Traces du Web et sphère publique transnationale. Approches méthodologiques et apports épistémologiques

Benjamin Loveluck, Laurent Beauguitte

▶ **To cite this version:**

Benjamin Loveluck, Laurent Beauguitte. Traces du Web et sphère publique transnationale. Approches méthodologiques et apports épistémologiques. 13e congrès de l'Association française de science politique, Jun 2015, Aix-en-Provence, France. halshs-01166486

HAL Id: halshs-01166486
https://shs.hal.science/halshs-01166486

Submitted on 22 Jun 2015

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.

Congrès AFSP Aix 2015

Section thématique 7
« Les appropriations méthodologiques d'internet dans la recherche sur des objets politiques »

Benjamin LOVELUCK, Université Paris Diderot, GIS-CIST (b.loveluck@gmail.com)
Laurent BEAUGUITTE, CNRS, UMR IDEES (beauguittelaurent@hotmail.com)

Traces du Web et sphère publique transnationale. Approches méthodologiques et apports épistémologiques

Cette communication propose une réflexion à la fois méthodologique et théorique sur la sphère publique transnationale telle qu'elle se déploie sur le Web, et sur les enseignements qu'il est possible de tirer à partir de la collecte et de l'analyse de deux formes de traces : les hyperliens et les flux RSS.

Nous nous appuierons pour cela sur deux projets ANR auxquels nous avons collaboré : le projet e-Diasporas (2009-2012) dirigé par Dana Diminescu, et le projet Géomédia (2013-2016) dirigé par Claude Grasland. Le projet e-Diasporas (www.e-diasporas.fr) a permis de mener des analyses topologiques de réseaux de sites diasporiques sur le Web, à partir d'une trentaine de corpus (Diminescu, 2012a ; Diminescu, 2012b). Pour ce projet, ont été en partie développés des outils devenus génériques (tels que Navicrawler ou encore Gephi) permettant de crawler, archiver, visualiser et analyser les réseaux d'hyperliens sur le Web, ainsi qu'une méthodologie quali-quanti de validation des corpus et de l'interprétation des cartes ainsi produites. L'objectif du projet Géomédia (http://geomedia.hypotheses.org), toujours en cours, est de mener une recherche sur les échanges contemporains de flux d'information médiatique au niveau international, à partir d'une base de données stockant les fils RSS associés aux articles publiés par une centaine de journaux en différents points du monde. Il s'agit d'en déduire deux types d'informations portant respectivement sur les flux et sur les événements internationaux, et ce travail a déjà été appliqué à une diversité de cas tels que la crise Syrienne, l'affaire Snowden (Giraud & Severo, 2013), l'étude des co-occurences pour mettre en évidence les structures globales (Beauguitte *et al.*, 2014) ou encore la médiatisation des séismes (Giraud *et al.*, 2013).

Tout d'abord, nous tâcherons de définir la notion de sphère publique transnationale, dans ses liens avec les médias et l'événement, et d'en situer les principaux enjeux. Notre contribution voudrait ensuite revenir, dans un deuxième temps, sur les difficultés concrètes que pose la collecte de données « nativement numériques », aussi bien s'agissant :

- de la mise en œuvre pratique ;
- du codage de l'information ;
- de la production de métadonnées permettant d'exploiter ces informations.

Enfin dans un troisième temps, nous souhaitons ouvrir une discussion sur ce que ces deux projets nous apprennent de la sphère publique transnationale, à partir de la formation ou de la transformation d'identités collectives diasporiques d'un côté, et de la circulation des événements de l'actualité internationale de l'autre. Il s'agit notamment d'interroger le sens qui peut être donné aux « traces du Web » à partir de deux démarches très différentes, et leurs apports épistémologiques respectifs.

1. Médias, événements et sphère publique transnationale. Définition et enjeux

De manière très générale, et bien qu'elle soit âprement débattue, la catégorie de « transnationalisme » dans la littérature en sciences sociales s'inscrit dans la réflexion sur la mondialisation des échanges. Elle met l'accent sur une interconnectivité accrue, sur un rôle moins décisif de l'Etat-nation et sur une importance déclinante des frontières inter-étatiques à différentes niveaux (économique, social, culturel etc.). Les origines de ce concept sont multiples, mais renvoient en premier lieu à des transformations dans le champ des relations internationales au cours des années 1970 et 1980. Celles-ci prirent acte sur le plan institutionnel de la multiplication des acteurs non-étatiques (institutions religieuses, entreprises multinationales etc.), entamant ainsi un déplacement de la perception stato-centrée des dynamiques géopolitiques (Rosenau, 1969 ; Keohane & Nye, 1973). En particulier, l'approche en termes d'économie politique, qu'elle soit d'inspiration réaliste (Keohane, 1984 ; Gilpin, 1987) ou critique (Strange, 1970 ; Cox, 1987 ; Strange, 1994), insista dans un contexte de crise pétrolière et de dérégulation des marchés sur le caractère de plus en plus déterritorialisés des flux commerciaux et financiers.

La catégorie de « sphère publique », depuis sa théorisation et son historicisation par Jürgen Habermas (Habermas, 1978), est également sujette à de nombreuses critiques et interprétations divergentes, à la fois sur la plan de sa représentativité et de sa légitimité politique et sur celui de sa capacité effective à influer sur les processus décisionnels, et pour le rôle d'idéal normatif que lui confère Habermas (Calhoun, 1992). Une définition a minima peut cependant en être retenue, selon laquelle il s'agit du lieu de formation de l'opinion publique – un espace situé entre la société civile et le pouvoir exécutif (initialement l'État) qui se caractérise sinon par l'argumentation dialogique à visée critique et délibérative (dans sa version habermasienne), du moins tout simplement par la communication et l'échange d'informations publiques (par opposition aux communications privées).

A la croisée de ces deux catégories, différentes perspectives disciplinaires ont été particulièrement influentes, à la fois pour redéfinir les contours du transnationalisme et son sens du point de vue de la sphère publique. La première recouvre à la fois les *cultural studies* et les *media studies.* La sociologie des migrations (notamment l'étude des diasporas) ainsi que l'étude des communautés religieuses, sont venus souligner que les interactions transfrontalières entre des communautés partageant une origine commune donnent également lieu à des sensibilités culturelles communes (Glick Schiller, Basch & Blanc-Szanton, 1992 ; Gilroy, 1993 ; Tölölyan, 1996 ; Portes, 1999 ; Schnapper, 2001). Ces recherches ont avancé dans un dialogue serré avec la sociologie des médias, qui s'est intéressée à la construction transnationale de publics à travers les flux médiatiques (Tomlinson, 1994 ; Mules, 1998). L'accent a été mis notamment sur l'essor de la télévision par satellite autorisant la diffusion de médias nationaux ou régionaux hors de leurs frontières (Sakr, 2001) et sur l'influence de chaînes d'information en continu telles que CNN sur l'agenda international (Flourney & Stewart, 1997 ; Volkmer, 1999) voire sur la « mondialisation de l'imaginaire » (Semprini, 2000), déclinant ainsi le thème mcluhanien du « village global » ou encore la perspective avancée par Anderson du rôle joué par les médias dans la constitution de « communautés imaginées » (Anderson, 1996).

La seconde perspective, très importante mais que nous laisserons de côté dans le cadre de ce travail, s'inscrit dans la sociologie des mouvements sociaux, et a notamment mis l'accent sur le rôle des mouvements contestataires ou de plaidoyer (*advocacy*) dans la formation de contre-pouvoirs ou même d'une simple conscience partagée sur le plan transnational (Guidry, Kennedy & Zald, 2000 ; Baeza, Bonnefoy & Thiollet, 2005 ; della Porta, Kriesi & Rucht, 2009 ; della Porta & Tarrow, 2005 ; Siméant, 2010).

Prenant acte de la tension inhérente à la rencontre entre transnationalisme et sphère publique, différentes hypothèses ont ainsi émergé concernant la « transnationalisation de la sphère publique », dont Nancy Fraser notamment a souligné le caractère problématique (Fraser, 2007 ; voir également Nash, 2014). Fraser rappelle que, indépendamment des limites déjà identifiées concernant sa

légitimité et de son *efficacité* politiques, la sphère publique fut initialement définie dans le cadre de l'État-nation et d'un « imaginaire politique westphalien implicite » (p. 9-14), et ne peut pas simplement être transposée à l'échelon international. Elle admet cependant qu'un espace public transnational existe déjà sous les effets des flux médiatiques, culturels, ou encore financiers transnationaux, des diasporas et du multiculturalisme, mais sous une forme qu'elle qualifie de « dépolitisée », dans la mesure où un cadre institutionnel permettant de déployer une véritable citoyenneté internationale lui fait encore défaut.

