



**HAL**  
open science

# Les flux RSS pour les études territoriales : une analyse de l'individu et ses espaces dans les actualités internationales du New York Times

Laurent Beauguitte, Marta Severo

## ► To cite this version:

Laurent Beauguitte, Marta Severo. Les flux RSS pour les études territoriales : une analyse de l'individu et ses espaces dans les actualités internationales du New York Times. CIST2014 - Fronts et frontières des sciences du territoire, Collège international des sciences du territoire (CIST), Mar 2014, Paris, France. pp.32-38. halshs-00968513v2

**HAL Id: halshs-00968513**

<https://shs.hal.science/halshs-00968513v2>

Submitted on 4 Jan 2017

**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.

# Les flux RSS pour les études territoriales : une analyse de l'individu et ses espaces dans les actualités internationales du *New York Times*

## AUTEURS

Laurent BEAUGUITTE, CIST (France)

Marta SEVERO, GERiICO (France)

## RÉSUMÉ

L'objectif de cet article exploratoire est de valider l'intérêt d'utiliser dans des études territoriales les flux RSS émis par les sites d'informations, à savoir une donnée simultanément pauvre et abondante, et librement accessible sur le web. Centrée sur la place et l'espace – les espaces – des individus, notre étude montre qu'une analyse qualitative et quantitative de ces flux permet de saisir en partie les logiques sous-jacentes relatives à la production et à la diffusion de l'information. La première partie de l'article décrit le corpus étudié en rappelant l'état de l'art de la recherche qui s'appuie sur les RSS. La deuxième partie présente la méthodologie utilisée et les principaux résultats concernant la place des individus dans la diffusion de l'actualité internationale. Enfin, une troisième partie s'intéresse à deux types d'individus identifiables dans les flux RSS : les stars mondiales et les anonymes.

## MOTS CLÉS

Actualité internationale, flux RSS, *New York Times*, individu, territoire

## ABSTRACT

This exploratory paper is meant to validate the interest of using RSS feeds of information websites (namely a poor and abundant information at the same time) in territorial studies. By focusing on the place and space –spaces– of individuals, our study shows that mixed qualitative and quantitative analysis of these feeds allows to identify some of the underlying logic for the production and dissemination of information. The first part describes the corpus by recalling the state of the art of research on RSS feeds. The second part presents the methodology and the main results on the place of individuals in international news. Finally, the third part focuses on two types of individuals identified in RSS feeds: global stars and anonymous. If the first category often appears on multiple scales, from local to global, anonymous remain mainly local witnesses or information providers.

## KEYWORDS

International news, RSS feeds, *New York Times*, individual, territory

## 1. FLUX RSS ET CHOIX DU CORPUS

Si l'analyse de discours d'information médiatique est désormais une pratique consolidée dans la recherche en sciences de la communication, la construction du corpus de données à analyser constitue encore une phase très délicate et coûteuse. Ces études sont généralement basées sur l'analyse des médias traditionnels tels que la presse et la télévision. Comme on le sait (Earl *et al.*, 2004), l'utilisation de la presse pour l'étude des différents objets de recherche (par exemple le territoire) peut soulever plusieurs critiques relatives à la collecte des données et à la sélection et la description du contenu des articles (McCarthy et McPhail, 1996). Cependant, nous sommes d'accord avec Earl *et al.*

(2004) et estimons que les données de la presse, en dépit de leurs limites, peuvent permettre d'étudier certains aspects que les chercheurs ne pouvaient guère approcher avec d'autres méthodes.

Une autre question importante liée à ce type de données est qu'elles ne peuvent être récupérées que dans des bases de données commerciales telles que Dow Jones Factiva, LexisNexis ou Europresse. L'utilisation de ces bases de données est non seulement coûteuse, mais soulève aussi plusieurs problèmes méthodologiques (comme la couverture discontinue des sources ou la non transparence sur la création des mots clés) et techniques (par exemple il n'est pas possible d'extraire plus d'une centaine d'articles en même temps).

