plus d'obtenir des informations de couleurs extraites directement des photographies. Une récente étude a permis de démontrer la précision métrique des résultats face aux scanners. Cependant, l'utilisation de ces méthodes nécessite une bonne maîtrise de l'appareil photo, un suivi scrupuleux des protocoles de prise de vues développés par le MAP ainsi qu'une compréhension globale des algorithmes utilisés pour générer les informations 3D. La limite est donc liée à l'utilisateur plutôt qu'au logiciel. Un premier prototype, mis en place à l'occasion d'un Projet Exploratoire Premier Soutien (PEPS) mené à Delphes en 2014, rassemblant mathématiciens, archéologues et chercheurs du MAP, a outrepassé les attentes. Il y a là une vraie rupture technologique. Cette solution permet aujourd'hui une numérisation 3D avec couleur en un temps et avec un coût par objet extrêmement réduits. Cet outil, issu d'une véritable compréhension des besoins de documentation associée aux contraintes imposées par le terrain ou le musée, est donc une solution concrète de démocratisation de la numérisation de collections patrimoniales et, au-delà, pour la communauté élargie des SHS. En 2015, une déclaration d'invention (DI) a été déposée. En outre, grâce à un soutien dans le cadre du programme prématuration du CNRS, le prototype existe désormais et une création d'entreprise, Mercurio, est d'actualité.

Comment accompagner ces démarches de valorisation en SHS et conforter leur pleine reconnaissance dans la politique de valorisation du CNRS ?

La Société coopérative d'intérêt collectif (SCIC) peut être une forme structurante de valorisation en SHS. Comme l'ont bien noté Isabelle Backouche et Sarah Gensburger, lors de l'élaboration du projet *Faire savoirs*, « aujourd'hui, la plupart des institutions de recherche et d'enseignement françaises offrent deux voies parallèles aux chercheurs : la recherche fondamentale et/ou la création d'une start-up pour valoriser, notamment financièrement, un projet auprès du reste de la société, au risque de limiter son activité de chercheur. La SCIC peut bien être une troisième voie qui permet aux chercheurs de continuer la recherche fondamentale, tout en irriguant de manière active et créative la société de leurs travaux et de leurs découvertes. »

La SCIC peut donc permettre aux scientifiques d'être acteurs de la valorisation et de l'innovation, au croisement du public et du privé, alliant intérêt collectif et viabilité économique. Ce

### Quelques chiffres sur les activités de valorisation portées en SHS au CNRS

- ▶ 40 start-up en activité
- ▶ 44 brevets et logiciels actifs
- ▶ 17 licences actives
- ▶ 8 laboratoires communs
- ▶ 2890 contrats de collaboration de recherche en cours

faisant, elle crée une dynamique prometteuse pour le transfert en sciences humaines et sociales comme pour la professionnalisation des étudiants de troisième cycle, en leur offrant de nouvelles perspectives professionnelles au sein de la SCIC. Celle-ci devient ainsi une pépinière, un incubateur très bien adapté aux spécificités des projets en SHS qui transfèrent un savoir-faire, une méthode.

Plus largement, la valorisation en SHS peut s'appuyer sur des outils et programmes mis en place par le CNRS. C'est le cas du programme de prématuration qui est un outil de détection de projets émergents à fort potentiel en termes d'innovation de rupture et de soutien aux premières étapes de leurs développements. La rupture peut être technologique, d'usage ou consister en une amélioration à forte valeur ajoutée notamment dans le domaine social, sans oublier la méthode et/ou le savoir-faire qui peuvent faire l'objet d'un transfert. L'InSHS a également mis en place des outils propres, tel le réseau des MSH, le CVT Athéna, le salon *Innovatives SHS* et le groupe de travail Innovation en SHS.

Le groupe de travail Innovation en SHS réunit des représentants de l'InSHS, de la Direction de l'innovation et des relations avec les entreprises (DIRE) du CNRS, de plusieurs services partenariat et valorisation en délégations régionales, de CNRS Innovation, ainsi que des directeurs de laboratoires rattachés à l'InSHS. Il vise à accompagner les projets portant sur une œuvre *a priori* protégée par le droit d'auteur ou sur des logiciels dans un parcours de valorisation adapté. Il promeut des bonnes pratiques transférables à d'autres domaines disciplinaires. Cet accompagnement se traduit par la conception de parcours adaptés de valorisation pour des projets qui n'ont pas vocation à être exploités via les processus habituels de valorisation — notamment sous forme de licence ou de brevet — et dont il est peu probable qu'ils généreront un retour financier, ce qui est souvent le cas lorsque de tels projets ciblent davantage l'« innovation d'usage », l'« innovation sociale » ou encore le transfert d'une méthode ou d'un savoir-faire.

Ce faisant, les travaux de ce groupe visent également à enrichir le modèle de déclaration d'invention/de logiciel élaboré par le CNRS dans le cadre de la valorisation des brevets et des logiciels pour fournir un outil qui soit adapté aux résultats scientifiques qualifiables d'œuvres de l'esprit — et donc susceptibles de faire l'objet d'une protection par le droit d'auteur — ou qualifiables de savoir-faire.

La valorisation de la recherche porteuse d'innovation en SHS : *to be continued*.

contact&info

▶ Maria-Teresa Pontois,  
InSHS

Maria-Teresa.PONTOIS@cnrs-dir.fr

## Les Humanités numériques 2.0

Alors qu'elles faisaient encore question et polémique il y a quelques années, les Humanités Numériques se sont imposées massivement dans tous les domaines des humanités et des SHS — à preuve les quarante-cinq dossiers reçus, souvent excellents, dans le cadre de l'appel lancé par l'InSHS à l'été 2018 pour le recrutement d'un postdoctorant. La modélisation 3D et la reconnaissance d'images aidées par l'intelligence artificielle deviennent des outils en archéologie; la conversion numérique dans des formats patrimoniaux ouverts et pérennes (le standard TEI) s'est imposée comme une norme philologique reprise par des dizaines de projets; les éditions enrichies de documents hypertextes et de compléments multimédia sont une réalité quotidienne en littérature et en philosophie; l'histoire numérique et l'histoire du numérique sont deux champs bien identifiés; les socialités et réseaux numériques sont couramment étudiés en sociologie tandis que l'économie numérique est devenue un domaine d'analyse à soi seul; les avancées en matière de reconnaissance d'écriture et de traitement de la voix et du langage naturel sont si importantes qu'elles ouvrent des domaines applicatifs originaux comme la traduction automatique ou l'analyse des controverses. Du côté des pratiques concrètes de recherche, des outils plus nombreux que jamais sont à portée de main, qu'il s'agisse de transcrire des manuscrits ou de fouiller des données massives par exemple. Du côté des résultats, les concepts fondamentaux des Humanités numériques, réseaux, systèmes complexes, interactions homme-machine, *deep learning*, éthique et politique du numérique, etc. font l'objet de travaux épistémologiques importants.

L'InSHS refuse de faire des Humanités Numériques une démarche enclavée et en promeut plutôt leur développement dans des champs variés. Il accompagne des équipes spécialisées (l'[Institut des Systèmes Complexes de Paris Île-de-France](#)<sup>1</sup> en particulier) et les Très Grandes Infrastructures de Recherche (TGIR) [Progedo](#) et [Huma-Num](#) mais soutient aussi de très nombreuses autres équipes fortement impliquées dans les Humanités Numériques, dont cet éditorial ne suffirait pas à faire la liste. Sa politique de soutien à la science ouverte en général, et au Centre pour l'édition électronique ouverte — [OpenEdition Center](#)<sup>2</sup> — en particulier, participe de cet accompagnement de la conversion numérique des pratiques scientifiques. Un écosystème allant de la formation des Ingénieurs et Techniciens (IT), des pratiques de recueil et de numérisation des données (en incluant la science collaborative) à la production de sites et de base de données pérennes est en train de se mettre en place. Au travers du consortium de la TGIR Huma-num et du soutien aux réseaux nationaux et internationaux, un appui aux initiatives *bottom-up* complète le dispositif, en favorisant les bonnes pratiques et l'appropriation d'outils et de techniques avancées. Il en va de même de l'appui de l'Institut aux méthodologies visant la modélisation et à une priorité émergente, l'Intelligence Artificielle : l'InSHS promeut des thèmes et des méthodes permettant à la recherche française d'être au meilleur niveau international.

