

## Two Paths of Glory

### Structural Positions and Trajectories of Websites Within their Topical Territory\*

**Dominique Cardon**

Orange Labs, TECH/SENSE  
38-40 rue du général Leclerc  
92794 Issy-Les-Moulineaux, France  
domi.cardon@orange-ftgroup.com

**Guilhem Fouetillou**

Linkfluence  
153 Boulevard Anatole France, Tour Pleyel  
93521 Saint-Denis, France  
guilhem.fouetillou@linkfluence.net

**Camille Roth**

CNRS/EHESS, CAMS  
190 avenue de France  
75013 Paris, France  
roth@ehess.fr

#### Abstract

We analyze the structure and processes of reputation building in a sample of Internet communities. Our approach uses a corpus of around 9,000 ‘active’ French websites and blogs which have been manually labeled as belonging to a specific topical territory, such as “cooking”, “crafts”, or “politics”. We propose a typology and map of the various structural positions that these websites may hold within their territory. This typology is essentially based on their authority status in terms of in- and out-going links and depending on whether those links come from/to outside their origin territory or not. Behind the unifying notion of “territory”, we may therefore distinguish several distinct topological structures for each topic. Building upon dynamic data spanning over a period of 10 months, we are then able to exhibit characteristic trajectories of engagement into each territory and, more broadly, reputation building on the web. In particular, we suggest that there are two opposing approaches in building up online authority: either by progressively developing reputation from within the territory – thereby moving ‘up’ through a typical online ladder – or by exploiting fame that had been acquired prior to one’s blogging activity. This descriptive model makes it possible to carry a sociological analysis of the variety of resources mobilized by bloggers to build their online reputation (relational authority, prestige, institutional support, commercial support, ...). We eventually illustrate both this phenomenon and specific successful trajectories to fame through a comprehensive study of a selection of topical territories.

#### Introduction

The development of personal media on the Internet revived the notion of “community” — discussions on “Internet communities” being so diverse that it is generally quite difficult to define its boundaries. Experts in a given field, political activists from the same movement, consumers of the same kind of product or brand, bloggers writing on the same kind of topic, open-source software developers, wine lovers or

do-it-yourselfers: all are eligible to be members of some “community”. The outstanding success of “community”-based language in denoting web collectives certainly stems from the key role played by the notion “virtual community” throughout the history of the Internet (Rheingold 1993). Originally, this language had been applied to the digital world as a reformulation of some of the notions that had been developed in the 1970s for the then-collapsed hippie communities (Turner 2006). The term later described collectives devoted to goals such as open-source software development or knowledge sharing, for instance Wikipedia (Benkler 2006). More loosely, the term also refers to aggregates of websites or blogs<sup>1</sup> which either share a common topic (semantic communities) (Li, Guo, and Zhao 2008; Wei 2004; Kumar et al. 2004; Chau and Xu 2007) or are linked in a relatively cohesive way (topological communities) (Kumar et al. 2003; Lin et al. 2006). Both types of communities may overlap (Adamic and Glance 2005; Uchida et al. 2009), yet, it is unclear whether members of such a community would be able to recognize that they belong to it — even when semantic and topological dimensions perfectly match.

These issues underline the challenge of *a priori* deciding whether a semantic and/or topological grouping of weblogs may be characterized as a “community”. In this paper we propose an empirical methodology which aims at describing the topological aspects of a group of topically-related blogs. In order to avoid typical misunderstandings linked to the use of the term “community”, we speak of a “territory” as a group of blogs sharing the same editorial goal, such as “politics” or “cooking” territories. We ambition to observe the topological configuration of websites within the same semantic territory. We indeed assume that the topical nature of these territories massively influences their topological organization, and that it is therefore possible to use these structural discrepancies for sociological inquiry. More precisely, we wish to exhibit which role and functions some topologi-

\*This work has been partially supported by the French National Agency of Research (ANR) through grant “Webfluence” #ANR-08-SYSC-009.

Copyright © 2011, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

<sup>1</sup>In what follows, for the sake of simplicity, we indifferently speak of “blogs” or “websites and blogs”. Indeed, 90% of our corpus is made of *individual* blogs; the remainder consists of collective blogs, institution and organization websites, and online medias. We decided to include the latter ones as they also belong to and play a role in the corresponding web territories, and are thus essential to the understanding of the relational structure of a semantic territory.

cal positions fulfill within a given territory, in particular with respect to exchanges towards and from other territories. In short, we aim at showing that a detailed relational analysis of the various components of a given semantic territory enables us to understand its formation, evolution and structuration — describing how some collective forms of online life can emerge, developing sometimes characteristics which are typical of a “community”, developing sometimes displaying “competitive” features which are closer to the market, or developing a more complex combination of both.

