



HAL
open science

Science bloggers on Hypotheses.org: Communities and co-authorship

Elsa Poupardin, Camille Prime-Claverie

► To cite this version:

Elsa Poupardin, Camille Prime-Claverie. Science bloggers on Hypotheses.org: Communities and co-authorship. PCST Conference, PCST, Apr 2023, Rotterdam, France. ⟨hal-04317925⟩

HAL Id: hal-04317925

<https://hal.science/hal-04317925v1>

Submitted on 1 Dec 2023

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 CC BY-NC-SA 4.0 - Attribution - Non-commercial use - ShareAlike - International License

Carnetiers et blogueurs de sciences : analyser des modèles et des communautés en construction

Camille Prime-Claverie, Dispositifs d'Information et de Communication à l'Ère du Numérique - Paris Île-de-France, Université Paris Nanterre : EA7339, cclaveri@parisnanterre.fr);

Elsa Poupardin, Centre d'Études et de Recherches Interdisciplinaires de l'UFR Lettres, Arts, Cinéma (CERILAC, URP 441), Université Paris Cité, elsa.poupardin@u-paris.fr)

Résumé :

Le blogging scientifique qui se développe sur *Hypotheses.org* depuis 2008 n'a jamais été étudié de manière quantitative. L'analyse de 389 089 contributions réparties sur 3540 blogs depuis la création de la plateforme a permis de cartographier les réseaux de carnetiers sur différentes périodes.

L'étude des co-présences des auteurs sur les blogs a permis de mettre en évidence des interactions, jusque-là invisibles par les méthodes scientométriques traditionnelles, aussi bien entre communautés disciplinaires qu'entre différentes strates d'écriture de la recherche en SHS (individus, programmes de recherche, institutions, séminaire, etc.).

Mots-clés : blogging scientifique, communication scientifique, participation numérique, cartographie

Introduction

L'analyse des citations des chercheurs et des co-signatures est largement utilisée pour documenter la structure d'une communauté scientifique et son développement, la manière dont les chercheurs discutent, collaborent ou s'ignorent. En particulier les cartographies de publications utilisées en bibliométrie (Small, 1973) offrent une visualisation de la « science déjà faite ». Elles s'appuient sur des corpus d'articles scientifiques publiés. Réalisées après le processus de validation par les pairs, après la publication des résultats et, selon les méthodes, après leur citation dans d'autres publications scientifiques, elles sont contraintes au temps long et à l'inertie du système de publication et de citation. C'est un outil puissant pour enquêter sur les réseaux scientifiques formels, mais qui n'offre que peu d'informations sur les réseaux informels associés, les cercles sociaux (Crane, 1969). L'analyse des réseaux sociaux, issue de la sociologie, moins souvent utilisée, permet d'avoir un aperçu de la communication informelle des chercheurs (Marion et al., 2003).

Les recherches sur la pratique du blogging académique ont peu exploré le fait que cette pratique, à l'interface entre communication formelle et informelle, permet de réaliser une autre cartographie des interactions entre chercheurs qui se situerait entre les deux premières. Elle permet de voir la « science en train de se faire », de révéler des liens entre les différentes strates de la recherche (individus, programmes de recherche, institutions, séminaire, etc.) et les différents types d'écriture de la recherche (Lefebvre, 2013).

Nous reviendrons dans un premier temps sur l'histoire et le développement du blogging scientifique à l'échelle internationale et présenterons les discours qui ont accompagné son développement, les espoirs qu'ils ont fait naître en lien avec les mouvements et les politiques d'*Open Access* et d'*Open Science*.

La variété des modèles proposés par les plateformes de blogging scientifique oblige à s'arrêter sur une définition plus restrictive du blogging scientifique pour questionner leur dimension collective. À partir de l'exemple de la plateforme d'*Hypotheses.org*, nous tenterons de voir ce que l'analyse de la participation à des carnets révèle des pratiques collaboratives et si on observe la création de véritables espaces d'échange. Au moyen de cartographies de co-présence des auteurs, nous verrons comment les

contributions des carnetiers se sont articulées les unes aux autres au cours de différentes périodes de fonctionnement de la plateforme¹.

I- Le blogging scientifique : une pratique multiforme

Pour Bora Zivkovic, co-fondateur de l'organisation *ScienceOnline* et de l'agrégateur de blogs *Science Seeker*, un blog scientifique est un blog qui répond à un ou plusieurs des critères suivants : « un blog écrit par un scientifique, un blog écrit par un écrivain ou un journaliste scientifique professionnel, un blog qui traite principalement de sujets scientifiques, un blog utilisé dans une classe de sciences comme outil d'enseignement, un blog utilisé pour diffuser des nouvelles et des communiqués de presse plus ou moins officiels par des sociétés, instituts, centres, universités, éditeurs, entreprises et autres organisations scientifiques » (Zivkovic, 2012). On regroupe donc sous ce terme une grande variété de textes et les contenus proposés varient de la présentation vulgarisée des travaux de collègues à des considérations méthodologiques ou des questions matérielles. Cette définition extrêmement large du blogging scientifique couplée à l'évolution rapide des pratiques (microblogging ou meso blogging) empêche d'avoir une vision précise des pratiques, et de leur évolution.

1. Historique et évolution des modèles économiques

Lancé à partir de 2005 aux États-Unis sur le modèle proposé par Dave et Greta Munger, le blogging scientifique n'est au départ envisagé que comme une initiative portée par des particuliers. Les articles discutent principalement les publications académiques validées par la communauté scientifique. Les blogs sont peu à peu recensés par des agrégateurs, comme *Research Blogging* (2007), qui accréditent leur contenu. En parallèle, les premières plateformes de blogs, comme *Scienceblogs* (2006), apparaissent. Elles hébergent des blogs créés sur la plateforme ou rapatriés après une vie plus ou moins longue qui désirent bénéficier d'une plus grande visibilité et d'un soutien technique (Zivkovic, 2012). Blanchard décrit une évolution similaire en France en 2006, avec l'ouverture de la plateforme du « café des sciences » (Blanchard, 2014). L'objectif d'une médiation directe entre chercheurs et grand public semble être la motivation principale des auteurs, les commentateurs s'intéressent à cette diffusion ouverte, sans filtre ni barrière qui permet un dialogue inédit. On envisage les échanges futurs avec un public diversifié (Brown Jarreau, 2015), les impacts divers que pourra avoir ce nouveau canal de communication sur les chercheurs et le large public (Kjellberg, 2010 ; Kjellberg et al., 2016 ; Marht & Puschmann, 2014).