D'immenses défis sont néanmoins au-devant de nous : d'une part, les pratiques des Humanités Numériques restent d'un niveau d'accès élevé, très coûteuses en temps et en ressources, exigeant souvent de profonds changements d'habitudes. La numérisation d'un ensemble de monnaies ou de partitions ou encore la constitution d'une base de données en économie sont des entreprises de longue haleine et l'exploitation des données recueillies exige une familiarité avec les interfaces, aux normes et aux pratiques du numérique. D'autre part, leur développement se fait parfois à un niveau individuel et en ordre dispersé : la mutualisation des outils et des méthodes, de même que la diffusion et la pérennisation des résultats et leur interopérabilité, restent souvent un problème, malgré les efforts d'Huma-num. Enfin, et surtout, un saut quantitatif est en train de se produire à un moment où la « dataification » du monde est un fait massif qui touche tous les domaines de la vie et du savoir — pensons juste à la récente ouverture au téléchargement et à la fouille des gigantesques dépôts de documents et de data de la Bibliothèque nationale de France. Car ce ne sont plus des textes, des bases ou des statistiques de taille modeste et facilement visualisables, mais des données en *big data* ou en *long data* qui sont à portée de main – données qui exigent des méthodes d'approche quantitative ou mixte (quali-quantitative), en impliquant le recours à des outils de statistique, de fouilles et de visualisations avancées et à une épistémologie nouvelle, qu'on a pu nommer « *distant reading* ». Pensons pour donner un seul exemple, à la manière dont l'ISC-PIF a été capable d'analyser des dizaines de millions de tweets de la campagne électorale pour cartographier l'état de l'opinion : derrière les Humanités Numériques se dessinent la diffusion dans la recherche des *data sciences* et l'émergence de sciences sociales computationnelles. Les efforts pour s'appropriier sans s'en effrayer ces Humanités Numériques de seconde génération sont encore largement à faire, même si certaines entreprises ont ouvert la voie et si l'intelligence artificielle offre bien des promesses. Quoi qu'il en soit, c'est ce tournant qu'il nous faut désormais négocier.

Alexandre Gefen, DAS InSHS

1. ISC-PIF, UPS3611, CNRS.

2. OpenEdition Center (USR2004, CNRS / EHESS / AMU / Université d'Avignon).

# Humanités numériques, humanités numérisées et humanités du numérique

Sociologue au sein du *Centre franco-allemand de recherches en sciences sociales de Berlin* (Centre Marc Bloch Berlin - CMB, USR3130, CNRS / MEAE), Camille Roth y a fondé un pôle Humanités Numériques / sciences sociales computationnelles mêlant sciences sociales, modélisation mathématique et informatique. Son programme de recherche est centré autour de la dynamique socio-sémantique de l'espace public numérique, de l'analyse de corpus textuels, des systèmes complexes sociaux et des réseaux de savoirs.

Les humanités numériques, traduction plus ou moins directe de l'anglais *digital humanities*, désignent un mouvement qui se structure autour de l'utilisation d'outils numériques en SHS. Le terme *digital* peut cependant être compris de trois manières différentes. Premièrement, en tant que numérisation : les humanités numériques ont alors essentiellement trait à la constitution et au traitement d'archives mises dans un format numérique (on pourrait parler de *digitized humanities*). Une deuxième acception est liée au développement de modèles numériques afin de décrire et expliquer les processus sous-jacents aux corpus étudiés (il s'agirait de *numerical humanities*). Enfin, les usages des outils numériques et, plus largement, les comportements des individus au sein des communautés en ligne font l'objet de ce que l'on pourrait nommer les humanités du numérique (*humanities of the digital*).

En pratique, ces trois acceptions correspondent à des épistémologies relativement distinctes et, partant, à des communautés scientifiques potentiellement disjointes. La première acception est peut-être la plus courante et recouvre en toute généralité la numérisation effective, le stockage, le traitement, la mise à disposition, la visualisation de collections de textes, de banques d'images, de documents multimédia d'origines variées. Ces tâches soulèvent en premier lieu des problématiques liées aux données et à leurs infrastructures : la construction des bases de données proprement dite (jusqu'à la définition de formats adéquats à leur réutilisation dans de nouveaux contextes), leur stockage (la pérennité des données est un élément important) et leur partage (dont le développement d'interfaces de visualisation et d'accès *ad hoc*). Elles requièrent aussi l'intégration de diverses méthodes mathématiques ou informatiques, dont notamment l'intelligence artificielle — qu'il s'agisse d'outils de bas niveau comme la reconnaissance optique de caractères (ROC) ou bien de plus haut niveau comme le traitement automatique des langues (TAL), l'analyse de réseaux sociaux (ARS) jusqu'aux approches récentes de curation de contenu collaborative. Cette intégration vise à provoquer l'émergence de nouvelles questions en sciences humaines et sociales, du fait des gains de productivité mais aussi des nouvelles manières de regarder les données. La littérature et la linguistique, notamment les *corpus linguistics*, ont ici tenu une position longtemps dominante. Ainsi, le champ jadis dénommé *humanities computing* avait pour ambition d'appliquer des approches quantitatives sur des corpus textuels (par exemple, la stylométrie utilise des marqueurs lexicaux, syntactiques ou sémantiques pour attribuer statistiquement des textes à des auteurs) et émerge dès la fin des années 40, bien avant que le terme *digital humanities* n'apparaisse. Ce champ s'est ensuite progressivement institutionnalisé, à partir de la fin des années 60, au travers de la création d'associations (*Association for Literary and Linguistic Computing*, *Association for Computers and the Humanities*), de revues (*Computers and the Huma-*

*nities* fondée en 1966 ou bien *Literary and Linguistic Computing* - LLC en 1986, qui s'est d'ailleurs récemment renommée *Digital Scholarship in the Humanities* - DSH) ou d'efforts de normalisation (comme la *Text Encoding Initiative* - TEI vers la fin des années 80). De nombreuses disciplines, au-delà des humanités et notamment au sein des sciences sociales, se sont entre temps approprié ces approches. C'est notamment le cas des sciences politiques et des sciences de la communication (avec, par exemple, l'analyse des discours au sein de corpus journalistiques numérisés<sup>1</sup>), de l'histoire (mentionnons ici la reconstruction de réseaux sociaux historiques grâce aux mentions dans des archives biographiques<sup>2</sup>) et, au-delà des corpus textuels, de l'archéologie ou de la géographie. Ces deux disciplines furent, en outre, pionnières dans l'utilisation d'outils numériques, particulièrement dans le champ déjà ancien de la *digital archeology* ou bien avec les systèmes d'information géographiques (SIG).

La deuxième acception (*numerical humanities*) met l'accent sur l'abstraction et la modélisation. Dans une certaine mesure, les corpus numérisés constituent indubitablement une abstraction du matériau empirique d'origine. Néanmoins, les approches proprement numériques ont spécifiquement pour but de décrire formellement certains des processus, potentiellement généraux, à l'origine des données observées. En ce sens, ces approches ont moins à voir avec l'automatisation de tâches autrement difficiles ou impossibles, comme le permettent la ROC ou les SIG, qu'avec le développement de modèles de comportement humain ou social en soi. Le cadre épistémologique diffère ainsi sensiblement car les bases de données ne sont plus exploitées comme des exemplaires singuliers propres à une étude de cas particulière, mais en tant qu'instances parmi d'autres et surtout interchangeables d'un phénomène générique (comme le seraient les données issues d'une expérience reproductible). L'intérêt croissant pour ces *numerical humanities* est illustré notamment par la multiplication des conférences et revues dédiées à ce que l'on appelle les « sciences sociales computationnelles ». Ce terme semble aujourd'hui regrouper une grande partie des chercheurs concernés par ce mouvement qui découlent de différents champs de formalisation en SHS, dont notamment la sociologie mathématique (déjà établie depuis les années 70 par le biais du *Journal of Mathematical Sociology* ou l'*International Network for Social Network Analysis*), la simulation sociale (ainsi le *Journal of Artificial Societies and Social Simulation* fondé à la fin des années 90) ou bien la modélisation des systèmes complexes sociaux (qui se développe fortement au début des années 2000, en interaction avec l'informatique, les mathématiques appliquées et la physico-statistique). On voit ainsi émerger, depuis quelques années, des revues et conférences mentionnant le terme *Computational Social Science*, voire *Social Computing* ou encore *Social Informatics*. Empiriquement, ces approches s'appuient essentiellement sur des corpus de données qui sont soit générés dans le cadre de

1. Van den Bos M. and Giffard H. 2016, Mining public discourse for emerging Dutch nationalism, in *Digital Humanities Quarterly*.

