



HAL
open science

La cartographie du Web : le lien social sur le Net

Marta Severo

► **To cite this version:**

| Marta Severo. La cartographie du Web : le lien social sur le Net. 2012. halshs-00678768

HAL Id: halshs-00678768

<https://shs.hal.science/halshs-00678768>

Preprint submitted on 13 Mar 2012

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.

La cartographie du Web : le lien social sur le Net

Marta Severo
GIS CIST , marta.severo<at>gis-cist.fr

Mars 2012



Introduction - Traces et méthodes numériques

Ces dernières années une nouvelle série de méthodes pour les sciences sociales est venue rivaliser avec les méthodes traditionnelles de la sociologie quantitative et qualitative : on les appelle *digital methods* ou méthodes numériques (Rogers, 2010[18]). Les méthodes qu'on regroupe sous cette étiquette sont très hétérogènes et encore peu intégrées. Cependant, elles partagent toutes le fait de se baser sur les traces numériques comme source d'information pour l'étude des phénomènes sociaux (Venturini et Latour, 2010[22]).

Le développement des méthodes numériques est relativement récent et a été favorisé par la diffusion des médias numériques dans un nombre croissant de branches de la société. Après une quinzaine d'années de gestation théorique débutée par les réflexions sur la « cyber-culture » (Negroponte, 1996[15]) et les communautés en ligne (Rheingold, 2000[17]; Turkle, 1995[21]), la première preuve de l'intérêt des méthodes numériques a été fournie par une importante étude de Google sur la diffusion de la pandémie grippale aux États-Unis¹. Grâce à l'analyse des logs des requêtes effectuées sur le moteur de recherche entre 2003 et 2008, les ingénieurs de Google ont pu identifier celles qui étaient les mieux corrélées avec les indicateurs traditionnels de surveillance de la grippe construits par le CDC (*Center for Disease Control*). Sur la base de ces requêtes, ils ont construit un indicateur de surveillance qui non seulement arrivait à une incroyable corrélation de 0.97 avec l'indicateur CDC, mais fournissait aussi des résultats une à deux semaines avant les rapports de surveillance CDC. Publié sur le journal *Nature*, ce cas a démontré que les traces numériques peuvent être employées pour étudier des phénomènes *offline* concernant des régions avec un nombre important d'internautes (Ginsberg *et al.*, 2009[8]).

1. www.google.org/flutrends (consulté le 7 mars 2012).

Outre l'analyse des données des moteurs de recherche, de nombreuses autres méthodes numériques sont récemment entrées dans la boîte à outils des sciences sociales (Lazer *et al.*, 2009[12]). Dans cet article, je présenterai la « cartographie sur le Web » (*web mapping*), une méthode qui s'est avérée à la fois robuste et utile dans l'étude des réseaux sociaux complexes.

1 La cartographie du web

La cartographie du web s'appuie sur l'idée que les liens hypertexte créés sur le web peuvent être utilisés comme proxy des liens sociaux. D'un point de vue pratique, il s'agit de retracer dans un graphe le réseau créé par les liens hypertextes présents sur un ensemble de pages web.

L'intérêt de cette technique dérive de deux régularités observées maintes fois dans la pratique de la création de liens hypertextes :

- les auteurs de sites web ne citent d'autres sites que s'ils partagent un intérêt thématique ou social ;
- les auteurs de sites web ne citent pas les sites qui ont un point de vue opposé au leur, même ceux traitant des mêmes thématiques.

Si le premier type de comportement n'est pas surprenant, le deuxième l'est davantage. Les personnes et les groupes ayant des positions opposées ont tendance à s'ignorer sur le web : ils ne se citent pas négativement, ils ne se citent pas tout court. Une des premières images révélant ce phénomène est une carte de la blogosphère politique aux États Unis (Adamic et Glance, 2005[1]). Sur cette carte, les liens entre les sites web considérés créent deux noyaux bien clustérisés, les blogs libéraux et les blogs conservateurs². Un autre exemple intéressant est tiré du travail de la Fondation Govcom étudiant les associations palestiniennes sur le web, avec la claire distinction entre Fatah et Hamas (Govcom.org, 2008[9]).

La pratique de ne pas lier ses opposants n'est pas profitable aux internautes qui voudraient connaître les différents points de vue sur le même sujet, mais elle est très utile aux chercheurs étudiant le web. En effet, elle assure que les hyperliens peuvent être pris comme indices d'autres types de liens *offline*. Autrement dit, sachant quels sites sont connectés *online*, nous pouvons inférer quelles personnes ou organisations sont liées *offline*.

Le web n'est donc pas organisé au hasard. En liant (ou en ne liant pas) leur discours à d'autres discours en ligne, les utilisateurs établissent des hiérarchies et des *clusters* (Gibson *et al.*, 1998[7]). Par conséquent, le web n'est pas un réseau unique, mais un réseau de réseaux, un graphe où des zones densément connectées sont séparées par des espaces relativement vides. Dans

2. Une étude récente de la blogosphère politique française a confirmé la même tendance. Les blogs sympathisants avec le même parti politique se citent et ne citent pas les autres. Ce phénomène a été vérifié notamment pour l'extrême droite. http://www.lemonde.fr/election-presidentielle-2012/visuel/2011/07/04/la-cartographie-de-la-blogosphere-politique_1544714_1471069.html (consulté le 7 mars 2012).

la plupart des cas, ces zones correspondent à des communautés thématiques : des grappes de personnes rassemblées autour d'intérêts et points de vue communs.

En considérant les relations entre sites web comme révélatrices des liaisons entre les acteurs sociaux que ces sites représentent, la cartographie du web peut alors être employée dans l'étude des réseaux sociaux. Bien qu'elle ne puisse fournir une description exhaustive de la nature et de la raison des liens, cette méthode peut aider les chercheurs à explorer les liens sociaux virtuels et à les représenter.

