

Social Science Computer Review

<http://ssc.sagepub.com/>

Mapping the Norwegian Blogosphere: Methodological Challenges in Internationalizing Internet Research

Hallvard Moe

Social Science Computer Review published online 21 September 2010

DOI: 10.1177/0894439310382511

The online version of this article can be found at:

<http://ssc.sagepub.com/content/early/2010/09/16/0894439310382511>

Published by:



<http://www.sagepublications.com>

Additional services and information for *Social Science Computer Review* can be found at:


Email Alerts: <http://ssc.sagepub.com/cgi/alerts>

Subscriptions: <http://ssc.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

Mapping the Norwegian Blogosphere: Methodological Challenges in Internationalizing Internet Research

Social Science Computer Review
000(00) 1-14
© The Author(s) 2010
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0894439310382511
http://sscr.sagepub.com


Hallvard Moe¹

Abstract

Even as the blog has become an established genre of computer-mediated communication, questions remain about how different blogs are from mass media, and what the transformative potential of blogging is. This article argues for the need for further explorations, especially outside the Anglo-American blogosphere. The article discusses key challenges in light of an ongoing research project aiming to gain insight into how blogging in Norway—a small democratic nation state with a correspondingly small language area—compares to other cases and to assess how online media participation matters for the structure of the public sphere. On this basis, the article presents preliminary findings from a mapping of the Norwegian blogosphere.

Keywords

blogs, blogosphere, hyperlink network analysis, methodology, Norway

Introduction

Even though the blog by 2010 has become an established genre of computer-mediated communication, it is still widely used to illustrate the disruptive potential of online media. According to such an argument, blogs represent the opposite of traditional mass media. Rather than being disseminated in controlled ways from a few to a passive mass, blogs create open, democratic dialogue among a networked public.¹ This view also remains widespread in academic writings (e.g. Feenberg, 2008; Kerbel, 2009). At the opposite end of the spectrum, scholars argue that blogs and related tools facilitate a fragmentation of the public sphere that could lead to balkanization of public debate, and consequently political polarization (Sunstein, 2007). Others underline that the average blogger does not take part in any substantial dialogue at all (e.g. Herring et al., 2005). The questions remain, then, how different blogs are from mass media, and what the transformative potential of blogging is. How can we approach an assessment of the ways in which the blogosphere matters for the public sphere?

¹University of Bergen, Bergen, Norway

Corresponding Author:

Hallvard Moe, University of Bergen, Infomedia, Bergen 5020, Norway
Email: Hallvard.Moe@uib.no

Although several studies aiming to answer such questions have emerged recently, there is a pertinent need for further explorations, I will argue, especially outside an Anglo-American context. Such an endeavor does, however, pose new methodological challenges. This article discusses challenges related to an ongoing research project that examines how blogging in Norway—a small social democratic nation state of 4,8 million inhabitants—compares to other countries and assesses how online media participation matters for the structure of the public sphere.

I first substantiate the need for further empirical analyses. This leads me to consider key methodological issues with hyperlink network analysis of the blogosphere, focusing on challenges with sampling and crawling blogs from a small language area. Next, I discuss findings from two separate data sets, illustrating how two political issues—the Data Retention Directive and sick leave—are dealt with in the Norwegian blogosphere. I argue that the Norwegian blogosphere can be described as open and free of authoritarian control; moderately networked; layered; with low levels of partisanship; and as housing young, alternative voices side-by-side with well-known politicians. In closing, I identify advantages and limitations of network analysis and point to paths for further research, including extending network mappings as well as adding complementary analytical approaches.

Which Blogosphere?

A blog is a frequently updated website consisting of chronologically dated entries in reverse order (e.g. Bruns & Jacobs, 2006). To provide necessary room for a wealth of subgenres, and allow for the genre to remain dynamic, it is fruitful to add that a blog is “an author-driven, asynchronous and informal genre of computer-mediated communication that uses various modalities and entails some interactivity” (Lomborg, 2009, p. 3). On this basis, and building on established models of content analysis, specific blogs can be positioned along a continuum related to three dimensions (Lomborg, 2009, p. 6): (a) Content axis: Internal—topical. Does it deal with personal themes, experiences, and emotions, or primarily topics of general interest? (b) Directional axis: Monological—dialogical. Is it primarily used for dissemination to an audience, or highly conversational and densely networked? (c) Style axis: Intimate—objective. Is it confessional and personal, or rather objective in tone?

Importantly, this definition does not at the outset single out political blogs. Studies of the public sphere and online media to an overwhelming degree tend to operate with a restricted notion of what topics are political or newsworthy, or what forms of communication are close enough to an ideal to count (Gerhards & Schäfer, 2010; Park, 2009 for recent examples). Public sphere research concentrates on blogs close to the objective and topical ends of the axes. Such a delineation, where the public is seen as dealing only with formal politics, is problematic. When looking for politically relevant communication, we need to consider also forms that content-wise are closer to the internal end of the scale than the topical, and the ones that combine intimate and objective styles. The point can be illustrated by looking at the most popular Norwegian blogs.