En 2010, le « *Pepsigate* » ébranle la plateforme américaine la plus en vue à l'époque (Raymond, 2010). La révélation d'un accord financier entre Pepsi et le propriétaire de *Scienceblogs.com* jette le doute sur la « scientificité » des contenus proposés dans les blogs et les auteurs quittent massivement la plateforme. On s'interroge sur les motivations des auteurs, on pèse ce qu'ils peuvent retirer de cette activité (Blanchard, 2017 ; Brown Jarreau, 2015 ; Dacos & Mounier, 2010 ; Gallezot & Deuff, 2010). Partout se pose la question des biais et des déviations apportés par cette nouvelle forme de médiation. On assiste à une série de rachats et de reconfigurations aux États-Unis.

La montée en puissance des blogs génère pendant toute cette période de nombreux travaux de recherche. Certains chercheurs examinent ce nouveau genre d'écriture (H. Zou & Hyland, 2020 ; H. (Joanna) Zou & Hyland, 2020), d'autres tentent de replacer l'écriture de billets de blogs dans le cycle de la recherche (Poupardin & Faury, 2018).

Aujourd'hui, les blogs de chercheurs autonomes, en France ou à l'étranger, sont rares. Les modèles économiques à l'œuvre derrière les plateformes de blogging déterminent largement le contenu et l'investissement des auteurs. La dénomination de blog recouvre en fait des sites extrêmement variés (du site institutionnel au journal en ligne) dans lesquels les acteurs usent de modalités d'écriture très

¹ Cette partie s'appuiera sur les résultats d'une étude menée dans le cadre du projet HYCAR financée par le GIS « réseau URFIST » à partir de données issues de la plateforme Hypotheses.org.

différentes. Dans de très nombreux cas, le modèle économique de la plateforme de blogging est relativement obscur et génère des suspicions sur les motivations des hébergeurs (Malkan, 2018).

Les plateformes peuvent être créées et hébergées par des médias généralistes (<https://blogs.mediapart.fr/marc-tertre/blogn> sur Mediapart), des institutions scientifiques (<https://blogs.lse.ac.uk> via la *London School of Economics*), des associations (<https://www.cafe-sciences.org>, café des sciences en France), le secteur public (<https://hypotheses.org>), ou des plateformes « indépendantes » (<https://www.over-blog.com/tag/sciences>).

Il est cependant notable que les éditeurs scientifiques, en particulier Springer, occupent aujourd'hui un rôle central dans le paysage du blogging. Les plateformes sont très souvent aujourd'hui explicitement adossées à une revue scientifique (<http://scilog.com> avec *Pour la science* en France, <https://scienceblogs.de> avec *Spektrum der Wissenschaft* en Allemagne, <http://blogs.nature.com> avec *Nature*). Il a été un temps possible de considérer les blogs comme pouvant potentiellement représenter un nouvel indicateur de la diffusion et de la réception de la science (Shema et al., 2015). L'investissement des éditeurs a bien entendu des répercussions sur le contenu proposé et on peut s'interroger sur la pertinence de certains indicateurs d'altmetrics conçus à partir de ces blogs.

Cette diversité des modèles économiques rassemblés sous le même intitulé de « blogging scientifique » empêche en fait d'avoir une vision précise des pratiques et des motivations des auteurs. Ainsi les rachats successifs de plateformes de blogging américaines (par exemple *Scienceblogs.com*) et les changements des politiques éditoriales masquent l'évolution spontanée des pratiques. La diversité de profils des bloggeurs ne permet pas non plus de travailler à une typologie précise des auteurs. Sur certaines plateformes, ils sont recrutés uniquement sur invitations (*Scienceblogs.com*), sur d'autres ils peuvent être issus du journalisme scientifique (Pierre Barthelemy - <https://www.lemonde.fr/passeurdessciences/> sur *Le Monde*) ou faire partie du service de communication de l'institution qui les héberge (un océan de découvertes sur <https://lejournale.cnrs.fr/nos-blogs>). Certaines plateformes permettent d'identifier clairement les différents contributeurs et leur institution de rattachement (*Hypotheses.org*, *Scienceblogs.de*) d'autres pas du tout (*Scienceblogs.com*). Enfin, les auteurs sont rémunérés (par la publicité) ou pas. Comment, dans ce cas, comparer les écritures et mesurer le poids de leur investissement dans leur activité professionnelle principale ? Comment reconnaître une spécificité au blogging, quand cette pratique si diversifiée ressemble à celle qu'on trouve sous d'autres dénominations² ?

2. *Hypotheses.org*

Pour ces raisons, il nous a semblé intéressant de centrer notre recherche sur la plateforme *Hypotheses.org*, qui occupe une position tout à fait singulière dans le paysage français des Sciences Humaines et Sociales. Développée par le Cléo, (une UMR du CNRS, de l'université d'Aix Marseille, d'Avignon et de l'EHESS) elle permet, depuis 2008, à tout acteur de la sphère scientifique, d'ouvrir un blog appelé sur cette plateforme « carnet de recherche ». De ce fait, elle héberge aujourd'hui la quasi-totalité des blogs de chercheurs dans ces disciplines. Les administrateurs de la plateforme assurent une certaine cohérence scientifique à ces contenus (procédure de validation, présence d'un conseil scientifique, etc.) qui a permis aux carnetiers de se saisir au choix des « carnets » comme outils de valorisation des travaux réalisés, lieux d'échange ou de débats auprès de ses pairs ou comme un média de communication vers le grand public. La plateforme, soutenue par les pouvoirs publics depuis sa création, assure « gratuitement » l'entretien de la plateforme aux auteurs et autorise des positionnements très variés.

Suite à la signature d'une convention avec *OpenEdition*, les membres du projet HYCAR ont pu travailler sur un corpus composé de l'intégralité des billets des blogs répertoriés au catalogue d'*Hypotheses.org* en août 2020. Ont été fournis le texte intégral de chaque billet, le login de son auteur, la langue utilisée,

² Les écrits de chercheurs dans *The Conversation* sont-ils si différents de ceux qui paraissent sur *Scienceblogs.com* ? Dans les deux cas les auteurs sont sollicités, les articles commandés et la présentation des billets n'est plus antéchronologique mais disciplinaire.

la date de publication et le nom du blog sur lequel le billet a été publié. Nous avons ajouté à ces données un ensemble de métadonnées sur chaque carnet disponible sur le catalogue³ proposé par OpenEdition et récupérées par scraping (type de carnet, responsable et adresse de publication, dates de mise en ligne et d'entrée au catalogue du carnet). Le corpus sur lequel s'appuie notre travail comporte 389 089 billets écrits par 12 893 auteurs différents (c'est-à-dire avec des logins différents) sur 3540 blogs.