1.1 Les outils d'exploration du web

La cartographie du web s'appuie sur deux types d'outil : les outils d'exploration du web et ceux de représentation du web.

En ce qui concerne l'exploration, les outils les plus employés sont les *crawlers*. Un *crawler* est un logiciel qui permet de naviguer dans une série de pages web et de tracer tous leurs liens hypertexte. Les *crawlers* peuvent être automatiques ou, plus rarement, manuels. En outre, de nouvelles méthodologies ont été récemment développées pour explorer les données des réseaux sociaux en ligne³. Dans cet article, nous détaillerons trois exemples d'exploration en présentant le pour et le contre de chacun.

En ce qui concerne la représentation, les graphes se sont imposés comme la forme de visualisation classique pour ce genre de données. Nous pouvons représenter la topologie des réseaux que nous avons retracé avec les *crawlers* à l'aide de logiciels de manipulation de graphes comme Gephi⁴. L'idée de base de ce type de visualisation est que chaque site web est représenté par un nœud et chaque lien hypertexte par un arc sur le graphe.

Ces graphes permettent de visualiser plusieurs caractéristiques des réseaux explorés, notamment :

- deux nœuds sont rapprochés par l'algorithme de spatialisation du graphe si les sites web qu'ils représentent sont directement ou indirectement liés (Jacomy *et al.*, 2011[11]) ;
- la taille des nœuds est proportionnelle à l'*in-degree* du site web, c'est-à-dire au nombre de liens qu'il reçoit des autres sites du réseau. Cette mesure peut être considérée comme représentative de l'autorité du site web dans le réseau ;
- la couleur des nœuds peut être changée pour montrer différentes catégories que le chercheur décide d'attribuer aux sites web pour décrire différents phénomènes.

3. Ces méthodologies sont généralement basées sur l'emploi des API (Interface de programmation) mises à disposition par les plateformes des réseaux sociaux en ligne.

4. <http://gephi.org/>

2 L'exemple du cyber-activisme pendant le printemps arabe

En début 2011, les pays arabes ont été le théâtre d'importants mouvements révolutionnaires qui ont entraîné la chute de Zine el-Abidine Ben Ali en Tunisie et de Hosni Moubarak en Égypte. Ces événements ont attiré l'attention non seulement de médias mais aussi de chercheurs qui se sont trouvés à observer de nouveaux phénomènes inattendus pour ces pays. Parmi ces phénomènes, l'émergence du cyber-activisme a sûrement été l'un des plus intéressants[4][5]. La cartographie du web s'est avérée particulièrement utile pour analyser ces phénomènes. Nous avons employé cette technique pour observer la blogosphère activiste liée à la révolte. Dans les paragraphes suivants, nous présenterons trois exemples d'analyse réalisés avec différentes méthodes exploratoires.

2.1 Le cas tunisien : limites et avantages d'un *crawl* automatique

D'abord, nous avons analysé les liaisons entre les blogs activistes tunisiens et les réseaux étrangers. Pour réaliser cette analyse, nous avons pris comme point de départ les blogs identifiés par la presse comme les déclencheurs de la révolte puis nous avons dressé la liste de tous les blogs qui avaient ajouté des hyperliens vers ces blogs en utilisant des outils de tracement de liens⁵. L'analyse se concentre sur les liaisons les plus stables construites dans les mois précédents la révolution et ne considère pas les réseaux générés par Facebook et Twitter pendant les jours de la révolution⁶.

Pour le réseau tunisien, l'analyse prend en considération une quinzaine de pages web tunisiennes (blogs, mais aussi Twitter et Facebook) identifiées par la presse internationale comme les moteurs de la révolte : notamment le blog, les comptes Twitter et Facebook de Slim Amamou, le site de Nawaat⁷, le blog *A tunisian girl* de Lina Ben Mhenni, Takriz⁸ et d'autres. Grâce au tracement de liens réalisé en janvier 2011, il a été possible de reconstruire les liaisons entre ces pages et 156 sites étrangers qui les ont cités.

Nous avons cherché à tracer les liens entre tous ces sites (tunisiens et étrangers) grâce à un outil de *crawl* automatique, *issuecrawler*⁹. Un *crawler* automatique utilise un script qui suit et répertorie tous les liens d'un site puis tous les liens des sites qu'il rencontre et ainsi de suite. Le principal avantage

5. SEO Open Explorer, Yahoo Explorer et la recherche avancée de Google.

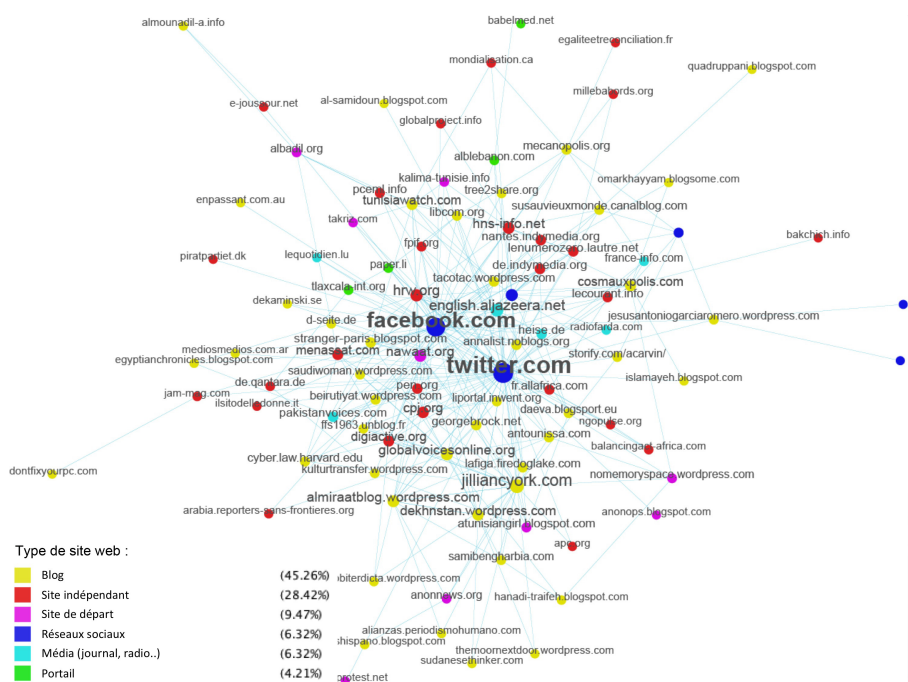
6. Ce choix est aussi dû aux limites techniques des outils qui permettent difficilement de *crawler* de nouveaux liens en temps réel.