According to one major blog ranking site, the top five blogs written in Norwegian by March 2010 were (a) NRKBeta: a highly topical (new media), dialogical and objective blog by the public service broadcaster’s R&D department; (b) VamPus’ Verden: a mostly topical (explicitly conservative political), largely dialogical blog mixing intimate and objective styles, by a 30-something female; (c) Sukkersött: a topical (gift cards), dialogical and rather objective blog maintained by a group of females; (d) Petunia: a mainly internal, dialogical and intimate blog by a 45-year-old female, largely based on her photos of everyday life; and (e) Revolusjonært roteloft: a largely dialogical blog with an intimate style, mixing internal and topical postings about “everything from poetry to personal stories in between politics and nerdy stuff” maintained by a 22-year-old female.² This list just gives a glimpse of the tip of the iceberg at one point in time. Still, it illustrates the wealth of popular blog subgenres and also shows that strictly political blogs are far from dominating the blogosphere.

Blogs make it easy for individuals to share information about a pressing issue publicly. They also stand out from traditional mass media with key features like links, diverse feedback mechanisms, easily accessible archives, and systems for signalling endorsement. Such characteristics have led to a body of work relevant for an assessment of online media's impingement on the structure of the public sphere. These studies analyze the blogosphere that is "the overall community of blogs and bloggers, which is interlinked by a large number of cross-references between individual blog entries" (Bruns & Jacobs, 2006, p. 5). If we are interested in the structure of the blogosphere as part of a public sphere, questions are where the borders are for that "community," how communal it is, and how its links are distributed and reciprocated. Referring to the second definitory axis, we are interested in to what extent blogs are monological or dialogical, and how different networks of blogs interrelate.

Hargittai, Gallo, and Kane (2008) offer one sound contribution. Testing the fragmenting potential of online media, they analyze linking practices among widely read liberal and republican blogs from the United States. Their findings confirm that bloggers are more likely to link to those who match their ideological persuasion, thus strengthening a pessimistic outlook. Hindman (2009) gives another contribution. He looks at diverse forms of Internet tools, including search engines and blogs, and argues that the Internet looks very much like the offline media world: audiences are no less concentrated, it is still extremely hard to get heard for those outside a small elite group, and highly educated White males write the handful of blogs that do get broad attention (also Wei, 2009).

Existing studies as these mentioned above are based on the U.S. context (e.g. Kenix, 2009; Tremayne, 2007 for further examples). The U.S. dominance is hardly surprising (e.g. Goggin & McLelland, 2009, p. 7). The U.S. focus does have obvious consequences for the generalizability of the findings. Like other cultural practices, blogging differs in different societies. There is a need for further empirical explorations. At the center of this issue, then, lies the questions of how networked and communal the blogosphere is in different social, political, and cultural contexts; how its different parts relate to each other; and how these structures compare with the workings of traditional offline mass media—how dissemination is balanced against dialogue. However, attempting to answer such questions in different societies raises a line of methodological issues.

Methodological Challenges

The ability to link pages and content is a defining feature of the web. Based on the idea that linking can yield useful information, link analysis has emerged as a widespread analytical method in many fields. Within the social sciences, two approaches can be traced to different frameworks (Park & Thelwall, 2006). Webometrics developed from information science (see Almind & Ingwersen, 1997). Hyperlink network analysis, however, is based on social network analysis. In the latter, a node is a website representing a social actor, and the relationship between two nodes is expressed by hyperlinks (Rogers, 2009a, pp. 13–14).

Data for hyperlink network analysis can be collected manually by observation, but due to high labor costs and the potential for coding errors, computer-assisted collection—in the form of a web-crawler system—is recommended (e.g. Park & Thelwall, 2006, p. 8). One obvious advantage with this method is its unobtrusiveness. A crawler "blindly" traverses the Internet without knowing the semantics of the data it is gathering (Bell, 2010, p. 2). This does, however, not mean it is shielded from common methodological biases. The researcher provides starting points for the crawler and sets parameters typically restricting the speed, breadth, and depth of the data gathering. In addition, the researcher may define types of data to be gathered and characterize content to be avoided (Bell, 2010, p. 2).

For studies of the structure of the blogosphere, one specific implementation of hyperlink network analysis is particularly helpful: issue network crawls. Such crawls start from "seeds," that is blog

posts' URLs defined by the researcher, follow available hyperlinks repeatedly in several steps, and identify "patterns of interconnections in the population of websites discovered in the process" (Bruns, 2007, p. 1). The result is a data set describing a network through measurements of density, centralization, the relation positioning of nodes, and their specific interconnections. Issue network crawl analyses can predict neither the actual use of the web nor the traffic generated by specific links. The assumption is rather that on aggregate, it is likely that users' "browsing patterns are going to resemble the patterns of interlinkage between key sites" with information of a particular topic (Bruns, 2007, p. 3). Search engines use the same patterns to rank sites. A central node in an issue network would thus presumably also be listed prominently in search results for that issue. Furthermore, for blogosphere studies, issue network crawls may offer insight into how bloggers, blog commentators, and external actors highlight information sources since most connections identified in the crawl will tend to be based on topical links in posts or comments (Bruns, 2007, p. 3).