Nicholson B. 2013, Exploring the methodological possibilities of digital newspaper archives, in *Media History*, 19 (1): 59–73.

2. Warren C. N., Shore D., Otis J., Wang L., Finegold M. and Shalizi C. 2016, Six degrees of francis bacon: A statistical method for reconstructing large historical social networks, in *Digital Humanities Quarterly*, 10 (3).



l'acception précédente (numérisation d'archives), soit disponibles directement au format numérique car issus de l'enregistrement de traces comportementales à grande échelle (les fameuses *Big data*). Les plateformes et communautés en ligne d'internet jouent ici un rôle prépondérant.

Cela nous amène à la troisième acception (*humanities of the digital*) liée à l'étude des mondes numériques. Ce type de terrain a lentement acquis une reconnaissance académique en étant progressivement considéré comme un contexte parmi d'autres d'expression de la vie sociale. Il s'agit à la fois de comprendre la vie quotidienne au sein des communautés virtuelles et l'intégration des outils de communication numérique dans la vie quotidienne.

À cet égard, les communautés en ligne peuvent être appréhendées de deux manières :

- comme un terrain d'étude en tant que tel, avec des groupes sociaux possédant leurs propres codes et leur propre grammaire d'interaction, configurés par un dispositif socio-technique particulier. C'est le cas de Wikipedia pris comme collectif de production de savoirs auto-organisés, avec sa propre bureaucratie ;
- comme une approximation (*proxy*) de la société toute entière. Ainsi, Twitter ou Facebook sont considérés comme des instances plus ou moins représentatives d'un certain espace public.

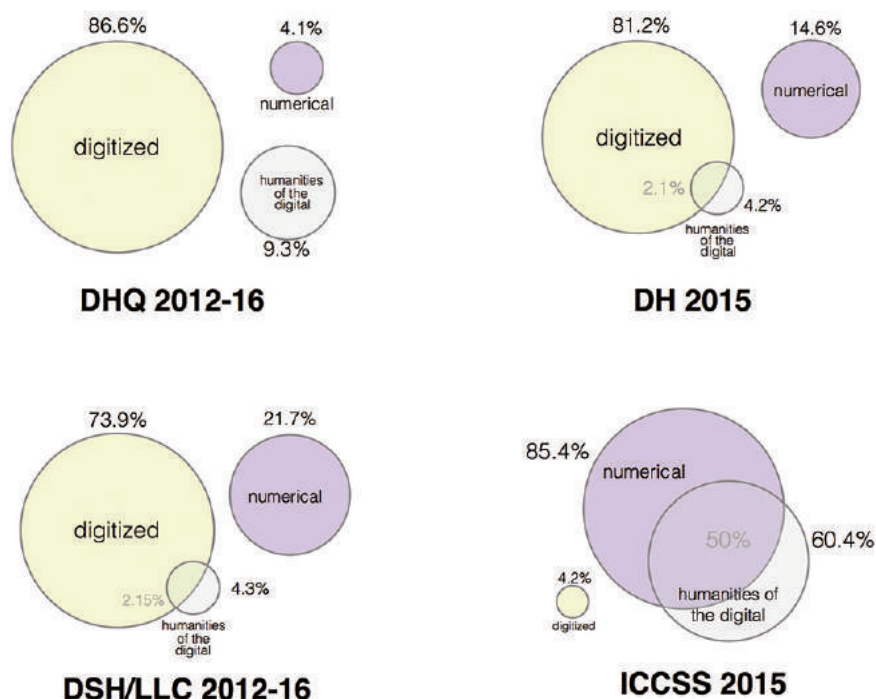
Cette acception semble orthogonale aux deux précédentes en ce qu'elle a un objet délimité et bien distinct, celui des interactions homme-machine ou des interactions humaines par le biais de machines. Néanmoins, elle touche aussi au cœur de ce qui fonde les *digitized* et *numerical humanities* : d'une part, parce que les plateformes apparaissent comme une source naturelle de corpus sous forme numérique (comme l'illustre, par exemple, le numéro spécial de *Computational Linguistics* de 2003 dédié au

web en tant que corpus textuel) ; d'autre part, parce que celles-ci semblent fournir l'essentiel des données exploitées par les modélisateurs<sup>3</sup>. L'interface entre les trois acceptions peut se prolonger lorsqu'il s'agit de comprendre les usages liés à la mise en ligne des archives numériques ou propres aux dynamiques virtuelles des communautés scientifiques.

Les humanités du numérique sont ainsi susceptibles de constituer un terrain de jeu intéressant chacune des deux premières acceptions, lesquelles possèdent en outre des points de vue distincts sur le rôle et le statut des archives et des corpus — particulièrement en regard de leur unicité (archives singulières vs bases de données répliquables à l'envi), de leur publicité (accès public à un bien commun vs difficulté de rendre publiques certaines données personnelles) et, plus largement, des phénomènes qu'il importe de mettre en évidence (descriptions idiographiques vs nomothétiques). Cette disjonction permet aussi d'expliquer, en partie, leur présence relative au sein des arènes de publication respectives : l'étude empirique de plus de deux cent cinquante études diversement affiliées à ces deux acceptions semble ainsi montrer qu'il existe deux communautés épistémiques peu liées (figure 1). L'une, explicitement affiliée à l'étiquette *digital humanities* (DH), accorde une place prépondérante à la première acception, alors que l'autre, sous la bannière *computational social science* (CSS), se nourrit principalement de la deuxième et de la troisième, qui se recouvrent par ailleurs significativement. Pour autant, il est possible de suggérer que les CSS pourraient bénéficier davantage de la variété des objets abordés par les DH et de leur connaissance de terrain, tandis que les DH pourraient profiter de l'attitude fermement nomologique et modélisatrice des CSS. Si les différentes communautés sont certes le produit de développements historiques distincts, leur attirail méthodologique partage en effet de nombreuses similitudes.

#### ► Référence

Hine C. 2004, Social research methods and the internet: A thematic review, in *Sociological Research Online*, 9 (2).



contact&info

► Camille Roth,  
CMB

roth@cmb.hu-berlin.de

Figure 1: Diagramme représentant la part des différentes acceptions au sein des études publiées dans des arènes faisant explicitement référence à l'étiquette « digital humanities » (revues DHQ et DSH/LLC de 2012 à 2016 et conférence DH à Sydney en 2015) ou à « computational social science » (conférence ICCSS à Helsinki en 2015).

3. Lazer D. and Radford J. 2017, Data ex machina: Introduction to big data, in *Annual Review of Sociology*, 43: 19–39.

# L'histoire dans un monde de données

Frédéric Clavert est historien, membre du Luxembourg Centre for Contemporary and Digital History (C<sup>2</sup>DH, Université de Luxembourg). Ses recherches portent sur trois grands domaines : l'histoire de l'organisation économique et monétaire du continent européen aux <sup>xx</sup><sup>e</sup> et <sup>xx</sup><sup>e</sup> siècles, la mémoire collective du passé à l'ère des grandes données et les relations entre les historiens et leurs sources primaires numériques.



sion de la conférence *Digital Humanities 2014*. Ce développement rapide des Humanités numériques, en un peu plus d'une décennie, peut être relié à trois phénomènes plus larges touchant nos sociétés : la mise en réseau (Internet et le web) ; la mise en données (la numérisation de nombreux livres, articles, archives, l'apparition de traces massives « nées numériques », etc.) ; la diffusion des ordinateurs individuels et, surtout, des *smartphones*.

Les sciences historiques ne peuvent éviter l'intégration de ces grandes évolutions. Historiens et historiennes n'ont d'ailleurs pas été en reste sur l'informatique et le numérique depuis la parution du premier article dans une revue francophone majeure évoquant l'usage d'un ordinateur<sup>6</sup>. Les années 1960, grande période de l'histoire quantitative sur données sérielles, ont abouti au fameux constat d'Emmanuel Le Roy Ladurie : « L'historien de demain sera programmeur ou il ne sera pas »<sup>7</sup>. Si l'arrivée de la micro-informatique a mis du temps à être digérée, l'usage de l'informatique par les historiens et historiennes reste bien présent, comme en atteste la création de la revue *Histoire & Mesure*<sup>8</sup>. Toutefois, on ne commence à parler de *digital history* que dans les années 1990, avec des projets liés plutôt à l'histoire publique — c'est-à-dire une histoire destinée à un large public — comme *Valley of the Shadows*. L'expression, traduite en « histoire numérique », fait florès en France dans les années 2010<sup>9</sup>.