7. Nawaat (www.nawaat.org/) est un blog tunisien collectif indépendant créé en 2004.

8. Takriz (www.takriz.com) est un projet collectif citoyen contre la censure sur Internet en Tunisie.

9. Une application développée par la Fondation Govcom Amsterdam, disponible gratuitement sur le web (<https://www.issuecrawler.net/>).

FIGURE 1 – Graphe des liens entre les blogs activistes tunisiens et les sites web qui les citent



Crawl : Issuecrawler, type de crawl : inter-actor, profondeur : 2 (février 2011).
 Visualisation : Gephi, algorithme ForceAtlas.

du *crawler* automatique est qu'il est disponible sur le web gratuitement et très facile à utiliser : il suffit de donner une liste de liens, de spécifier le type et la profondeur du *crawl* et lui fournit en réponse un réseau.

Nous avons classé les sites par type. En considérant le type de site, nous pouvons observer que la plupart des sites qui établissent des liens vers les blogs de la révolution sont des blogs ou des sites « indépendants », par exemple des médias citoyens, des ONG militant pour la liberté d'expression ou pour les droits humains ou d'autres types d'acteurs alternatifs (figure 1).

Ce réseau est clairement dominé par les plateformes Facebook et Twitter. Par ailleurs nous ne pouvons pas identifier vraiment une clusterisation par pays. Cette carte est intéressante aussi parce qu'elle permet de souligner les limites du *crawl* automatique.

Nous avons vu que les sites web mentionnent surtout des sites qui partagent leur thématique et leur point de vue. Il existe cependant des exceptions cruciales à cette règle, en particulier les auteurs de sites web citent souvent des sites qu'ils estiment être des autorités ou des ressources de valeur, même si ces sites ont un focus thématique plus large que le leur.

Le *clustering* thématique est une force puissante du web, mais à elle s'oppose une force tout aussi puissante : la loi de puissance ou *power law*. Selon cette loi, une petite minorité de sites, qui constitue la couche haute du web, reçoit la large majorité des liens hypertextes (Barabási *et al.*, 2000[3]).

La loi de puissance a plusieurs conséquences importantes, mais celle qui nous intéresse ici est ce qu'on appelle l'effet petit monde : le fait que tous les *clusters* thématiques sont connectés par l'intermédiaire des sites de la couche haute (Milgram 1967[14], revu par Watts et Strogatz, 1998[23] et par Barabási et Albert, 1999[2]). Le problème de l'effet petit monde est qu'il ne permet pas de s'appuyer exclusivement sur des *crawlers* automatiques. Les *crawlers* automatiques seraient une méthode parfaite pour explorer un *cluster web*, mais à cause de l'effet petit monde ils sont rapidement aspirés vers la couche haute du web puis vers une infinité d'autres *clusters* toujours plus éloignés du point de départ.

Cependant, en prenant en compte leurs limites, les *crawlers* automatiques peuvent être utilisés comme outils pour tracer tous les liens à l'intérieur d'un corpus de sites web déjà défini. Il faut savoir que si on sélectionne des sites de la couche haute (comme Facebook et Twitter dans notre exemple), ils seront amenés à dominer sur le réseau et nous verrons difficilement les réseaux thématiques.

2.2 Le cas égyptien : le *crawl* manuel

Le deuxième exemple concerne la révolte égyptienne. En connaissant la structure dense et réflexive de la blogosphère égyptienne, quasi exclusivement en arabe et avec peu de liaisons avec les sites étrangers (Radsch, 2008[16]), nous avons privilégié une analyse de la structure locale de la blogosphère activiste égyptienne, c'est-à-dire du réseau qui relie les principaux blogs et sites web actifs pendant la révolution (25 janvier 2011 - 11 février 2011).

Dans ce cas, nous voulions vraiment tracer un réseau thématique et, par conséquent, nous n'avons pas eu d'autre possibilité que de définir les limites de ce réseau par une exploration manuelle. Pour ce faire, nous avons employé un *crawler* manuel, une extension du navigateur Firefox appelée *Navicrawler*¹⁰.