A major methodological challenge for such issue network analyses of the blogosphere stems from the basic fact that different, often self-written, computer programs have been used in different projects. This has implications for comparative potentials, as well as for prospects of replicating designs (Park & Thelwall, 2006, p. 8). To seek to amend this, the current study uses a publicly available crawler system—IssueCrawler—to perform a per-page co-link analysis.³ By 2010, a range of studies, also of the structure of the blogosphere, has used IssueCrawler.⁴ This does, however, not mean a basic replication is readily available. Further methodological concerns can be illustrated by the question of sampling.

Li and Walejko (2008, p. 282ff) identify four problems with sampling blogs. The first problem is spam blogs or splogs. Splogs are advertisements in the form of unconnected and repetitive content used to increase page ranking. Second, blogs with access restrictions cause problems. Such sites may show up in the sample without enabling analysis, or they may be difficult to locate if search engines do not index them. Third, it is well known that a great portion of blogs that is created quickly gets abandoned. Such abandoned blogs may distort the sample. Last but not the least, blogs are increasingly embedded in new contexts like social networking sites or mainstream news outlets. In addition, blogging is morphing, for instance into the so-called micro-blogging through services like Twitter.

These four general issues are to different degrees relevant for the current study. First, although research from the United States has estimated that one in five blogs is a splog (Morrissey, 2005), no corresponding study is available for Norway. One may, however, assume that blogs as spam are less prevalent in a small market and language area. The seeds which are the actual blog posts used as starting points for the data collection, have been manually checked for the presence of splogs. Moreover, although the crawl may run into spam, the co-link analysis means these will only show up as nodes in the final network if at least two other nodes link to them.

There is no reason to assume the presence of access-restricted blogs to be smaller in the Norwegian blogosphere than in the United States. Such blogs are, however, avoided in the samples of seeds in the same way as splogs. In addition, since the current study concerns the role of blogging in public-mediated communication, private blogs can be defined as less relevant.

The third problem identified by Li and Walejko (2008), abandoned blogs, is more directly applicable for my research interest. It is notoriously difficult to measure the blogosphere. Few statistics exist of Norwegian blogs. Some estimate that the number of bloggers in the country has risen from around 30,000 in 2005 to 300,000 in 2009. Others estimate 350,000, whereas yet others calculate as many as 450,000 (Saxvik, 2009; Tveit, 2009).

Studies have found that nearly two thirds of surveyed blogs are left without updates for 2 months. Recently, in Denmark, blog portal site *overskrift.dk* reported that by late 2009, roughly half of all blogs were abandoned, that is with no posts in the last 3 months (Lund, 2010). If transferred to Norway, one may estimate the number of active blogs to be approximately 200,000. Abandoned blogs have been avoided in the list of seeds for the current exploration. The co-link analysis also

serves to limit the inclusion of zombie sites. Still, there is a possibility that blogs that are no longer updated get included in the data sets. This is, however, not necessarily negative: blog posts or ensuing discussions may serve as points of reference even though a site has been left static. If abandoned blogs were filtered out from the crawl, the analysis would potentially be biased.

The fourth issue, that of the amorphous nature of blogs, also poses a challenge to the current study. National leading tabloid newspapers like VG have integrated blog services in their online offers. In addition, the use of social networking sites, especially Facebook, is widespread in Norway. Although the sampled seeds are all “traditional” blogs, the data collection is open to include links to and from all forms of websites, including social networking sites. As such, given the explorative nature of the study, a task could be to identify the occurrence of such nodes in the issue networks, and use those as a basis for further analysis.

When relating these general methodological issues to the current study, the need for contextual awareness becomes clear. The weight of the different problems not only depends on research interest, forms of data collection and analytical design but also to a substantial degree varies with the context of the research. Key characteristics of a country—size, information and communication technologies (ICT) use, language, and degree of state censorship—not only set it apart from other cases, but also impinge on methodological considerations. If mapping the blogosphere in a state with a repressive political regime, for instance, the problem of nonindexed or semi-restricted blogs might be pressing, whereas the mass of abandoned blogs might be less problematic if uptake of ICTs has been slower. This means we need to add a methodological argument to the empirical argument for internationalizing Internet studies. This point is also crucial for another part of the methodological design: the identification and selection of seeds for the crawls.

The seeds are sampled using blog aggregators.⁵ Since aggregators differ for instance in the algorithms used to rank blogs, and in the total number of indexed blogs, the sampling method poses constraints for comparisons. In recent years, Technorati has emerged as a widely used site for sampling. For instance, Bruns (2007) used the site’s search function to identify blog posts on a specific political issue. Bruns also used Technorati’s authority settings to generate different lists of seeds. The result was three sets of URLs used as starting points for crawling the Australian blogosphere.

Although Technorati indexed nearly 113 million blogs by mid-2008, it is of little use for studies of non-English communication. Searches for Norwegian-language blogs or phrases in Norwegian are futile. This is hardly surprising but has consequences for the comparability across studies. There is no aggregator of a similar standing to Technorati that indexes Norwegian-language blogs. However, other options exist. Twingly is a Swedish-based blog aggregator focusing on Europe. It offers searches of blogs in 30 different languages, including Norwegian. Further, Bloggurat is an aggregator indexing Norwegian-language blogs. The current study used these 2 to create a list of seeds by filtering out the unique hits from the top 100 lists generated by similar keyword searches for both. Although this sampling method does not cover all Norwegian blogs, the combination of two aggregators should limit discrepancies caused by different ranking mechanisms (Li & Walejko, 2008, p. 285). More fundamentally, the design illustrates how different contexts pose different methodological challenges, underlining the need to gather experiences with studies in diverse settings.