Cette approche identifie donc un processus de déplacement à l'échelle supranationale de la perception des enjeux politiques et de leur critique, qui constituerait le contrepoint des activités de plus en plus transfrontalières d'institutions supranationales, inter-étatiques, para-étatiques ou non-étatiques, mais traduirait également une conscience accrue de la dimension intrinsèquement globale des différents phénomènes (économiques, politiques, sociaux, environnementaux etc.) auxquels sont confrontés les sociétés contemporaines. Elle rejoint ainsi pour partie les propositions renouant avec l'idéalisme kantien et assumant la dimension normative initiale de l'idée de sphère publique dans le cadre des théories de la démocratie, et invitant à l'extension et à l'adaptation des institutions – en particulier des pratiques démocratiques – afin de mettre en place un cadre de citoyenneté et de gouvernance en phase avec cette nouvelle dynamique « cosmopolitique », et aux nouveaux risques et épreuves auxquels la « deuxième modernité » nous expose sur un plan global (Archibugi & Held, 1995 ; Held, 1995 ; Beck, 2008 ; Beck, 2000). Pour Fraser cependant, un tel programme implique de repenser de manière radicale à la fois la question de la légitimité ou inclusivité de la sphère publique, et celle de son efficacité ou capacité politique, dans le contexte d'un « monde post-westphalien » (Fraser, 2007, p. 19-24).

Une hypothèse plus « faible » de la sphère publique transnationale, qui ne s'appuie pas sur l'idée d'une opinion publique critique au sens habermasien, peut cependant également être retenue, et demande à être testée empiriquement. Il s'agit là d'une forme différente de l'engagement dans l'espace public, fondé non pas sur la délibération voire la mobilisation contestataire, mais d'abord sur des convergences culturelles ou des émotion partagées. Cette perspective renvoie pour partie aux analyses de Luc Boltanski sur « la souffrance à distance » qui comporte des injonctions morales et en définitive politiques mais pouvant venir « fermer » la discussion (Boltanski, 1993). Elle fait également écho au travaux de Daniel Dayan et Elihu Katz sur la dimension de communion, au sens anthropologique, induit par certains types d'événements médiatiques (Dayan & Katz, 1996)[1]. Ou encore à la proposition de Jocelyne Arquembourg concernant l'avènement d'une « sensibilité mondiale » (Arquembourg, 2006). Selon Arquembourg, la réaction émotionnelle doit être appréhendée comme une forme de jugement, « le signe d'un partage collectif des valeurs ». Analysant le cas du tsunami en Asie en 2004, elle note « une évolution qui a fait passer d'une conception de l'événement international logé dans une relation d'intériorité/extériorité à une conception globale où les spécificités des contextes politiques locaux se trouvent subsumées par l'universalisme des "causes" », tout en soulignant que « si au travers de ce phénomène certaines situations peuvent gagner une universalité qui les fait accéder au rang de "causes" sur le plan international, débouchant sur des actions pratiques (la mise en œuvre d'une surveillance technique des séismes dans des régions d'Asie où elle n'existait pas, par exemple), elle comporte aussi le risque d'amalgamer ou d'occulter des réalités plus problématiques précisément à cause de leur spécificité locale ». Arquembourg conclut : « Il semble qu'en réalité la notion même d'espace public de débat empruntée à Habermas soit malmenée par la manière dont les événements internationaux contemporains font aujourd'hui émerger des *sensibilités* qui n'ont plus rien à voir avec la rationalité de l'argumentation politique mais qui n'en incorporent pas moins des jugements et des points de vue

[1] « Des événements tels que les Jeux olympiques, le voyage d'Anouar el-Sadate à Jérusalem, les funérailles de J.-F. Kennedy ont en effet donné naissance à un nouveau genre narratif. Jouant des possibilités de diffusion simultanée offertes par les médias électroniques, les récits à l'extrême présent qui constituent ce genre unifient de vastes aires géographiques en y suscitant une temporalité commune. » (Dayan & Katz, 1996, p. 1).

normatifs » (p. 20). Ces émotions partagées – mais pas nécessairement sur un même mode, selon une même intensité ou en fonction d'un même cadrage – peuvent susciter à leur tour des jugements et des attentes politiques convergentes, par exemple en termes de sécurité. La sélection, l'intensité, le cadrage des informations, en suscitant certains types d'émotions, peuvent présenter des affinités avec des formes d'action politique spécifiques. L'événement médiatique, ici, se présente donc comme le catalyseur d'un espace d'attention commune, qui rend ensuite possible l'« activation » d'un public (Cefaï & Pasquier, 2003), la forme de ce dernier étant donc tributaire du dispositif de médiation qui lui donne naissance.

Mais nous pouvons également nous interroger sur l'existence d'espaces dotés d'une certaine autonomie face à l'événement et face à la sphère journalistique, tels qu'ils se déploient notamment sur internet et tels qu'ils participent à la constitution d'une « sphère publique en réseau » (Benkler, 2009). En effet, la question du transnationalisme a acquis une acuité nouvelle avec l'avènement d'internet. Celui-ci, en tant que vecteur essentiel d'une « société de réseaux », est fréquemment mentionné comme l'un des facteurs clés de cette évolution. S'il fut initialement invoqué comme l'instrument d'une démocratie dialogique transnationale (Bohman, 2004), des perspectives plus mesurées et adoptant une approche plus pragmatique apparaissent désormais plus convaincantes. Selon Ingrid Volkmer par exemple, le réseau permet d'investir de manière continue et ubiquitaire (sur le train, dans la rue etc.) non seulement des espaces de sociabilité restreinte (réseaux familiaux, communautés d'amis ou d'intérêt etc.) mais également des espaces publics ou semi-publics « subjectivement vécus » (Volkmer, 2014, p. 1-4), où l'information circule à différents niveaux, dans des interactions complexes entre médias – petits et grands, de masse, communautaires, associatifs ou encore militants etc. – et réseaux socionumériques. Ainsi les « horizons médiatiques » convergent-ils plus ou moins, et les « sphères d'interdépendance communicationnelles » affectent-elles de manière différenciée notre conception de la citoyenneté, ainsi que nos pratiques et nos engagements vis-à-vis de la sphère publique. Pour suivre Noortje Marres, il s'agit de saisir comment non seulement les événements eux-mêmes mais également la matérialité des dispositifs communicationnels orientent et donnent forme à des publics (Marres, 2007 ; Marres & Lezaun, 2011), que le Web permet de tracer et de circonscrire (Marres & Rogers, 2005), et qui peuvent acquérir une dimension transnationale et éventuellement manifester un engagement au sein d'une controverse dont les enjeux dépassent le cadre local ou national.

Nous voudrions donc montrer dans la suite de cet article comment les méthodes d'investigation des traces du Web permettent de circonscrire certaines dimensions de ce que nous pouvons qualifier de *sensibilités communes*. Pour cela, nous mobilisons deux approches complémentaires déjà évoquées, la sociologie des diasporas et la sociologie des médias, abordées à partir de méthodes numériques qui seront explicitées dans la partie suivante. Loin des visions très larges et peu étayées empiriquement d'un cosmopolitisme en devenir, l'hypothèse à vérifier est qu'une multiplicité d'opinions publiques peuvent être identifiées, qui partagent soit un point de vue commun, soit un cadrage commun des enjeux, au-delà ou en-deçà des frontières nationales. Les diasporas et les flux médiatiques, nous l'avons rappelé, sont parmi les principaux phénomènes invoqués pour caractériser l'émergence d'une sphère publique transnationale. La question qui nous intéresse est donc de savoir dans quelle mesure les traces du Web nous permettent d'en objectiver les manifestations, ou du moins d'approcher les formes concrètes d'interdépendances communicationnelles et d'en saisir les aspects saillants[2].

[2] Notons donc à cet égard que les questions importantes de la *légitimité* de ces perspectives partagées et de leurs *effets* politiques sur les institutions locales, nationales et internationales qui ont un rôle décisionnel se situent hors du champ de notre étude.

2. Données du web et science politique : enjeux méthodologiques

L'internet est un réseau mondial, en expansion continue et il est somme toute logique que les sciences humaines et sociales s'intéressent de plus en plus à cet objet, qu'il soit considéré comme un matériau supplémentaire permettant de questionner des problématiques disciplinaires classiques ou comme un objet révélateur de pratiques inédites et nécessitant donc une approche disciplinaire dédiée (les *digital humanities*). Quelle que soit la posture épistémologique de départ, l'utilisation de données du web suppose trois étapes où les enjeux techniques, méthodologiques et conceptuels s'imbriquent constamment : la collecte des données, leur filtrage et enfin leur catégorisation. Comme toute donnée archivistique, et la forme numérique en cela ne présente qu'une originalité très relative, toute étude suppose un prisme d'analyse et une sélection réfléchie.