### Related work

The literature dealing with topological aspects of blogspace is immediately relevant. Many studies previously emphasized the strong internal connectivity of blogspace (Kumar et al. 2003), the intertwining between network structure and information diffusion phenomena (Adar et al. 2004), or explored authority-related phenomena (Marlow 2004). Beyond those works, some authors focused on the role played by links in establishing one’s position within blogspace. Looking first at comment networks, some scholars have underlined the low number of comments and thus weak interactivity of blogs (Herring et al. 2004); by contrast, others (Mishne and Glance 2006) have shown that, while the number of comments follows a power-law (Shirky 2003), comments were playing a decisive role in constructing conversations on blogspace (Efimova and Moor 2005). On the whole, these studies suggested that distinct types of links had various importance on the ecosystem of blogs. Perhaps more importantly, some asserting that dynamical aspects were crucial in qualifying the relationship between two blogs (Lento et al. 2006), specifically suggesting that reciprocity was key before concluding an effective social link between two blogs (Lin et al. 2006). At a more “meso” level, some implemented “community” detection algorithms from the structure of comment networks (Chin and Chignell 2006; Adar et al. 2004). (Herring et al. 2005) described blogs as a very hierarchized space where a small core of blogs having a high readership and actively citing each other stands on top of a large cloud of smaller islands of relatively peripheral and isolated blogs. Finally, another stream of research has aimed at jointly appraising the topological and semantic structure of blogspace, demonstrating how the two dimensions were correlated and impacting each other (Adamic and Glance 2005; Cointet and Roth 2009; Uchida et al. 2009; Cardon et al. 2011).

Our work additionally builds upon the state-of-the-art related to the construction of reputation and authority. This question has been addressed in a variety of disciplines. Sociology of culture, for one, insisted on the role of some social bodies in establishing and certifying cultural fame, or legitimating artistic value — social bodies such as professional critics, cultural institutions and award-related events (Bourdieu 1992). Sociology of medias especially described the various audience-related strategies which could be used to ensure a visible and central position in media space (Couldry and Markham 2007). The sociology of the Internet emphasized the role of citation links in holding authority positions on the web (Stoica, Couronne, and Beuscart 2010) or at-

tempted at quantifying (and nuancing) the online reputation of amateur musicians (Beuscart 2008). To some extent, these three kinds of approaches will be useful to characterize the various forms of visibility which some blogs may acquire within their topical territory. We will specifically distinguish two dimensions of reputation — internal and external — depending on whether it is built upon recognition acquired from peers (i.e., from the same territory), or it results from an already strong external visibility which is rapidly likely to attract an audience beyond the origin territory.

### Relational and/or topical territories

Adopting a topological viewpoint on web territories leads to construe them as *informational territories*, where proximities between sites and groups of sites follow from the structure of hyperlinks. Here, the relational proximity of websites defines their “territorial” affiliation and position; in this respect, sites are similar if they have resembling patterns of incoming and outgoing links (Ghitalla et al. 2004). Yet, this territory is also a dynamic space where websites struggle to hold central positions, for instance in order to gain visibility on the web, search engines, etc. In a more or less strategic manner, bloggers are building up their website visibility by creating links towards other blogs so that, in turn, they could receive incoming links. Additionally, they carefully choose their cited websites among those of a target territory in order to increase their relational proximity to this territory, while they refuse to cite and link some other types of blogs to remain farther from contents they might oppose. Put briefly, they are using links to establish their position in a thematic space that they share with others (Nardi, Schiano, and Gumbrecht 2004).

### The *linkfluence* dataset

In addition to this topological approach, it is possible to adopt a purely semantic viewpoint by associating blogs depending on the topics of their posts. In this case, one can define a typology of blogs based on the similarity of the “main” topic they address. For this research, we used a collection of websites consisting of several topical territories of the French blogosphere. Since 2007, these territories have been carefully and constantly updated by the librarians of *linkfluence*<sup>2</sup>, a start-up specialized in providing opinion studies on weblogs territories. It is mainly made of blogs and media websites (traditional and pure plays) which have a certain publication activity and allow users to post comments.

More practically, the corpus has been collected according to the following steps, using a proprietary crawler developed by *linkfluence*. The first step consisted of an initial topical seed extracted from Google and the open directory, with topics ranging from “cooking” to “IT talk” or “economics” and “politics”. This initial list of websites has then been progressively extended by following relevant hyperlinks stemming from these websites. For each newly discovered website, the language was automatically detected and inbound links were computed. French websites which (i) publish at least once a month, (ii) have comments (excluding spam)

<sup>2</sup><http://us.linkfluence.net/>

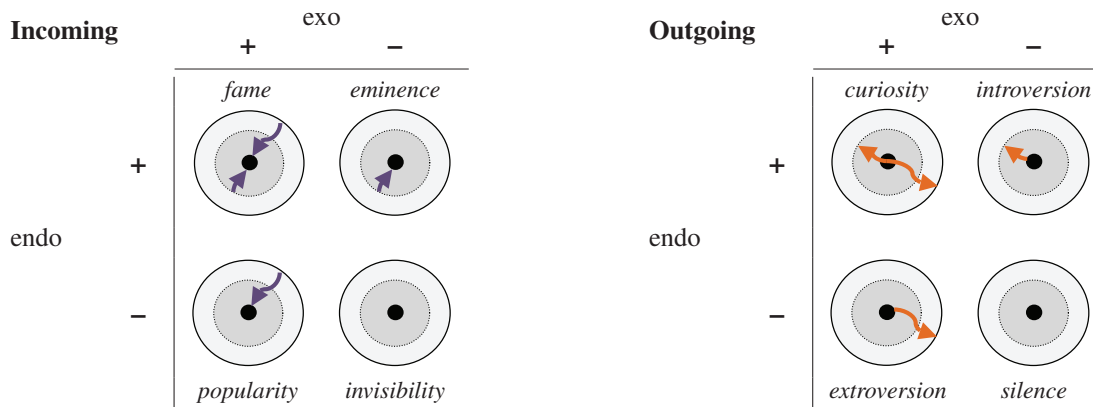


Figure 1: Position configurations for *incoming* (left) and *outgoing* (right) links. The internal and external gray disks respectively denote links coming from the territory and from outside.