Pour ces raisons, nous proposons d'utiliser un autre type de données qui semble plus adapté à des études territoriales. Nous testons l'intérêt d'utiliser les flux RSS fournis par les sites d'information, notamment par les journaux quotidiens. Les RSS sont censés avoir trois grands avantages : ils sont libres de droit, ils peuvent être archivés et tagués sans limites, ils sont généralement fournis en temps réel. Cependant, les RSS sont un objet encore peu étudié. Si on retrouve plusieurs recherches sur leurs aspects techniques (Hmedeh *et al.*, 2011), leur valeur informationnelle est rarement analysée (Thewall et Prabowo, 2007). Un précédent important est la recherche de Franck Rebillard dans le contexte de l'Observatoire TransMédia où les RSS sont employés comme source d'information pour tracer des événements (Marty *et al.*, 2012).

Dans le cadre du projet ANR Corpus Geomedia (2013-2016), nous proposons de construire une base de données stockant les flux RSS associés à des articles publiés dans une centaine de journaux dans différentes parties du monde et d'en extraire plusieurs types d'informations. Le contenu d'information d'un flux RSS (titre et résumé de quelques lignes) est évidemment plus limité que celui d'un article de presse. Les tags thématiques ou géographiques ne sont pas systématiquement fournis par les journaux. Il faut donc, en aval de la collecte des flux RSS, mettre en place des procédures de codage du contenu et vérifier si l'ensemble constitué par le titre et le résumé conduit à des codages géographiques et thématiques conformes à ceux qu'aurait produit une analyse intégrale du contenu de l'article. Pour ce faire, nous menons des études de cas pour tester la validité des RSS comme source d'informations dans des études territoriales internationales. Dans cet article nous travaillons exclusivement sur le flux RSS de la rubrique « International » du *New York Times* et nous considérons la période du 1<sup>er</sup> janvier au 31 mars 2013. Dans ce corpus, nous analysons la présentation des individus et leur rapport avec l'espace dans l'actualité internationale.

## **2. LES INDIVIDUS ET LEURS ÉCHELLES : HYPOTHÈSES, MÉTHODE ET RÉSULTATS GLOBAUX**

Pourquoi s'intéresser à l'individu dans l'actualité internationale ? Notre hypothèse principale, fondée sur les conditions de production de l'information, est liée à la nécessité pour les journalistes de personnaliser l'information en général et l'information internationale en particulier (Gans, 1979). Intéresser un lecteur d'un pays *x* à ce qui se passe dans un pays *y* suppose de mettre en place des stratégies permettant de rendre l'information compréhensible : jouer sur l'identification est l'une des stratégies employées dans la sphère médiatique (Ostgaard, 1965 ; Mercier, 2006).

Suivant des études antérieures, nous nous attendons à trouver trois grands types d'individus : les stars mondiales, les élites et les gens ordinaires ("*common people*", voir Galtung et Ruge, 1965). Les caractéristiques attendues sont les suivantes :

- la star mondiale est toujours nommée, elle apparaît à plusieurs reprises dans l'actualité et son action se déploie sur une large gamme d'échelles ;
- la personne ordinaire n'apparaît qu'une fois et n'est pas nommée. Elle est ancrée localement mais, pour justifier sa place dans l'information internationale, son cas est considéré comme le symbole d'une situation plus vaste. Elle peut également jouer le rôle de source ou de témoin d'une actualité ;
- les élites apparaissent à plusieurs reprises, sont généralement nommées, leur action se déploie sur une gamme restreinte d'échelles (État ou relation bilatérale entre États).

Compte tenu de la brièveté de ce texte, nous testerons les deux premières hypothèses à l'aide des flux RSS.

L'ensemble des 1 304 éléments du flux RSS de la rubrique internationale émis par le *NYT* entre le 1<sup>er</sup> janvier et le 31 mars 2013 a été codé par les deux auteurs : y a-t-il présence d'un individu dans le titre et/ou l'item du flux ? L'individu est-il nommé ou anonyme ? Est-il possible de déterminer l'échelle de l'actualité (locale, nationale, mondiale, etc.) ? Ce dernier item a été le plus délicat à coder : de rares flux sont a-spatiaux ("*Oscar Pistorius, the double amputee Olympic runner accused of murder, will be allowed to travel abroad under strict conditions*", item du 28 mars 2013), d'autres mêlent plusieurs espaces de référence ("*The United States government has petitioned the Irish High Court to extradite an Algerian man living in Ireland on terrorism charges relating to a plot to kill a Swedish artist*", item du 1<sup>er</sup> mars 2013). Les catégories utilisées sont les suivantes : 0 (pas d'espace), local (infra-national), national, inter-national (relation entre deux ou trois États), régional (supra-national), global.