2.1 La collecte des traces numériques : principes et enjeux

Il est possible de considérer les données de l'internet comme une collection non ordonnée d'archives numériques en expansion continue et actualisée de manière permanente. Là où l'archive traditionnelle, institutionnelle ou non, obéit à des règles relativement constantes – même si les politiques et les pratiques d'archivage changent selon les contextes, les époques et les préoccupations des contemporains –, le web génère de manière a priori non contrôlée ses propres principes de conservation des données, des contenus et des traces numériques. Qu'il s'agisse des hyperliens étudiés dans l'ANR *e-Diaporas* ou des flux RSS de journaux collectés dans l'ANR *Corpus-Géomédia*, les traces sont susceptibles de changer brutalement : tel administrateur de site changera la liste de ses liens favoris, et les suppressions et ajouts de liens ne seront pas documentés ; tel quotidien décidera de cesser les flux RSS, peut-être pour privilégier Facebook ou Twitter.

La collecte des données doit donc être pensée en intégrant ces changements brutaux et toujours possibles de politique des traces numériques. Il s'agira donc de mettre en œuvre une méthode de collecte thématiquement pertinente (pourquoi retenir tel journal et écarter tel autre par exemple?) mais suffisamment large pour que le corpus soit résilient à des changements localisés et brutaux. La logique thématiquement contestable du « *big data* » (recueillir le maximum de données possibles) s'explique alors aisément : plus les données collectées sont nombreuses, plus la probabilité de pouvoir construire ensuite un corpus cohérent en fonction de telle ou telle problématique est élevée. Inversement, ne collecter que ce qui paraît pertinent *a priori* ne peut que fragiliser des projets dont l'enjeu est de saisir un phénomène avec un minimum de profondeur temporelle.

Une voie possible – et actuellement mise en œuvre dans le projet Corpus-Géomédia – est de se fixer des objectifs relativement précis en terme de composition du corpus collecté (type de contenu, origine géographique, débit journalier minimal etc.) puis de récolter au-delà de ces objectifs initiaux pour assurer une relative stabilité temporelle au corpus constitué. Ainsi, si dans une optique médiatique, il est prévu de récupérer les données numériques d'un quotidien de référence par pays étudié, il est de fait préférable de collecter les données pour au moins deux quotidiens… Seul problème lié à cette sécurisation de l'intérêt thématique du corpus collecté, plus le corpus est volumineux et plus les processus de nettoyage de la base créée seront longs et coûteux.

Dans le cas d'e-Diasporas, la collecte des sites et la constitution du corpus est réalisée à mesure que le chercheur explore et parcourt le Web à la recherche de sites pertinents. Pour cela, le chercheur s'appuie d'une part sur son expertise préalable du sujet, qui lui permet d'évaluer dans quelle mesure un site rencontré correspond à la définition d'un « site diasporique » ou d'un « site frontière »[3] et

[3] Les *sites diasporiques* sont des sites créés et/ou administrés par des migrants, ou qui traitent de migrations et de diasporas : blogs, sites associatifs, sites institutionnels, portails, forums etc. ; les *sites frontières* n'appartiennent pas à la

mérite à ce titre d'être retenu, et sur laquelle il se fonde également pour coder et classifier ces différents sites. Soulignons donc que la nature des données collectées dépendra pour partie de l'interprétation qui sera faite de ces définitions initiales assez générales – une problématique classique, à laquelle les méthodes numériques n'échappent pas, et que le projet e-Diasporas s'est efforcé de rendre explicite.

D'autre part, le chercheur est « équipé » d'outils adéquats, et en particulier du logiciel Navicrawler (Jacomy & Ghitalla, 2007) ; celui-ci se présente comme un module pour le navigateur Firefox, qui permet d'enregistrer les liens des pages visitées ainsi que les liens sortants des pages visitées, et d'annoter les sites collectés à l'aide de tags. Ce travail semi-automatisé permet au chercheur de conserver un contrôle sur la constitution de son corpus, et ainsi de toujours rester attentif au contexte des hyperliens conservés. Cet ensemble de données se présente comme un graphe, dont les pages Web constituent les nœuds et dont les hyperliens constituent les liens. Il pourra ensuite être exporté, de manière à être traité et représenté par des logiciels d'analyse et de visualisation de graphes.

2.2 De la donnée « brute » à la donnée exploitable

Les traces numériques collectées sont toujours des données « sales » (Bastard et al, 2013), qui nécessitent des pré-traitements avant de pouvoir être exploitées.

S'agissant des corpus d'hyperliens du projet e-Diasporas, l'une des principales difficultés consiste à vérifier le sens des liens collectés (en particulier les liens sortants des pages visitées), et de s'assurer que ne se sont pas glissés dans le corpus des liens inutiles pour l'analyse, ou qui risqueraient d'en fausser le sens : liens publicitaires, crédits techniques etc. Un travail de vérification de la dimension la plus automatisée du processus se révèle donc indispensable, afin de pouvoir déterminer les limites et la représentativité de l'échantillon sélectionné.

Enfin, le passage crucial à la visualisation, à travers l'utilisation du logiciel Gephi (en grande partie développé dans le cadre de cette ANR) qui permet de générer des graphes à partir de ces données, impose également un certain nombre de précautions. Le logiciel s'appuie en effet sur des algorithmes de spatialisation (voir notamment Jacomy *et al.*, 2014) et permet de modifier les paramètres visuels du graphe (taille, forme, couleur des nœuds et des liens), influant ainsi sur l'interprétation des données ; il s'avère donc nécessaire de garder à l'esprit les partis pris de cet outil lors de l'analyse, afin de ne pas surinterpréter les dimensions topologiques du réseau ainsi produit (Bahoken *et al.*, 2013).

Dans le cadre des flux RSS collectés par l'ANR Corpus-Géomédia, trois problèmes techniques plus ou moins délicats se sont posés à ce jour : la datation des données, l'existence de doublons et enfin les problèmes d'encodage de caractères.

Le premier problème est plus délicat qu'il n'y paraît : l'outil de collecte construit dans le cadre du projet affecte à chaque item (i.e. chaque flux est constitué d'un ensemble d'items correspondant le plus souvent, mais pas toujours, à un article de la version papier du journal) une date et une heure de collecte. Ces informations sont en heure française. Or le décalage horaire peut être conséquent lorsque l'on capte les RSS de journaux néo-zélandais par exemple. En outre, la date de publication renseignée dans le flux RSS lui-même n'est pas toujours fiable. L'étude fine de la diffusion d'une information dans la sphère médiatique mondiale serait *a priori* possible au sein du corpus étudié mais affecter de manière automatique une heure précise et homogène à l'ensemble des items fait partie des problèmes informatiques non encore résolus à ce jour.

---

diaspora étudiée, mais présentent de fortes connections avec elle : par exemple, des sites gouvernementaux ou des sites de médias du pays d'origine.

La nécessité d'un dédoublonnage (total ou partiel) n'était pas apparu parmi les problèmes possibles avant les premiers tests menés sur la base : pour certains journaux, certains items sont strictement identiques (titre, description, lien url etc). Or cette redondance de l'information ne touche ni tous les journaux ni même tous les items d'un même journal. On pourra par exemple avoir durant deux heures tous les items d'un journal en double puis une récolte « normale ». Or une analyse quantitative ne prenant pas en compte cette redondance incontrôlée donnerait des résultats biaisés. Des chaînes de traitement ont dû donc être mise au point pour contrôler ce paramètre.

Dernier problème lié aux items des flux RSS, l'encodage des caractères. Si les items anglophones ne posent généralement pas de problème, les items francophones, lusophones et espagnols posent de récurrents problèmes liés à la présence de tildes, d'accents et cédilles, etc. L'uniformisation du codage des items collectés en utf-8, la non-utilisation de tableurs générant de redoutables problèmes d'affichage (qu'il s'agisse d'Excel ou d'Open Office) a permis de régler ce problème technique.

Si le corpus est construit de manière à éviter toute rupture intempestive de contenu, si les données sont nettoyées avant toute exploration, reste à leur adjoindre un certain nombre de métadonnées liées tant à l'origine des items qu'à leur contenu sémantique.