3. Évolution de la production sur *Hypotheses.org*

La croissance de la plateforme *Hypotheses.org* a été continue et rapide entre 2008 et 2020 (*OpenEdition team*, 2021)⁴. Dans notre corpus, le nombre de carnets ouverts est en forte augmentation chaque année jusqu'en 2013, puis s'établit autour de 400 nouveaux carnets par an. Le nombre d'auteurs distincts qui participent à la plateforme chaque année augmente également jusqu'en 2016 et se stabilise autour de 3500 les années suivantes. La baisse de production qui apparaît sur la figure 1 est trompeuse. Elle est due en partie aux nouveaux carnets créés qui ne figurent pas encore au catalogue.

Dans ce corpus le français est la langue dominante (58,4% des billets) suivi par l'allemand (19,4%) et l'anglais (14,3%). Le portugais (3,2%), l'espagnol (3,1%) et l'italien (1,1%) sont en moindre quantité. Dans cette étude nous choisissons de garder toutes les langues et de faire une analyse macro de la plateforme. La répartition selon les pays de production est présentée sur la figure 1 bis⁵.

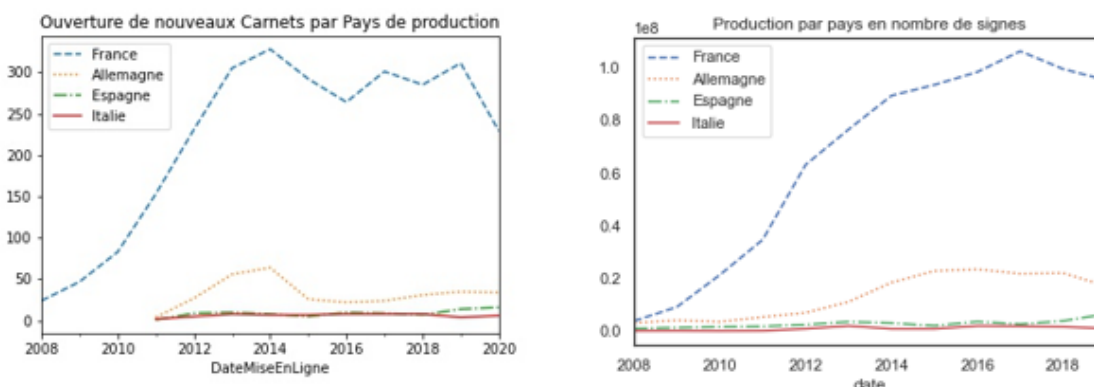


Figure 1 : Ouverture de nouveaux carnets par pays Figure 1 bis : Évolution du nombre de signes par pays

Durant la période étudiée, les auteurs peuvent avoir écrit dans un seul carnet (88,6% des auteurs), deux carnets (7,8%), trois carnets (2%) ou dans plus de trois carnets (1,6%).

La plateforme annonce la diversité des écrits qu'elle abrite en proposant aux créateurs des carnets de choisir, s'ils le désirent, un type parmi les 17 proposés : carnets de thèse, de vulgarisation, de chercheurs, de projet de recherche, de master, etc. Les types évoluent dans le temps et certains types ont été ajoutés au fil de l'eau. Le contenu qu'ils sont censés receler n'est décrit nulle part. Ils proposent donc une variété importante de style d'écrits et de participation, le créateur d'un carnet pouvant interpréter les intitulés à sa guise. Un carnet de master peut être écrit par un seul enseignant ou par des promotions successives d'étudiants, un carnet de structure de recherche peut être alimenté par un « community manager » isolé aussi bien que par une équipe rédactionnelle nombreuse. De ces types, nous ne pouvons déduire ni la taille, ni la forme des billets puisqu'elle est extrêmement variable : 28,49% des billets ont moins de 2000 signes, 31,18% entre 2000 et 4000, 26,88% entre 4000 et 10000 et 13,44% plus de 10 000 signes. On

³ <https://www.openedition.org/catalogue-notebooks> à la date du 24 janvier 2022

⁴ Certains billets de carnets rapatriés sur *Hypotheses.org*, apparaissent avec une date antérieure à 2008 car ils ont conservé la date de publication originelle.

⁵ Pour voir l'évolution pour toutes les langues voir la dispositive 5 sur <https://doi.org/10.5281/ZENODO.7113196>

ne peut affecter une taille de billet à un type de blogs ou à une période donnée tant la dispersion est constante.

Tout au plus peut-on constater une production plus importante sur les carnets de « structure de recherche » et de « programme de recherche » ces dernières années, devançant largement celle des « carnets de chercheurs » à partir de 2014 (figure 2)⁶.

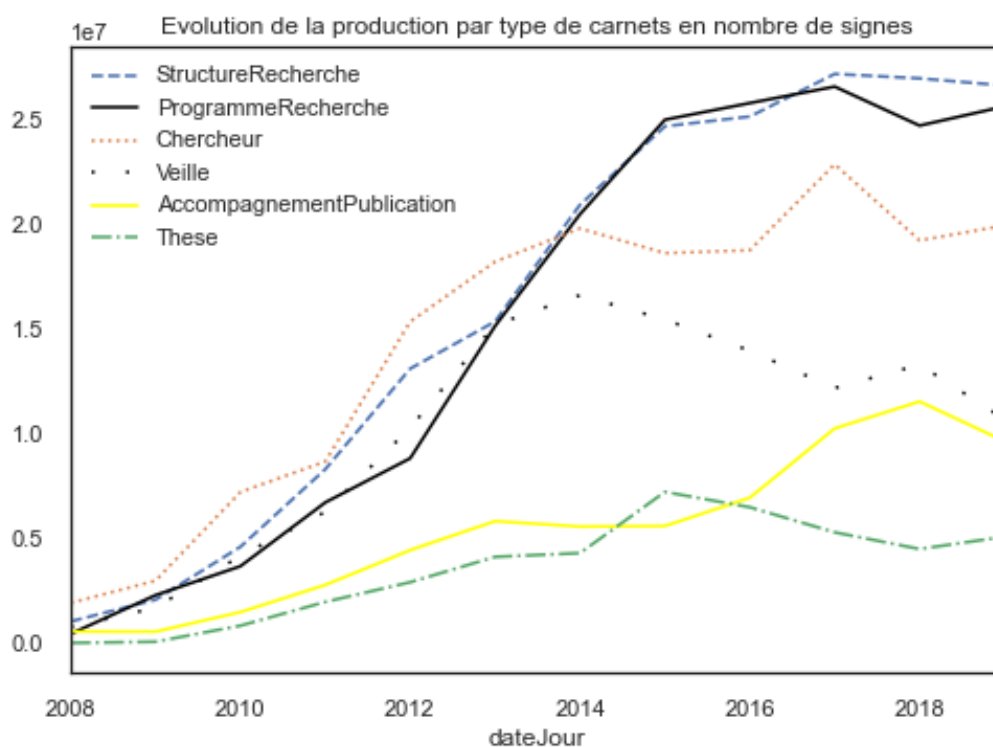


Figure 2 : Évolution de la production (nbre de signes) par type de carnets sur Hypotheses.org