L'étude du corpus montre une information internationale fortement individualisée : plus de la moitié des articles impliquent au moins un individu (698 sur 1 304) et ces individus sont le plus souvent des personnes clairement identifiées (75 % des individus ont un nom).

Le tableau ci-dessous donne la répartition obtenue lorsque l'on croise identité/anonymat et échelle.

	0	local	national	inter-national	régional	global	total
Anonyme	3	42	57	52	4	13	171
Nommé	26	49	236	148	17	51	527
Total	29	91	293	200	21	64	698

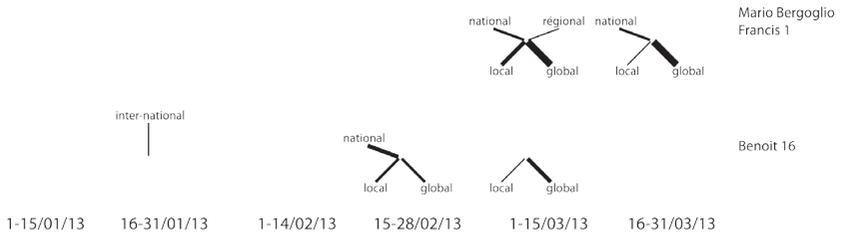
La forte domination des échelles nationales et internationales (entendues au sens de relations entre deux ou trois États) doit être soulignée et ce que les individus soient anonymes ou non. Et, contrairement à notre hypothèse, l'individu anonyme peut être l'objet d'une information locale, nationale ou globale. Le nombre élevé d'anonymes au niveau global s'explique par nos choix de codage : tout observateur lié aux Nations Unies et rapportant une situation locale a été considéré comme lié à une actualité globale ("*A United Nations investigator, whose report on the C.I.A.'s use of secret detentions and interrogation techniques is to be released, called on the United States to publicly*", item du 5 mars 2013).

### 3. STARS GLOBALES ET ANONYMES : FRÉQUENCE ET ESPACES D'APPARITION

Si le lien entre individu et espace est moins flagrant qu'attendu, la fréquence des apparitions, elle, conforte nos hypothèses et rejoint d'ailleurs les conclusions des études relatives à la distribution des sujets d'actualités (Smyrnaioi *et al.*, 2010 ; Marty *et al.*, 2010) : une poignée d'individus est citée régulièrement, la grande majorité apparaît une ou deux fois avant de disparaître de l'actualité. Si 334 noms différents peuvent être identifiés dans les titres et les items, 249 apparaissent une seule fois. À l'autre extrémité, les personnalités du monde politique dominent très largement (32 occurrences pour Obama et Chavez, suivis de Benoît XVI, François I<sup>er</sup>, Netanyahu, Bashar el Assad, etc.). Sur les 20 personnes évoquées au moins 5 fois durant ces trois mois, seul l'athlète Oscar Pistorius est extérieur au monde politique. La capacité à se maintenir comme sujet ou acteur de l'actualité est très inégalement distribuée : les 9/10<sup>e</sup> des individus évoqués (anonymes ou non) le sont deux fois maximum durant la période de trois mois considérée.

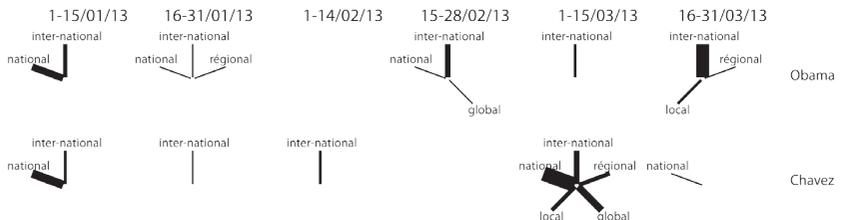
Lors de la période étudiée, la démission de Benoît XVI et l'élection du pape François I<sup>er</sup> donnent lieu à un grand nombre d'éléments RSS (19 liés à la démission du premier et 20 liés à l'élection du second) et la multiplicité des échelles mobilisées est impressionnante : les actualités varient du monde (*the Catholic Church*) au local (*St. Peters Square*) en mobilisant toutes les échelles intermédiaires (le Vatican comme État, l'Amérique latine comme région, et le monde en général).