Norwegian media and communication researchers have tended to approach blogs from either journalism studies or the humanities. There are valuable studies of how mainstream online news actors incorporate new possibilities for interactions and participation (e.g. Engebretsen, 2006) and studies of relations between bloggers and the press (Øvrebø, 2006). However, although a Norwegian-based scholar wrote the first English-language introductory textbook on blogs (Rettberg, 2008), no study mapping the Norwegian blogosphere exists. Such a mapping could provide insight into the ways in which blogging in Norway compares to other cases and also help build a basis for discussion of how this dimension of online media participation matters for the structure of the public sphere.

Two Impressions of the Norwegian Blogosphere

To shed light on the question of how the Norwegian blogosphere is structured, two different issue networks were mapped with two identically set up IssueCrawler operations.⁶

The first case is the debate about whether Norway should implement the European Union's Data Retention Directive. The Directive regulates archiving of traffic data for telephony and Internet use. The rationale is prevention and prosecution of (organized) crime and terrorism (EC, 2005). Somewhat like the U.S. Patriot Act, the Directive triggers debate about fundamental privacy issues held against the fight against crime and terrorism. During autumn of 2009, the issue was debated publicly, as well as within several political parties, in Norway. By early 2010, the question of implementation in Norwegian law was given a formal public hearing. The second case is a broader issue without a clear time frame. It concerns debate about increases in the use of sick leave in Norway. A cornerstone of the welfare state system, the rights held by employees during illness has come under increased pressure as politicians debate measures to lower the numbers to control public spending.

The two issues were both high on the Norwegian political agenda during winter 2010. More importantly, the two issues differ in several aspects: the first is time- and case-specific, technical and juridical, and would presumably fit well with topics debated among net savvy users. The sick leave issue, however, does not refer to one specific document or decision, it is less technical, but rather touches on a long-standing debate that affects the everyday lives of most citizens.

As Figure 1 shows, the crawl for the Data Retention Directive issue generates a low-density network ($\Delta = 0.194$), indicating limited cohesion (Wasserman & Faust, 1994, p. 129), without any clear clusters. The actor ranking list puts several well-known blogs close to the center of the network: the Twingly rated number 5-blog *Revolusjonært roteloft*, along with several topical blogs on political, social, and cultural issues written in a largely objective style by middle aged males (e.g. *Vox Populi* and *Hjorten*). These are individual bloggers operating outside traditional media or established political parties.

There is also a presence of liberal and conservative political voices among the highest-ranking sites. These include explicitly political individual bloggers—prominently *Vampus* and *Kapitalismus*. Furthermore, the website of a liberal think tank (*Civita*) and the website of a liberal-conservative print quarterly (*Minerva*) both appear as central nodes. This is hardly surprising, since the topic in question touches on basic liberal political rights. The presence of another set of central nodes is perhaps more unexpected: top politicians from the social left, currently in government. The two vice-leaders of the Socialist Left Party maintain blogs (*Rettvenstre* and *Bardvegar*), along with a third party and parliament member (*Heikki*). These three, as well as the party website (*SV*), show up prominently in the network. If ranking the core nodes according to how many links they receive from different sites, 7 of the top 15 nodes can be identified as male bloggers, whereas merely 2 are females.

At the periphery of the Data Retention Directive network depicted in Figure 1, we find a range of sites acting as "resources": mainstream news sites (e.g. *VG* and *Dagbladet*), general social networking sites, official channels of information (*Datatilsynet*), as well as interest groups (*IKT-Norge*), and informational sites (*Europaveien*).

The actual linking between the nodes in the network shows some striking patterns. First, although several of the core nodes demonstrate a balance between in- and out-links, signaling reciprocity and a dialogical mode of communication, some core nodes (e.g. *Vox Populi* and *Vampus*) stand out with no out-links to the network whatsoever. As such, they appear as outlets for dissemination. Second, looking at the explicitly political sites along with the ones written by established politicians, they link extensively across the political spectrum. Socialist Left vice-leader *Bård Vegar Solhjell*, for instance, links to his party colleagues, but also to *Minerva*, the vice-leader of the Liberal Party's youth organization (*Jonasliberal*) and to the leader of a lobbyist for technology businesses within



Figure 1. Data Retention Directive network map generated from IssueCrawler by govcom.org at February 4, 2010. Cluster map (size of nodes by in-links and out-links).

the Confederation of Norwegian Enterprise (Paulchaffey). Likewise, the other Socialist Left vice-leader links to the blog of a Conservative Member of Parliament (Konservativ), and several other bloggers on the political right. Importantly, the conservative nodes do not reciprocate these links.