and (iii) have at least 3 inbound links in the whole corpus were, in turn, visited and categorized or excluded by a specialized librarian at *linkfluence*. This large-scale and tentatively comprehensive snowball process was iterated until no new relevant websites were discovered, leading to a corpus of about 12 000 blogs. Since Feb 1st, 2010, all articles and posts have been archived, resulting in about 2m posts over 10 months of data, discriminating text contents from hyperlinks, and comment links from citation and blogrolls links. From this corpus, 9 000 websites and around 200 000 posts feature at least one link to another post within the corpus. We subsequently focus on this subnetwork.

### A model of structural positions based on incoming and outgoing links

We essentially wish to qualify blogs through their pattern of outgoing and incoming links, both within and outside of their topical territory. To this end, we propose a structural model featuring 16 typical positions which blogs may hold within their territory.

Incoming links denote attention devoted by others to a given blog. They are reputation indicators and may come from the territory (we speak of “*endo*” links) or from outside of it (“*exo*” links). A blog is said to be dominant for some type of link if it receives at least as many such links than a proportion  $p$  of the blogs of its territory. We fix this proportion at  $p = 0.8$  for the whole paper. In other words, a *dominant* blog in “incoming *endo* links” receives at least as many links from its territory as 80% of the other blogs of that territory. We arbitrarily fixed the threshold of dominance at  $p = 0.8$  as a reference to the Pareto rule of 80-20 in case of heterogeneous wealth distribution, applied here to the case of links. We checked that small variations around this value of 0.8, such as using  $p = 0.9$  or  $p = 0.7$ , roughly leads to the same main qualitative conclusions.

We thereby define a double dichotomy made of four categories, depending on whether a blog is dominant in “incoming *endo* links”, or not, and whether it is dominant in “incoming *exo* links”, or not. We introduce and use conventional terms to denote these categories. These terms are es-

entially convenient shortcuts, and little attention should be paid to their potential academic connotations. See the left panel of Figure 1. Thus, we say that a site is a *fame* if it receives many more links from everywhere, both from its territory and from outside. A site is *popular* if it attracts a lot of *exo* links, without eliciting much *endo* attention – in this respect, this definition of “popularity” matches an observation, stemming from sociology of medias, that some individuals get a high recognition from the external audience without being recognized by their peers (Chalvon-Demersay and Pasquier 1990). By contrast, a site is an *eminence* if it attracts *endo* links without getting many *exo* links. Finally, ghost sites are *invisible* to both their territory and the outer world, and receive few to no links at all.

Outgoing links show how much a website is engaging with the web. Along the same lines, we assume that a site may send links towards (*endo*) or outside of (*exo*) its territory; see the right panel of Figure 1. We speak of *curiosity* when a site sends links everywhere; *extroversion* when it mostly sends *exo* links. On the contrary, we speak of *introversion* when a site mostly sends *endo* links towards its territory. Eventually, a site is *silent* when it sends few or no links at all.

By crossing these two double dichotomies, we get a logical map of 16 possible topological positions within a territory, as summarized on Figure 2.

### Territorial maps

We applied this position model on six territoires, as defined by the *linkfluence* corpus: *agora*, *beauty/fashion*, *cooking*, *crafts*, *IT-talk*, and *politics*. We considered them as distinct case studies. These territories were chosen both because of their significant size (around a thousand blogs) and, in a more arbitrary fashion, because they deal with issues which are often seen as typical of personal media.<sup>3</sup>

<sup>3</sup>*agora* and *politics* are two distinct territories: *agora* is devoted to society issues and how these issues are being discussed in the public space. Blogs from this territory comment on news, acting more or less as editorialists. *politics*, on the other hand, is exclusively made of politicians and political parties activists, supporters,

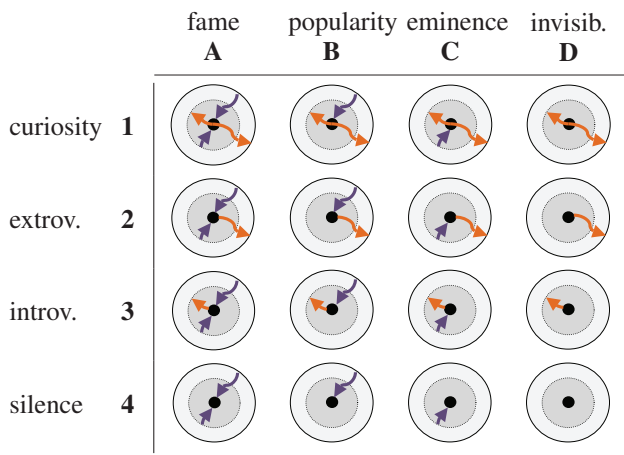


Figure 2: Theoretical positions within a given territory. For convenience, we refer to slots with their alphabetical column  $\times$  numerical row name, e.g. the bottom left slot is A4.