Figure 1. Benoît XVI et François I<sup>er</sup> dans les flux RSS international du NYT



La figure 1 montre comment se distribuent les apparitions papales dans les flux RSS en fonction des échelles mobilisées (les articles a-spatiaux ne sont pas représentés ici). Dès son apparition dans l'actualité, le futur François I<sup>er</sup> est sujet ou acteur d'une actualité multiscalaire. La démission de son prédécesseur est, elle, évoquée tantôt au niveau local, tantôt au niveau global.

Figure 2. Obama et Chavez dans le flux RSS international du NYT



À titre de comparaison, la figure 2 montre les échelles de référence des deux personnes les plus présentes dans l'actualité internationale du New York Times de janvier à mars 2013, Barak Obama et Hugo Chavez.

Obama est durant la période considérée la seule personnalité qui apparaisse à une telle fréquence et avec une telle pluralité d'échelles. Hugo Chavez est aussi présent mais selon des modalités différentes : son hospitalisation à Cuba crée une actualité internationale (Vénézuéla, Cuba), les conséquences de son décès sont évoquées à toutes les échelles possibles, et enfin sa succession devient un problème national avant que son nom ne disparaisse de l'actualité.

Le cas Pistorius est intéressant car il permet de saisir en partie la temporalité d'une actualité : s'il apparaît toujours à une échelle locale – explicite ou non (cf. *supra*) –, deux temps émergent : un premier temps, très resserré, où 5 items le citent (du 20 au 22 février) et qui correspond à l'apparition du fait divers dans l'actualité ; un second temps, moins riche (2 items seulement) et plus lâche (un item le 8 mars et un le 28 mars) correspondant à des rebondissements mineurs de l'affaire. Prendre une période plus longue permettrait de mettre en évidence la construction narrative en feuilleton des actualités liées aux faits divers, construction expliquant que certains éléments (le lieu notamment) ne soient pas systématiquement rappelés.

Si l'on s'intéresse aux *common people*, l'un des marqueurs les plus dignes d'une étude approfondie des flux RSS concerne sûrement le fait de nommer ou non les personnes, notamment lorsque celles-ci appartiennent à l'élite politique d'un État. Ainsi il semble, mais des études approfondies seraient nécessaires pour confirmer ce fait, que si les responsables politiques américains sont systématiquement nommés, ce n'est le cas que pour les chefs d'État et les ministres étrangers de certaines nationalités (Chine et Royaume-Uni notamment). Les responsables ministériels d'autres États ne sont nommés qu'exceptionnellement. Dans notre échantillon, 21 articles concernent des autorités étrangères non nommées ("*The foreign ministers of Japan and the Philippines*", "*Greek Prime Minister*"). Une étude comparative avec les flux émis par d'autres journaux d'origine différente donnerait sans doute des résultats intéressants.

Par ailleurs, si on décide de ne prendre en considération que les articles qui concernent des personnes ordinaires en leur appliquant la classification de Gans (1979) des personnes « inconnues » pour étudier leur relation avec l'espace. On peut ainsi distinguer : les manifestants, les « émeutiers » et les grévistes mentionnés dans 14 articles ; les fonctionnaires – policiers, soldats, juges et bureaucrates fédéraux – mentionnés dans 39 articles ; les victimes dans 30 articles ; les contrevenants présumés et réels des lois (souvent dans des contextes de guerre), catégorie la plus représentée avec 46 articles. Il y a également un groupe d'individus qui apparaît difficilement catégorisable, comme par exemple les cardinaux liés à l'élection du pape (nous avons dû créer une catégorie « autre » pour 15 articles).

La relation la plus évidente qui émerge est entre l'échelle locale, d'un côté, et activistes et victimes. Une actualité liée à une protestation, ou un fait violent qui touche une victime, est souvent située de manière précise dans l'espace. Une exception est liée au cas du garçon russe trouvé mort au Texas qui rend l'actualité internationale ("*Adopted*

*Boys Death in U.S. Stirs Outrage in Russia. The authorities in Texas are careful about commenting on a 3-year-olds death, with investigations continuing*").

Pour les autres catégories, il n'est pas toujours possible d'identifier des connexions aussi claires. Par exemple, les fonctionnaires peuvent être liés à toutes les échelles et dans des contextes très différents. À noter qu'ils sont souvent cités comme source de l'information ("... *an official in Dar es Salaam said*", "... *a top official with the World Health Organization said*" ou "*A judge said that...*") et, dans ce cas précis, il faut prendre en compte le phénomène de l'anonymat des sources ("... *said a military official who declined to be named*") (Lambert, 2001).

