



HAL
open science

Où sont les données de la recherche ?

Cécile Delay-Artous

► **To cite this version:**

Cécile Delay-Artous. Où sont les données de la recherche? : Essai de cartographie. 2015. halshs-01369745

HAL Id: halshs-01369745

<https://shs.hal.science/halshs-01369745>

Preprint submitted on 21 Sep 2016

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

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



Distributed under a Creative Commons Attribution - NonCommercial| 4.0 International License

Où sont les données de la recherche ? Essai de cartographie

DELAY-ARTOUS Cécile

*Laboratoire Dynamiques économiques et sociales des transports,
IFSTTAR, Marne-la-Vallée. France.*

Résumé :

La question émergente en France des données de la recherche se situe dans un cadre institutionnel foisonnant mais rigide, délicat à cerner. La recherche est aussi financée et évaluée au niveau européen. Cette organisation nationale et européenne se double d'un aspect international inhérent à la recherche et aux échanges d'informations rapides et répétés, accélérés par le développement d'Internet. Le labyrinthe institutionnel franco-européen se superpose ainsi avec le millefeuille international et disciplinaire du monde de la recherche. Enfin, la proximité de deux mouvements qui ne sont pourtant pas synonyme, l'Open Access et l'Open Data, vient encore troubler la compréhension de ce panorama.

Il n'est donc pas aisé de comprendre les rôles de chacun des acteurs quant aux données de la recherche. C'est à une clarification de ce paysage que nous nous proposons de participer, en initiant une cartographie des initiatives et acteurs visibles en France concernant les données des sciences humaines et sociales.

Mots clef : Open data, Open access, données de la recherche, France, sciences humaines et sociales, valorisation de la recherche, politique d'ouverture, acteur de la recherche, recherche publique.

1. Introduction

En France, la question émergente des données de la recherche (Fayet, 2013) se situe dans un cadre institutionnel foisonnant mais rigide, toujours délicat à cerner (Soulé, 2014). Comme le montrent les descriptions sur le site du Ministère de l'Enseignement Supérieur et de la Recherche (MESR, 2015), ce cadre change au gré des réformes (dernière en date, la loi n°2013-660 du 22 juillet 2013 relative à l'enseignement supérieur et à la recherche) ou des réorganisations ; par exemple, le TGE Adonis est devenu en juillet 2013 la Très Grande Infrastructure de Recherche (TGIR) Huma-Num, en fusionnant avec la TGIR Corpus-IR

La recherche est aussi financée et évaluée au niveau européen, au sein de l'Espace Européen de la Recherche, comme le montre l'existence de programmes européens pour la recherche, dont Horizon 2020, dit H20-20, entré en vigueur en janvier 2014, et qui prend en compte la gestion des données de la recherche. Le cadre institutionnel et administratif de la recherche publique française se structure donc également à l'échelle de l'Europe.

Cette organisation nationale et européenne se double d'un aspect international inhérent à la recherche et aux échanges d'informations rapides et répétés, accélérés par le développement d'Internet. Le labyrinthe institutionnel franco-européen se superpose ainsi avec le millefeuille international et disciplinaire du monde de la recherche.

Il n'est donc pas aisé de comprendre les rôles de chacun des acteurs quant aux données de la recherche (Gaillard, 2014). Comment, au sein d'un institut de recherche finalisée comme L'Institut Français des Sciences et Technologies des Transports, de l'Aménagement et des Réseaux (IFSTTAR) où nous avons effectué un stage durant l'été 2014 et sommes en poste actuellement, mettre en place un projet de partage des données de recherche ? Aux questions : vers qui se tourner ? Qui finance leur gestion ? Quelles règles encadrent leurs usages ? Que doit-on faire des données déjà accumulées ? Que prévoir pour les projets de recherche en cours ? La réponse est invariablement : tout dépend – du contexte, des besoins, du type de données, du statut de l'organisme ou du chercheur, de la discipline concernée, etc. Pourtant, le programme-cadre Horizon 2020 esquisse à travers un projet pilote une future obligation de gestion de ces données

Où sont les données de la recherche ? Essai de cartographie.

(UE, 2013), qui pourrait bien s'élargir et se pérenniser. C'est pour tenter de répondre à ces questions pratiques que cette étude a débuté, en juin 2014.

2. Le périmètre de l'étude

Au sein de ce paysage complexe et sur la question des données de la recherche, quelques acteurs émergent et semblent incontournables. Ainsi, souvent sous forme de sigles ou d'acronymes, fleurissent dans la littérature –articles académiques en sciences de l'information ou de la presse nationale quotidienne, ouvrages et sites Web destinés aux chercheurs ou aux professionnels de l'IST et du soutien de la recherche- les noms de H2020, OpenAire, ERC, Jisc et DCC, TGIR Huma-Num et PROGEDO, ou encore DataCite et Consortium Couperin. Mais qui sont-ils ? Et quels rôles occupent-ils du point de vue des données de la recherche ?

C'est à une clarification de ce paysage que nous nous proposons de participer, en élaborant une cartographie des initiatives et acteurs de la recherche publique visibles en France et concernant les données des sciences humaines et sociales (SHS). Nous avons tracé ce périmètre restreint parce qu'il est nécessaire de se limiter, mais surtout parce que c'est là que le retard français est apparu le plus important (Silberman, 1999 ; Chenu, 2011).

Néanmoins, ces limites se sont révélées délicates à suivre : il est impossible de s'arrêter aux initiatives purement françaises car nombre de projets en France sont des émanations de projets européens ou internationaux. On rencontre le même problème pour dessiner les contours des initiatives par disciplines : si l'on se limite aux SHS, que faire des projets interdisciplinaires ? D'autant que les nouveaux projets se multiplient (voir par exemple le site www.donneesdelarecherche.fr, onglet « initiatives et projets »), et qu'il semble hasardeux de se prétendre exhaustif.