Compared to the Data Retention Directive network, the crawl of the sick leave issue generates a somewhat denser network ($\Delta = 0.234$), as illustrated by Figure 2, with a clear cluster of central nodes and a more dispersed periphery. Among the central nodes are several of the highest-ranking blogs from the Data Retention Directive crawl: *Revolusjonært roteloft*, *Vox Populi*, *Hjorten*, as well as *Vampus*. The latter three are highly topical blogs written chiefly in an objective style. *Revolusjonært roteloft*, however, stands out—like it does on the Data Retention Directive map—but is joined here by another set of core nodes (*Undreverset*, *Serendipitycat*, *Under stjernerne*, *Reality Challenged*). Like *Revolusjonært roteloft*, these blogs are thematically and stylistically broad in scope, adding observations from everyday life, hobbies, school, or work to posts on social and political issues. That is, they combine personal themes, experiences, and emotions with topics of general interest and mix a confessional style with more objective and argumentative modes of address.

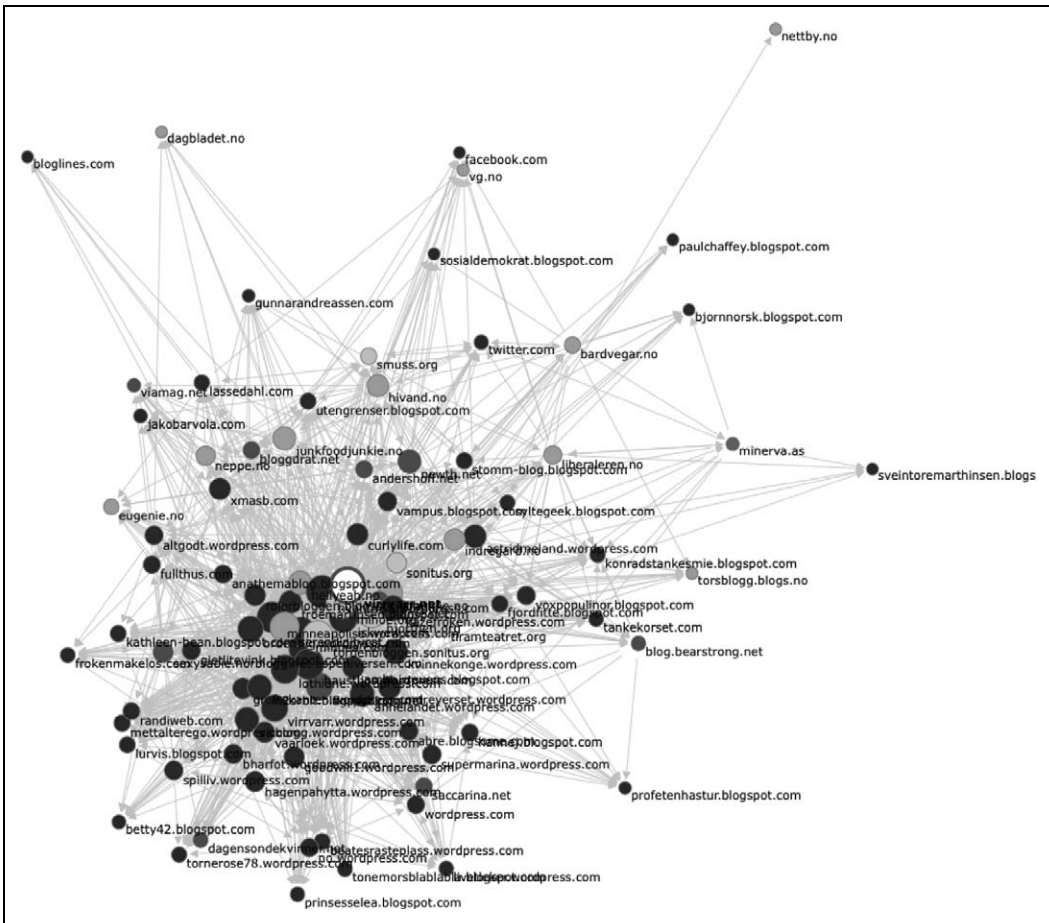


Figure 2. Sick leave issue network map generated from IssueCrawler by govcom.org at February 4, 2010. Cluster map (size of nodes by in-links and out-links).

Specifically, on the issues of sick leave, health care, and the welfare system, some of these blogs (Revolusjonært roteloft, Under stjernene, and Rolerbloggen) draw on personal experiences with emotional illness and mental health care services. Finally, the core network on sick leave stands out in terms of gender composition: Ranked according to number of in-links from different sites, a clear majority of the nodes (9 of the top 15) can be identified as female bloggers, with only three identified as male.

The sick leave issue network is also interesting since the centrality of these less objective blogs go hand in hand with the absence of self-defined political outlets as prevalent in the Data Retention Directive issue network. Although Socialist Left representative Bård Vegar Solhjell and the liberal-conservative site Minerva both appear at the periphery, none of the established voices from either side of the political spectrum gets included close to the center. The layout of the remaining nodes in the periphery also differs from the Data Retention Directive case: a larger portion consists of individual blogs, whereas the prominent institutional informational sources are fewer and, literally, farther between.

The actual linking between the nodes in the central cluster of the sick leave issue network also serves as a sign of its density. Importantly, there is a pattern of reciprocity in the linkages among

the nodes that indicate a dialogical communicative form, without a clear set of sources which do not link to the network, but receive many links from other nodes. The question, then, is what these two data sets can tell us about the structure of the Norwegian blogosphere.