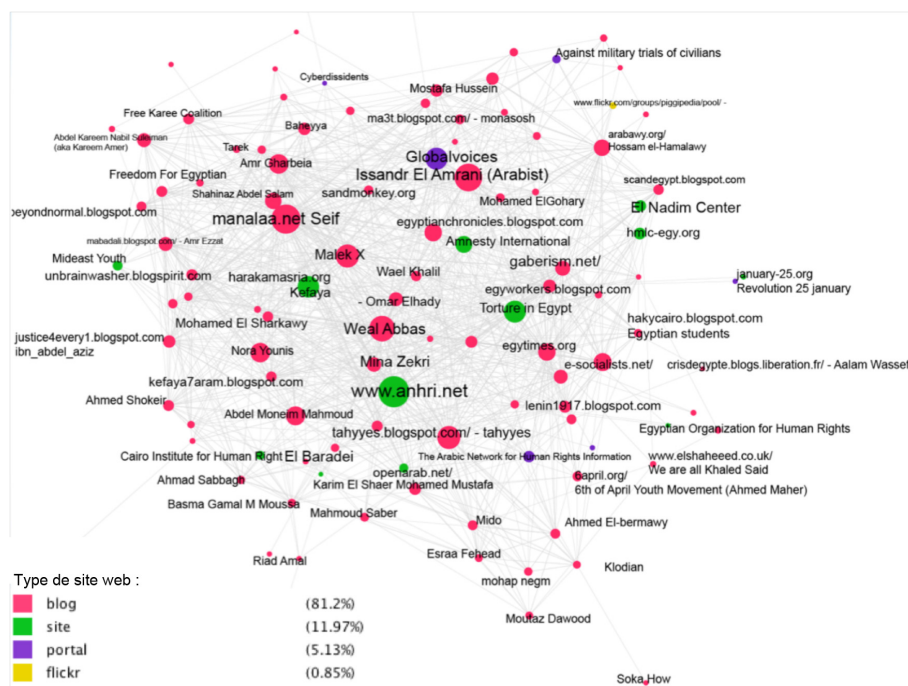
Navicrawler accomplit deux tâches :

- il suit la navigation de son utilisateur en gardant trace de tous les sites qu'il visite et de tous les hyperliens contenus dans leurs pages ;
- pour tous les sites visités, il permet à l'utilisateur de décider si le site doit être inclus ou exclu du réseau thématique.

Après avoir inclus ou exclu un nombre suffisamment large de sites web, il est possible de lancer un *crawl* automatique, qui aura comme points de départ les sites inclus et comme limites les sites exclus. Si le travail manuel

10. <http://webatlas.fr/wp/navicrawler/>.

FIGURE 2 – Graphe de la blogosphère activiste égyptienne par type de site web



Crawl : 20 juin 2011. Visualisation avec Gephi (Force Atlas).

de définition de limites a été méticuleux, le *crawler* devrait trouver tous les sites qui composent le *cluster* thématique sans sortir du *cluster*. Ce travail peut être itéré plusieurs fois pour garantir la qualité du réseau résultant.

En suivant cette procédure, nous avons défini un noyau de 116 nœuds correspondant à autant d'acteurs¹¹. Ces nœuds ont été catégorisés selon plusieurs paramètres (type, langue, localisation, activité principale de l'auteur). À travers l'analyse des stratégies de connexion entre les sites web, nous avons identifié la topologie spécifique du réseau local du cyber-activisme égyptien, en montrant qui en sont les acteurs centraux et périphériques, s'il existe des agrégats et qui joue le rôle d'autorité et de *hub* (figure 2). Cette technique d'analyse de la blogosphère n'a aucune ambition d'exhaustivité, son but est d'identifier un noyau représentatif des blogs activistes et les liaisons entre eux pour pouvoir offrir une vision générale du phénomène et identifier les directions de recherche les plus fécondes.

Ce qui émerge en général est une structure très dense et interconnec-

11. Plusieurs sites web gérés par un seul acteur (personne, institution, association ou autres) correspondent à un seul nœud.

tée. Les catégorisations ont rarement mis en évidence des agrégats locaux de sites web, c'est-à-dire des groupes de nœuds densément liés qui peuvent être identifiés comme sous-unités. En considérant la catégorisation par type de site web, nous avons observé que 80% des nœuds correspondent à des blogs, les autres sont pour la plupart des sites web d'institutions et d'ONG (*Amnesty International*, *El Nadim Center*...). Le corpus contient aussi quelques portails internationaux comme *Globalvoice.com* (le principal agrégateur international de bloggeurs activistes) et *Cyberdissidents.org* (un projet américain qui identifie les principaux bloggeurs dissidents dans plusieurs pays) et des portails thématiques liés à l'activisme égyptien, comme *Tortureinegypt.org*¹², *January-25.org* et *Tahrirdiaries.wordpress.com*. Il est important de noter que nous avons choisi de ne pas intégrer dans le corpus les réseaux sociaux (bien que 66% des bloggeurs identifiés aient aussi un compte Twitter). Cela signifie que l'image de la blogosphère que l'on va étudier ne sera pas un instantané des réseaux Internet mais plutôt une image consolidée dans la mesure où les liens tissés entre blogs ou sites web sont plus solides que les rapides notifications diffusées sur Twitter ou Facebook. Nous identifions les structures pérennes du cyber-activisme égyptien plutôt que les changements rapides et quasi-quotidiens des réseaux sociaux¹³.

L'observation de la distribution des degrés dans le graphe permet d'étudier quelques phénomènes intéressants. La taille des nœuds est basée sur l'*in-degree* : cette mesure donne le nombre de liens entrant dans un site web, c'est-à-dire le nombre de fois qu'un site web est cité par d'autres. L'*in-degree* peut donc être considéré comme un indice de l'autorité sur le web ; plus un nœud est cité, plus les autres nœuds reconnaissent son intérêt.