A purely qualitative analysis may plainly focus on the contents of each slot, describing which blogs are fulfilling which role within a territory. For instance, “A1” denotes strongly dominant blogs, featuring a correlation between all types of dominances: incoming/outgoing endo/exo links. Rather, we are interested in using this approach to show how certain configurations are remarkable in some territories (sometimes even all of them). For example, we may ask whether there are territories where “A1” (famous and curious) is much more populated than should randomly happen if there were no correlation at all between various types of dominances. In other words, any deviation from the random uniform setting indicates a more marked propensity to combine at least two types of dominances (or non-dominances).

To this end, we compute the relative presence of blogs in some slot with respect to the expected population of that slot *if* blogs were uniformly distributed among all four types of dominant/non-dominant 80%-20% quantiles. We plot the results on heat maps as shown on Figure 3.

By doing so, we are able to highlight positions which are typical of the positional structure of a given semantic territory. We may also, at a higher level, exhibit possible regularities between various topical territories. More precisely, these position maps show regularities and shared features among territories, as well as significant discrepancies between them.

### Over or under-representation of some positions

**Invisible blogs** To interpret these maps, we may first focus on the “D” column corresponding to invisible blogs. They have a poor rank in the territory’s hierarchy and are not particularly recognized by other blogs. “D4” (invisible and silent) is always over-represented, while “D1” (invisible and curious) is so in 5 territories out of 6; on the contrary,

advocates; their editorial agenda is essentially focused on the political arena.

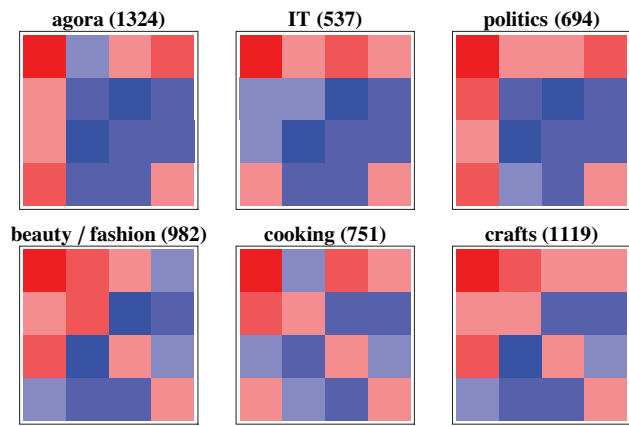
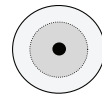
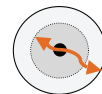


Figure 3: Position maps for each of the 6 territories (figures inside brackets indicate the number of blogs considered for each territory). A given slot will be in red or blue depending on whether it is respectively more or less populated than expected – paler colors correspond to weak deviation from the null model (from a factor of 1 to 2), while darkest colors correspond to at least one order of magnitude of deviation.

“D2” and “D3” are always under-represented (and generally weakly populated).



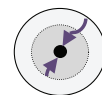
The slot “D4” gathers a considerable number of blogs from our corpus, whichever the territory (around 60% of all blogs). A large part of blogspace is inactive or weakly active (Herring et al. 2004). We observe below that a large number of sites sleep or die after a more or less lengthy activity period by eventually (re)joining “D4”.



By contrast, “D1” corresponds to blogs which send links everywhere without necessarily receiving a large amount of links in return. This position is typical of starting blogs attempting to get introduced into a territory by abiding to the particular economics of visibility, which requires to first cite before getting cited (Nardi, Schiano, and Gumbrecht 2004; Cardon and Delaunay-Teterel 2006).

The constant under-population of “D2” and “D3” (invisible and extrovert/introvert), on the other hand, shows that blogs are less likely to cite either only insiders or only outsiders to the territory: assuming that invisible blogs are essentially starting blogs, it suggests that these blogs might still be unaware of the boundaries of their territory (or its importance), thereby establishing either no links (“D4”), or links anywhere (“D1”), both inside and outside the territory.

**Silent blogs** This corresponds to the 4th row, of which the above-mentioned “D4” is also a member.



Let us focus on the “A4” position which corresponds to silent famous blogs and is relatively often over-represented, in at least 4 cases out of 6. In all territories, it gathers sites and blogs animated by some of the most important institutional and commercial actors of the territory’s topic — famous enough to receive links from inside and outside without endeavor-

ing at redistributing outgoing links. This dominance without reciprocity is typical of forms of statutory authority. In this respect, within the “*agora*” territory, we indeed find the main French state institutions (ministries, the Constitutional Council, etc.), NGOs, labor unions and traditional media websites. The “A4” slot of “politics” is made of statewide, prominent politicians who are usually weakly active on the web.

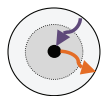
As such, the constant under-representation of “B4” (popular and silent) and “C4” (eminent and silent) suggests that silent yet famous blogs should generally be of the above type (“A4”) — the task of acquiring reputation without creating links appearing otherwise to be somewhat difficult.

**Famous blogs** “A1” (famous and curious) gathers the most active celebrities of a territory. Blogs holding this position in a territory are both its elite and its stars, in that they are successfully exchanging links with, and outside of, their territory, respectively. This category is strongly over-represented in all territories, indicating a typical figure of the blogspace structure and, perhaps, suggesting that these nodes experience a joint endo/exo incoming/outgoing link reinforcement dynamics within these networks.