Nous allons ici nous pencher sur un angle particulier de l'histoire numérique : celui des archives. Matières premières de l'historien, les archives ou sources primaires sont les traces laissées dans notre présent par le passé : issues des administrations, des entreprises, des organisations, familiales, publiques ou privées, elles sont le cœur du métier d'historien, sa matière première.

Que change l'ère numérique aux sources primaires des historiens ? Qu'elles soient numérisées ou nées numériques, les archives sont désormais mises en données et mises en réseau. Abondantes ou rares, elles peuvent faire l'objet de nouveaux modes de lecture. Si ces deux points ouvrent de grandes possibilités de recherche, ces dernières ne doivent pas occulter néanmoins les ombres projetées par la mise en données et l'ordre illusoire qu'elle peut instaurer.

## La mise en données des sources primaires

Mettre en données revient à transformer un phénomène en un format quantifiable à des fins d'analyse par le nombre<sup>10</sup>. Si on

Depuis la parution de *A Companion to Digital Humanities*<sup>1</sup> en 2004, le développement des Humanités numériques, champ interdisciplinaire à la croisée des sciences informatiques et des sciences humaines et sociales, a été très rapide. Dans le cas français<sup>2</sup>, un premier THATCamp (*The Humanities and Technology Camp*)<sup>3</sup> est organisé en 2010 à Paris : l'événement, précédé par la création de la *liste de diffusion [DH]*<sup>4</sup>, se termine par l'écriture collective d'un *Manifeste des digital humanities*<sup>5</sup> consacrant la naissance d'une communauté francophone des Humanités numériques. Quatre ans plus tard, à Lausanne, l'association francophone pour les Humanités numériques, *Humanistica*, voit le jour à l'occa-

1. Siemens R., Unsworth J. et Schreibman S. 2004, *A Companion to Digital Humanities*, Blackwell Publishing Professional.
2. La première conférence internationale appelée *Digital Humanities* a été organisée à la Sorbonne en 2006.
3. Les THATCamps ont été créés par le Roy Rosenzweig Center for History and New Media en 2008.
4. Cette liste de diffusion est toujours active, comme service de l'association Humanistica.
5. Collectif, « Manifeste des Digital humanities », 2010.
6. Furet F. et Daumard A. 1959, « Méthodes de l'Histoire sociale : les Archives notariales et la Mécanographie », in *Annales ESC* 14 (4) : 676-693.
7. Le Roy Ladurie E. 1968, « La fin des érudits », in *Le Nouvel Observateur*. Le Roy Ladurie ne parlait néanmoins que de l'histoire quantitative.
8. Genet J.-P. 1986, « Histoire, Informatique, Mesure », in *Histoire & Mesure* 1 : 7-18.
9. Clavert F. et Noiret S. 2013, L'histoire contemporaine à l'ère numérique, P.I.E.-Peter Lang.
10. Voir le dossier « Le métier d'historien à l'ère numérique : nouveaux outils, nouvelle épistémologie ? », in *Revue d'histoire moderne et contemporaine*, 2011-4bis.
10. Mayer-Schönberger V. et Cukier K. 2013, *Big data: a revolution that will transform how we live, work, and think*, Houghton Mifflin Harcourt, p. 725.



peut parler de mise en données avant l'arrivée de l'informatique, l'ère numérique lui donne une portée sans précédent. Dans le cas des sciences humaines et sociales, les projets de numérisation massive comme *Gallica* (BNF) donnent aujourd'hui un accès à d'immenses corpus de sources primaires directement en ligne. À ces programmes de numérisation, s'ajoute la production de sources primaires nées numériques. Les archives du web en sont, pour le moment, l'exemple le plus souvent cité, en attendant la possibilité d'exploiter les archives nées numériques des organisations publiques et privées, soumises à un délai avant consultation.

Grâce à ces données, l'originalité de notre époque est de pouvoir ainsi constituer des corpus importants de sources primaires numérisées ou nées numériques, en fonction des problématiques que les chercheurs et chercheuses souhaitent investiguer. La mise en données de ces sources, leur volume important, si ce n'est massif, pose néanmoins un problème fondamental : comment lire des sources primaires tellement nombreuses que l'historien ne peut humainement les lire ?

## De nouveaux modes de lectures des sources primaires

Pour faire face à ce défi, pour investiguer les possibilités qu'offrent de nouvelles méthodes, les historiens, notamment, ont répondu par la notion de lecture distante. Si l'on doit l'expression à Franco Moretti, d'autres chercheurs — tels Hervé Le Bras ou Claire Lemerrier — ont utilisé le même type de méthodes<sup>11</sup> à la même époque et, pour certains, dès les années 1960<sup>12</sup>. La lecture distante revient à demander à l'ordinateur de lire les sources primaires mises en données pour nous, grâce à des méthodes informatiques et statistiques très variées. L'utilisation de logiciels comme *IRaMuTeQ* ou *Gephi* rentrent dans ce cadre : le premier permet de déduire d'un corpus de texte des « profils » permettant de saisir les grands thèmes qui y sont abordés ; le second engendre des visualisations fondées sur la sociologie des réseaux sociaux et, ainsi, une analyse des liens entre des « entités » (personnes, lieux, monuments, concepts, etc.) présentes dans les sources historiques.

Cette lecture distante, toutefois, ne peut se passer de la lecture proche. Cette dernière est, d'une certaine manière, le mode classique de lecture en histoire : l'analyse « humaine » des archives, de leur contenu et de leur cohérence. Ces deux modes de lectures sont interdépendants, les allers et retours entre l'un et l'autre étant souvent nécessaires pour comprendre le corpus ainsi créé.

Mise en données et lecture distante sont aussi transformatrices des sources : les projets *Numapresse* et *Impresso* montrent que la numérisation massive de la presse et son analyse par divers algorithmes de fouille de texte transforment cette source numérisée en source nouvelle, au sens où elle nous autorise à opérer des recherches inédites, par exemple sur la circulation de l'information au XIX<sup>e</sup> siècle.

## Un ordre illusoire ?

Travailler sur des gigaoctets de données historiques peut donner l'impression d'avoir affaire à l'intégralité des sources sur un sujet donné, sources bien ordonnées dans des bases surorganisées. Ce sentiment d'ordre est un effet d'illusion qui masque ce qui n'est pas numérisé<sup>13</sup>, ces sources que nous risquons de ne plus consulter car elles ne sont pas (encore) en ligne.

Un autre risque touche aux inégalités d'accès aux données : si *Google Books* ou *Gallica* sont, partiellement ou pleinement, accessibles en ligne, de nombreuses sources numérisées ne le sont qu'auprès d'opérateurs commerciaux, souvent très chers. Le risque d'une division entre *has* et *has not* du numérique dans les sciences historiques est réel. À ces inégalités d'accès s'ajoutent une inégalité en termes de production des données, donc des sources historiques futures<sup>14</sup> : il existe aujourd'hui des régions produisant peu de données, qui seront faiblement documentées dans le futur.

## Pour une histoire numérique critique et réflexive

Face à ces enjeux méthodologiques — et ils ne sont pas tous détaillés ici —, l'histoire numérique a aujourd'hui besoin de ce qui est le socle méthodologique des sciences historiques : esprit critique et réflexivité, qui nous permettront de renouveler la critique des sources et de leur mise en données et d'y associer celle des algorithmes utilisés par nos logiciels. Cette critique nécessite une réflexion sur notre propre travail, et, notamment, sur notre relation à l'archive. Si Arlette Farge<sup>15</sup>, il y a près de 30 ans, a très bien su saisir cette dernière, il est temps aujourd'hui de l'interroger à nouveau en questionnant le goût de l'archive à l'ère numérique<sup>16</sup>.

contact&info

► Frédéric Clavert,  
C<sup>2</sup>DH  
frederic.clavert@uni.lu

Crédit image de fond : La « galaxie de la recherche » en Suisse. 45 000 chercheurs connectés s'ils ont participé ensemble à un projet financé par le FNS entre 2006 et 2015  
© Martin Grandjean. Tiré de Grandjean M. 2017, « Complex Network Visualisation for the History of Interdisciplinarity: Mapping Research Funding in Switzerland », in *Digital Humanities 2017* : 711-713.

11. Moretti F. 2007, *Graphs, Maps, Trees: Abstract Models for Literary History*, Verso.

Le Bras H. 2000, *Essai de géométrie sociale*, Odile Jacob, 2000.

Lemerrier C. 2005, « Analyse de réseaux et histoire », in *Revue d'histoire moderne et contemporaine* (2) : 88–112.