II- Cartographie

L'objectif de notre cartographie est de mettre en évidence, en analysant les co-présences des auteurs sur les carnets, des liens réels entre les auteurs et plus largement des liens entre les carnets et leurs communautés. Les liens révélés par cette analyse signalent une proximité moins grande qu'une co-signature d'article, mais peuvent révéler d'autres réseaux formels ou informels. On peut supposer que les liens qui unissent les personnes écrivant sur un même carnet dépassent ceux qui unissent deux auteurs signant des articles dans une même revue. Toute personne qui intervient sur un carnet connaît au moins son créateur/administrateur/ propriétaire, ce n'est pas le cas quand deux auteurs participent à une même revue. La participation à un même carnet peut signaler une proximité institutionnelle (appartenance à un même labo par exemple), une proximité thématique ou être plus difficile à tracer (utilisation des mêmes outils de recherche, mobilisation autour d'un événement scientifique ou non, etc.).

Les études antérieures portant sur la blogosphère ont pris en considération les liens ou la navigation entre les carnets (Debaz, 2010 ; Mercklé, 2019) pour révéler communautés de pratiques ou communautés épistémiques (Stassin, 2016). Concernant *Hypotheses.org*, deux types de cartographies ont pour le moment été réalisés⁷.

⁶ Pour voir l'évolution de toutes les catégories de carnets sur hypotheses.org, voir la dispositive 6 sur <https://doi.org/10.5281/ZENODO.7113196>

⁷ Voir pour plus de détail (Mercklé, 2011)

La première, intitulée « Hypothesosphère », tente en 2010 puis en 2011 de dessiner une carte du réseau des circulations possibles à l'intérieur de la plateforme. En utilisant les liens hypertextes présents sur 91 d'entre eux (Debaz, 2019), Debaz met ainsi en évidence les renvois nombreux au carnet Leo, et souligne l'importance pour les carnetiers du caractère institutionnel de la plateforme. Il note également certains liens croisés, sans doute des liens de lecture.

Une deuxième étude, intitulée « Hypersphère », repère et liste les liens « réellement » empruntés entre deux carnets. Pierre Mercklé utilise les « logs » du serveur d'*Hypotheses.org* afin de mesurer le trafic d'un carnet à un autre à partir des informations sur les « referers », autrement dit l'adresse de la page d'où arrive le visiteur. Une première cartographie est réalisée entre janvier et avril 2011, une seconde durant 15 jours en novembre 2018. Cette dernière étude plus importante considère 2 044 carnets, 980 374 pages et analyse la circulation sur 13 226 liens entre ces carnets (densité=6,597). Les liens navigationnels sont pondérés par le nombre de personnes qui les ont empruntés (Mercklé, 2019).

Paradoxalement, la deuxième méthode semble donner une meilleure vision des circulations entre les carnets. Cette cartographie des lectures et des citations révèle une lecture des uns par les autres, probablement une proximité scientifique ou méthodologique, mais un degré d'interaction relativement faible.

1. Point méthodologique pour la production des cartes

Nous avons choisi de réaliser plusieurs cartographies des carnets hébergés sur *Hypotheses.org* à des périodes différentes. De l'ouverture de la plateforme jusqu'à la fin de l'année 2019, nous avons sélectionné des durées de 4 ans. Nous avons retenu des dates qui correspondent à des tournants pour la plateforme : 2012 pour la création de la plateforme allemande, 2016 correspond à une chute de la production en nombre de sites pour la France et l'Allemagne.

Ce découpage est nécessaire pour appréhender des périodes très différentes dans la vie longue d'*Hypotheses.org* qui héberge des carnets aux durées de vie très variables. Certains carnets n'ont existé que très temporairement à certains moments du développement, ce qui conforte le choix que nous avons fait de réaliser une analyse par période.

Pour chaque période de notre étude, nous avons construit la matrice des co-présences des auteurs sur les carnets de la manière suivante. L'ensemble des manipulations de données a été réalisé en utilisant la librairie panda du langage de programmation python sur la distribution Anaconda.

- Suppression des billets des auteurs dits « génériques ». Dans notre corpus, les logins des auteurs sont très souvent proches de leur véritable nom. Nous avons cependant identifié quelques logins employés indifféremment sur plusieurs carnets sans pouvoir garantir qu'ils désignent une même et unique autorité d'un carnet à l'autre (Rédaction, Administrateur, etc.). Les billets de ces logins d'auteur « génériques » ne seront pas pris en compte pour la production des cartographies, soit 2,2% des billets. Le nouveau corpus contient 380 299 billets publiés sur 3521 carnets différents.
- Création de la matrice des auteurs par carnet. Il s'agit d'une matrice de présence-absence (binaire) qui ne prend pas en considération le poids de la participation de chaque auteur sur les carnets. Ce choix nous permet de mettre au même niveau les auteurs principaux et secondaires des carnets afin de repérer également des liens faibles entre les carnets. Les auteurs publient en fait rarement sur un grand nombre de carnets, même une participation anecdotique révèle un investissement inhabituel. Ne considérer que les liens forts reviendrait à retracer les réseaux de carnets individuels construits au fil du temps.
- Création de la matrice de co-présence d'auteurs sur les carnets. En multipliant la matrice des auteurs par carnet par sa transposée, on obtient ainsi une matrice d'adjacence pondérée qui

recense pour deux carnets du graphe le nombre d'auteurs en commun. Les matrices d'auteurs par carnet étant relativement creuses, les calculs peuvent s'opérer aisément.

	période 1 : 2008-2011	période 2 : 2012-2015	période 3 : 20016-2019	Ensemble du corpus
nombre de carnets actifs	509	1950	2933	3540
nombre de logins différents (hors génériques)	1156	5793	8128	12 879
nombre de carnets reliés	146	887	1203	1743
% carnets en réseau	28,7%	45,5%	41,0%	49%
nombre de liens entre les carnets	164	2006	1953	3378

Tableau 1 : Caractéristiques des cartes selon les périodes

Nous avons utilisé Gephi avec l'algorithme ForceAtlas pour le placement. Deux carnets sont reliés entre eux s'ils partagent au moins un login d'auteur en commun dans les articles qu'ils hébergent. Le trait est d'autant plus épais que le nombre d'auteurs commun est important. La taille du nœud est proportionnelle à son degré, c'est-à-dire au nombre de liens qu'il a avec d'autres nœuds. La couleur est d'autant plus foncée que le degré est important.