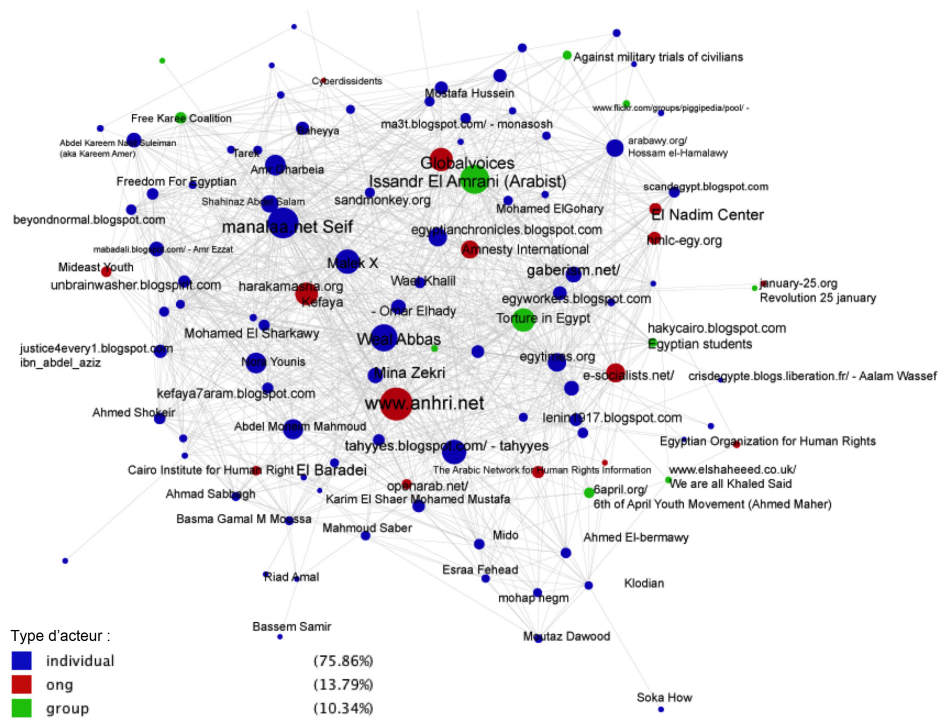
Si l'on considère la typologie des acteurs (individu, ONG, groupe), la majorité des sites web étant des blogs, ils sont gérés par des individus (76%), les autres appartiennent à des associations, à des institutions (14%) ou à des groupes informels (10%). On remarque dans la carte (figure 3) que les autorités ne sont pas seulement les « blog-stars » comme Wael Abbas (<http://misrdigital.blogspot.com/> et <http://misrdigital.com/>), Seif (<http://www.manalaa.net>) et Malek-X (<http://malek-x.net/>)¹⁴, mais aussi des groupes et des associations : le réseau arabe pour l'information sur les droits humains (*Arabic Network for Human Rights Information*, www.anhri.net) ; le portail *Torture in Egypt* (tortureinegypt.net), groupe de bloggeurs militants contre la torture ; le blog *Arabist* (<http://www.arabist.net>), géré par le journaliste Issandr El Amrani qui publie les billets de plusieurs bloggeurs concernant la politique dans les pays arabes ; et *Globalvoice.com*. Ces

12. Site web créé en 2007 par Noha Atef pour dénoncer des cas de torture en Égypte. Il est considéré comme un de premier cas d'*information activism*.

13. Pour une analyse de l'activité Twitter liée à la révolution égyptienne voir Severo, 2011[19].

14. Ces bloggeurs étaient identifiés comme activistes bien avant le 25 janvier 2011. Voir www.cyberdissidents.org (consulté le 30 juin 2011).

FIGURE 3 – Graphe de la blogosphère activiste égyptienne par type d'acteur



Crawl : 20 juin 2011. Visualisation avec Gephi (Force Atlas).

sites représentent des acteurs « historiques » (existant avant la révolution) du cyber-activisme égyptien. L'activisme, sous forme de soutien au mouvement Kifaya ou comme mobilisation autour de droits humains, s'était déjà approprié internet comme espace de protestation. Comme le souligne Courtney Radsch (2008[16]) dans son analyse de l'évolution de la blogosphère égyptienne, les blogueurs égyptiens sont devenus très tôt des activistes¹⁵.

Cette carte met donc en évidence la longévité de la blogosphère activiste égyptienne. Les protestations en ligne de janvier et février 2011 ont pu s'appuyer sur un réseau de blogueurs et de groupes d'activistes bien interconnecté et consolidé par des années de lutte pour les droits humains. Loin d'être des liaisons éphémères occasionnées par les événements de la Place Tahrir, les liens entre les blogs activistes égyptiens étaient déjà établis.

Par conséquent, la cartographie du web a permis de capturer la solidité de la blogosphère égyptienne et l'expérience des acteurs. À travers l'analyse des billets des blogs et l'observation de leur contenu pendant et après la révolution¹⁶, nous n'avons pas vu la création d'un nouveau réseau d'activistes (bien que plusieurs blogs aient été ouverts pendant cette période¹⁷), mais nous avons plutôt observé un changement du motif de la mobilisation : la lutte contre la torture a été remplacée par le soutien à Khaled Saïd ; le soutien à Khaled Saïd a été remplacé par la bataille contre Moubarak ; et après la chute de Moubarak, l'armée est devenue l'ennemie. Encore plus intéressant, si l'émergence de nouveaux motifs de lutte influence le contenu et souvent l'iconographie de blogs, elle affecte peu les liaisons entre les blogs. L'observation diachronique de la blogosphère pendant 5 mois a mis en évidence la « mort » de certains sites¹⁸ et la naissance d'autres¹⁹, mais ces événements dynamiques se sont passés à la périphérie de la blogosphère, alors que le « centre » est resté presque inaltéré autour des projets les plus stables et des causes les plus persistantes.

Une autre topologie qui mérite d'être décrite concerne la distribution des sites web par langue employée (figure 4). La carte basée sur cette ca-

15. "And despite the myriad of blog types the fact remains that Egypt's blogosphere is distinct from its regional counterparts, where the majority of blogs are personal. "In most of other Arab countries blogs are personal not activist, Egypt is exceptional", says Abd Al Moneim Mahmoud" (Radsch, 2008 : 11).