12. Garelli P. et Gardin J.-C. 1961, « Étude par ordinateurs des établissements assyriens en Cappadoce », in *Annales ESC* 16 (5) : p. 837-876.

13. Milligan I. 2013, « Illusory Order: Online Databases, Optical Character Recognition, and Canadian History, 1997–2010 », in *Canadian Historical Review* 94 (4) : 540-569.

Putnam L. 2016, « The Transnational and the Text-Searchable: Digitized Sources and the Shadows They Cast », in *The American Historical Review* 121 (2) : 377-402.

14. C'est l'une des interrogations mises en avant dans Manning P. 2013, *Big data in history*, Palgrave Pivot.

15. Farge A. 1997 (1989), *Le goût de l'archive*, Seuil.

16. Voir le projet collectif en cours autour du goût de l'archive à l'ère numérique, initié par Caroline Muller et Frédéric Clavert.



## Humanités numériques et études littéraires : une question d'interfaces

Professeur à l'Université de Yale, Christophe Schuway est spécialisé dans la littérature et la culture de la France du XVIII<sup>e</sup> siècle. Ses recherches portent principalement sur le marché du livre, la mode et le marketing à l'ère classique, la circulation de l'information, la hiérarchisation et les transferts culturels (notamment entre l'Angleterre et la France). Il est également actif dans le domaine des sciences humaines numériques et est l'auteur d'un ouvrage à paraître intitulé *Interfaces : une approche concrète des humanités numériques littéraires* (2019).

*The Digital Humanities Bust*, littéralement *Le bide des humanités numériques*, est le titre d'un pamphlet paru l'année dernière dans les *Chronicles of Higher Education*, une nouvelle attaque contre un adversaire à la mode. En Europe ou aux États-Unis, des articles mêlent ainsi régulièrement scepticisme bon teint et philosophie pour dénoncer le dangereux gaspillage de ressources que provoqueraient les humanités numériques : « Après une décennie d'investissements et de battage médiatique, qu'est-ce que le champ a accompli ? Pas grand-chose<sup>1</sup> ».

De la part d'un professeur de littérature, la déclaration est surprenante. L'informatique a radicalement modifié les pratiques du domaine, transformant aussi bien les questions posées que les manières d'y répondre. De quoi cet article est-il alors le symptôme ? Peut-être d'un problème de définition. Les humanités numériques convoquent encore un imaginaire bien éloigné des apports concrets de l'informatique aux études littéraires.

### Pour une autre perspective

L'expression « humanités numériques » évoque souvent un mélange flou de graphiques spectaculaires, d'analyses automatisées aux résultats abscons ou évidents et de dissertations sur les promesses du *big data*. De fait, plusieurs spécialistes en la matière limitent le champ à un certain type de pratiques et d'approches, favorisant particulièrement celles qui compensent le complexe d'infériorité de la littérature face aux Sciences/Technologie/Informatique/Mathématiques. Une telle définition est féconde en articles théoriques et déclarations tonitruantes. Elle attire également l'attention des chercheurs et des organismes de financement.

Elle occulte toutefois les réalisations et les apports concrets de la discipline : les vingt dernières années ont vu se développer pléthore de ressources numériques qui modifient profondément la recherche en littérature. En apprécier la richesse et l'intérêt nécessite toutefois d'adopter un autre point de vue. L'apport du numérique à la littérature réside bien moins dans des fantasmes d'interprétations automatisées que dans son rôle le plus fondamental, à savoir, celui d'interface. En informatique, ce terme désigne principalement la manière dont sont représentées les données ainsi que les outils (ou points d'accès) permettant leur manipulation. Cette représentation revêt un enjeu tout particulier pour les études littéraires : les interfaces proposent de nouvelles expériences de lecture, suscitent des découvertes inattendues et inspirent d'autres approches de la littérature.

Pour beaucoup, cette question de l'interface demeure pourtant un enjeu purement esthétique (au sens le plus réducteur du terme), par opposition aux données elles-mêmes. Le soin apporté à la (re)présentation de ces dernières serait une distraction, un tribut payé au confort de l'utilisateur. C'est oublier que, les données n'étant qu'une suite de 1 et de 0, leur représentation suppose nécessairement une interprétation programmée par l'être humain, ce que Johanna Drucker formule dans *Graphesis* : « Les données n'ont pas de formes visuelles inhérentes qui pré-détermineraient leur représentation graphique<sup>2</sup>. »

Que l'on travaille avec un terminal en mode texte ou dans un environnement graphique stylisé, il s'agit toujours d'une interface et aucune représentation n'est plus vraie (ou moins illusoire) qu'une autre. Dans notre société qui se méfie des images (l'allégorie de la caverne n'est jamais loin<sup>3</sup>), il sera toujours bien vu de se méfier des dispositifs graphiques. En l'occurrence, l'illusion consiste précisément à croire que l'on peut faire des humanités numériques sans penser à leur représentation et aux outils permettant aux chercheurs d'y accéder. Ces dernières années, des chercheurs en *design* ont heureusement redonné une place centrale à l'interface, en exposant son importance et son potentiel ; d'autres ont souligné les dangers latents qui pèsent sur cette question, sans toutefois proposer d'alternatives.

Certes, penser les choses à partir de l'interface peut paraître contre-intuitif puisque la représentation des données présuppose leur disponibilité. Invertissons la perspective : ces données peuvent bien exister, tant qu'on ne sait pas comment les représenter, elles ne servent à rien. Faisons un pas de plus : la réflexion sur l'interface doit intervenir au tout début d'un projet. Il ne s'agit certainement pas de négliger les défis techniques et politiques, en particulier, la pérennité et l'interopérabilité des données<sup>4</sup>, mais ces considérations ne peuvent prendre le pas sur ce que sera, concrètement, le projet, y compris lorsque celui-ci est expérimental.

Le refus de prendre la question des interfaces au sérieux n'est pas sans conséquence : en France (mais aussi en Suisse), les organismes chargés de pérenniser les ressources numériques privilégient largement la sauvegarde des données aux dépens de celle des interfaces, ces dernières étant plus complexes à entretenir. On se soucie de conserver le texte d'une édition en ligne, mais pas la manière dont il apparaît à l'écran, bien que le dispositif imaginé par les concepteurs d'une édition critique (mise en scène du texte, appareil critique, interactions, etc.) soit porteur de sens, au même titre que le texte (les données) lui-même.

1. « After a decade of investment and hype, what has the field accomplished? Not much ». Brennan T. 2017, « The Digital Humanities Bust », in *Chronicles of Higher Education*.

2. « All information visualizations are metrics expressed as graphics. Visualizations are always interpretations – data does not have an inherent visual form that merely gives rise to a graphic expression », Drucker J. 2014, *Graphesis. Visual Forms of Knowledge Production*, Harvard University Press, p. 7.

3. Voir Vial S. 2013, *L'Être et l'écran, comment le numérique change la perception*, PUF, chap. 1, « La technique comme système ».

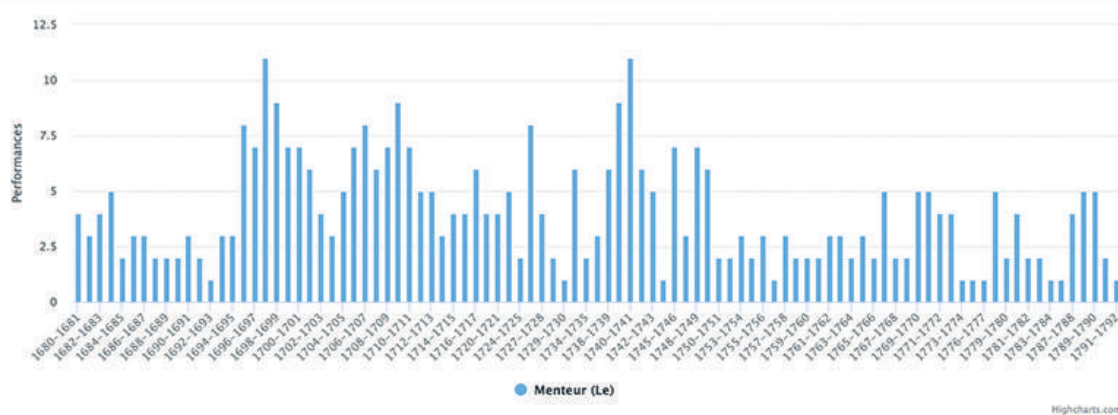
4. L'espace d'un article ne permet pas d'en traiter, mais les API sont également une interface, de machine à machine, permettant par ailleurs de multiplier les interfaces en fonction des cas d'utilisation.