Discussion

Although the scope of the analysis presented here is limited and the findings preliminary, the crawls inform a discussion of some key characteristics of the Norwegian blogosphere, necessary for an assessment of its impact on the structure of the general public sphere.

First of all, when reviewing the findings, one is compelled to look at the very fundament for blogging in Norway. The country is a constitutional monarchy with a parliamentary system of government. It has been governed by a labor party-led center-left coalition since 2005. Norway ranks among the wealthiest countries in the world, with the largest capital reserve per capita of any nation, due mostly to its oil and gas industry. The country's media system is characterized by early development of the press, a shift away from political pluralism toward a neutral commercial press, strong institutionalized professionalism, and strong state intervention (Hallin & Mancini, 2004). ICTs are comparatively widespread. In 2008, Norwegians spent 65 min online during an average day (MedieNorge/SSB, 2009). By 2009, 88% had access to a PC at home, and 86% of Norwegians had Internet access at home. During the first 3 months of 2009, 91% of the population said to have used Internet services in one form or another (SSB, 2009). The online version of the country's largest tabloid newspaper—VG—reached just short of 1.3 million users per day by early 2010 (TNS Gallup, 2010). By February 2010, Alexa ranked it as the country's number 5 website, below Google's Norwegian site, Facebook, Google.com, and YouTube. The Norwegian blogosphere, then, is situated in a wealthy state with high penetration of ICTs, and freedom of religion and speech. Within such a societal framework, blogging is an open practice free of state censorship.

On this very basic level, the Norwegian blogosphere stands out from others, subject to more oppressive regimes. John Kelly and Bruce Etling (2008) have undertaken a mapping of the Persian-language blogosphere in Iran. They identified four major network formations or poles in the blogosphere of which politically radical blogs make out one portion. Despite state authorities' attempts to stop bloggers and block their sites, Kelly and Etling argue that the blogosphere "may represent the most open public communications platform for political discourse" in Iran today (Kelly & Etling, 2008, p. 2). It is worth noting these basic features, since they have a clear impact on the role of blogging in the public sphere (Moe, 2010).

A second, related observation based on the mapping of the Norwegian blogosphere concerns the presence of alternative voices. This is a key argument for the democratic potential of online media in general (e.g. Bohman, 2004), and blogs in particular (e.g. Kenix, 2009): the extent to which the participants in the blogosphere, and their views, differ from those in the mainstream media would signal a broadening of public debate. The issue networks illustrate that the blogosphere in Norway indeed does provide outlets for alternative voices. *Revolusjonært roteloft* is perhaps the clearest example. Appearing as a central node in both mappings, the blog—whose name translates as *Revolutionary Attic*—combines personal and topical issues and styles, and is consistently dialogical. As a young female student, the blogger behind it also stands out from the dominant voices in mainstream political debate.

Such examples notwithstanding, the mapping gives an overall impression of the Norwegian blogosphere on these issues as fairly conventional. This goes for voices—the prominence of established politicians is a case in point—as well as modes of communication. In their study of the Australian political blogosphere, Bruns and Adams (2009, pp. 88–89) found a considerable left-wing inclination with strong polarization on specific issues coupled with activist, progressive causes, in clear opposition to a mainstream media largely supportive of the conservative federal government. The

current analysis does not support a similarly clear picture. Activist causes can be said to be important, for instance for understanding the communication about the Data Retention Directive, but the Norwegian blogosphere is not accurately described within such movements or the political left–right dimension. Although the specific character of the Data Retention Directive issue—with extensive internal opposition in every major political party—goes some way to explain this finding, part of the explanation also lies in the fact that a labor-led government has ruled the country since 2005. In such a political climate, the face of the alternative is obviously different for instance from Australia in 2007. In the Norwegian blogosphere, conservative–liberal voices might very well be seen to illustrate the disruptive potential of online media for the public sphere.

A third point, then, has to do with the level of partisanship. Represented by separation into clear clusters of interlinked nodes, bipartisanship is a key issue in mappings especially of the U.S. blogosphere. In terms of clustering, the mappings undertaken here do not show steep divisions between politically diverging groups.

Findings from the South Korean blogosphere could help explain the mechanisms at play. Han Woo Park and Randy Kluver (2009) looked at the linking practices of South Korean politicians' personal blogs to see how online behavior reflects offline political culture. They found cultural values and patterns to reproduce themselves: the linking practices tend to follow “the expectations established by centuries of Korean political traditions” (Park & Kluver, 2009, p. 260). In that sense, the authors argue, “social and political relations remain largely untouched by the technology, in spite of the vast potential” (Park & Kluver, 2009, p. 261). Transferred to the Norwegian case, the multiparty system of the Norwegian parliamentary democracy does not facilitate partisanship between two distinct camps, but rather extends the culture of coalition building and consensus seeking into the online realm.

Based on the analyses undertaken here, partisanship appears as neither a key point of reference, nor an obvious identifier among Norwegian bloggers. Still, some central blogs are explicitly political. The finding from the Data Retention Directive issue network that bloggers—also established, high profile politicians—from opposite sides of the political spectrum appear relatively close to each other, however, signals a level of cross-partisan communication. Judging from this aspect, while the blogosphere does facilitate segmentation of the public sphere, the study does not strengthen a claim of balkanization of public debate.