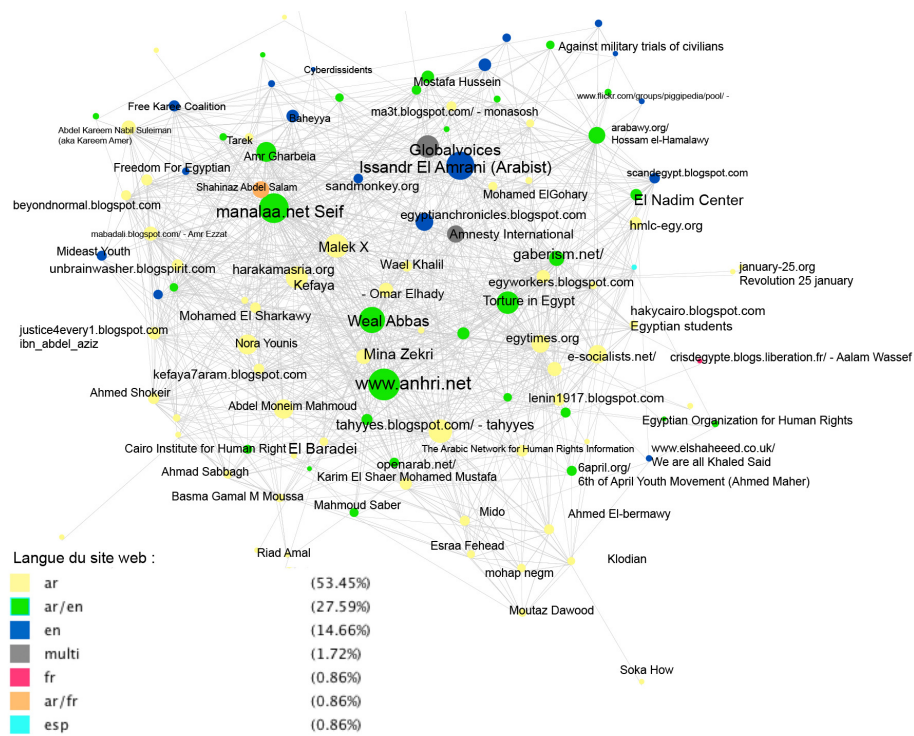
16. Nous avons observé les blogs du corpus du 25 janvier au 20 juin 2011, jour où le dernier *crawl* a été lancé.

17. Souvent les nouveaux blogs ouverts après la révolution ne manifestent pas une vraie activité. Après un ou deux billets, ils sont abandonnés.

18. Certains blogueurs « passent » aux réseaux sociaux (<http://halaho2a.blogspot.com>, <http://kefaya7aram.blogspot.com/>, <http://beyondnormal.blogspot.com/>, <http://norayounis.com/>, <http://wagdyghoneim.net>) ou des sites web concernant des motifs de lutte moins liés à l'actualité, comme le site pour la libération du blogueur Kareem (<http://www.freekareem.org/>).

19. <http://6april.org>, <http://crisdegypte.blogs.liberation.fr/>, <http://tahrirdiaries.wordpress.com/>, <http://www.flickr.com/groups/piggipedia/pool/>, <http://www.january-25.org/>.

FIGURE 4 – Graphe de la blogosphère activiste égyptienne par langue employée



Crawl : 20 juin 2011. Visualisation avec Gephi (Force Atlas).

tégorisation est la seule qui montre des phénomènes de clusterisation. Les blogs exclusivement en arabe (53%) et les blogs exclusivement en anglais (15%) constituent deux groupes bien distincts, mis en réseau par des sites multilingues (notamment *Globalvoice.com*). Il ne faut pas oublier que la blogosphère égyptienne était animée initialement par des blogs en anglais. C'est seulement dans un second temps, avec les frères Gharbeia, que se construit une blogosphère en langue arabe directement adressée au large public égyptien. Toutefois, les autorités du réseau, même dans la communauté arabe, restent des sites qui proposent des contenus autant en arabe qu'en anglais (28%). Ces sites jouent souvent le rôle de pont (*bridge*) et garantissent ainsi la liaison avec d'autres réseaux externes à la blogosphère activiste (Lynch, 2007[13]; Etling *et al.*, 2010[6]). L'analyse de cette carte souligne l'importance de la question linguistique qui reste encore un des facteurs influençant le plus la forme du réseau, ses limites et ses liens avec d'autres réseaux.

2.3 Le mot-clic « #right2vote » : la cartographie d'un réseau social en ligne

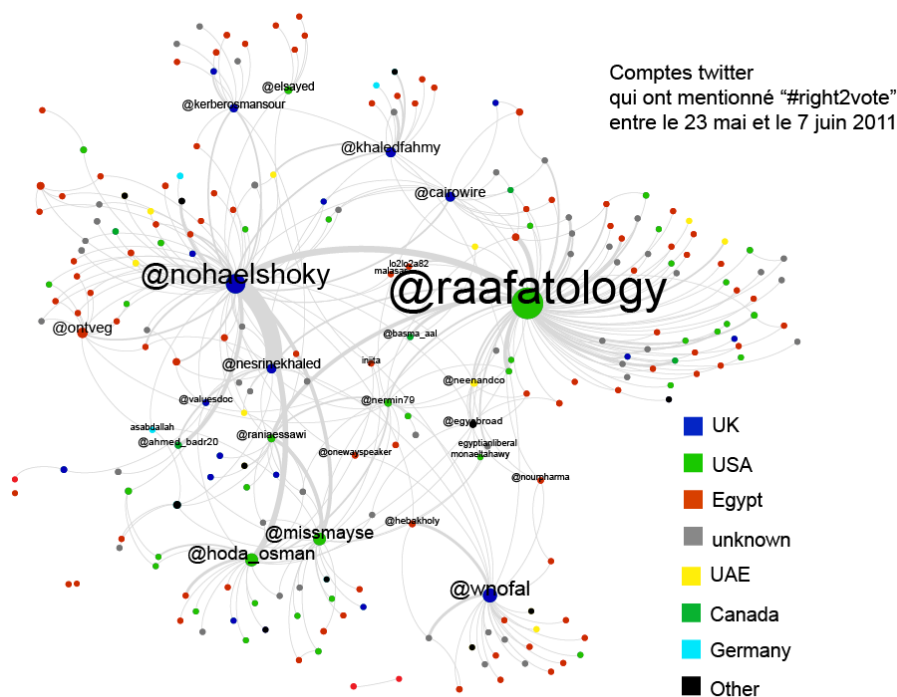
Le dernier exemple concerne la récente mobilisation des Égyptiens de l'étranger pour le droit de vote²⁰. Pendant la révolte égyptienne, non seulement les bloggeurs activistes se sont mobilisés, mais aussi les Égyptiens émigrés ont montré leur solidarité envers les compatriotes. À travers les sites web et les réseaux sociaux, les Égyptiens de l'étranger se sont mobilisés pour pouvoir voter aux prochaines élections. Nous avons étudié cette mobilisation grâce à la cartographie du web. Dans ce cas, nous avons appliqué cette technique à un type particulier de réseau web, c'est-à-dire le réseau des utilisateurs Twitter²¹, où les liens sont produits par l'activité de *retweet*. Nous avons analysé en particulier l'usage du mot-clic (appelé *hashtag* en anglais) « #right2vote ».