## Menteur (Le)

Auteur  
Corneille  
(Pierre)Discours  
versActes  
5Genre  
comédieMusique /  
Danse /  
Machie  
nonPrologue  
oui

## PERFORMANCES (381)



Visualisation tirée d'une base de données sur l'histoire de la Comédie-Française. Cette interface offre une vision instantanée de la fortune d'une pièce durant le XVIII<sup>e</sup> siècle, invitant tous les chercheurs à intégrer ce type d'informations à leurs réflexions © Cfregisters.org

## Trois apports concrets

Lorsque l'on mesure le potentiel des interfaces, chaque réalisation numérique devient l'occasion de construire un nouveau rapport aux objets de recherche, de susciter d'autres lectures et d'autres questions. Ainsi que l'écrit Stéphane Vial, chaque projet est l'occasion « d'imaginer des régimes d'expérience eux-mêmes inédits [...]. La tâche du design numérique est donc celle de l'exploitation créatrice [de l'interface] en vue de créer [...] des effets d'enchantement factitif de l'expérience<sup>5</sup> ».

On pourrait classer les apports du numérique en trois catégories.

L'interface des **éditions en ligne** suscite de nouvelles lectures des textes et invite à repenser l'organisation de l'appareil critique. Mentionnons d'abord les bibliothèques en ligne telles que *Gallica*, *Google Books* ou *Open library*. En donnant accès à une reproduction photographique de l'ouvrage, même imparfaite, elles sont à l'origine d'une évolution épistémologique majeure, en particulier pour les siècles anciens. Le lecteur multiplie la consultation d'ouvrages dans leurs mises en pages d'origine et profite des indications données par cette matérialité, des informations que les éditions critiques modernes occultent. En outre, la consultation quasi instantanée de reproductions photographiques de livres, la possibilité de passer d'un document à l'autre sans bouger de sa chaise, tout cela révèle des phénomènes d'ensemble

fondamentaux passés inaperçus. La numérisation et la recherche plein texte à l'intérieur d'immenses corpus constituent par ailleurs des avancées significatives, chaque recherche révélant potentiellement des ouvrages oubliés et des sources inconnues. Enfin, la conception d'une interface est l'occasion d'inciter à d'autres modes de consommation du texte. La présentation d'une édition numérique d'un livre ancien peut être l'occasion d'inciter le lecteur à redécouvrir l'ouvrage en adoptant des pratiques de lectures pour lesquelles l'ouvrage était conçu.

Les **bases de données** constituent des outils de recherche incontournables, fournissant instantanément des réponses qui exigeaient autrefois des mois de travail. Pour connaître le succès d'une pièce de théâtre au fil des siècles, il n'est plus nécessaire de compiler de lourds registres papier, il suffit d'interroger la base correspondante. L'interface revêt ici un rôle fondamental puisqu'elle décide des interrogations possibles et, donc, de l'utilité des données. Que l'on pense à un catalogue de bibliothèque : au format papier, il servait à localiser des livres. Au format numérique, les différentes interrogations possibles en font un outil capable de répondre à des questions de littérature (par exemple, la vogue d'un sujet sur une période donnée). La manière dont sont présentés les résultats n'est pas non plus sans conséquence. Chaque recherche mêle potentiellement des sources mineures ou oubliées à des références bien connues, relativisant ainsi le canon et suscitant des découvertes aussi inattendues que fructueuses.

5. Vial, *op. cit.*, chap. 6.

Graph these comma-separated phrases:  ☐ case-insensitive

between 1650 and 2000 from the corpus French with smoothing of 3 Search lots of books



L'outil Ngram permet de visualiser la fréquence d'un terme dans un large échantillon de textes, suscitant de nouvelles pistes de recherche  
© Google

Enfin, les **graphiques et approches quantitatives** de la littérature constituent également des interfaces : il s'agit de visualisations du texte qui permettent d'aborder un texte ou un ensemble de textes sous un autre angle, à l'instar du *distant reading* de Franco Moretti. Qu'elles soient quantitatives ou

relationnelles, les représentations visuelles offrent d'autres manières d'approcher les corpus et les collections et de déceler des phénomènes difficiles à observer en lecture intensive. De manière générale, la puissance de calcul invite à multiplier les interrogations et à explorer des hypothèses nouvelles... pour peu que les données existent, certes, mais surtout que l'interface le permette.

L'apport des humanités numériques aux études littéraires tient ainsi fondamentalement aux nouveaux rapports que les interfaces construisent avec nos objets, soit parce qu'elles créent de nouvelles expériences ou qu'elles offrent de nouvelles visualisations, soit parce qu'elles nous permettent d'accéder à des informations complexes. Les apports des humanités numériques à la littérature sont en réalité omniprésents : ils se dissimulent dans les questions posées, dans l'ampleur des corpus explorés et dans la multiplicité des approches et des données utilisées.

## L'ALCORAN DE LOUIS XIV

1695

Édition critique en ligne d'un pamphlet anti-Louis XIV

Présentation du projet Présentation de l'œuvre Télécharger le texte

### Dialogue des morts entre



Innocent XI  
Pape



Mazarin  
Cardinal

Rencontrant le cardinal Mazarin sur les bords du fleuve Styx, lui parle ainsi.

Mazarini, Mazarini, est-ce vous que je vois dans ces sombres lieux ? Ne me trompe-je point ? Non, c'est lui-même ! Le voici qui vient à moi.

Chi sei vostra signoria ? Qui êtes-vous ?

Il santissimo Padre Odescalchi, Innocentio XI.

O ! Il patrone, Santissimo Padre Innocentio XI. O meraviglia ! Ô merveilles ! Ma è egli vero ? mais est-il bien vrai ?

Si Signore, per certo.

Ha ! Santissimo et benévole Padre Odescalchi, ben venuto ! Soyez donc le bienvenu, dans ces contrées noires. Il y a près de 30 ans que je vous attends, avec l'impatience du plus malheureux de tous les damnés.

Per Dio. Il s'est bien passé des choses à Rome depuis votre départ du bois de Vincennes, car Clement X que vous avez connu et qui était de vos bons amis, ayant laissé par sa mort la Chaire Apostolique vacante, je fus élu pour lui succéder sous le nom d'Innocent XI. Mais ce ne fut pas sans

### contact&info

► Christophe Schuwey,  
Yale University  
christophe.schuwey@yale.edu

Édition critique d'un pamphlet du XVIII<sup>e</sup> siècle. L'interface souligne la théâtralité et l'humour du texte original © Kathrina LaPorta / Dartmouth College.



## Huma-Num : au-delà des humanités numériques, une infrastructure pour les sciences humaines et sociales

Olivier Baude et Stéphane Pouyllau sont respectivement directeur et directeur technique de la très grande infrastructure de recherche (TGIR) *Huma-Num* (UMS3598, CNRS / AMU / Campus Condorcet). Olivier Baude est professeur de sciences du langage ; il mène des recherches dans le domaine de la linguistique variationniste sur corpus et des humanités numériques. Stéphane Pouyllau est ingénieur de recherche au CNRS, spécialisé dans les domaines des humanités numériques, de l'informatisation des données de la recherche en sciences humaines et sociales et en information scientifique et technique. La TGIR Huma-Num repose sur une équipe d'une quinzaine d'ingénieurs hautement spécialisée.



Utilisation de la 3D en histoire : la restitution du travail à la chaîne chez Renault en 1924 © Stéphane Pouyllau

### La longue histoire des infrastructures pour la recherche en sciences humaines et sociales

Créée il y a cinq ans et héritière des infrastructures Adonis et IR-Corpus, la très grande infrastructure de recherche (TGIR) Huma-Num est aujourd'hui un succès. Elle propose une organisation originale, qui s'appuie largement sur les communautés des sciences humaines et sociales, et déploie un ensemble de services numériques raisonnés offrant des moyens de stockage, de traitements et de diffusion de corpus, jeux de données, collections de documents et bases de données. Huma-Num est aujourd'hui l'un des piliers des humanités numériques en France et dans le monde. La TGIR développe ses activités au sein de plusieurs réseaux d'infrastructures européennes (DARIAH, CLARIN et demain OPERAS), vers les pays d'Amérique du sud, avec le Canada au niveau de la Francophonie et *via* le réseau des Écoles Françaises à l'étranger.