Lors de l'élaboration première de cette cartographie, au cours de lectures, rencontres et conférences effectuées entre mars et octobre 2014, afin de constituer la bibliographie du mémoire de fin d'étude de l'INTD

(Delay-Artous, 2014), le choix a été fait de retenir toute appellation nommant des projets, acteurs et initiatives concernant les données de recherche, ayant un impact en France, et concernant au moins une science dite humaine ou sociale. L'objectif initial de cette investigation était donc d'essayer de répondre à la question : quels acteurs pour quelles initiatives ?

Souvent confondu ou identifié un peu rapidement à l'*Open Data*, qui concerne l'ouverture des données publiques (Tur, 2015 ; Chignard, 2013), l'ouverture des données de la recherche superpose aux difficultés de tout partage de données celles inhérentes au monde de la recherche, et se situe plutôt dans le mouvement de l'*Open access*. Afin de ne pas alimenter cette confusion, nous avons exclu de la cartographie les acteurs de l'Open Data français, tel que la mission Etalab et sa plateforme de données publiques data.gouv.fr, bien que celle-ci accepte en principe les données issues de la recherche. Après analyse des données présentes et des métadonnées les décrivant (Delay-Artous, 2014), il nous a semblé que, comme les sites et plateformes de collectivités proposant un accès à une sélection de données publiques, data.gouv.fr ne permettait pas de documenter les *datasets* issus de la recherche selon les standards internationaux nécessaires au monde académique.

Le champ de l'*Open Access* (OA) concerne en premier lieu, au moins du point de vue chronologique, les publications scientifiques. Mais ce secteur de l'ouverture des résultats scientifiques constitue un sujet d'étude à lui seul. Nous avons donc choisi de ne pas inclure les acteurs et projets concernant exclusivement l'édition scientifique. Mais cette limite s'est rapidement révélée fragile. En effet, les éditeurs scientifiques eux-mêmes se préoccupent des données : depuis longtemps quand il s'agit des informations qui viennent étayer des résultats dévoilés au sein d'un article de science exacte, par exemple dans la revue britannique Nature ; mais plus récemment, des éditeurs comme Elsevier se sont mis à proposer des services de gestion et/ou de recherche, partage, réutilisation...de données pour les chercheurs (Hologne, 2014). C'est donc au cas par cas que nous avons intégré -ou non- un item traitant des publications scientifiques au sein de notre liste. Et par exemple, OpenDOAR, l'annuaire d'archives et entrepôts académiques ouverts, recense parmi ses 2938 entrées, principalement des sites permettant le dépôt -et l'accès- de publications.

Où sont les données de la recherche ? Essai de cartographie.

Mais il liste également 144 entrepôts de données (ou archivants entre autres des *datasets*). Nous l'avons donc inclus dans notre liste d'objets.

En tâchant de poursuivre ce travail tout au long de l'année 2015, au sein de l'IFSTTAR, nous avons orienté une partie de nos lectures vers les informations concernant les données géographiques et de transport. Ce qui a ouvert de nouvelles questions en termes de typologie des initiatives concernant les données. Est-il intéressant d'intégrer à la cartographie la directive européenne INSPIRE (UE, 2007), concernant l'ensemble des données publiques géographiques ou celle sur les données des transports publics ? Nous avons, provisoirement du moins, opté pour le rejet de cette idée, afin de conserver un aspect pluridisciplinaire SHS et résolument académique à cet outil de travail.

Des interrogations du même ordre continuent de se poser, notamment en ce qui concerne le champ juridique, qui constitue une part importante des questions incontournables lors d'un projet d'ouverture de données de recherche. Et entre autres, faut-il intégrer la Commission nationale de l'informatique et des libertés (CNIL) à la liste, acteur central en France quant à la protection des données personnelles ? Le domaine d'action de la CNIL déborde très largement celui des données, ce qui, pour le moment, a fait pencher la balance en faveur de son exclusion de notre champ de travail ; mais la question reste en suspens, car au quotidien, les aspects juridiques constituent un nœud particulièrement important des soucis rencontrés pour la construction d'un entrepôt institutionnel de données de recherche.

3. La typologie proposée

3.1. Critères d'inclusion

Nous souhaitons, une fois cette liste établie, livrer une typologie des initiatives et relier chacune d'elle au(x) acteur(s) qui les portent ou s'y adossent. Nous imaginons, a priori, pouvoir classer aisément, en fonction de leur statut, les initiatives repérées en deux catégories de premier niveau : les acteurs, d'une part, et les initiatives et projets» d'autre part.

Pourtant, c'est dès cette première opération sémantique que les difficultés sont apparues. Par exemple, où classer DataCite, nom rencontré à de multiples reprises, tant lors de lectures que durant des conversations avec des chercheurs ou des professionnels de l'Information Scientifique et Technique (IST) ? C'est à la fois un acteur, en tant que consortium international, et une initiative qui consiste à accompagner la mise en place d'identifiants pérennes de type DOI pour les jeux de données scientifiques. Plus compliqué, encore : le cas d'Horizon 2020. Il s'agit d'un programme-cadre. C'est-à-dire à la fois une initiative européenne bornée dans le temps, et un important acteur du financement de la recherche en Europe, porteur de multiples initiatives...