Le 23 mai 2011 « #right2vote » apparaît sur Twitter. Au travers du mot-clic, l'utilisateur peut attribuer un sujet à un message. Le mot-clic est alors un moyen que les membres du réseau ont d'attirer l'attention sur un thème et d'en débattre en concaténant facilement les prises de parole. Le mot-clic « #right2vote » voit sa première utilisation le 23 mai 2011. En quelques semaines (entre le 23 mai et le 7 juin 2011), 227 utilisateurs l'emploient en plus de 500 *tweets*. Nous avons archivé ces *tweets* et nous les avons analy-

20. Cette étude a été effectuée dans le cadre du projet « e-diaspora Atlas » réalisé par le programme Tic-Migrations (Fondation Maison des sciences de l'homme, Paris). Voir Severo et Zuolo, 2012[20].

21. Twitter (www.twitter.com) est un outil de réseau social et de microblogage qui permet à l'utilisateur d'envoyer gratuitement des messages brefs (140 caractères) appelés *tweets* (« gazouillis ») par Internet, par messagerie instantanée ou par SMS. Un utilisateur peut écrire un message original ou peut renvoyer (*retwitter*) un message publié par un autre utilisateur pour en faciliter la diffusion parmi ses abonnés.

FIGURE 5 – Graphe des comptes Twitter qui ont mentionné « #right2vote » entre le 23 mai et le 7 juin 2011 colorés par pays de domicile



sés dans le détail²². En suivant l'activité de *tweet* et *retweet*, nous avons pu identifier un groupe d'utilisateurs concernés par le sujet, suivre leurs débats et reconstruire leurs liaisons. Ces relations ont été représentées dans un graphe où les points représentent les utilisateurs (des comptes Twitter du type <http://www.Twitter.com/#/nomutilisateur>), les liens correspondent aux mentions d'une *tweet* faite dans la *tweet* d'un autre utilisateur et la taille des nœuds à la quantité de mentions. Si le nom de l'utilisateur est précédé d'un @, cela signifie qu'il a été mentionné au moins une fois.

La figure 5 inclut 227 nœuds et 352 liens. Nous avons organisé ce corpus en catégories sur la base des déclarations de l'utilisateur dans son descriptif Twitter. Ce sont les catégorisations par langue d'écriture des *tweets* et par pays de domicile qui ont donné les résultats les plus intéressants.

En ce qui concerne la langue d'écriture, le contenu de la plupart des sites

22. Le choix d'utiliser cette méthodologie fondée sur une analyse qualitative et approfondie des *tweets* répond à un besoin de débanaliser Twitter déjà manifesté par la *Digital Method Initiative* (2009[10]).

est en anglais (44%) ou arabe et anglais (28%), 20% des sites seulement sont en arabe. Même des utilisateurs qui alimentent leur site web avec des contenus en arabe préfèrent l'anglais pour leur Twitter²³. Ce choix répond sans doute au besoin d'accroître la visibilité de la cause auprès des migrants des différents pays (mais aussi des membres de la deuxième génération qui pourraient avoir oublié l'arabe), des médias et de la communauté internationale.

En ce qui concerne le pays de domicile des utilisateurs, il n'a pas été possible de l'identifier pour 22% d'entre eux ; pour les autres, il apparaît qu'ils résident majoritairement (soit 40% des comptes Twitter) en Égypte. Pourtant, les autorités du graphe, c'est-à-dire les comptes qui ont produit les *tweets* qui ont été les plus repris, appartiennent clairement à des Égyptiens de l'étranger, notamment du Royaume-Uni (8%) et des États-Unis (7%). Comme ce graphe le fait apparaître, la mobilisation est coordonnée par Raafatology, un blogueur qui réside à New York, par Noha El Shoky, une anthropologue qui habite au Royaume-Uni, par Hoda Osman, une journaliste qui travaille à New York et d'autres qui ont réussi à créer le réseau nécessaire pour donner une visibilité à leur cause.

À travers des données Twitter, nous avons observé l'émergence d'une communauté égyptienne transnationale qui débat et se bat pour une cause commune. Cet exemple nous permet de souligner la différence entre une cartographie du web « classique » qui explore les hyperliens entre sites web et une analyse des liaisons occasionnées par les réseaux sociaux en ligne. Si la cartographie du web « classique » nous a permis de représenter un réseau stable, le réseau de blogueurs égyptiens consolidé par des années de lutte pour les droits humains, l'étude des échanges sur Twitter nous a permis de tracer en temps réel les liaisons entre les Égyptiens à l'étranger et notamment entre les migrants activistes autour de la question du droit de vote. Ce fait met l'accent sur la diversité des temporalités entre le web et les réseaux sociaux (dans ce cas Twitter). Ajouter un lien à un site web, en particulier à un site web associatif ou institutionnel, est une opération pas du tout triviale : elle demande un effort décisionnel et opérationnel et, par conséquent, un certain délai. Les Égyptiens se sont mobilisés pour la révolte sur les réseaux sociaux numériques. C'est pourquoi, la cartographie du web s'adapte mal à l'observation de mouvements rapides comme les liaisons éphémères occasionnées par les événements de la Place Tahrir. Dans ce type d'analyse, elle peut montrer des tendances qui pourront être confirmées seulement après une certaine durée.