Dans une telle cartographie, un blog isolé peut parfaitement avoir un grand nombre d'auteurs, c'est-à-dire rassembler une vraie communauté qui, du fait de sa taille, peut n'avoir rien publié dans d'autres carnets.

Le tableau 1 présente le nombre de carnets reliés pour chaque période. Il apparaît que sur l'ensemble des périodes considérées, le nombre de carnets isolés est très important. En rappelant que 49% des carnets du corpus n'ont qu'un seul auteur et que 88,6% des auteurs ne publient que sur un seul carnet, on comprend aisément que la grande partie des carnets individuels n'apparaît de facto pas sur les cartographies. Il existe toutefois une petite communauté d'auteurs très investis qui alimentent à fois des carnets personnels et collectifs.

2. Composante fortement connexe sur la période 2008-2011

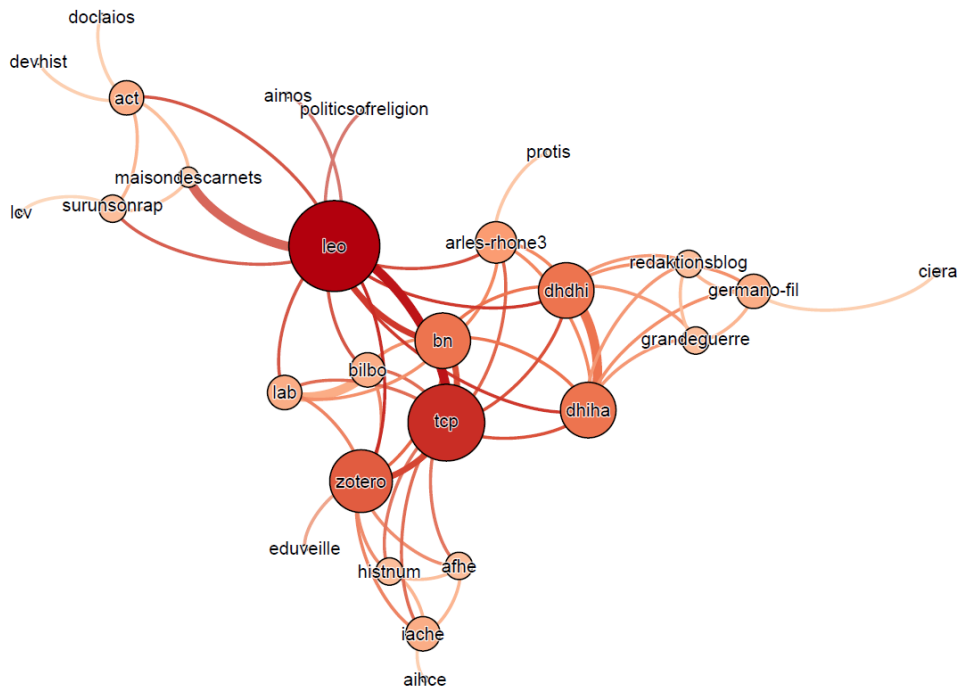


Figure 3 : Représentation de la plus grande CFC sur la période 1

La plus grande composante fortement connexe (CFC) sur la première période relie 27 carnets (figure 3). Les trois nœuds les plus fortement connectés, les *hubs*, sont des carnets français (*Leo*, *Tcp* et *Zotero*). *Leo* (*L'édition électronique ouverte*), le carnet d'*OpenEdition*, est sans surprise, le plus important. Il regroupe sur la première période 26 auteurs différents dont 8 participent à d'autres carnets.

Dans les premiers temps de l'histoire de la plateforme, les liens institutionnels entre auteurs ont largement contribué à établir de nombreux ponts entre les carnets. *Leo* est le carnet du portail qui héberge *Hypotheses.org*. Les personnes qui y travaillent peuvent être investies dans différents projets de la plateforme. Il est donc naturel de trouver des auteurs en commun entre *Leo* et les autres carnets portés par le Cléo : la *Maison des carnets* qui explique comment créer son carnet et les différentes règles à respecter, le carnet du programme de recherche *Bilbo* ouvert en 2011 et actif jusqu'en 2014, et le carnet d'*OpenEdition Lab*⁸ créé pour mieux rendre compte de l'activité de recherche et de développement du Cléo. Auteurs en commun ne signifie pas pour autant que la thématique des carnets le soit. En effet, parmi les auteurs des carnets du Cléo, certains y sont souvent employés un temps comme ingénieur et participent également à d'autres carnets directement liés à leurs propres activités de recherche. Parmi eux, mentionnons Loïc Le Pape, Karim Hammou ou Céline Barthonnat, tous les trois actifs sur *Leo* et respectivement sur les carnets *Politicsofreligion*, *Surunsonrap* et *Aimos*. Soulignons ici que *Bn* (Blogo-numericus), le carnet de Marin Dacos, directeur du Centre pour l'édition électronique ouverte, se retrouve en position assez centrale puisque son auteur intervient à la fois sur des carnets portés sur le Cléo (*Léo*, *Lab*, *Bilbo*) ou sur des carnets collectifs comme *Zotero* et *TCP*. À l'époque cependant, ce carnet n'était pas hébergé sur *Hypotheses.org*. Il a été « rapatrié » en 2015 avec tous ses anciens billets⁹. Des liens institutionnels apparaissent aussi du côté allemand entre le carnet de structure de recherche de l'Institut historique allemand de Paris, *Dhdhi* portant sur la thématique des humanités numériques et les

⁸ Ce carnet n'est créé qu'en 2014 mais il reprend l'ensemble des billets édités sur le carnet *Bilbo*, d'où sa présence sur la carte de la première période.

⁹ Adresse URL du blog de Marin Dacos avant 2015 -<http://blog.homo-numericus.net/>

Figure 4 : La plus grande CFC sur la période 2

Un deuxième nœud important de la carte est le carnet *Réflexivités* qui expérimente une écriture collective. Il accueille des billets individuels ou collectifs, écrits spécifiquement ou parus précédemment sur d'autres blogs, des extraits de « grands textes » français et étrangers. Il est relié de fait avec d'autres blogs, plus ou moins personnels, mais qui ambitionnent d'essayer de nouvelles formes d'écriture ou de médiation, de remplir de nouvelles fonctions (carnet *Enthese, Dynamiques*). Sur ce carnet se retrouvent un grand nombre des auteurs les plus « productifs » d'*Hypotheses.org*. Il s'est constitué autour d'une thématique interdisciplinaire.