Le statut de l'entité ne suffisant manifestement pas, un deuxième critère retenu peut être celui de la pérennité. Si un «actant» - pour reprendre la terminologie de Bruno Latour (Latour, 2006), ce terme permettant de réconcilier acteurs, objets et organisations en une même entité, qui se définit par sa capacité à peser sur le déroulement d'une action- est borné dans le temps, et ce dès sa conception, identifions le plutôt à une initiative. Au contraire, si l'actant ne se définit pas par une date de clôture, il sera catégorisé comme acteur. Mais dans certains cas, comme ceux mentionnés plus haut, ce critère n'est pas opérant à lui seul.

Le troisième critère retenu est le fait d'être posé comme un outil (un entrepôt, un modèle d'archivage ...) et/ou un service (d'aide à la gestion des identifiants pérennes, d'archivage à long terme...) proposé par un acteur ; dans ce cas, il est plus aisé de l'identifier à une initiative. Notons que ces critères peuvent se révéler éphémères et nécessitent une veille constante : certains projets, nés avec une date de clôture, se sont pérennisés. C'est en croisant ces critères que la liste d'entrées de la cartographie s'est affinée.

3.2. Critères d'exclusion

Puis, afin de limiter la croissance de cette liste et vu le nombre croissant de nouveaux projets dans ce secteur, une liste d'exclusion a été établie : certains actants de la recherche et de l'IST pourraient légitimement figurer dans la cartographie, mais non sans alourdir ce qui se veut un outil de travail, par exemple le schéma de Dublincore, schéma de métadonnées

Où sont les données de la recherche ? Essai de cartographie.

générique qui permet de décrire des ressources numériques ou physiques, ou les acteurs incontournables de la recherche en France comme le ministère de l'Enseignement Supérieur et de la Recherche et le CNRS, et à l'échelle européenne, l'EER (Espace Européen de la Recherche). A contrario, ont été retenus des actants qui œuvrent pour la recherche, bien au-delà des données, comme l'initiative nationale BSN (Bibliothèque Scientifique Numérique), dont un segment (le segment 10) concerne directement les *Research data*, ou données de la recherche.

4. La cartographie et sa représentation visuelle

Malgré ces limites, la cartographie établie comprenait en décembre 2014 trente-six entrées et quarante en juillet 2015. Ce foisonnement témoigne d'une grande activité et d'un dynamisme certain. Mais il est aussi une des causes du retard avec lequel la recherche française se saisit de la gestion de ses données. En effet, si plusieurs semaines sont nécessaires avant toute orientation au sein de ce paysage, quels chercheurs prendront le temps de mettre de côté leurs activités pour se repérer avant d'organiser un partage de données ?

Pour éclairer le caractère éclaté de cet environnement, une visualisation était nécessaire. Après quelques tentatives de représentations centrées sur les liens entre « acteurs » et « projets », qui ne mettaient en lumière que le foisonnement d'items et de liens de différentes natures, c'est le plan cartésien qui s'est imposé. En effet, le nombre d'acteurs comme d'initiatives, et les liens très variés que ceux-ci entretiennent entre acteurs ou entre acteurs et initiatives, voire entre initiatives, obligeait à multiplier les traits représentant ces correspondances. Les schémas obtenus étaient confus et peu satisfaisant pour une analyse, comme le montre la figure 1.

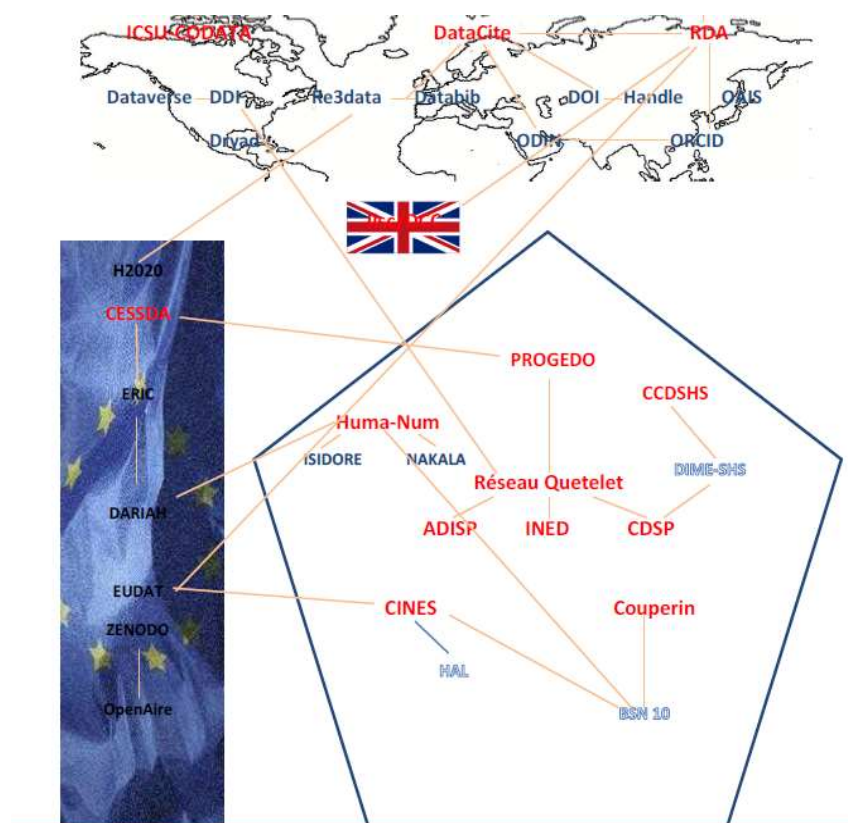


Figure 1 : premier essai de schématisation des liens entre acteurs et initiatives des données de la recherche en SHS

Optant donc pour une autre forme de représentation, la figure 2 propose une lecture qui s'abstrait u caractère réticulaire mais permet de rendre compte des statuts hybrides de nombreux actants, et des subtilités de positionnements.