Le pari était de taille quand, entre 2005 et 2007, le CNRS (et alors son département SHS) et le ministère de l'enseignement supérieur et de la recherche lancèrent plusieurs initiatives tels que les centres de ressources numériques ou le très grand équipement

Adonis qui furent les bases de la structuration du paysage national des infrastructures pour les humanités numériques. Le développement de ce type d'infrastructures nécessite des temps longs car les différentes disciplines des sciences humaines et sociales ne s'approprient pas de la même façon, ni au même rythme, les enjeux, les méthodes et les outils du numérique.

Une rapide approche épistémologique permet de mieux cerner l'importance des infrastructures dans le tournant numérique des SHS à l'heure actuelle.

Si le concept d'humanités numériques peut paraître récent, il est néanmoins lié à une longue tradition des instruments d'aide à la recherche : dès 1963, avec la Fondation Maison des sciences de l'homme ; dans les années 1960-1970, à travers les grandes bibliothèques thématiques (IRHT, CEGET) ; dans les années 1970-1980, sur les grands plateaux tels que la Maison de l'Orient méditerranéen ancien, désormais *Maison de l'Orient et de la Méditerranée* (MOM, FR3747, CNRS / Université Lumière Lyon 2), fondée en 1975 par Jean Pouilloux. Ces instruments se sont développés au sein d'une révolution technologique. Si la micro-informatique et la publication assistée par ordinateur ont pu, au cœur des



La modélisation des connaissances en sciences humaines et sociales © Stéphane Pouyllau

années 1980-1990, donner l'illusion que l'informatique serait désormais gérable à l'échelle individuelle, la diversification des méthodes et outils, encouragée par l'invention du Web au début des années 1990, a rapidement montré qu'elle nécessitait des infrastructures stables. Le câblage en haut débit et les services de base de l'internet furent les premières d'entre elles : les e-mails, les protocoles d'interrogation des données, puis le Web ont été les briques indispensables pour la diffusion et l'appropriation par les enseignants-chercheurs de ces outils de travail. C'est sur ce socle qu'ont été construites, collectivement, les secondes : les infrastructures de services pour les humanités numériques.

En l'état actuel, les données réutilisables et « interopérables » — c'est le terme consacré — sont encore trop limitées et les efforts sont principalement orientés vers l'ouverture des publications scientifiques finalisées (HAL-SHS, OpenEdition Center, Persée, Cairn, etc.). Or, depuis plus de dix ans, l'explosion de la numérisation des programmes de recherche en SHS, initiée par la numérisation des bibliothèques, puis étendue à l'ensemble des méthodes et outils de la recherche, a provoqué un tournant numérique qui fait évoluer profondément les rapports aux données des enseignants-chercheurs, doctorants, étudiants. Si l'effort a été puissant pour la publication électronique, les archives ouvertes et, en général, pour les résultats de la recherche, les corpus de sources primaires et secondaires, les archives, les fonds de bibliothèques de recherche et de laboratoires et, surtout, les données utilisées par les enseignants-chercheurs n'ont pas été traitées avec autant d'efficacité et de réussite. Toutefois, les difficiles expériences des années 2005-2015 ont montré l'importance de l'appropriation par les communautés des questions liées aux formats ouverts, aux méthodes d'interopérabilité et aux standards pour l'échange et la publication des données et corpus de documents. De fait, une prise de conscience a émergé dans les communautés scientifiques sur la nécessité d'utiliser ou de promouvoir l'utilisation de standards scientifiques et techniques

dans la constitution de jeux de données. C'est dans ce contexte qu'Huma-Num est née en 2013.

## Huma-Num au service de la recherche

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

Elle structure, par l'intermédiaire de consortiums regroupant des acteurs du monde de la recherche et d'un réseau de points de présence dans les maisons des sciences de l'Homme (MSH), l'accompagnement des communautés scientifiques SHS en matière d'infrastructure numérique pour les données de la recherche. Ancrée au cœur des humanités numériques, elle met en œuvre une infrastructure permettant aux communautés SHS de développer, de réaliser et de préserver sur le long terme les résultats des 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 »<sup>1</sup> 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).

À cette mission principale, s'ajoute une mission de prospective sur l'évolution des besoins technologiques et documentaires des activités de recherche en SHS et le développement de services numériques visant à accompagner, par l'innovation, les méthodes de recherche en SHS. Dans le cadre de cette mission, la TGIR prospecte dans le domaine stratégique de l'utilisation et de la réutilisation des données de la recherche, comme « sources » et comme productions scientifiques, dans les publications scientifiques, en interaction avec les acteurs SHS actifs dans ce

1. FAIR est l'acronyme anglais de « *Findable, accessible, interoperable and reusable* » : trouvables, accessibles, interopérables et réutilisables.



domaine. Enfin, la TGIR Huma-Num porte et coordonne la participation française dans deux infrastructures européennes de types ERIC (*European Research Infrastructure Consortium*). L'ERIC DARIAH (*Digital Research Infrastructure for the Arts and Humanities*) couvre toutes les disciplines des SHS et du patrimoine prises dans un sens large. L'ERIC CLARIN (*Common Language Resources and Technologies Infrastructure*) se consacre aux ressources du langage. La TGIR est également engagée dans différents projets européens financés dans le cadre du programme Européen Horizon-2020. Huma-Num suit les développements du projet EOSC (*European Open Science Cloud*) et est également engagé dans des coopérations bilatérales hors d'Europe (Canada, Argentine,...).

Son infrastructure de services numériques accueille aujourd'hui un très grand nombre de projets, plus de 300 programmes de recherche, qui y développent leurs jeux de données (plus de 500 à ce jour), y publient leurs corpus scientifiques sous différentes formes (sites web, etc.). De nombreux projets emploient ses outils de traitement des données (des environnements R aux outils de photogrammétrie) et plus de deux mille enseignants-chercheurs utilisent ses services de stockage ; c'est plus d'un péta-octet (1000 téraoctets) qui y est actuellement stocké. L'une des spécificités d'Huma-Num est de proposer à la fois une infrastructure de base à destination des programmes de recherches — qui peut être vu comme un *cloud* pour les SHS (mais hébergé sur des serveurs publics et sur le territoire national) — et des services innovants. Dans un sens, la première partie de l'infrastructure peut être considérée comme un dispositif de première nécessité pour l'enseignant-chercheur et son programme de recherche. Ce dispositif doit néanmoins être d'une fiabilité à toute épreuve et répondre aux évolutions des pratiques des chercheurs. Ainsi, le développement d'usages de données volumineuses a conduit la TGIR Huma-Num à mettre récemment en place un service de stockage distribué et sécurisé, la Huma-Num Box, qui connaît un succès très rapide.

Mais Huma-Num offre aussi des services qui correspondent aux besoins des enseignants-chercheurs en les inscrivant dans les limites des disciplines académiques et scientifiques SHS et en proposant des dispositifs d'amélioration du partage et de la diffusion des données. Ce second volet est largement nourri par les travaux menés par les enseignant-chercheurs dans les différents consortiums d'Huma-Num et dans le réseau d'acteurs que forme les Maisons des sciences de l'Homme. C'est par exemple le rôle de NAKALA, le service numérique de stockage, de diffusion et de partage de documents et données numériques proposé par Huma-Num. NAKALA permet de documenter les fichiers avec des jeux de métadonnées issues des grands standards internationaux utilisés par les SHS à l'échelle mondiale. Des programmes de recherche tels que ceux du Centre Franco-Egyptien d'Etude des Temples de Karnak (CFEETK, USR3172, CNRS / Ministère des Antiquités d'Égypte) ou encore les corpus filmiques de l'Institut interdisciplinaire d'anthropologie du contemporain (IIAC, UMR8177, CNRS / EHESS) s'appuient sur NAKALA pour être publiés, diffusés et conservés. NAKALA accueille déjà plus de 100 000 données en tous types de formats standards.

Huma-Num assiste les programmes de recherche à toutes les étapes y compris lors de l'ébauche de projets. Le fait de définir ces corpus ou ces bases de données dès leur constitution, dans une logique de partage disciplinaire et de science ouverte, permet de partager bien plus que des données primaires. Ce sont les publications, les thèmes de recherche scientifiques, les

échanges au niveau international, qui en sont améliorés. Ainsi, signaler par le moteur de recherche ISIDORE non seulement ses publications mais aussi ses corpus et bases de données, et même les divers documents liés à un projet et déposés dans NAKALA, offre l'opportunité de maîtriser la diffusion de ses travaux et de les rendre très largement disponibles. ISIDORE, premier moteur de recherche académique en SHS au monde, est alimenté par plus de 6 000 sources de données provenant du monde entier. Accessible en trois langues, il est utilisé par plus de 800 000 personnes par an.