2.3 Taggage et catégorisation

Une fois le corpus créé, les données collectées et harmonisées, la dernière étape consiste à les enrichir afin de permettre une exploitation thématique. La création d'ontologies et de métadonnées apparaît comme une étape indispensable, notamment si le corpus a vocation à être ensuite rendu accessible à d'autres équipes de recherche voire au grand public. Cette exigence *a priori* élémentaire reste pourtant difficile à satisfaire dans la mesure où l'aspect lié au partage des données se pose souvent en fin de projet, quand les crédits sont plus ou moins épuisés et quand l'énergie des personnes impliquées se tournent déjà vers les projets futurs...

La création d'ontologies spatialisées est le plus souvent faite sur des corpus stables (Hill, 2006), voire à couverture géographique limitée (cf notamment les travaux menés à Pau sur le corpus de la MDIR – Sallaberry, 2014). Dans le cas de Géomédia, le caractère nécessairement évolutif du corpus constitué, tout comme les compétences des personnes impliquées dans le projet, ont réduit l'ambition initiale pour fournir trois types d'informations complémentaires aux données nettoyées : des informations relatives au média émetteur, des informations relatives aux lieux et parfois aux acteurs cités, et enfin des informations relatives à des événements précis. L'objectif est toujours de faciliter la prise en main du corpus pour des études thématiques tant pour des personnes impliquées dans le projet que pour des personnes extérieures au projet.

La première catégorisation produite concerne les médias émetteurs définis en termes qualitatifs (type de journal, nationalité) et en termes quantitatifs (débit moyen d'items). En ce qui concerne les lieux cités, un dictionnaire a permis de reconnaître de façon automatique la présence d'entités étatiques au sein d'un item. Des tests actuellement en cours permettront de connaître précisément la marge d'erreur en fonction des journaux – l'objectif étant de la maintenir sous le seuil des 5 %. Lorsque l'acteur permet d'identifier l'État (ex. Obama et États-Unis dans la presse mondiale, Hollande et France dans la presse française ou européenne), il a été inclus dans le dictionnaire. En ce qui concerne les événements repérables dans les items, seuls les événements ayant un spectre lexical réduit peuvent être identifiés ; dans le cas contraire, le nombre de faux positifs devient trop important. Si le cas d'Ebola (cf *infra*) a donné lieu à des premières explorations prometteuses, d'autres thèmes pourtant *a priori* facilement identifiables (ex. séismes) posent des problèmes qu'une procédure automatique ne peut résoudre aisément (Grasland et Le Texier, 2014). Ainsi, si le champ lexical des séismes est relativement réduit (Richter, *earthquake*), il apparaît d'une part dans certains médias des individus nommés Richter et sans lien avec un quelconque tremblement de terre et,

d'autre part, certains événements liés à la vie politique notamment, sont fréquemment décrits dans le langage journalistique comme des séismes (ex score du Front national en France).

Pour le projet e-Diasporas, une partie des métadonnées a été produite en cours de collection : le système de taggage incorporé à l'outil Navicrawler permet en effet de réaliser un codage « dynamique » des données. Comme nous l'avons déjà souligné cependant, ce travail d'enrichissement des données implique des opérations de codage manuel, qui s'inscrivent pleinement dans le cadre des méthodologies classiquement mobilisées dans les sciences sociales. D'autres métadonnées ont cependant été générées après coup, grâce à une « boîte à outils » mise à disposition des chercheurs : interrogation des bases de données de bureaux d'enregistrements de noms de domaines afin de recueillir des informations sur le propriétaire d'un nom de domaine (en particulier sa localisation géographique ou celle du serveur hébergeant le site), logiciels d'analyse textuelle afin d'extraire des entités nommées (personnes, lieux, organisations), ou encore logiciels de reconnaissance des langues utilisées (afin de donner des indications sur le multilinguisme potentiel d'un site ou d'une communauté diasporique).

Là encore cependant, un travail de vérification et de prise en compte de certaines incertitudes a dû être effectué, révélant des problèmes qui, pour certains, n'ont pas pu être efficacement surmontés. Par exemple, la géolocalisation affichée d'un serveur est-elle fiable, ou bien s'agit-il d'un proxy ? Il devient alors difficile d'établir un lien systématique entre une e-Diaspora et les structures matérielles (localisation physique de l'administrateur, de l'hébergement et des serveurs) qui la supportent.

3. Données du Web et sphère publique transnationale

3.1 Identités collectives et e-Diasporas

Les travaux sur l'utilisation par les migrants des technologies numériques d'information et de communication ont permis de situer le phénomène dans le cadre du paradigme de la « présence connectée » (Licoppe, 2004), en soulignant que l'expérience migratoire s'inscrit désormais dans des logiques de « présence à distance » multiples, qui combinent différents niveaux de proximité ou de distance physique (Diminescu, 2005 ; Diminescu, 2010), et qui ont des incidences profondes sur la formation des identités individuelles et collectives. Afin d'en approcher certaines des dimensions, le projet e-Diasporas Atlas s'est intéressé aux réseaux de sites diasporiques sur le Web, combinant des méthodologies d'analyse de réseaux avec d'autre formes d'expertise sociologique en termes d'aires culturelles. En effet, en étudiant les réseaux d'hyperliens il a été possible de mettre au jour des structurations différentes selon les corpus étudiés, et d'en explorer les significations en fonction des diasporas auxquelles ils appartiennent (Diminescu, 2012a ; Diminescu, 2012b).

Les moyens de communication constituent une entrée classique dans l'étude des diasporas, en particulier avec la mobilisation des échanges épistolaires (Elliott, Gerber & Sinke, 2006). De même, différentes contributions récentes ont cherché à comprendre l'articulation entre nouveaux médias et formation des identités diasporiques contemporaines (Georgiou, 2006 ; Alonso & Oiarzabal, 2010). La présence des migrants sur le Web a donné lieu à de multiples formes d'activité, et certains travaux se sont plus particulièrement intéressés aux sites Web diasporiques, en s'appuyant notamment sur des entretiens et/ou des analyses de discours (Brinkerhoff, 2009 ; Trandafoiu, 2013). Mais s'ils ont permis d'éclairer le phénomène, ces approches se concentrent sur une petite quantité de sites et ne permettent pas d'appréhender la caractéristique principale des sites diasporiques : leur inscription dans un *réseau*, ce qui a des implications décisives du point de vue de leur *représentation* (c'est-à-dire de leur visibilité publique en tant qu'entités collectives), mais également du point de vue des *formations identitaires* collectives. Le Web, en effet, a transformé la manière dont les diasporas se forment, évoluent, et se représentent en tant que diasporas (Diminescu & Loveluck, 2014).

Les enseignements sociologiques tirés des analyses d'hyperliens, des traces relationnelles que ceux-ci manifestent et de la signification des associations ainsi produites, s'inscrivent dans un champ de recherche déjà bien établi (Park, 2003 ; Thelwall, 2004 ; de Maeyer, 2013 ; Rogers, 2012). Dans les cas étudiés au sein du projet e-Diasporas, la topologie générale des réseaux diasporiques (leur degré de cohésion, les sous-réseaux, le rôle charnière joué par certains nœuds du réseau etc.) mais aussi des stratégies de *mise en visibilité* ou au contraire de *dissimulation* de certaines composantes des identités diasporiques ont pu être mises en lumière, en analysant les effets d'association générés par l'établissement d'hyperliens. A ce titre, un réseau de sites diasporiques constitue l'une des incarnations du projet commun qui apporte cohésion à des communautés dont les individus sont par nature dispersés.

Par exemple, la survivance sur le Web de la sphère culturelle Yougoslave après l'implosion de ce pays peut être identifiée, ainsi que ses interactions avec les « nouvelles » entités nationales issues de cette situation (Mazzucchelli, 2012). Des informations sur la langue et le contenu de cette « yougosphère » peuvent être apportées, ainsi que les sites ayant acquis un rôle central dans le réseau ou ceux qui, à l'inverse, en sont détachés. Autre exemple, l'analyse des sites de la diaspora nationaliste hindoue, en grande partie administrés depuis les États-Unis, a permis de mettre l'accent sur les liens établis avec certains extrémistes conservateurs juifs américains, avec qui ils partagent des orientations islamophobes, mais aussi les « trous » dans le réseau qui dénotent des stratégies d'évitement (Therwath, 2012).