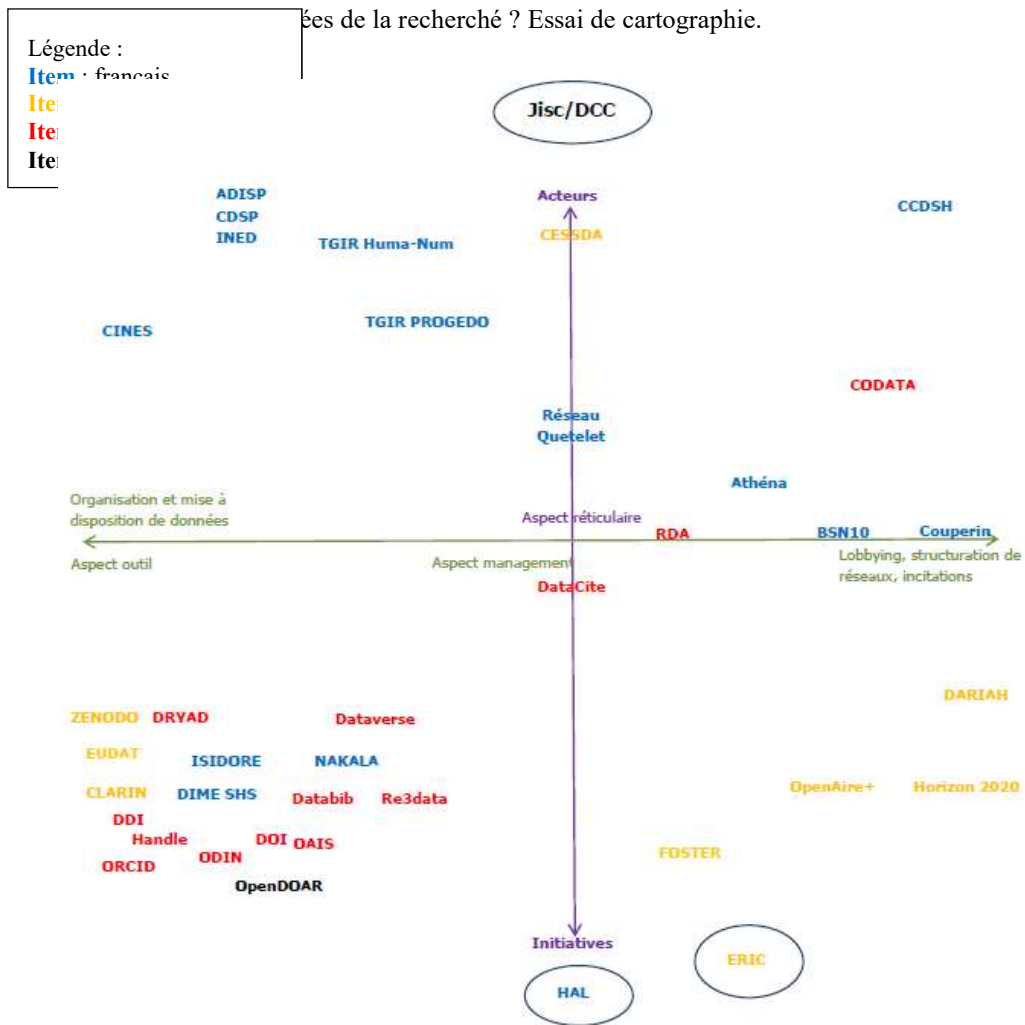


Figure 2: Acteurs et initiatives en SHS en France

En fonction de leurs rôles et objectifs, ce schéma représente les actants des données de la recherche autour de deux axes perpendiculaires. Les objets présentés hors cadre sont extérieurs, soit au territoire national (pas d'effet direct en France), soit aux données de la recherche proprement dites (qui concerne plutôt les publications, par exemple). Notons que cette catégorisation correspond à une « stylisation de la réalité » sous forme d'idéaux-types (Schnapper, 1999).

Sur cette figure, l'abscisse représente le type d'action ou de réalisation d'un actant. Plus il tend à créer des outils, plus il se trouve vers l'ouest. Au contraire, plus il œuvre pour la mise en place de politiques ou organise des actions de lobbying ou d'incitation, plus il sera placé à l'est. S'il est proche de l'origine, et donc au croisement avec l'ordonnée, cela indique une action managériale quant aux données.

En ordonnée, c'est le type d'actant qui est caractérisé. Un actant situé vers le nord répond aux caractéristiques d'un acteur sur la question des données de la recherche. Un actant situé vers le sud ressemble à une initiative ou à un projet. Au croisement avec l'abscisse se situent les actants qui fonctionnent en réseau.

Notons que ces positions ne sont pas figées : les évolutions sont rapides, et bouleversent parfois la position d'un actant. Pour ne signaler qu'un changement déjà initié, les catalogues d'entrepôts Databib et Re3data après être devenu des services de DataCite début 2015, devraient fusionner avant la fin de l'année 2015 (Ruiz, 2014 ; Strasser, 2015).

5. Analyse et premiers enseignements

Un premier décompte nous apprend que nous sommes en présence de quinze actants (dont un hors champ, le britannique Jisc, et son service dédié aux données, le DCC) dont les caractéristiques sont celles d'acteurs, pour vingt-quatre qui ressemble à des initiatives ou projets (dont deux hors champ, le statut juridique européen ERIC et HAL, l'archive ouverte française dédiée prioritairement aux publications). Cette balance purement comptable paraît cohérente, un acteur pouvant bien sûr porter ou piloter plusieurs initiatives.