À l'inverse, le troisième nœud qui regroupe les auteurs qui participent à *Devhist* est complètement disciplinaire. *Devhist* est un carnet de méthodologie qui se veut « un espace de formation, de réflexions collectives sur les méthodes en histoire, de veille et de partage de références sur l'historiographie, la méthodologie et l'épistémologie de l'histoire ». Les carnets qui l'entourent sont, dans leur écrasante majorité, tenus par des historiens.

À côté de ces trois grands nœuds, une myriade de petits regroupements désigne des communautés diverses qui peuvent signaler une grosse institution dont les membres ont ouvert des carnets ou un petit groupe (laboratoire) qui s'est saisi d'*Hypotheses.org*. C'est le cas par exemple du laboratoire Intru.

Un exemple : le cas du réseau autour d'Intru

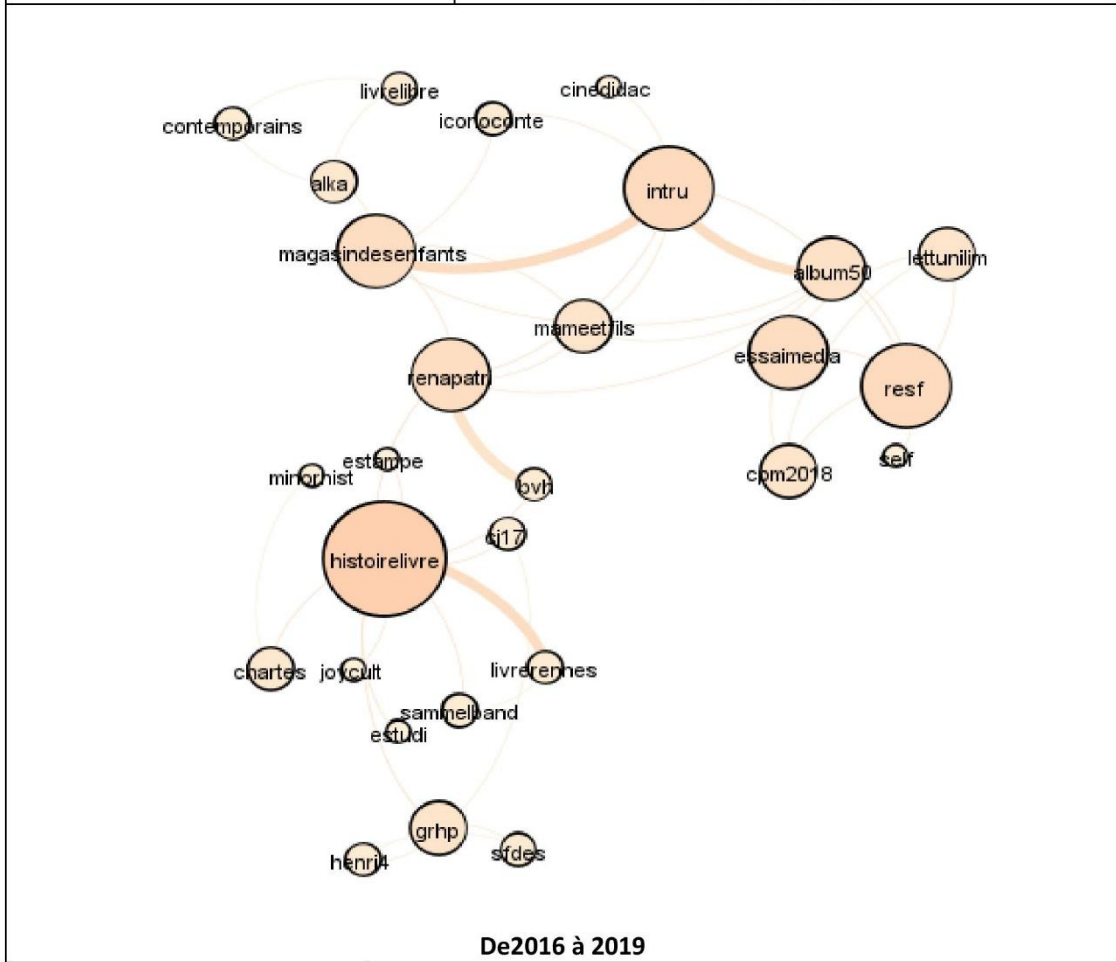
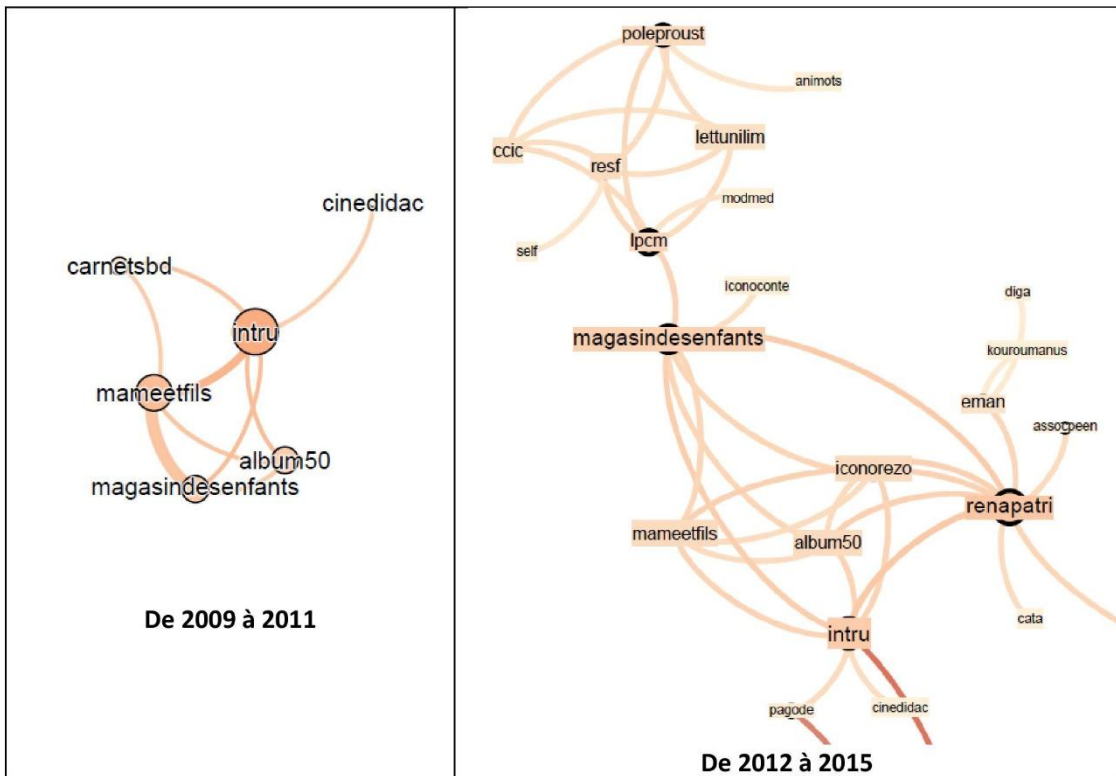
Les carnets liés à un laboratoire (Intru) sont sur une thématique assez large (culture graphique). Le laboratoire Intru réunit « des chercheurs et des chercheuses issus de l'histoire de l'art et de l'architecture, de la littérature, la philosophie, l'esthétique de la bande dessinée, l'histoire de la photographie, du cinéma, du design ».