Enfin dernier exemple, l'étude des réseaux de sites « diasporiques » palestiniens sur le Web (en tenant compte de la complexité et de la charge politique du terme diaspora dans ce contexte), a permis de mettre en évidence une oscillation entre déterritorialisation et reterritorialisation, « un glissement thématique et démographique d'organisations de communautés Palestiniennes à l'étranger vers un réseau de solidarité transnationale centré sur les droits palestiniens et le mouvement du Boycott » (Ben-David, 2012, p. 459) – c'est-à-dire vers des ONG et des associations militant pour la « cause » palestinienne (qui mettent en avant les droits de l'homme et le mouvement de boycott), plutôt que vers des organisations qui revendiqueraient avant tout l'appartenance à une « communauté » palestinienne (et pour qui l'existence de la diaspora et le droit au retour constituent l'enjeu principal)[4]. Ainsi, les liens entre acteurs diasporiques et acteurs non-diasporiques (organisations de la société civile, associations, fondations etc. dont les activités sont tournées vers le problème palestinien) s'avèrent plus marqués que les liens entre les différentes sous-communauté diasporiques elles-mêmes (voir Figure 1).

[4] « The shift from the network's starting-points as diaspora actors to its resulting transnational solidarity network is perhaps indicative of the lack of ties between diaspora actors among themselves and in relation to the center of gravity. Rather, they are brought together by their strong ties to non-diaspora civil-society organizations in their host societies, which are dedicated more to the Palestinian cause and less to issues related to the Palestinian diaspora.

The network reveals aspects of the dynamics of the Palestinian diaspora that emerge on the Web: it is no longer defined around Palestine as a *place of origin*, but is instead constructed around Palestine as a *point of reference*; its organization is less around a network of familial, social and transactional ties between *communities* of Palestinians who have been dispersed to many places in the world, and more around global *advocacy networks* that transcend their immediate social networks; and its members are no longer only *Palestinians abroad*, but also *natives of the host countries* who identify with the Palestinian cause. » (*ibid.* p. 470).

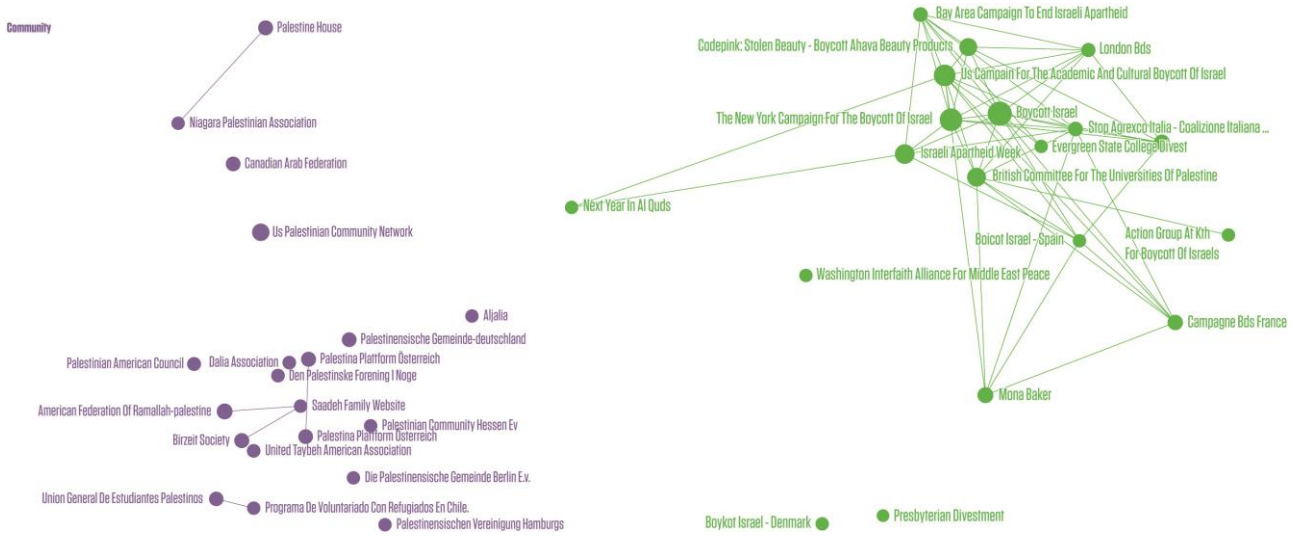

Figure 1 – Ben-David, 2012 : sous-graphes de réseaux de sites diasporiques palestiniens. A gauche ceux dont l'orientation est avant tout « communautaire », à droite la priorité au mouvement de boycott

Ces résultats viennent compléter et affiner des travaux antérieurs, qui ont montré une corrélation entre les réseaux d'hyperliens et le cadrage des enjeux. Les ONG israéliennes œuvrant localement pour le processus de paix se situent ainsi dans un réseau distinct de celui des acteurs transnationaux (en particulier s'agissant de problèmes précis telles que le débat autour de l'érection d'une « barrière de sécurité » entre Israël et les territoires palestiniens), en dépit de positions militantes convergentes – et ce en raison de la nécessaire prise en compte du contexte local par les ONG israéliennes, qui doivent composer avec un cadrage avant tout national ou régional, et les spécificités de l'opinion publique israélienne (en particulier le rejet de toute interférence par les institutions internationales) et risquent de voir leurs positions disqualifiées si elles s'alignent sur celles des acteurs transnationaux (Rogers & Ben-David, 2008).

Ce type d'analyses permet ainsi de mettre au jour, du point de vue des réseaux d'hyperliens sur le Web, différentes dimensions de la sphère internationale dans toute sa complexité. Celle-ci est faite de sensibilités culturelles qui perdurent au-delà de la disparition des Etats-nations (la « yougosphère »), d'affinités idéologiques et religieuses qui traduisent des positions communes ou au contraire cherchent à dissimuler des intérêts partagés (la diaspora nationaliste hindoue), ou encore de sous-réseaux qui, bien que leurs positions politiques ou leurs origines culturelles puissent converger, se déploient pour partie sur un plan transnational, et pour partie sur un plan local et régional (les ONG et les réseaux diasporiques palestiniens).

3.2 L'événement géomédiatique et la sphère publique transnationale : le cas d'Ebola

L'étude de l'actualité internationale et des hiérarchies qu'elle produit a une longue histoire dont il peut être utile de rappeler les principaux résultats (Beauguitte, 2014) : l'information internationale reflète en grande partie les inégalités entre territoires et, à événement dramatique d'importance égale, la logique du mort kilomètre (plus c'est lointain et moins cela est important) reste un facteur explicatif majeur (pour une étude globale, voir Wu, 2000). À cela s'ajoute une très nette tendance à l'auto-centrisme de l'actualité internationale : l'État où se situe le média émetteur (et donc son lectorat) joue un rôle déterminant (voir le cas nord-américain dans Saleem, 2007 ; et les pics d'apparition de l'Afrique dans les médias français dans Robinet, 2013).

Une étude actuellement menée sur l'épidémie d'Ebola montre que l'intensité de la couverture médiatique ne suit que très imparfaitement le déroulement de l'événement. Si le cas Ebola a été

choisi comme terrain d'étude, c'est en raison de son caractère relativement borné dans le temps (la date du premier cas détecté par l'OMS est connue) et dans l'espace. La couverture de l'épidémie a été mondiale seulement lorsque les victimes ont été des personnes originaires des pays développés et plus encore quand des cas (isolés rappelons-le) ont été détectés dans ces mêmes pays. Comparer les courbes du nombre de décès et de cas comptabilisés par l'OMS durant la période de l'épidémie et le nombre d'items RSS liés à l'épidémie permet de mettre en évidence ces logiques médiatiques en partie déconnectées des logiques sanitaires.

La Figure 2 présente les résidus d'un modèle de régression cherchant à expliquer le nombre d'items évoquant Ebola par journal et par semaine en fonction du nombre de cas annoncé par l'OMS durant la même période. L'intensité des carrés correspond au nombre d'items émis par le journal durant la semaine considérée. Plus le cercle est grand, plus le résidu est élevé ; la couleur indique si le résidu est positif (rouge : on parle plus d'Ebola qu'attendu étant donné le nombre de cas) ou négatif (bleu : on parle moins d'Ebola qu'attendu étant donné le nombre de cas).