Nous soulignons les premiers enseignements que l'on peut tirer de la lecture de ce schéma :

- La présence directement sur les axes de six actants signale bien que nous sommes en présence de statuts hybrides, tant leurs rôles et activités sont difficiles à positionner.
- L'aspect le plus frappant de ce schéma est de dessiner un paysage constitué en majorité :

Où sont les données de la recherche ? Essai de cartographie.

- D'initiatives multidisciplinaires, souvent internationales (13 items dans le quadrant *sud-ouest* dont 9 internationales) ;
- D'acteurs français émanant des autorités compétentes pour la recherche (les 6 items occupant le quadrant *nord-ouest* sont nationaux).
- Parmi les projets de type initiatives, on trouve aussi bien des entrepôts (comme ZENODO ou DRYAD) et des catalogues d'entrepôts (Databib et Re3data, OpenDOAR), que des outils comme des schémas de métadonnées (DDI, norme de description de données d'enquêtes en SHS), ou des projets portant en germe plusieurs outils (DIME SHS).
- En comparaison, l'Europe de la recherche est plus présente au sein du quadrant *sud-est* ; elle se situe davantage dans une logique d'initiatives incitant à l'ouverture des données, notamment avec un huitième plan cadre, héritier du 7e Plan cadre pour la recherche et le développement (PCRD), intitulé Horizon 2020 et l'outil que constitue OpenAire+ pour l'Open Research Data en Europe.

6. Conclusion

La cartographie que nous avons dressée est foisonnante et il est encore tôt pour savoir quelles initiatives seront à la fois pérennes et retenues dans les pratiques des laboratoires. Le contexte spécifique de chaque champ scientifique et son déploiement international sera certainement également déterminant.

Pour poursuivre et approfondir ce travail, il serait intéressant d'imaginer un outil collaboratif qui permettrait à un réseau de professionnels de l'IST d'alimenter et mettre à jour, au fur et à mesure des changements de statuts et création de projets, la liste qui compose cette cartographie. En datant les représentations schématiques qui en émaneraient, nous pourrions comparer les situations à différentes périodes, et observer d'éventuelles évolutions et tendances. Le dialogue entre professionnels aiderait également à trancher les points en suspens, comme l'entrée de nouveaux actants dans la cartographie, et le positionnement qui leur serait dévolu dans le schéma.

7. Glossaire des acteurs et initiatives de la cartographie par ordre alphabétique

<p>ADISP. Archives de Données Issues de la Statistique Publique, unité mixte de recherche (UMR) CNRS-EHESS-ENS, service de données du Centre Maurice Halbwachs (CMH), archive et diffuse essentiellement des données issues de la statistique publique française. http://www.cmh.ens.fr/greco/adisp.php http://www.reseau-quetelet.cnrs.fr/</p>
<p>Alliance Athena. Alliance THEmatique NAtionale des Sciences Humaines et Sociales, alliance de la SNRI du Ministère chargé de la Recherche, groupe de concertation pour coordonner les priorités de la recherche. http://www.allianceathena.fr/</p>
<p>BSN Segment 10. Bibliothèque Scientifique Numérique, TGIR créée par le MESR, gérée par le CNRS, la CPU et le CGE, plateforme de coopération pour l'accès à la littérature scientifique, notamment numérique. BSN 10 : segment concernant les données de la recherche. http://www.bibliothequescientifiquenumerique.fr/</p>
<p>CCDSHS. Comité de concertation pour les données en sciences humaines et sociales, auprès des ministères français chargés de l'économie, de l'emploi, de l'éducation nationale et de la recherche, placé sous la présidence du ministre en charge de la recherche ; propose des orientations pour la politique publique de données pour les SHS. http://www.reseau-quetelet.cnrs.fr/spip/article.php3?id_article=237</p>
<p>CDSP. Centre de Données Socio-Politiques, Unité Mixte de Service (UMS) 828, Sciences Po-CNRS, et membre du Réseau Quetelet, facilite l'accès des chercheurs et des étudiants à des données socio-politiques. http://cdsp.sciences-po.fr/ http://www.bequali.fr/bequali/</p>
<p>CESSDA. <i>Council of European Social Science Data Archives</i>, réseau européen des archives de données pour la recherche en sciences sociales, infrastructure de l'European Strategy Forum on Research Infrastructures (ESFRI), statut temporaire de société à droits limités, loi norvégienne, envisagé comme futur ERIC, œuvre sur les questions d'accès aux données de la statistique publique et l'homogénéisation des licences d'accès en Europe http://www.cessda.net/</p>