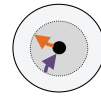
Slots “A2” & “A3” (famous extrovert/introvert) represent other figures of fame, and are fairly over-represented (5/6 for “A2”, 4/6 for “A3”). They are similar and correspond to a slightly distinct form of fame, where bloggers are either getting famous in a territory while referring extensively to the exterior of the territory (“A2”) or where bloggers are firmly embedded within a territory yet still acquiring reputation from the outside (“A3”). Similarly, on the curiosity line (1st row), “B1” (popular and curious) and “C1” (eminent and curious) are also fairly over-represented (4/6 for “B1”, 5/6 for “C1”) and display a behavior which resembles partially that of “A1”.

**The middle of the map** Another interesting area relates to the middle of the position matrix. Two positions are constantly and strongly under-represented: “B3” (popular and introvert) and “C2” (eminent and extrovert). The logical structure of our model may plausibly lead to meaningless and unrealistic positions — here, these two slots display conflicting and therefore unlikely reputation and outgoing link patterns.



“B2” (popular and extrovert) is quite peculiar however, as it gathers blogs which are only dominant as citing and being cited outside of their territory. It features fundamentally topical blogs which are poorly embedded in the topological fabric of their topical territory. This position is only over-represented in ego-centered territories such as “*beauty/fashion*”, “*crafts*” and “*cooking*”. Qualitative observation reveals two types of blogs: (i) either amateur blogs which are close to a diary and thus not entirely focused on their topic: typically, a cooking blog which someday discusses travel issues, some other day recounts a personal story or (ii) blogs which are reflexive and meta-discursive upon their assigned territory:

for instance, diet blogs in the “*cooking*” territory, or design blogs providing a general reflexion on fashion trends in the “*beauty/fashion*” territory, etc. Both types are likely to point to and be pointed from outside their assigned topical territory.



The “C3” slot (eminent and introvert) is only populated by websites whose reputation is strictly internal and strongly symmetric. It is also over-represented in ego-centered territories.

Qualitative examination reveals here that these blogs are animated by passionate amateurs who are very active and strongly engaged in discussing and exchanging links with their topical territory — the number of comments per post is high. For instance, “*cooking*” bloggers mutually follow recipes of others and post pictures of their achievements to the attention of others. Scrapbooking blogs exchange tips, embroider blogs swap alphabet embroidery cards with their peers. In an almost underground fashion, these amateur territories are very active while being invisible to others, neither looking at nor being seen from the outside.

### Cross-territory examination

A sensible axis of differentiation between territories could hence be related to the distribution of blogs in the central slots — specifically by separating territories where slots “B2”–“C3” are under-represented (“*agora*”, “*IT*” and “*politics*”) from those where these slots are over-represented (“*beauty/fashion*”, “*cooking*” and “*crafts*”).

We may use this axis to sketch two classes of territories. A first class (on top on Fig. 3) would be mainly structured around curiosity and fame (row “1” and column “A”), all other positions being under-represented (except for, of course, “D4”, which are invisible and silent blogs). This is typical of topics related to issues of general interest, attracting the attention of a broad audience. One feature of these territories may be that they exert a marked *external attractivity* from a non-specialized audience. A second class (bottom of Fig. 3) would also be structured around curiosity and fame, but would also fill intermediary “middle-of-the-map” positions. These territories would be more focused on their interior, and feature more exchanges between central and less central blogs, around topics which are much more specialized and not necessarily appealing to all audiences. Beauty, cooking, fashion and crafts (knitting, embroidery, scrapbooking, etc.) certainly qualify. To summarize, this class would also feature a marked *internal attractivity*, including a much higher proportion of self-centered blogs — and in this respect, they would exhibit a behavior which is closer to the usual understanding of “community”.

Besides, the “A4” slot (famous and silent) offers another relevant axis of differentiation: it is strongly over-represented in the fundamentally public territories “*agora*” and “*politics*”, weakly over-represented in “*IT*” and “*cooking*” and under-represented in “*beauty/fashion*” and “*crafts*”. As said above, this position relates to the presence of externally famous actors in the territory. Anecdotally, the weak over-representation of “A4” in “*cooking*” may be explained