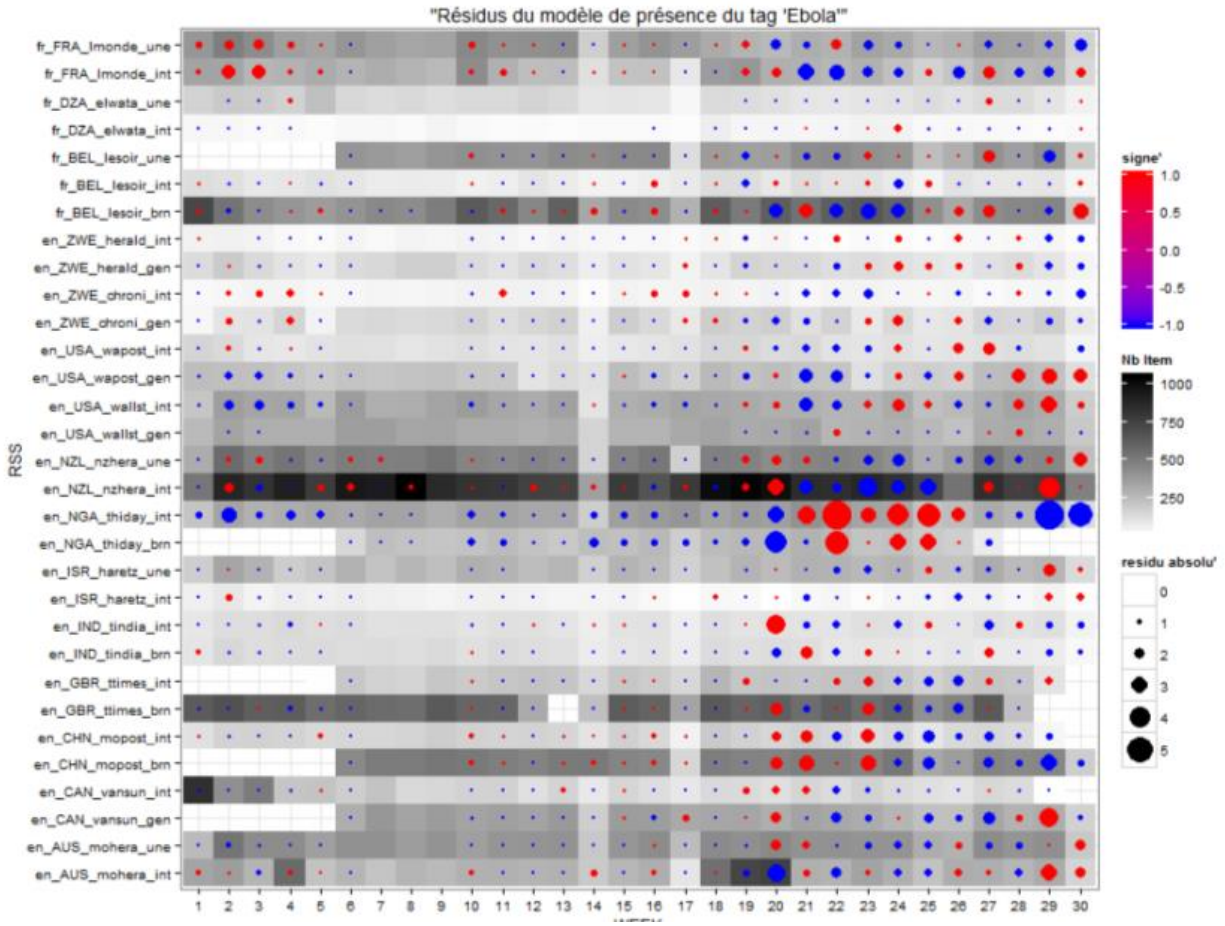

Figure 2 : Liens entre situation épidémique et couverture médiatique. ***Extrait du working paper Ebola, ANR Corpus-Géomédia, Claude Grasland et Hugues Pecout, 2014***

Les codes utilisés à gauche de la figure permettent de repérer la langue du flux (en_ anglais, fr_ français), le pays émetteur noté en code-iso3 num (ex GBR pour Royaume-Uni), le nom abrégé du journal (ex wapost = Washington Post) et enfin la catégorie du flux (int pour international, brn pour breaking news, gen pour général et une pour une). Il faut noter que cette dernière catégorisation est celle déclarée par le média émetteur.

L'analyse des flux RSS se situe du côté de la publication d'actualités, et nous interdit de nous prononcer sur la question de la *production* de ces actualités (par exemple, le rôle des agences de

presse dans la formation et pour certaines perspectives critiques le *formatage* d'une sphère publique transnationale ; cf. Boyd-Barrett & Palmer, 1979 ; Boyd-Barrett & Rantanen, 1998 ; Marchetti, 2002), ni sur la *réception* des informations ainsi diffusées. Cependant, le traitement médiatique en tant que discours public institutionnalisé fournit en soi des indications sur les intentions qui président à la représentation des événements : sélection des informations, intensité de la couverture médiatique, cadrage proposé etc. Ainsi l'étude des résidus des flux RSS des journaux anglophones (codes USA, NZL, CAN, GBR) montre qu'ils deviennent fortement positifs durant les deux premières semaines d'octobre qui voient le virus arriver aux USA et produire une contamination sur sol américain. Si l'épidémie fait l'objet d'une attention globale, elle ne le fait pas au même rythme selon le média émetteur, sa localisation et surtout celle de son lectorat.

Conclusion

Les résultats obtenus au cours de ces deux projets, s'ils n'autorisent pas à se prononcer sur les carences voire les apories politiques ou institutionnelles de l'avènement d'un espace public transnational, permettent cependant d'apporter des éclairages sur les processus de formation de l'opinion en interrogeant deux formes entrecroisées de « communautés dispersées » – les communautés diasporiques et les publics médiatiques.

Ils permettent notamment d'éclairer les rapports de pouvoir qui traversent l'espace transnational, et qui participent à structurer les publics en son sein. Ainsi les réseaux diasporiques sur le Web viennent-ils matérialiser l'affirmation d'une identité partagée, face et au-delà du système des États-nations : la vitalité de la « Yougosphère » après la disparition du pays sur le plan administratif ; à l'inverse, les liens qui peuvent s'établir entre différentes formes de nationalisme à caractère religieux par l'intermédiaire de communautés diasporiques formant entre elles des alliances objectives. Ainsi également la circulation des informations médiatiques sur le plan international, telle qu'elle a pu être retracée par l'intermédiaire des flux RSS, permet-elle d'affiner ce que Paul Virilio appelle pour sa part la « synchronisation des affects » (Virilio, 2007) face à un événement, en particulier d'ordre catastrophique. Le cas de l'épidémie d'Ebola est à cet égard éloquent, puisqu'il permet de mettre en évidence, de manière précise, l'autonomisation du champ médiatique : alors que ce problème grave de santé publique touchant des pays du Sud s'intensifie graduellement, ce n'est que tardivement qu'il bénéficie véritablement d'un pic d'attention dans les pays du Nord, lorsque ceux-ci se retrouvent directement impliqués à travers leurs ressortissants. Il confirme ainsi, tout en les précisant, les asymétries dans la circulation internationale de l'information déjà mentionnées (Beauguitte, 2014 ; voir également Mattelart, 2014), et les biais ainsi établis dans la formation de sensibilités communes.

Les conclusions auxquelles nous pouvons parvenir sont donc partielles, mais participent à représenter une « sphère publique en réseau » complexe, et constituée d'une multiplicité de couches qui présentent chacune des spécificités en termes d'ancrage géographique et du degré d'hybridité entre différentes aires culturelles. Certains espaces médiatiques « nativement numériques », tels que les réseaux de sites étudiés au sein d'e-Diasporas, révèlent des communautés dont l'identité spatiale est devenue multiple, qui sont parvenues à constituer des espaces dotés d'une certaine autonomie par rapport aux découpages institutionnels, administratif ou frontaliers, mais à l'intérieur desquels les échanges ne sont pas agencés au hasard : comme nous l'apprend l'analyse de réseau, ces espaces sont en effet structurés de manière chaque fois particulière, dévoilant les orientations, les partis pris, et parfois les structures de pouvoir qui se reflètent dans les liens qui existent ou non entre différentes entités du réseau (ainsi que les entités qui n'apparaissent pas du tout dans le réseau)[5].

[5] En outre, ces analyses demanderaient à être complétées par des travaux permettant de mesurer non seulement la présence d'espaces transnationaux (sous la forme de réseaux d'hyperliens, ou de la convergence de traitements médiatiques), mais la manière dont différents dispositifs socio-techniques (plateformes du web, applications mobiles etc.) filtrent et surdéterminent l'*accès* à ces espaces. Richard Rogers *et al.* par exemple ont pointé le rôle joué par les

Néanmoins, ils se présentent comme des espaces où peuvent émerger des enjeux publics de nature transnationale.