<p>CINES. Centre Informatique National de l'Enseignement Supérieur, Etablissement Public à caractère Administratif national, tutelle du MESR, propose des moyens exceptionnels à la communauté scientifique dans ses missions stratégiques nationales pour le calcul numérique intensif et l'archivage pérenne de données électroniques. https://www.cines.fr/</p>
<p>CLARIN. <i>Common Language Resources and Technology Infrastructure</i>, projet européen de construction d'infrastructure pour rendre les ressources et techniques linguistiques aisément disponibles pour l'ensemble de la communauté européenne des SHS, porté par l'ESFRI, puis par CLARIN ERIC depuis 2012. http://eudat.eu/common-language-resources-and-technology-infrastructure</p>
<p>CODATA. <i>COmmittee on DATA for Science and Technology</i>, organisation internationale regroupant 21 pays, 16 unions scientifiques internationales, 2 organisations scientifiques co-optées et 13 organisations partenaires (Industrie, Gouvernement et Universités), membre de l'ICSU, International Council for Science ; encourage à l'échelle mondiale la compilation, l'évaluation et la diffusion de données numériques fiables d'importance pour la science et à la technologie. http://www.codata.org/ http://codata-france.org/</p>
<p>Couperin. Consortium unifié des établissements universitaires et de recherche pour l'accès aux publications numériques, association loi 1901, loi française, réseau de négociation et d'expertise des ressources documentaires électronique pour l'IST. http://www.couperin.org/</p>
<p>DARIAH. <i>Digital Research Infrastructure for the Arts and Humanities</i>, infrastructure numérique européenne de recherche pour les SHS sous statut ERIC depuis 2014, réseau d'acteurs pour les TIC dans les SHS, ayant Huma-num comme nœud français. www.dariah.fr http://www.huma-num.fr/international/dariah</p>
<p>Databib. Catalogue d'entrepôts en ligne de données de recherche, en cours de fusion avec Re3data en 2015 sous la direction de DataCite. http://databib.org/index.php</p>
<p>DataCite. International Data Citation Initiative, organisme à but non lucratif, consortium international d'institutions, agence d'enregistrement de DOI, membre de la Fondation Internationale de DOI (IDF), promeut la citation des jeux de données de recherche par l'attribution d'identifiants pérennes. http://www.datacite.org/</p>

<p>Dataverse Network project. Projet de réseau Dataverse, logé à l'Institut des sciences sociales quantitatives (IQSS) à l'Université de Harvard. Dataverse Network : logiciel libre, application open source pour publier, partager, référencer, extraire et analyser les données de recherche, utilise les standards de description Dublincore et/ou DDI. http://thedata.org/</p>
<p>DCC. <i>Digital Curation Center</i>, émanation du Jisc, dédié à l'archivage, la gestion, le partage des données de la recherche pour l'enseignement supérieur et la recherche au Royaume Uni, membre de la RDA et de CODATA. http://www.jisc.ac.uk/dcc</p>
<p>DDI. <i>Data Documentation Initiative</i>, projet international (programme de l'<i>University of Michigan</i>, géré par la <i>DDI Alliance</i>, initié en 1995. Standard de documentation technique pour décrire et conserver les informations statistiques et plus globalement les informations et données d'enquêtes en sciences humaines et sociales. http://www.ddialliance.org/</p>
<p>DIME-SHS. Données Infrastructures et Méthodes d'Enquête en Sciences Humaines et Sociales, équipement EQUIPEX de la vague 2010, porté par Sciences Po, l'INED et le GIS Quetelet entre autre, vise à créer une structure nationale pour la collecte, l'enrichissement et la diffusion des données pour la recherche en SHS. http://www.sciencespo.fr/dime-shs/</p>
<p>DOI. <i>Digital object identifier</i>, mécanisme d'identification de ressources numériques (cas particulier de Handle), porté par l'<i>International DOI Foundation</i> (IDF), dont DataCite et CrossRef sont membres et <i>registration agencies</i> (organismes d'enregistrement). Norme ISO depuis 2012. http://www.doi.org/</p>
<p>DRYAD. <i>Dryad Digital Repository</i>, entrepôt de données de recherche géré par un organisme à but non lucratif américain, de tous domaines de recherche, sous conditions d'être associées avec un article de revue à comité de lecture. Utilise les licences <i>Creative Commons</i> v3.0, le <i>waiver</i> CC0 et les DOI. http://datadryad.org/</p>

<p>ERIC. <i>European Research Infrastructure Consortium</i>, soit Consortium européen pour les Infrastructures de Recherche, cadre juridique applicable à ce type d'infrastructures adopté par la Commission Européenne en 2009 pour faciliter l'établissement et l'utilisation communes d'installations de recherche d'intérêt européen par plusieurs États membres associés au 7^e PCRD. http://www.dgdr.cnrs.fr/dsfim/fiscalite/Outils%20TVA/M%C3%A9mento_TVA%20.pdf http://europa.eu/legislation_summaries/research_innovation/general_framework/ri0005_fr.htm</p>
<p>EUDAT. Initiative européenne cofinancée par le 7^e Plan-cadre recherche et dev. (PCRD ou Framework Programme 7, FP7) de la Communauté Européenne puis par son successeur H2020, infrastructure de données de recherche et services accessibles via le portail EUDAT pour tout chercheur européen. Membre de RDA. CLARIN, le Cines et le Jisc participant. http://www.eudat.eu/</p>
<p>FOSTER. <i>Facilitate Open Science Training for European Research</i>, Projet FP7 2014-2016, programme de formation à l'échelle européenne d'aide aux chercheurs, étudiants de troisième cycle, bibliothécaires et autres parties prenantes pour l'intégration des approches Open Access aux méthodologies de recherche pré-existantes. https://www.fosteropenscience.eu/project/</p>
<p>HAL. Hyper Articles en Ligne, archive ouverte française permettant aux chercheurs de déposer articles et manuscrits dans une base à accès ouvert, hébergée gérée par le Centre pour la communication scientifique directe (CCSD) du CNRS. http://hal.archives-ouvertes.fr/ http://cache.media.enseignementsup-recherche.gouv.fr/file/HAL/93/3/01_Convention_HAL_246933.pdf</p>
<p>Handle. <i>Handle System</i>, ou Système Handle, spécification technique pour attribuer, gérer et résoudre des identifiants persistants attribués à des objets numériques et à d'autres ressources Internet, co-développé par B. Kahn avec le soutien initial de la <i>Defense Advanced Research Projects Agency</i> (DARPA), géré par la <i>Corporation for National Research Initiatives</i> (CNRI), organisme à but non lucratif américain. http://www.handle.net/ Defense Advanced Research Projects Agency http://www.cnri.reston.va.us/</p>