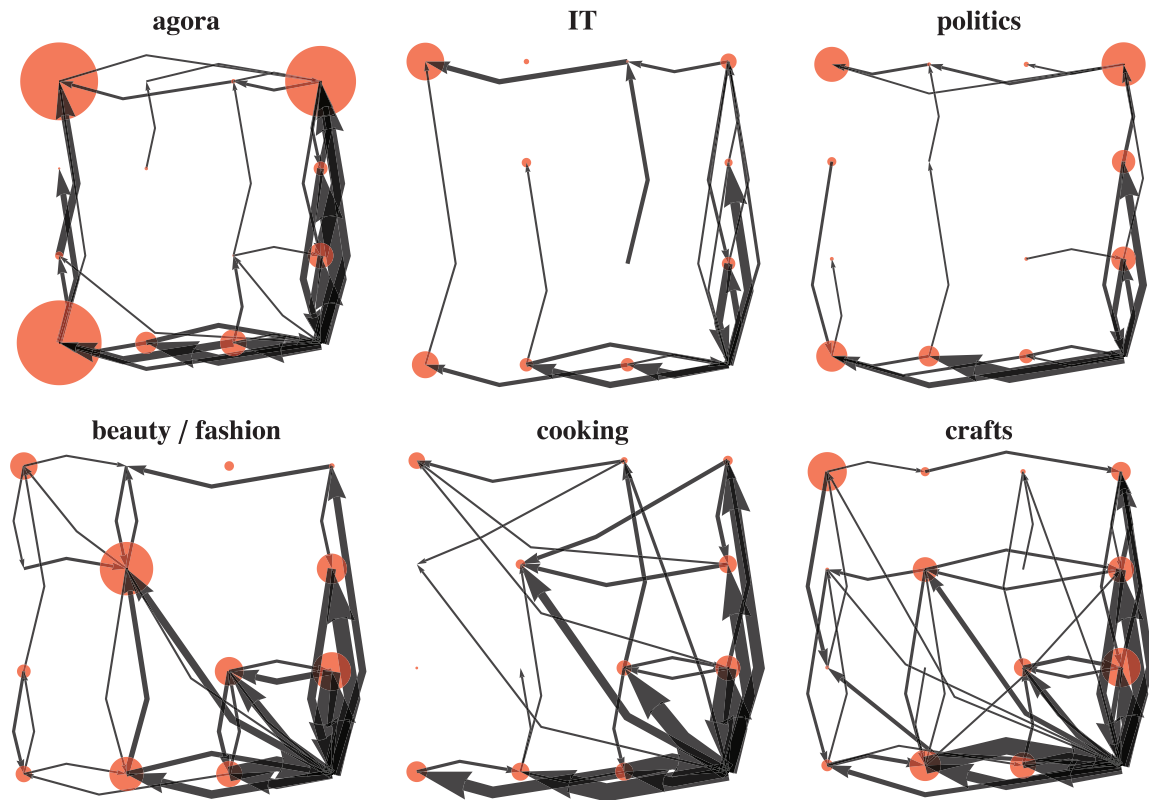


Figure 4: Migration maps. Red disks represent the size of the stable population in a given slot, while arrow sizes denote the number of blogs which migrated from a slot to another, among blogs which are present in both the first and second periods. Migrations consisting of a unique blog are not represented.

by the recent incursion of professionals of fooding and culinary medias into the territory — the space of amateur cooking blogs being more and more a marketing target.

Except for “*cooking*”, “C3” (eminent and introvert) is under-represented whenever “A4” is over-represented, and vice versa. This evidences an opposition between relational authority (“C3”) and statutory authority (“A4”).

### Trajectories

How could this position model shed light on the evolution of blogs within a territory? Are there specific paths of glory for each topical territory? The analysis of ascending and descending trajectories throughout the various possible positions may be helpful in describing the organizational and functional specifics of each topical territory. To this end, we carried a longitudinal analysis by comparing the populations of each position between two successive 5-months periods: February-June 2010 and July-November 2010.

Figure 4 represents migrations for each territory. We exclude descending trajectories to “D4”, i.e. invisible silent blogs, which are certainly numerous but probably not of striking interest: after some period of activity, possibly rewarded by external recognition, bloggers are getting weary and withdraw progressively, creating and receiving less and less links.

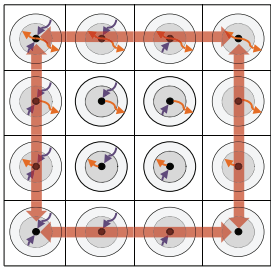
Rather, we wish to comment on ascending trajectories. We use the above dichotomy between “external attractivity” and “internal attractivity” territories as a guide to the interpretation of these trajectories. In particular, we observe that the first class essentially follows paths along the external square (D4-D1-A1-A4) and do not go through the intermediary positions of the map. By contrast, paths followed by blogs belonging to “internal attractivity” territories are more often movements through the center of the map (along D4-D3-C3-C4 and D4-D2-B2-B4). Without carrying a statistical analysis of the underlying patterns — the aim of a further study — we would like to show how the existence of typical paths can represent useful tools towards qualitatively analyzing the organization of these territories, especially in terms of reputation building and progressive embedding within a topical territory.

We schematized the most salient elements of these various types of trajectories in the following small diagrams.

### From curiosity to fame

We may notice two types of ascending trajectories in “external attractivity” territories (“*agora*”, “*IT*”, “*politics*”, on top of Fig. 4). The first one concerns actors who were already enjoying a strong popularity/reputation in the public opinion *prior* to their arrival on the web: medias, companies, statewide politicians, NGOs, etc. They need not en-

gage in an explicit citation strategy. They therefore circulate on the 4th row of our map, which enables invisible silent blogs (“D4”) to become famous silent blogs (“A4”). Specialized institutions appearing as experts in a given territory (information center, professional federation, association, etc.) often become *eminences* and go to “C4”.



A careful examination of the list of migrating blogs reveals however that, even for such externally renowned actors, it may not necessarily be easy to hold the “A4” position: they could indeed quickly move back to invisibility (“D4”) if they fail to feed their site with updated information. It is typically what happens to a

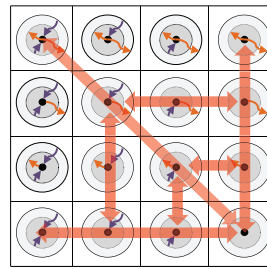
number of statewide politicians who stop caring for their blog once their campaign is over.