Ses membres ont ouvert différents carnets sur la première période signalés sur le carnet du laboratoire¹⁰ mais ils publient également dans des carnets moins signalés sur le laboratoire et qui correspondent à des projets de recherche (figure 5). Ainsi, Cécile Boulaire publie des billets dans *Mame & fils*, qui est un carnet d'un programme de recherche financé par l'A.N.R. qu'elle coordonne. Elle intervient également dans le carnet *Magasin des Enfants* (de l'association française de recherches sur les livres et objets culturels de l'enfance). Laurent Gerbier, un autre membre du laboratoire qui publie dans *Intru*, collabore également à *Carnetsbd, Album50, Mameetfils*. *Cinedidac* est le login derrière lequel intervient Valérie Vignaux, maître de conférences à l'Université de Tours qui héberge Intru de 2006 à 2017. Il est utilisé dans le blog *Cinedidac* et le carnet *Intru*.

Entre 2012 et 2016, on passe de six carnets, liés les uns aux autres pendant la première période, à 22. Les anciens carnets sont encore là, certains ont disparu pour renaître sous un nouveau nom ou avec de nouvelles fonctions. On passe progressivement d'un réseau dont on peut expliquer la cohérence en notant que ses auteurs appartiennent à un même laboratoire, à un réseau dont la cohérence tient à une thématique partagée. Le réseau s'agrandit et regroupe de nouveaux acteurs, des carnets d'institutions plus imposantes (*Lpcm*, carnet de l'Association internationale des chercheurs en « Littératures populaires et culture médiatique ») ou *Histoirelivre*, animé par la Société bibliographique de France.

L'évolution du réseau autour d'Intru peut s'interpréter comme un bon exemple de la manière dont une communauté trouve à s'exprimer sur une plateforme commune. On ne peut pas préjuger de ce qui est cause et conséquence de l'existence forte de cette communauté sur *Hypotheses.org*. Cette présence est-elle le fait de quelques convaincus qui poussent à l'ouverture de nouveaux carnets au fur et à mesure que leur carrière se développe ou est-ce qu'attirés par la notoriété de certains blogs d'autres viennent les rejoindre ? Le réseau en tout cas s'étend continuellement entre 2009 et 2020 et relie progressivement de nouveaux carnets auxquels les « anciens » participent.

¹⁰ *Album '50'*, carnet de Cécile Boulaire, *Iconoconte*, carnet de François Fièvre, *Territoire des images*, carnet de Raphaële Bertho, *Territoires de l'album*, carnet de Christophe Meunier.



L'analyse des cartographies de co-présences, réalisées sur des périodes courtes, montre que le réseau des collaborations change constamment : des carnets centraux à certains moments parce qu'ils regroupent des interventions de beaucoup d'acteurs sur la plateforme peuvent à une autre période se trouver en périphérie (Leo par exemple). On observe également que la plasticité de la plateforme permet des configurations d'acteurs très différents. Les réseaux sont très connectés en Allemagne, et beaucoup moins en France. Les réseaux de co-écriture se constituent autour d'une discipline, de thématiques transversales, d'institutions, ou de publications en ligne.

Certaines disciplines des SHS sont plus représentées dans ces cartes que d'autres, surtout dans les deux premières périodes. Les Sciences de l'Information et l'Histoire sont, sans doute les plus touchées par la transformation rapide de leurs méthodes de recherche et de leur sujet d'étude, du fait du développement important des outils numériques. Ces disciplines profitent des carnets, plus que les autres, pour discuter et échanger. Les blogs sont d'autant plus adaptés à ces révolutions disciplinaires qu'ils sont entourés d'un discours et d'un imaginaire qui les auréolent d'une capacité à révolutionner le fonctionnement de la recherche, grâce à l'ouverture et aux échanges qu'ils permettent avec le grand public ou d'autres collègues.

En France, les porteurs du développement des humanités numériques promeuvent dès 2010 une « science ouverte » qui correspond aux valeurs d'ouverture défendues par les blogueurs. Sans surprise, les signataires du Manifeste des Digital Humanities (Dacos, 2011) qui assistent au Thatcamp, sont les premiers à échanger et co-écrire sur les carnets. Le point 9 du Manifeste appelle à « l'accès libre aux données et aux métadonnées » et cette revendication du libre accès se mêle à celle de médiation sans tiers, de discussion ouverte que promet le blogging en 2010. Les discours sur la liberté offerte par les carnets s'entremêlent avec les potentialités ouvertes par les nouveaux outils numériques et résonnent avec l'espérance d'une transformation non seulement des institutions de recherche (évaluation en particulier), mais également du positionnement de chaque chercheur envers ses travaux. D'où l'importance des carnets sur la réflexivité, sur la méthodologie, sur l'épistémologie dans les premiers temps.

Hypotheses.org, parce qu'elle offre aux auteurs une grande liberté (création de plusieurs carnets, suppression libre, configuration diverse, possibilité de plusieurs types de carnet) les rend propriétaires de leur espace. C'est ce qui garantit des écritures « multicarnet », des invitations de collègues et en général une adaptation de l'investissement au cours du temps. Si certains auteurs ont plusieurs carnets en même temps, une grande majorité passe de l'un à l'autre générant des liens entre communautés.