<p>Horizon 2020. Abrégé en H2020, programme-cadre européen de financement de la recherche et de l'innovation, démarré le 1er janvier 2014 pour 7 ans, successeur du 7^e programme-cadre de recherche et développement (P.C.R.D) ou en abrégé de l'anglais FP7 (<i>Framework Programme 7</i>). Nouveauté H2020 : la Commission européenne généralise l'accès libre aux publications de recherche et introduit dans certains cas l'accès libre aux données issues des recherches qu'il finance; encourage à la rédaction de plan de gestion des données (<i>Data Management Plan</i>, DMP). http://www.horizon2020.gouv.fr/ http://ec.europa.eu/programmes/horizon2020/</p>
<p>Huma-Num. Née en 2013 de la fusion du TGE Adonis et de la TGIR Corpus IR, très grande infrastructure de recherche (TGIR) portée par L'UMS 3598 associant CNRS, Université d'Aix-Marseille et Campus Condorcet, visant à faciliter le tournant numérique de la recherche en SHS. http://www.huma-num.fr/</p>
<p>INED. Institut national d'études démographiques, établissement public à caractère scientifique et technologique, producteur d'enquêtes socio-démographiques, depuis 2000 met à disposition ses enquêtes auprès de la communauté scientifique notamment via le réseau Quetelet. http://www.ined.fr/ http://www.reseau-quetelet.cnrs.fr/spip/rubrique.php3?id_rubrique=58</p>
<p>ISIDORE. Intégration de services, interconnexion de données de la recherche et de l'enseignement, service de la TGIR Huma-num qui collecte par « moissonnage », enrichit et offre un accès unifié aux documents et données numériques des SHS, plus gros projet d'<i>open data</i> scientifique en France, propose l'accès à plus de trois millions de documents numériques. http://www.rechercheisidore.fr/apropos http://www.huma-num.fr/services-et-outils/signaler</p>
<p>JISC. <i>Joint Information Systems Committee</i>, organisme semi-public britannique pour le numérique dans l'éducation et l'enseignement supérieur, fondé en 1993 par les agences de financements pour l'éducation des plus de 16 ans au Royaume-Uni, créateur du DCC, pilote du programme de financement <i>Managing Research Data</i>. http://www.jisc.ac.uk/ http://www.ambafrance-uk.org/Les-publications-en-acces-libre</p>

<p>NAKALA. Service d'exposition de données de la TGIR Huma-Num, soit : accès aux données elles-mêmes (identifiants pérennes, sécurisation, accès permanent...) et présentation des métadonnées (Web de données et entrepôt RDF, accès interopérable avec OAI-PMH...) pour les producteurs de données numériques scientifiques des SHS. http://www.huma-num.fr/service/nakala</p>
<p>OAIS. <i>Open Archival Information System</i>, soit système ouvert d'archivage de l'information, norme ISO, modèle conceptuel 'archivage électronique, décrit les responsabilités, les fonctions et les rapports avec son environnement d'un système d'archivage électronique pour assurer la pérennisation de l'information numérique. http://www.archivesdefrance.culture.gouv.fr/gerer/archives-electroniques/standard/norme-oais-iso-14721/ http://www.archivesdefrance.culture.gouv.fr/static/4940</p>
<p>ODIN. <i>ORCID and DataCite Interoperability Network</i>, projet financé par la Commission européenne au titre du FP7, doit permettre l'interopérabilité des réseaux ORCID et DataCite pour identifier de façon unique des scientifiques et des jeux de données. http://odin-project.eu/</p>
<p>OpenAIRE et OpenAIRE+. <i>Open Access Infrastructure for Research in Europe</i>, projet européen dont le but est d'accompagner l'obligation de dépôt en accès libre décidée par la Commission Européenne et le Conseil Européen de la Recherche (ERC), prolongé par OpenAIREplus qui élargit son périmètre aux données de la recherche. https://www.openaire.eu/</p>
<p>OpenDOAR. <i>The Directory of Open Access Repositories</i>, catalogue d'entrepôts Open access interrogeable, service de SHERPA, créé par le britannique Center for Research Communication et financé par le Jisc. http://www.opendoar.org/ Et http://www.opendoar.org/about.html</p>
<p>ORCID. <i>Open Researcher and Contributor ID</i>, organisme à but non lucratif américain, registre d'identifiants de chercheurs et méthode transparente pour relier activités et résultats de la recherche à ces identifiants. http://orcid.org/</p>
<p>PROGEDO. TGIR appuyé sur l'UMS Quetelet-PROGEDO (responsabilité conjointe du CNRS et de l'EHESS), assure la mise en œuvre d'une politique publique de données pour les SHS dont les grandes orientations sont proposées par le CCDSHS et coordonnées au sein d'un Groupement d'intérêt Scientifique « réseau Quetelet », en partenariat avec plusieurs grands établissements. http://www.progedo.fr/ http://www.reseau-quetelet.cnrs.fr/</p>