Another more difficult trajectory towards the all-dominant position “A1” is also possible in these territories. It concerns actors who have little to no reputation prior to their online presence: local politicians, independent IT experts, sport pundits and citizen bloggers. Here, a typical path consists in actively ascending the “D” column from invisibility to curiosity, citing more and more blogs from and outside of one’s territory. For some happy fews, this strategy is followed by another laterally ascending path along the first row, being progressively cited by *endo* blogs (eminence) then, moving forward, by *exo* blogs (popularity) and eventually up to “A1”. For instance citizen blogs, who are producing political commentaries related to news, are quickly leaving their circle of peers and go establish links with politicians, parties, media sites and a wide variety of sites outside of their origin territory.

We hence observe two main paths of glory in “external attractivity” territories: in the first one, actors are famous before their online presence; in the second one, bloggers become famous because they received recognition from their online activity, and perhaps even because they eventually received recognition from the first type of blogs. In both cases, capitalizing upon an exclusively internal presence (“C3”) is singularly rare — the term of “community” is likely to be maladapted to these territories. In these territories, there is a strong structural interdependence between, on one side, medias, companies and institutions, and, on the other side, blogspace pro-ams who became famous by orienting their activity directly towards the first type of actors. In “*agora*” for instance, this interdependence is evidenced by the presence of blog platforms on traditional media websites, who attempt at attracting the most “influential” amateur or expert bloggers — thereby helping further professional websites to move from “A4” to “A1”.

### From eminence to fame

This type of process is much less marked in “internal attractivity” territories (“*beauty/fashion*”, “*cooking*” and “*crafts*”, bottom of Fig. 4).



The most significant trajectories are rather located within an internal square circulating along D4-D3-C3-C4 and D4-D2-B2-B4. Trajectories along the fourth row (D4-A4, from silence to fame) are however still possible for websites enjoying some form of external recognition, for instance for gastronomy profes-

sionals, restaurants or fashion entrepreneurs who have a blog. Most migrations within these territories indeed involve a move through:

- either “C3”, by accomplishing a patient work of discussions with peers,
- or “B2”, which are trajectories of specialization or “un-specialization”, with respect to the topic of the territory, by blogs which are usually multi-topic blogs (often taking the form of a diary).

Blogs which, in each territory, hold the “A1” slot, reached a fame status which goes beyond the plain local eminence of “C3” blogs. Most of them are pro-am, at the boundary between professional and amateur blogs. Finally, in these territories too, there are still very few migrations to “A1” from “C3”, i.e. from locally famous pro-am blogs. On the whole, the scarcity of these trajectories, for both types of territories, illustrates the difficulty of obtaining an external recognition even when there is internal recognition. Contrarily to some prophecies regarding the online advent of amateurs, those only rarely reach fame. More concretely, we observe two stages of amateurism in the “*cooking*” territory. First, some amateur cooks (in “C3”), who specialize their blog when they stop narrating family or travel stories, focus on posting familial recipes, borrowing other bloggers’ recipes and sharing relatively casual pictures of plates. Second, some almost-professional cooks, who create original recipes specifically for the other bloggers (i.e. towards the territory), claim copyright on the recipes and make polished pictures of their achievements, are often endorsed by women magazines and engage into sponsoring activities or partnerships with agro-industrial businesses or restaurants. While there are mutual recognition and exchanges between these two universes, it is relatively rare that an amateur cooker suddenly moves to “A1”.

### Concluding remarks

The present paper aimed at providing a few insights on the topological variability of the topical territories of the blogosphere. The position model that we proposed should be understood as an interpretative framework to appraise salient differentiation facts regarding both the structure and evolution of topical territories. Several further directions of research can be emphasized. Beyond the handful of case studies which could be discussed here, we should generalize and extend the present methodology to a broader selection of territories and other corpuses, relying upon different collection methods. Our approach could also be useful to provide

recommendations for online community animation based on categories of blog trajectories within their territory.

Finally, in terms of sociological research, this framework may be useful to exhibit the relationships between amateur and professional modes of digital expression. The two types of trajectories that we underlined correspond to territories where companies and institutions are relatively present. In this respect, our approach may help shedding light on the emergence of new types of interdependencies between communities and markets in these online territories.