Huma-Num propose aujourd'hui une gamme complète de services d'aide à la recherche s'inscrivant pleinement dans le contexte de la science ouverte. L'annonce d'un plan national par le ministère de l'enseignement supérieur, de la recherche et de l'innovation a été l'occasion de le rappeler.

## Les défis d'Huma-Num et de ses partenaires

Les travaux menés depuis deux ans pour consolider Huma-Num permettent désormais aux sciences humaines et sociales de s'appuyer sur des services d'une très grande stabilité. L'un des enjeux des prochaines années porte sur la pérennisation et la réutilisation des données numériques dans le contexte d'une recherche s'ouvrant de plus en plus sur la société. Le mouvement *Open Science* touche l'ensemble de l'enseignement supérieur et de la recherche au niveau mondial et les SHS sont, en France, en bonne position pour réussir ce défi. Il faudra parallèlement permettre l'évolution des pratiques des chercheurs dans les relations entre outils et données numériques, dans une perspective de diffusion et de réutilisation. En effet, la publication d'une base de données ou d'un corpus enrichi n'est pas un processus comparable à une édition scientifique papier. L'obsolescence des systèmes informatiques rend impossible la cession du travail fait comme l'édition papier semble l'autoriser. Si les chercheurs font confiance — une fois les normes établies — aux éditeurs, bibliothécaires et archivistes pour garantir l'accès à la connaissance, il faut élaborer les solutions et les processus qui produiront ce même niveau de confiance pour les données numériques. C'est le principal défi de la TGIR Huma-Num qui doit mieux tenir compte des besoins et des méthodes de travail des communautés SHS, tout en proposant un écosystème qui dépasse la simple mise à disposition de « puissance numérique » (outils, services, accompagnement).

L'évolution des services de la TGIR doit ainsi rendre plus compréhensibles les contraintes liées aux documents numériques et les enjeux de l'interopérabilité dans une démarche scientifique. Il s'agit non seulement de faciliter les pratiques actuelles, mais aussi de mettre en œuvre une réelle accumulation de données disponibles.

Pour répondre à cet ambitieux objectif, la TGIR ne peut agir seule. Elle doit définir une stratégie d'actions tenant compte de sa situation singulière dans le paysage national, dans un contexte qui la place à l'échelon intermédiaire d'un paysage national pluriel et des grandes instances internationales.

### contact&info

► Olivier Baude

olivier.baude@huma-num.fr

Stéphane Pouyllau

stephane.pouyllau@huma-num.fr

Huma-Num



# UN CARNET À LA UNE



## La 1ère Guerre vue de Paris

À travers le carnet *La première guerre vue de Paris*, l'historienne Odile Gaultier-Voituriez propose une édition critique des lettres envoyées par Étienne de Nalèche, directeur du *Journal des Débats*, à son ami Pierre Lebaudy durant la première guerre mondiale.

Débutée en septembre 1914, cette correspondance privée composée de 1020 lettres dura jusqu'en 1919. Étienne de Nalèche n'y aborde que peu d'éléments de sa vie personnelle. Il livre plutôt à son correspondant une série de commentaires et d'anecdotes sur la vie politique parisienne et sur la vie du journal que ce dernier finance. Observateur bien informé, il raconte dans le détail les coulisses de la [politique intérieure](#), du [journalisme parisien](#), différents scandales comme l'[affaire Caillaux](#) ou les réactions parisiennes à l'évolution du conflit. Il commente aussi l'actualité politique étrangère telle que le [mouvement révolutionnaire russe de 1917](#) ou encore les [revendications de création d'un État juif autonome en Palestine](#).

Cette correspondance offre un témoignage précieux *via* le point de vue socialement situé qu'elle donne à voir : celui d'un aristocrate parisien, ancien diplomate, journaliste au cœur des réseaux de pouvoir s'adressant à un très riche industriel sucrier.

D'un point de vue historique, cette correspondance présente l'intérêt de constituer une source d'information sur le vécu de la guerre très différente des récits de poilus par exemple. Ce sont les réseaux de sociabilité politiques et mondains qui peuvent être reconstitués grâce à elle ; les cercles d'informations et d'influence de la vie politique.

Chacune des 1020 lettres est publiée le jour de son centième anniversaire. Les archives du carnet rendent donc compte de la régularité de cette correspondance dont 1917 et 1918 ont été les années les plus riches avec presque une lettre par jour. Le carnet de recherche permet aussi à l'auteur de mettre en avant son travail d'édition critique et d'annotation de ces lettres, fournissant des indications sur les personnes citées et leurs relations, les erreurs potentielles d'Étienne de Nalèche ou encore des précisions sur les événements évoqués dans les lettres. Illustrant parfaitement les complémentarités possibles entre le livre et le format blog, ce carnet constitue également le support de publication en ligne de la matière à l'origine de l'ouvrage d'Odile Gaultier-Voituriez *Chronique cachée de la Grande Guerre* (2017), que les lectrices et lecteurs intéressés par ce sujet pourront également consulter.

Céline Guilleux

### La 1ère Guerre vue de Paris

La correspondance d'Étienne de Nalèche, directeur du Journal des Débats, à Pierre Lebaudy, 1914-1919

ACCUEIL À PROPOS CRÉDITS SOURCES ET BIBLIOGRAPHIE



#### Accueil

Ce carnet consiste en une édition critique de la correspondance d'Étienne de Nalèche, directeur du Journal des Débats, à Pierre Lebaudy, industriel sucrier, pendant la Grande Guerre (1914-1919). Annoté et indexé, ce corpus de 1020 lettres offre un riche contenu d'intérêt général : politique intérieure, relations internationales et guerre. L'intérêt et l'apport majeurs de cette correspondance résident dans la richesse du récit et de l'analyse quotidienne de la vie parisienne pendant la guerre de 1914, vue par un observateur de qualité à la plume alerte et à la croisée de l'opinion et des réseaux de pouvoir. Il est complété par :

- un portrait des deux protagonistes : Nalèche, grand notable parisien, directeur d'un journal reconnu, républicain libéral lié aux milieux d'affaires et Lebaudy, riche industriel sucrier et financier gendarme du Journal ;
- une analyse du statut des lettres qui deviennent, à partir du retour de Lebaudy à Paris fin 1916, une chronique chargée de garder mémoire de la guerre ;
- une étude de l'important réseau de sociabilité de Nalèche, constitué par l'élite parisienne ;
- un éclairage de la vision du journalisme de Nalèche et de sa méthode de travail qui repose sur les informations, très riches et recoupées, qui lui arrivent par son réseau.



#### contact&info

- Odile Gaultier-Voituriez,  
Cevipof

[odile.gaultiervoituriez@sciencespo.fr](mailto:odile.gaultiervoituriez@sciencespo.fr)

- Pour en savoir plus

<https://naleche.hypotheses.org>  
<https://www.openedition.org/16237>

#### contact&info

- François Pacaud  
OpenEdition

[francois.pacaud@openedition.org](mailto:francois.pacaud@openedition.org)

- Pour en savoir plus

<http://www.openedition.org>  
<http://cleo.openedition.org>

# la lettre de l'InSHS

- **Directeur de la publication** François-Joseph Ruggiu
- **Directrice de la rédaction** Marie Gaille
- **Responsable éditoriale** Armelle Leclerc [armelle.leclerc@cnrs-dir.fr](mailto:armelle.leclerc@cnrs-dir.fr)
- **Conception graphique** Sandrine Clérisse & Bruno Roulet, Secteur de l'imprimé PMA
- **Graphisme Bandeau** Valérie Pierre, direction de la Communication CNRS
- **Crédits images Bandeau**  
© Photothèque du CNRS / Hervé Théry, Émilie Maj, Caroline Rose, Kaksonen
- **Pour consulter la lettre en ligne**  
[www.cnrs.fr/inshs/Lettres-information-INSHS/lettres-informationINSHS.htm](http://www.cnrs.fr/inshs/Lettres-information-INSHS/lettres-informationINSHS.htm)
- **S'abonner / se désabonner**
- **Pour accéder aux autres actualités de l'InSHS**  
[www.cnrs.fr/inshs](http://www.cnrs.fr/inshs)
- **Retrouvez l'InSHS sur Twitter** [@inshs\\_cnrs](https://twitter.com/inshs_cnrs)

**Institut des sciences humaines et sociales CNRS**

• 3 rue Michel-Ange 75794 Paris cedex 16 •

ISSN : 2272-0243