La plateforme largement occupée par les chercheurs en leur nom propre ou pour un projet de recherche a été investie peu à peu par les institutions. Mais derrière elles, ce sont bien souvent toujours les mêmes chercheurs qui écrivent. Le lien qui, de fait, existe ainsi entre carnets formels et informels entre les voix officielles et militantes rend la structure vivante. Même si le poids des carnets en réseau diminue légèrement au cours de l'existence d'*Hypotheses.org*. (45,5 % des carnets sont liés durant la deuxième période et seulement 41% durant la troisième), ce sont ces liens d'écriture, qu'on ne trouve pas ailleurs (ni dans les publications officielles ni dans les écrits associatifs), entre structures et individus qui font de la plateforme un lieu extrêmement intéressant à étudier si on veut, par exemple, repérer la dissémination d'idées ou d'auteurs.

Aujourd'hui l'originalité d'*Hypotheses.org* par rapport aux autres plateformes de blogging tient à la grande diversité d'écrits qu'elle héberge et aux formes d'investissement des auteurs qu'elle autorise. Les sciences dures ne sont pas représentées sur *Hypotheses.org*, mais l'absence de modèle éditorial, tel qu'il existe ailleurs, permet de vrais croisements d'écrits, de points de vue.

Bibliographie

Blanchard, A. (2014). Quand la culture scientifique s'affranchit sur le web : L'exemple des blogs de science en français (2003-2014). 17.

- Blanchard, A. (2017). Les blogs de science dans la recherche et la médiation scientifique : Pourquoi, comment et pour qui ? Éditions du Cercle de la Librairie. <https://www.cairn.info/les-sciences-en-bibliotheque--9782765415251-page-265.htm>
- Brown Jarreau, P. (2015). Science bloggers' self-perceived communication roles. *Journal of Science Communication*, 14(04), A02. <https://doi.org/10.22323/2.14040202>
- Crane, D. (1969). Social Structure in a Group of Scientists : A Test of the « Invisible College » Hypothesis. *American Sociological Review*, 34(3), 335-352. <https://doi.org/10.2307/2092499>
- Dacos, M. (2011, mars 26). Manifeste des Digital humanities – THATCamp Paris. ThatCamp Paris. <https://tcp.hypotheses.org/318>
- Dacos, M., & Mounier, P. (2010). Les carnets de recherche en ligne, espace d'une conversation scientifique décentrée. In *Lieux de savoir, T.2, Gestes et supports du travail savant* (p. N/A). Albin Michel. https://archivesic.ccsd.cnrs.fr/sic_00439849
- Debaz, J. (2010, décembre 8). Hypothésosphère [Billet]. *Socio-informatique et argumentation*. <https://socioargu.hypotheses.org/1921>
- Debaz, J. (2019, janvier 15). Qu'est-ce qui se trame entre les carnets de recherche ? « L'Hypothésosphère » en 2018 [Billet]. *Socio-informatique et argumentation*. <https://socioargu.hypotheses.org/5775>
- Gallezot, G., & Deuff, O. L. (2010). Chercheurs 2.0 ? *Les Cahiers du numérique*, 5(2), 15-31.
- Kjellberg, S. (2010). I am a blogging researcher : Motivations for blogging in a scholarly context. *First Monday*. <https://doi.org/10.5210/fm.v15i8.2962>
- Kjellberg, S., Haider, J., & Sundin, O. (2016). Researchers' use of social network sites : A scoping review. *Library & Information Science Research*, 38(3), 224-234. <https://doi.org/10.1016/j.lisr.2016.08.008>
- Lefebvre, M. (2013). L'infra-ordinaire de la recherche. Écritures scientifiques personnelles, archives et mémoire de la recherche. *Sciences de la société*, 89, 3-17. <https://doi.org/10.4000/sds.203>
- Malkan, S. (2018, décembre 6). Hank Campbell's Maze of Monsanto- Loving Science Blgos- -ION Publications LLC Archives. U.S. Right to Know. <https://usrtk.org/tag/ion-publications-llc/>
- Marht, M., & Puschmann, C. (2014). Science blogging : An exploratory study of motives, styles, and audience reactions. *Journal of Science Communication*, 13(3). https://jcom.sissa.it/archive/13/03/JCOM_1303_2014_A05
- Marion, L. S., Garfield, E., Hargens, L. L., Lievrouw, L. A., White, H. D., & Wilson, C. S. (2003). Social network analysis and citation network analysis : Complementary approaches to the study of scientific communication. Sponsored by SIG MET. *Proceedings of the American Society for Information Science and Technology*, 40(1), 486-487. <https://doi.org/10.1002/meet.1450400186>
- Mercklé, P. (2011, juin 7). L'Hyposphère. De la sociologie et plein d'autres choses. <http://pierremerckle.fr/2011/06/lhyposphere/>
- Mercklé, P. (2019, août). Hyposphère 2. 10 ans d'Hypotheses. <https://halshs.archives-ouvertes.fr/halshs-02270049>
- OpenEdition team. (2021, avril 7). Hypothèses : Histoire d'une collaboration internationale [Billet]. L'Édition électronique ouverte. <https://leo.hypotheses.org/17463>
- Poupardin, E., & Fauray, M. (2018). Hypotheses : L'inscription d'une pratique de communication dans l'activité de recherche. *Revue française des sciences de l'information et de la communication*, 15. <https://doi.org/10.4000/rfsic.4877>
- Raymond, D. (2010, juillet 26). Blogosphère en crise : Le PepsiGate. Agence Science-Press. <https://www.sciencepresse.qc.ca/actualite/2010/07/26/blogosphere-crise-pepsigate>
- Shema, H., Bar-Ilan, J., & Thelwall, M. (2015). How is research blogged? A content analysis approach. *Journal of the Association for Information Science and Technology*, 66(6), 1136-1149. <https://doi.org/10.1002/asi.23239>
- Small, H. (1973). Co-citation in the scientific literature : A new measure of the relationship between two documents. *Journal of the American Society for Information Science*, 24(4), 265-269. <https://doi.org/10.1002/asi.4630240406>

- Stassin, B. (2016). *La blogosphère info-doc : Une communauté de savoir, une mosaïque de médiations : étude des dynamiques sociales, socio-thématiques et discursives d'un réseau de blogueurs experts*. Cépaduès-éditions.
- Zivkovic, B. (2012, juillet 10). Science Blogs—Definition, and a history. Scientific American Blog Network. <https://blogs.scientificamerican.com/a-blog-around-the-clock/science-blogs-definition-and-a-history>
- Zou, H., & Hyland, K. (2020). Academic blogging : Writers' views on interacting with readers. *Iberica*, 2020(39), 267-294.
- Zou, H. (Joanna), & Hyland, K. (2020). "Think about how fascinating this is" : Engagement in academic blogs across disciplines. *Journal of English for Academic Purposes*, 43, 100809. <https://doi.org/10.1016/j.jeap.2019.10080>