This brings us to the differences between the two issue networks. Among the central nodes, there is clear variation both in terms of the kind of sites and the bloggers. The Data Retention Directive network has a majority of male bloggers among its core nodes and shows a strong presence of topical blogs with an objective style. The sick leave network, however, is denser, includes more diverse, not strictly political, blogs, and is dominated by female writers. To a certain extent, such differences can be understood with reference to the issues at hand. The Data Retention Directive touches on basic civic rights but also entails complicated legal arguments. As such, it requires a great deal of expert knowledge. Thus, it attracts activists, expert professionals and academics, and also lays out internal party oppositions. The issue of sick leave, on the other hand, is more evidently of general interest, requires little expert knowledge, invites contributions based on personal experience, and as such mobilizes laypersons in addition to experts.

The differences also illustrate the strength of the methodological set up: it does neither discriminate between explicitly political and nonpolitical blogs, nor restrict the analysis to communication following a predefined argumentative mode. Especially the mapping of the sick leave debate makes evident how openness to a wide range of subgenres and communicative forms helps us get a fuller understanding of the construction of the blogosphere.

The two mappings also demonstrate that it is hard to grasp the Norwegian blogosphere if focusing on one kind of issue. The blogosphere is layered, and its different networks only partly overlap. Each mapping based on an issue crawl merely provides insight into one layer. To get a fuller

understanding of the structure, we are well served by an approach that gives attention to diverse issues of political relevance. The findings of the current study describe the Norwegian blogosphere as open and free of authoritarian control; moderately networked; layered; with low levels of partisanship; and housing young, alternative voices side-by-side with established well-known politicians.

Conclusion

This article started with the question of how blogging matters for public debate. I have argued that the still diverging answers to this question remain based on insight into some specific blogospheres, chiefly a U.S.-based, English-language one. The challenge for research is not only empirical, as different social, political, and cultural contexts also entail different methodological challenges. I have discussed their relevance for a preliminary mapping of the Norwegian blogosphere.

The findings of this mapping, based on hyperlink network analysis, yield valuable insight into the workings and layout of the blogosphere as a part of the mediated public sphere. The analytical approach does, however, have some important limitations.

First, a hyperlink network analysis cannot tell us how blogs are read by actual users, what bloggers' and commentators' motivations are for including or avoiding specific links, or how the habit of blog reading matters for individuals' overall media consumption. To answer such questions, we need to add content analysis and reception studies (e.g. Benkler, Shaw, & Stodden, 2010; Wojcieszak, 2010).

The second limiting aspect of the methodological set up is more specific and serves to remind us of the amorphous state of the blogosphere and online media in general. Both data sets show interlinkages to major networking sites like Facebook. To label Facebook as just another node does hardly cover the wealth of content the site houses. A similar argument pertains to micro-blogging services like Twitter, which also appear in the data sets of the two issue networks. This speaks to the need for expanding the scope of the hyperlink analysis, for instance by combining crawls starting from traditional blogs with analyses based on budding forms of networked media.

Third, the crawling method also has some more direct limitations based on what it does not show us. This concerns filtering. Filtering is necessary for the co-link analysis to work. Sites like google.com get a massive amount of generic, often automatically generated links, irrelevant for a mapping of a specific issue. Blocking such sites from the co-link analysis should help secure a relevant and consistent data set. Yet, especially the recent growth of commercial blog services poses thorny questions of blocking. In the Norwegian case, blogg.no is a hugely popular blog provider and portal, especially among younger users. Its standardized blog templates, coupled with the portal's massive linking to individual blogs' and blog posts, threaten to swamp an issue network map by generating such massive amounts of generic links that results are heavily skewed. Mainstream media outlets' blog platforms can pose similar problems.⁷ The point is that the constantly changing forms of blogging together with increased commercialization pose challenges to the methodological set up of hyperlink network analysis.

Lastly, the reliance on links risks ignoring sparsely networked sites. One such example from the case of the Data Retention Directive is a site set up by the Ministry of Transport and Communications to instigate online debate in connection with the formal public hearing. The site used the WordPress publishing application, and the Ministry called it a blog. Still, the site only had one single post—an invitation to use the commentary field for discussion. The Ministry also stated it would not participate in the discussion. In practice, then, the site is not a blog, but meant to function as a limited message board. As the case was for the Data Retention Directive data set, a crawl-based issue network map might not depict such sites. Nevertheless, content analysis could clearly shed further light on the ways in which the topic is publicly debated on the Internet.

In sum, a comprehensive and detailed map of the blogosphere will require not only analysis of content and uses but also an extended network analysis beyond traditional blogs. Importantly, when embarking on such studies, we need to also scrutinize the connections—or lack thereof—to mainstream media. That is, we need to acknowledge that the Internet is an ubiquitous part of public life.

Declaration of Conflicting Interests

The author declared no conflicts of interest with respect to the authorship and/or publication of this article.

Funding

The author received financial support from The Research Council of Norway for the research of this article.