<p>RDA. Research Data Alliance, projet international de promotion du partage des données de la recherche, communauté “<i>open science</i>” sans frontière de pays ou de disciplines, partageant l’objectif “<i>data sharing without barriers</i>”, basée sur la participation des individus, créée à l’initiative de 3 institutions, Commission européenne, <i>US National Science Foundation and National Institute of Standards and Technology</i>, et <i>Australian Government’s Department of Innovation</i>Australie. https://rd-alliance.org/</p>
<p>Re3data.org. <i>Registry of research data repositories</i>, registre ou catalogue d’entrepôts de données de la recherche, financé par la German Research Foundation. http://www.re3data.org/</p>
<p>Réseau Quetelet. Composantes de PROGEDO, coordonne les activités d’archivage, de documentation et de diffusion des données en sciences humaines et sociales du CMH-ADISP, du CDSPet du service des enquêtes de l’Ined, organisée par le CCDSHS ; GIS « Réseau Quetelet » fondé par le CNRS, l’INED, EHESS, l’université de Caen-Basse Normandie, l’Université des sciences et technologies de Lille 1, la Fondation nationale des sciences politiques, l’EEP et l’IRDES, assure la coordination des orientations proposées par le CCDSHS. http://www.reseau-quetelet.cnrs.fr/spip/</p>
<p>ZENODO. Entrepôt de données et publications issus de la « longue traine » des résultats de recherches, multidisciplinaire et tout formats, hébergé par le CERN ; créé par le CERN et OpenAire, financé par la CE. http://zenodo.org/</p>

Bibliographie

- Chenu, Alain (2011). « Introduction », in *La France dans les comparaisons internationales*, Presses de Sciences Po, Paris, France, p. 9-17. Disponible en ligne : www.cairn.info/la-france-dans-les-comparaisons-internationales--9782724612189-page-9.htm. (Consulté le 28/03/2015).
- Chignard, Simon (2012). *Open Data : Comprendre l’ouverture des données publiques*. FYP, Paris, France.
- Delay-Artous, Cécile (2014). *Open, or Not Open, Research Data ? Jeux d’acteurs et questions d’accès à l’IFSTTAR*, Mémoire INTD. Disponible en ligne :

Où sont les données de la recherche ? Essai de cartographie.

http://memsic.ccsd.cnrs.fr/mem_01128833/document (Consulté le 15/03/2015).

Fayet, Sylvie (2013). « “Données” de la recherche, les mal-nommées ». URFIST Info, 15 novembre 2013. Disponible en ligne : <http://urfistinfo.hypotheses.org/2581> (Consulté le 4/06/2014).

France (2013). Loi n°2013-660 du 22 juillet 2013 relative à l’enseignement supérieur et à la recherche. Disponible en ligne : <http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT00027735009> (Consulté le 12/07/2015).

France (2014). Ministère de l’enseignement supérieur et de la Recherche. L’état de l’Enseignement supérieur et de la Recherche, n°7, Mars 2014. Disponible en ligne : http://cache.media.enseignementsup-recherche.gouv.fr/file/EESR_2014/60/7/EESR7_316607.pdf (Consulté le 8/06/2015).

Gaillard, Rémi (2014). « De l’Open data à l’Open research data : quelle(s) politique(s) pour les données de recherche ? ». Mémoire, Enssib. Disponible en ligne : <http://www.enssib.fr/bibliotheque-numerique/notices/64131-de-l-open-data-a-l-open-research-data-quelles-politiques-pour-les-donnees-de-recherche> (Consulté le 10/05/2014).

Hologne, Odile (2014). Les données de la recherche dans l’écosystème des publications scientifiques. Présentation de la journée d’étude du 12 février 2014. GFII, Paris, France. Disponible en ligne : http://www.gfii.fr/uploads/docs/GFII_Hologne_Intro.pdf (Consulté le 12/07/2015).

Latour, Bruno (2006). *Changer de société. Refaire de la sociologie*, La Découverte, Paris, France.

Silberman, Roxanne (1999). *Rapport Silberman : Les sciences sociales et leurs données*, Ministère de l’Education Nationale, de la Recherche et de la Technologie. Disponible en ligne : <http://www.education.gouv.fr/cid1925/les-sciences-sociales-et-leurs-donnees.html> (Consulté le 15/05/2014).

Soulé, Véronique (2014). Interview de Pierre-Paul Zalio, président de l’ENS Cachan : « Le système français est complexe et illisible. » in Libération du 19/10/2014.

- Union Européenne (2007). Directive 2007/2/CE du Parlement Européen et du Conseil du 14 mars 2007 établissant une infrastructure d'information géographique dans la Communauté européenne (INSPIRE). Journal Officiel de l'Union Européenne du 25/04/2007. Disponible en ligne : <http://www.developpement-durable.gouv.fr/IMG/pdf/Inspire.pdf> (Consulté le 25/04/2014).
- Union Européenne (2013). Lignes directrices pour la gestion des données dans Horizon 2020. Version 1.0. Décembre 2013. Trad. INIST-CNRS. Disponible en ligne : http://openaccess.inist.fr/IMG/pdf/14086_lignes_directrices_la_horizon_2020_tr_fr_version-oct2014.pdf (Consulté le 8/06/2015).
- Ruiz, Pedro (2014). DataCite, re3data.org, and Databib Announce Collaboration. Disponible en ligne : <https://www.datacite.org/news/datacite-re3dataorg-and-databib-announce-collaboration.html> (consulté le 9/08/2014).
- Strasser, Carly (2015). Merger of Databib and re3data.org, first version of API available. Disponible en ligne : <https://www.datacite.org/news/merger-databib-and-re3dataorg-first-version-api-available.html> (Consulté le 22/06/2015).
- Schnapper, Domonique (1999). *La compréhension sociologique. Démarche de l'analyse typologique*. PUF, Paris, France.
- Tur, Alexandre (2015). Bibliothèque, open science, open data et données de la recherche au Canada : quels enjeux ? Enssiblab, 21 juillet 2015. Disponible en ligne : <http://www.enssib.fr/recherche/enssiblab/les-billets-denssiblab/bibliotheque-de-recherche-open-access-open-data-gestion> (Consulté le 27/07/2015).