## References

- Adamic, L. A., and Glance, N. 2005. The political blogosphere and the 2004 U.S. election: divided they blog. In *LinkKDD '05: Proc. 3rd Intl. Workshop on Link discovery*, 36–43. New York, NY, USA: ACM Press.
- Adar, E.; Zhang, L.; Adamic, L. A.; and Lukose, R. M. 2004. Implicit structure and the dynamics of blogspace. In *Proc. Workshop Weblogging Ecosystem, 13th Intl. Conf. on WWW*.
- Benkler, Y. 2006. *The Wealth of Networks: How Social Production Transforms Markets and Freedom*. New Haven, Conn: Yale University Press.
- Beuscart, J.-S. 2008. Sociabilité en ligne, notoriété virtuelle et carrier artistique. les usages de myspace par les musiciens autoproduits. *Réseaux* 152:139–168.
- Bourdieu, P. 1992. *Les règles de l'art. Genèse et structure du champ littéraire*. Paris: Seuil.
- Cardon, D., and Delaunay-Teterel, H. 2006. La production de soi comme technique relationnelle. Un essai de typologie des blogs par leurs publics. *Réseaux*.
- Cardon, D.; Fouetillou, G.; Lerondeau, C.; and Prieur, C. 2011. Esquisse de géographie de la blogosphère politique (2007-2009). In Greffet, F., ed., *Continuerlalutte.com*. Paris: Presses de Science Po.
- Chalvon-Demersay, S., and Pasquier, D. 1990. *Drôles de stars, la télévision des animateurs*. Paris: Aubier.
- Chau, M., and Xu, J. 2007. Mining communities and their relationships in blogs: A study of online hate groups. *International Journal of Human-Computer Studies* 65(1):57–70.
- Chin, A., and Chignell, M. 2006. A social hypertext model for finding community in blogs. In *Proc. Hypertext and Hypermedia Conference '06*. New York: ACM. 11–22.
- Cointet, J.-P., and Roth, C. 2009. Socio-semantic dynamics in a blog network. In *IEEE Intl. Conf. Social Computing*, 114–121.
- Couldry, N., and Markham, T. 2007. Celebrity culture and public connection: bridge or chasm? *Intl. Journal of Cultural Studies* 10(4):403–422.
- Efimova, L., and Moor, A. D. 2005. Beyond personal webpublishing: An exploratory study of conversational blogging practices. In *Proc. 38th HICCS'05*.
- Ghitalla, F.; Boullier, D.; Neau, A.; Douarin, L. L.; and Guoskhou, P. 2004. *L'outré-lecture. Manipuler, s'appropriier, interpréter le web*. Coll. Etudes et Recherches. Paris: Bibliothèque Publique d'Information - Centre Georges Pompidou.
- Herring, S. C.; Scheidt, L. A.; Bonus, S.; and Wright, E. 2004. Bridging the gap: A genre analysis of weblogs. In *Proc. 37th Annual HICSS Conference*.
- Herring, S. C.; Kouper, I.; Paolillo, J. C.; Scheidt, L. A.; Tyworth, M.; Welsch, P.; Wright, E.; and Yu, N. 2005. Conversations in the blogosphere. an analysis “from the bottom up”. In *Proc. 38th HICCS'05*.
- Kumar, R.; Novak, J.; Raghavan, P.; and Tomkins, A. 2003. On the bursty evolution of blogspace. In *WWW '03: Proc. 12th Intl. Conf. on World Wide Web*, 568–576. New York, NY, USA: ACM.
- Kumar, R.; Novak, J.; Raghavan, P.; and Tomkins, A. 2004. Structure and evolution of blogspace. *Communications of the ACM* 47(12):35–39.
- Lento, T.; Welser, H. T.; Gu, L.; and Smith, M. 2006. The ties that blog: Examining the relationship between social ties and continued participation in the wallop weblogging system. In *Proc. 3rd annual workshop on the Weblogging Ecosystem: aggregation, Analysis and Dynamics*.
- Li, X.; Guo, L.; and Zhao, Y. 2008. Tag-based social interest discovery. In *Proceeding of the 17th international conference on World Wide Web*, 675–684. ACM.
- Lin, Y.-R.; Sundaram, H.; Chi, Y.; Tatemura, J.; and Tseng, B. 2006. Discovery of blog communities based on mutual awareness. In *Proc. 3rd annual workshop on the Weblogging Ecosystem: aggregation, Analysis and Dynamics*.
- Marlow, C. 2004. Audience, structure and authority in the weblog community. In *Proc. Intl. Communication Association Conf.*
- Mishne, G., and Glance, N. 2006. Leave a reply : An analysis of weblog comments. In *Proc. 3rd annual workshop on the Weblogging Ecosystem: aggregation, Analysis and Dynamics*.
- Nardi, B. A.; Schiano, D. J.; and Gumbrecht, M. 2004. Blogging as Social Activity or ‘Would You Let 900 Million People Read Your Diary?’ In *Proc. CSCW'04 Computer-Supported Cooperative Work*.
- Rheingold, H. 1993. *Virtual Communities: Homesteading on the Electronic Frontier*. New York: Addison-Wesley.
- Shirky, C. 2003. Power laws, weblogs and inequality. [http://www.shirky.com/writings/powerlaw\\_weblog.html](http://www.shirky.com/writings/powerlaw_weblog.html).
- Stoica, A.; Couronne, T.; and Beuscart, J.-S. 2010. To be a star is not only metaphoric: from popularity to social linkage. *Proc. ICWSM'10 4th. Intl. Conf. Weblogs & Social Media*.
- Turner, F. 2006. *From Counterculture to Cyberculture. Steward Brand, the Whole Earth Network, and the Rise of Digital Utopianism*. Chicago: The University of Chicago Press.
- Uchida, M.; Shibata, N.; Kajikawa, Y.; Takeda, Y.; Shirayama, S.; and Matsushima, K. 2009. Identifying the large-scale structure of the blogosphere. *Advances in Complex Systems* 12(2):207–219.
- Wei, C. 2004. Formation of norms in a blog community. *Into the Blogosphere; Rhetoric, Community and Culture of Weblogs*.