Les logiques médiatiques traditionnelles, au contraire, semblent en grande partie ancrées dans des logiques nationales – et ce y compris lorsque l'information est traitée et diffusée sur le Web. Si quelques chaînes mondiales existent, les journaux quotidiens restent des révélateurs d'identités collectives plus nationales que transnationales, et les logiques de couverture médiatique demeurent fortement tributaires des asymétries économiques, politiques ou sociales qui structurent l'espace internationale. Ainsi par exemple l'épidémie d'Ebola n'a-t-elle pas été élevée au rang de problème public transnational, engageant à la fois une réponse émotionnelle et une responsabilité sur le plan international – du moins sur ce plan médiatique.

Des logiques complémentaires et parfois contradictoires sont donc à l'œuvre. D'un côté en effet une poussée « cosmopolite », prenant acte à la fois de l'intensification des échanges informationnels, démographiques, monétaires ou culturels ainsi que du caractère transnational d'un nombre croissant de problèmes économiques, sociaux, sanitaires, écologiques et politiques, pour conclure à la nécessité normative d'un dépassement de l'ancrage national-étatique de la sphère publique ; les communautés e-diasporiques se présentent ainsi comme des espaces intermédiaires mais aussi des passerelles, qui viennent saper les cadrages strictement nationaux des enjeux. De l'autre, et en s'intéressant comme nous l'avons fait au traitement d'une catastrophe épidémique appelant des efforts concertés et une responsabilité globale, il est important de souligner qu'au sein de la sphère publique internationale subsiste une multiplicité de décalages, de frontières, de mise en visibilité différenciées voire d'absences ; en particulier, il est nécessaire de rappeler encore et toujours non seulement la survivance de « systèmes médiatiques nationaux » (Flew & Waisbord, 2015), mais aussi les biais de *sélection* et de *cadrage* inhérents aux institutions journalistiques. Ceux-ci viennent affaiblir ou du moins sérieusement nuancer ce que Ulrich Beck qualifie de « communauté imaginée du risque globalisé » (Beck, 2011) ; cette communauté imaginée ou « conscience partagée » se révèle être à plusieurs vitesses. Les méthodes numériques mises en place, qu'il s'agisse du projet e-Diasporas ou Géomédia, permettent d'évaluer au plus près à la fois la réalité de nouveaux espaces transnationaux (telles que les diasporas sur le web), et les limites précises du traitement différencié de l'actualité internationale (à travers les flux RSS des journaux à travers le monde).

Références bibliographiques

ALONSO, A. & OIARZABAL, P.J. (dir.), (2010), *Diasporas in the New Media Age. Identity, Politics, and Community*, Reno, NV, University of Nevada Press.

ANDERSON, B. (1996), *L'Imaginaire national. Réflexions sur l'origine et l'essor du nationalisme*, Paris, La Découverte.

ARCHIBUGI, D. & HELD, D. (dir.), (1995), *Cosmopolitan Democracy. An Agenda for a New World Order*, Cambridge, Polity Press

ARQUEMBOURG, J. (2006), « De l'événement international à l'événement global : émergence et manifestations d'une sensibilité mondiale », *Hermès* n° 46, pp. 13-21

BAEZA, C., BONNEFOY, L. & THIOLLET, H. (2005), « L'invention de la contestation transnationale par les forums et sommets : la naissance d'un "espace public mondial" ? », *Raisons politiques* n° 19, pp. 25-43.

moteurs de recherche pour dessiner, à travers le tri opéré dans les résultats de requête, un espace « renationalisé » du Web (qui peut cependant inclure des sites parfois très éloignés géographiquement et culturellement), après une phase initiale caractérisée par les utopies déterritorialisées du cyberespace (Rogers, 2013).

BAHOKEN F., BEAUGUITTE L. & LHOMME S. (2013), « La visualisation des réseaux : enjeux, principes et perspectives », Papier méthodologique du groupe FMR [https://halshs.archives-ouvertes.fr/halshs-00839905].

BEAUGUITTE L. (2014), « Les déterminants de l'actualité internationale – synthèse bibliographique », *Blog Corpus-Géomédia* [http://geomedia.hypotheses.org/71, consulté le 10 mars 2015].

BEAUGUITTE L., SEVERO M. & PECOUT H. (2014), « Do International News Reflect World Structure? A Network Approach », Communication présentée au colloque EUSN, Barcelone [http://geomedia.hypotheses.org/176 consulté le 10 mars 2015].

BECK, U. (2000), « The cosmopolitan perspective: sociology of the second age of modernity », *The British journal of sociology* 51(1), pp. 79-105.

BECK, U. (2008), *La Société du risque. Sur la voie d'une autre modernité*, Paris, Flammarion.

BECK, U. (2011), « Cosmopolitanism as imagined communities of global risk », *American Behavioral Scientist* 55(10), pp. 1346-1361.

BEN-DAVID, A. (2012), « The Palestinian diaspora on the Web: between de-territorialization and re-territorialization », *Social Science Information* 51(4), pp. 459-474.

BENKLER, Y. (2009 [2006]), *La Richesse des réseaux. Marchés et libertés à l'heure du partage social*, Lyon, Presses universitaires de Lyon.

BOHMAN, J. (2004), « Expanding dialogue. The Internet, the public sphere and prospects for transnational democracy », *The Sociological Review* 52(Supplement 1), pp. 131-155.

BOLTANSKI, L. (1993), *La Souffrance à distance. Morale humanitaire, médias et politique*, Paris, Métailié.

BOYD-BARRETT, O. & PALMER, M. (1979), *Le Trafic des nouvelles. Les agences mondiales d'information*, Paris, A. Moreau.

BOYD-BARRETT, O. & RANTANEN, T. (dir.), (1998), *The Globalization of News*, London, Thousand Oaks, CA and New Delhi, Sage.

BRINKERHOFF, J.M. (2009), *Digital Diasporas. Identity and Transnational Engagement*, Cambridge and New York, Cambridge University Press.

CALHOUN, C.J. (dir.) (1992), *Habermas and the Public Sphere*, Cambridge, MA, MIT Press.

CEFAÏ, D. & PASQUIER, D. (2003), « Introduction », in *Les Sens du public. Publics politiques, publics médiatiques*, sous la direction de D. Cefaï & D. Pasquier, Paris, Presses universitaires de France, pp. 13-59.

COX, R.W. (1987), *Production, Power, and World Order. Social Forces in the Making of History*, New York, Columbia University Press.

DAYAN, D. & KATZ, E. (1996), *La Télévision cérémonielle. Anthropologie et histoire en direct*, Paris, Presses universitaires de France.

DELLA PORTA, D. & TARROW, S.G. (dir.), (2005), *Transnational Protest and Global Activism*, Lanham, MD, Rowman & Littlefield.

DELLA PORTA, D., KRIESI, H., & RUCHT, D. (dir.), (2009), *Social Movements in a Globalizing World*, 2nd expanded edition, Basingstoke, Palgrave Macmillan.

DIMINESCU, D. (2005), « Le migrant connecté : pour un manifeste épistémologique », *Migrations Société* 17(102), pp. 275-292.

DIMINESCU, D. (dir.) (2010), « Dossier : les migrants connectés. TIC, mobilités et migrations », *Réseaux* n° 159.

DIMINESCU, D. (dir.) (2012a), *e-Diasporas Atlas. Explorations and Cartography of Diasporas on Digital Networks*, Paris, Ed. de la Maison des Sciences de l'Homme.

DIMINESCU, D. (dir.) (2012b), « Special issue : diasporas on the Web », *Social Science Information* 51(4).

DIMINESCU, D. & LOVELUCK, B. (2014), « Traces of dispersion: online media and diasporic identities », *Crossings: Journal of Migration & Culture* 5(1), pp. 23-39.

ELLIOTT, B.S., GERBER, D.A., & SINKE, S.M. (dir.), (2006), *Letters Across Borders. The Epistolary Practices of International Migrants*, New York and Basingstoke, Palgrave Macmillan.

FLEW, T. & WAISBORD, S. (2015), « The ongoing significance of national media systems in the context of media globalization », *Media, Culture & Society* 37(4), pp. 620-636.

FLOURNEY, D.M. & STEWART, R.K. (1997), *CNN. Making News in the Global Market*, Luton, University of Luton Press.

FRASER, N. (2007), « Transnationalizing the public sphere. On the legitimacy and efficacy of public opinion in a post-westphalian world », *Theory, Culture & Society* 24(4), pp. 7-30.

GEORGIOU, M. (2006), *Diaspora, Identity and the Media. Diasporic Transnationalism and Mediated Spatialities*, Creskill, NJ, Hampton Press.

GILPIN, R. (1987), *The Political Economy of International Relations*, Princeton, NJ, Princeton University Press.

GILROY, P. (1993), *The Black Atlantic. Modernity and Double Consciousness*, Cambridge, MA, Harvard University Press.