23. Cela est vrai surtout pour les blogueurs. Même si la blogosphère activiste égyptienne est née en langue anglaise, elle est aujourd'hui dominée par la langue arabe.

Conclusion

Dans ce texte nous avons proposé une nouvelle méthodologie, la cartographie du web, pour décrire le réseau d'acteurs du cyber-activisme lié au printemps arabe. Cette méthode nous a paru particulièrement adaptée pour des analyses des liens sociaux virtuels, notamment dans des contextes multilingues. Même si les cartes du web ne sont pas facilement lisibles à première vue, elles nous ont permis d'observer un certain nombre de phénomènes généraux qui peuvent aider le chercheur qui s'intéresse aux liaisons numériques à avoir une vision d'ensemble d'un phénomène sur le web et à trouver les directions de recherche les plus intéressantes pour ses analyses futures.

Références

- [1] Lada ADAMIC et Natalie GLANCE : The Political Blogosphere and the 2004 U.S. Election : Divided They Blog. *Proceedings of the 3rd international workshop on Link discovery*, pages 36–43, 2005.
- [2] Albert-Laszlo BARABÁSI et Réka ALBERT : Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.
- [3] Albert-Laszlo BARABÁSI, Réka ALBERT et Jeong HAWOONG : Scale-free characteristics of random networks : The topology of the world-wide web. *Physica A*, 281(1):69–77, 2000.
- [4] Maire BÉNILDE : La révolution arabe, fille de l'Internet ? *Information 2.0*, 2011. <http://blog.mondediplo.net/2011-02-15-La-revolution-arabe-fille-de-l-Internet> (consulté le 8 août 2011).
- [5] Christophe DESHAYES : La révolution tunisienne n'est pas une révolution Internet, c'est une révolution à l'heure de l'Internet. *Révolutionnaires du numérique*, 2011.
- [6] Bruce ETLING, John KELLY et John PALFREY : Mapping the Arabic blogosphere : politics and dissent online. *New Media & Society*, 12(8): 1225, 2010.
- [7] David GIBSON, John KLEINBERG et Prabhakar RAGHAVAN : Inferring Web communities from link topology. *In Proceedings of the ninth ACM conference on Hypertext and hypermedia : links, objects, time and space-structure in hypermedia systems*, pages 225–234, 1998.
- [8] Jeremy GINSBERG, Matthew H. MOHEBBI, Rajan S. PATEL, Lynnette BRAMMER, Mark S. SMOLINSKI et Larry BRILLIANT : Detecting influenza epidemics using search engine query data. *Nature*, 457(7232): 1012–1014, 2009.
- [9] GOVCOM.ORG : Mapping the palestinian web space, 2007. http://www.govcom.org/pisp_maps1.html (consulté le 8 août 2011).

- [10] Digital Method INITIATIVE : For the ppl of Iran - #iranelection RT. 2009. https://movies.issuecrawler.net/for_the_ppl_of_iran.html (consulté le 7 mars 2012).
- [11] Mathieu JACOMY, Sébastien HEYMANN, Tommaso VENTURINI et Mathieu BASTIAN : ForceAtlas2, a graph layout algorithm for handy network visualization. *Médialab whitepapers*, 2011. <http://www.medialab.sciences-po.fr/en/publications-en/> (consulté le 12 février 2012).
- [12] D. LAZER et al. : Life in the network : the coming age of computational social science. *Science*, 323(5915):721–723, 2009.
- [13] Mark LYNCH : Blogging the new Arab public. *Arab Media & Society*, 1(1), 2007.
- [14] Stanley MILGRAM : The Small World Problem. *Psychology Today*, 1(1):60–67, 1967.
- [15] Nicholas NEGROPONTE : *Being Digital*. Vintage, New York, 1996.
- [16] Courtney RADSCH : Core to Commonplace : The evolution of Egypt’s blogosphere. *Arab Media & Society*, 2008. <http://www.arabmediasociety.com/?article=692> (consulté le 9 août 2011).
- [17] Howard RHEINGOLD : *The Virtual Community : Homesteading on the Electronic Frontier*. The MIT Press, Cambridge Mass., 2000.
- [18] Richard ROGERS : Internet Research : The Question of Method. *Journal of Information Technology and Politics*, 7(2-3):241–260, 2010.
- [19] Marta SEVERO, Timothée GIRAUD et Nicolas DOUAY : Citizen protest in online networks : The case of the China’s bloody map. *Proceedings 7th UK Social Networks Conference*, 2011. http://hal.archives-ouvertes.fr/docs/00/67/55/17/PDF/BloodyMap_Severo_inviato.pdf.
- [20] Marta SEVERO et Eleonora ZUOLO : *E-diaspora*, chapitre Egyptian e-diaspora : limites et perspectives. Fondation Maison des Sciences de l’Homme, Paris, 2012.
- [21] Sherry TURKLE : *Life on the Screen : Identity in the Age of the Internet*. Simon & Schuster, New York, 1995.
- [22] Tommaso VENTURINI et Bruno LATOUR : The Social Fabric : Digital Traces and Quali-quantitative Methods. *Proceedings of Future En Seine*, 2009.
- [23] Duncan J. WATTS et Steven STROGATZ : Collective dynamics of ‘small-world’ networks. *Nature*, 393(6684):440–442, 1998.