## CONCLUSION

L'objectif de cette étude était double : s'intéresser à la place de l'individu dans l'actualité internationale ; montrer l'intérêt d'une analyse qualitative d'une information simultanément abondante et pauvre, les flux RSS, dans des études territoriales. L'individu, anonyme ou non, est mobilisé très fréquemment par les journalistes pour mettre en scène l'actualité internationale. Une poignée d'individus est susceptible de revenir fréquemment dans l'actualité et à des échelles variées tandis que la très grande majorité fait une apparition rapide, souvent anonyme et généralement mono-scalaire. Dans le cas de faits divers impliquant des célébrités, une construction en épisodes apparaît. Si le flux RSS est une donnée qualitativement pauvre, il permet pourtant de saisir diverses logiques relatives à la construction et à la diffusion de l'actualité. La facilité de sa collecte pour des médias divers est un atout pour l'étude comparative de médias de différentes origines.

## RÉFÉRENCES

- Earl J., Martin A., McCarthy J. D., Soule S., 2004, "The Use of Newspaper Data in the Study of Collective Action", *Annual Review of Sociology*, 30(1), pp. 65–80 [doi:10.1146/annurev.soc.30.012703.110603].
- Galtung J., Holmboe M., 1965, "The Structure of Foreign News", *Journal of Peace Research*, 2(1), pp. 64-91.
- Gans H., 1979, *Deciding what's news: A study of CBS evening news, NBC nightly news, Newsweek, and Time*, New York, Pantheon Books.
- Gill K., 2005, "Blogging, RSS and the Information Landscape: A Look At Online News", in WWW2005 [en ligne [http://faculty.washington.edu/kegill/pub/gill\\_www2005\\_rss.pdf](http://faculty.washington.edu/kegill/pub/gill_www2005_rss.pdf)].
- Hmedeh Z., Travers N., Vouzoukidou N., 2011, "Everything you would like to know about RSS feeds and you afraid to ask", *Base de Données Avancées Morocco* [en ligne [http://hal.archives-ouvertes.fr/docs/00/73/72/43/PDF/art\\_2158.pdf](http://hal.archives-ouvertes.fr/docs/00/73/72/43/PDF/art_2158.pdf)].
- Lambert F., 2001, *Figures de l'anonymat : médias et société*, Paris, L'Harmattan.
- Marty E., Rebillard F., Pouchot S., Lafouge T., 2012, « Diversité et concentration de l'information sur le web », *Réseaux*, 176(6), pp. 27 [doi:10.3917/res.176.0027].
- Marty E., Rebillard F., Smyrniaos N., Touboul A., 2010, « Variété et distribution des sujets d'actualité sur Internet. Une analyse quantitative de l'information en ligne », *Mots. Les langages du politique*, 93, pp. 107-126.
- McCarthy J., Mcphail C., 1996, "Images of Protest: Dimensions of Selection Bias in Media Coverage of Washington, 1982 and 1991", *American sociological review*, 61(3), pp. 478-499 [en ligne [http://www.jstor.org/stable/10\\_2307/2096360](http://www.jstor.org/stable/10_2307/2096360)].
- Mercier A., 2006, « Logiques journalistiques et lecture événementielle des faits d'actualité », *Hermès*, 46, pp. 23-35.
- Ostgaard E., 1965, "Factors Influencing the Flow of News", *Journal of Peace Research*, 2, pp. 39-63.

Smyrniaios N., Marty E., Rebillard F., 2010, "Does the 'Long Tail' apply to online news? A quantitative analysis of French-speaking websites", *New Media and Society*, 12(8), pp. 1244-1261.

Thelwall M., Prabowo R., 2007, "Identifying and characterizing public science-related fears from RSS feeds", *Journal of the American Society for Information Science and Technology*, 58(3), pp. 379-390 [en ligne <http://onlinelibrary.wiley.com/doi/10.1002/asi.20504/full>].

## **LES AUTEURS**

**Laurent Beauguitte**

CIST

beauguittelaurent@hotmail.com

**Marta Severo**

Équipe GERiiCO

Université de Lille 3

marta.severo@univ-lille3.fr