GIRAUD, T. & SEVERO, M. (2013), « Le périple d'Edward Snowden. Analyse quali-quantitative d'un évènement médiatique international », *Netcom. Networks and Communication Studies* 27(3-4), pp. 385-410.

GIRAUD, T., GRASLAND, C., LAMARCHE-PERRIN, R., DEMAZEAU, Y. & VINCENT, J.-M. (2013), « Identification of International Media Events by Spatial and Temporal Aggregation of RSS Flows of Newspapers. Application to the Case of the Syrian Civil War between May 2011 and December 2012 », *Proceedings of the 18th European Colloquium on Theoretical and Quantitative Geography (ECTQG'13)*, Dourdan, France, p. 112-114.

GLICK SCHILLER, N., BASCH, L. & BLANC-SZANTON, C. (1992), « Transnationalism: a new analytic framework for understanding migration », *Annals of the New York Academy of Sciences* 645, pp. 1-24.

GRASLAND, C. & LE TEXIER, M. (2014), « Les tremblements de terre comme outil d'étalonnage des flux RSS », *Blog Corpus-Géomédia* [http://geomedia.hypotheses.org/264, consulté le 10 mars 2015].

GUIDRY, J.A., KENNEDY, M.D., & ZALD, M.N. (dir.), (2000), *Globalizations and Social Movements. Culture, Power, and the Transnational Public Sphere*, Ann Arbor, MI, University of Michigan Press.

HABERMAS, J. (1978), *L'Espace public. Archéologie de la Publicité comme dimension constitutive de la société bourgeoise*, Paris, Payot.

HELD, D. (1995), *Democracy and the Global Order. From the Modern State to Cosmopolitan Governance*, Cambridge, Polity Press.

HENDLER, J., SHADBOLT, N., HALL, W., BERNERS-LEE, T. et WEITZNER, D. (2008), « Web science: an interdisciplinary approach to understanding the Web », *Communications of the ACM* 51(7), pp. 60-69.

HILL, L.L. (2006), *Georeferencing: The geographic associations of information (Digital Libraries and Electronic Publishing)*, Cambridge, MA, MIT Press.

JACOMY, M. & GHITALLA, F. (2007), *Méthodologies d'analyse de corpus en sciences humaines à l'aide du Navicrawler. Rapport final,* Paris, Fondation de la Maison des Sciences de l'Homme, programme TIC-Migrations [http://webatlas.fr/wp/share/navicrawler/Guide_m%E9thodo_NC2007.pdf].

JACOMY, M. et al (2014), « ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software », *PloS one* 9(6).

KEOHANE, R.O. (1984), *After Hegemony. Cooperation and Discord in the World Political Economy*, Princeton, NJ, Princeton University Press.

KEOHANE, R.O. & NYE, J.S. (dir.), (1973), *Transnational Relations and World Politics*, Cambridge, MA, Harvard University Press.

LICOPPE, C. (2004), « Connected presence: the emergence of a new repertoire for managing social relationships in a changing communication technoscape », *Environment and Planning D: Society and Space* 22, pp. 135-156.

DE MAEYER, J. (2013), « Towards a hyperlinked society: a critical review of link studies », *New Media & Society* 15(5), pp. 737-751.

MARCHETTI, D. (2002), « L'internationale des images », *Actes de la recherche en sciences sociales* 145, pp. 71-83.

MARRES, N. (2007), « The issues deserve more credit: pragmatist contributions to the study of public involvement in controversy », *Social Studies of Science* 37(5), pp. 759-780.

MARRES, N. & LEZAUN, J. (2011), « Materials and devices of the public: an introduction », *Economy and Society* 40(4), pp. 489-509.

MARRES, N. & ROGERS, R. (2005), « Recipe for tracing the fate of issues and their publics on the Web », in *Making Things Public. Atmospheres of Democracy*, sous la direction de B. Latour & P. Weibel, Cambridge, MA, MIT Press, pp. 922-935.

MATTELART, T. (2014), « Les enjeux de la circulation internationale de l'information », *Revue française des sciences de l'information et de la communication* n° 5.

MAZZUCCHELLI, F. (2012), « What remains of Yugoslavia? From the geopolitical space of Yugoslavia to the virtual space of the Web Yugosphere », *Social Science Information* 51(4), pp. 631-648.

MULES, W. (1998), « Media publics and the transnational public sphere », *Critical Arts* 12(1-2), pp. 24-44.

NASH, K. (dir.) (2014), *Transnationalizing the Public Sphere*, Cambridge and Malden, MA, Polity Press.

PARK, H.W. (2003), « Hyperlink network analysis: a new method for the study of social structure on the Web », *Connections* 25(1), pp. 49-61.

PORTES, A. (1999), « La mondialisation par le bas. L'émergence des communautés transnationales », *Actes de la recherche en sciences sociales* 129, pp. 15-25.

PRICE, M.E. et DAYAN, D. (dir.) (2008), *Owning the Olympics. Narratives of the new China*, Ann Arbor, MI, University of Michigan Press.

ROBINET, F. (2013), « La fabrique médiatique des événements internationaux : Afrique, un continent en marge (1994-2008) ? », *Le Temps des médias*, 20(1), p. 152-167

ROGERS, R. & BEN-DAVID, A. (2008), « The Palestinian-Israeli peace process and transnational issue networks: the complicated place of the Israeli NGO », *New Media & Society* 10(3), pp. 497-528.

ROGERS, R. (2012), « Mapping and the politics of Web space », *Theory, Culture & Society* 29(4-5), pp. 193-219.

ROGERS, R. (2013), « Ch. 6 – National Web studies », in *Digital Methods,* Cambridge, MA, MIT Press, pp. 125-151.

SEMPRINI, A. (2000), *CNN et la mondialisation de l'imaginaire,* Paris, CNRS éditions.

ROSENAU, J.N. (dir.) (1969), *Linkage Politics. Essays on the Convergence of National and International Systems,* New York, Free Press.

SAKR, N. (2001), *Satellite Realms. Transnational Television, Globalization and the Middle East,* London, IB Tauris.

SALEEM N. (2007), « U.S. Media Framing of Foreign Countries Image: An Analytical Perspective », *Canadian Journal of Media Studies*, 2(1), p. 130-162.

SALLABERR C. (2014), *Recherche d'information géographique dans des corpus textuels*, Londres, ISTE éditions/Hermès, collection SIG.

SCHNAPPER, D. (2001), « De l'Etat-nation au monde transnational. Du sens et de l'utilité du concept de diaspora », *Revue Européenne des Migrations Internationales* 17(2), pp. 9-36.

SEVERO, M. et GIRAUD, T. (2014), « Le périple d'Edward Snowden: analyse quali-quantitative d'un événement médiatique international », *Netcom networks and communication studies* 27(3-4)

SIMEANT, J. (2010), « Transnationalisation de l'action collective », in *Penser les mouvements sociaux. Conflits sociaux et contestation dans les sociétés contemporaines,* sous la direction de O. Fillieule, E. Agrikoliansky, & I. Sommier, Paris, La Découverte, pp. 121-144.

STRANGE, S. (1970), « International economics and international relations: a case of mutual neglect », *International Affairs* 46(2), pp. 304-315.

STRANGE, S. (1994), *States and Markets,* 2nd edition, London and New York, Continuum.

THELWALL, M. (2004), *Link Analysis. An Information Science Approach,* Amsterdam, Elsevier Academic Press.

THERWATH, I. (2012), « Cyber-*hindutva*: Hindu nationalism, the diaspora and the Web », *Social Science Information* 51(4), pp. 551-577.

TOMLINSON, J. (1994), « Mass communication and the idea of a global public sphere », *Journal of International Communication* 1(2), pp. 57-70.

TÖLÖLYAN, K. (1996), « Rethinking diaspora(s): stateless power in the transnational moment », *Diaspora: A Journal of Transnational Studies* 5(1), pp. 3-36.

TRANDAFOIU, R. (2013), *Diaspora Online. Identity Politics and Romanian Migrants*, New York and Oxford, Berghahn.

VIRILIO, P. (2007), *L'Université du désastre,* Paris, Galilée.

VOLKMER, I. (1999), *News in the Global Sphere. A Study of CNN and its Impact on Global Communication,* Eastleigh, UK, University of Luton Press.

VOLKMER, I. (2014), *The Global Public Sphere. Public Communication in the Age of Reflective Interdependence,* Cambridge and Malden, MA, Polity Press.

WU D. (2000), « Systemic Determinants of International News Coverage: A Comparison of 38 Countries », *Journal of Communication*, 50(2), p. 110-130.