Notes

1. A public is a body made up of members of a democratic polity. Acting collectively, they should control the rule of their society. When the public communicates about the collective control of societal rule, it constructs a public sphere. In complex modern societies, this communication is to a substantial extent dependent on mediation. Mass media should function as the infrastructure of the public sphere, ideally by making important information available for all, and enabling citizens to communicate with each other about relevant issues. That is, the media should both disseminate information and facilitate dialogue.
2. List accessed March 10, 2010 at <http://www.twingly.com/top100?lang=no>. Information about the authors of the five blogs is based on their own presentations. So are the descriptions of content and styles, which in addition build on close reading of the blogs in question.
3. Details about the operation are available at <http://www.govcom.org>. In short, govcom.org Director Richard Rogers describes Issuecrawler as “a server-side web network location software. Input URLs into the Issue-Crawler, and the software crawls the URLs, captures page/site outlinks, performs co-link analysis, and outputs the results in lists as well as visualizations” (Rogers, 2009b, p. 116, see Bruns, 2007 for a more detailed description of features).
4. See <http://www.govcom.org> for updated lists of further studies. The value of this body of work is heightened by the fact that all data sets are available to registered users in an online archive.
5. Blog aggregators are “websites that collect, organize, and publish links to large numbers of blogs using web crawlers to locate blogs or by receiving pings from blogs” (Li & Walejko, 2008, p. 285). They offer lists of blogs based on popularity, authority, or activity, and facilitate searches of blogs.
6. The co-link analyses were set up by page, without privileged starting points, with a crawl depth of two (meaning links are collected in two steps from the starting points), and two iterations (meaning the co-link analysis is undertaken for both steps). The numbers of unique URLs used as seeds were 126 in the Data Retention Directive crawl and 127 in the sick leave crawl.
7. One example is leading tabloid VG’s vgb.no. An alternate crawl of the sick leave issue showed a distinct cluster of a handful vgb blogs. At closer inspection, however, the connections between these blogs turned out to be solely constituted by template-based links from individual blogs to the vgb blog portal site; none of the bloggers actually inserted links to each other.

References

- Almind, T. C., & Ingwersen, P. (1997). Informetric analysis on the World Wide Web. *Journal of Documentation*, 4, 404-426.
- Bell, G. B. (2010). The dangers of webcrawled datasets. *First Monday*, 2.
- Benkler, Y., Shaw, A., & Stodden, V. (2010). *A tale of two blogospheres: Practices on the left and right*. Berkman Center Research Publications. Cambridge, MA: Harvard University.

- Bohman, J. (2004). Expanding dialogue: The Internet, the public sphere and prospects for transnational democracy. In N. Crossley & J. M. Roberts (Eds.), *After Habermas: New perspectives on the public sphere* (pp. 131-155). Oxford, UK: Blackwell Publishing.
- Bruns, A. (2007). Methodologies for mapping the political blogosphere: An exploration using the IssueCrawler research tool. *First Monday*, 5.
- Bruns, A., & Adams, D. (2009). Mapping the Australian political blogosphere. In A. Russell & N. Echaibi (Eds.), *International blogging: Identity, politics, and networked publics* (pp. 85-111). New York, NY: Peter Lang.
- Bruns, A., & Jacobs, J. (2006). Introduction. In A. Bruns & J. Jacobs (Eds.), *Uses of blogs* (pp. 1-9). New York, NY: Peter Lang.
- EC. (2005). Press release: Data retention directive. *Europa*. Retrieved March 10, 2010, from <http://europa.eu/rapid/pressReleasesAction.do?reference=MEMO/05/328&format=HTML&aged=0&language=EN&guiLanguage=en>
- Engebretsen, M. (2006). Shallow and static or deep and dynamic? Studying the state of online journalism in Scandinavia. *Nordicom Review*, 1, 3-16.
- Feenberg, A. (2008). Et demokratisk Internett? [A Democratic Internet?]. In B. Gentikow & E. G. Skogseth (Eds.), *Medier og demokrati [Media and Democracy]* (pp. 209-229). Oslo, Norway: Scandinavian Academic Press/Spartacus.
- Gerhards, J., & Schäfer, M. S. (2010). Is the Internet a better public sphere? Comparing old and new media in the USA and Germany. *New Media & Society*, 1, 143-160.
- Goggin, G., & McLelland, M. (2009). Internationalizing Internet studies: Beyond anglophone paradigms. In G. Goggin & M. McLelland (Eds.), *Internationalizing Internet studies: Beyond anglophone paradigms* (pp. 3-18). New York, NY: Routledge.
- Hallin, D. C., & Mancini, P. (2004). *Comparing media systems—Three models of media and politics*. Cambridge, UK: Cambridge University Press.
- Hargittai, E., Gallo, J., & Kane, M. (2008). Cross-ideological discussions among conservative and liberal bloggers. *Public Choice*, 134, 67-86.
- Herring, S. C., Kouper, I., Paolillo, J. C., Scheidt, L. A., Tyworth, M., Welsch, P., Wright, E. & Yu, N. (2005). Conversations in the blogosphere: An analysis “from the bottom up.” *Proceeding from the Thirty-Eighth Hawai'i Conference on System Sciences*. Los Alamitos: IEEE Press.
- Hindman, M. (2009). *The myth of digital democracy*. Princeton, NJ: Princeton University Press.
- Kelly, J., & Etling, B. (2008). *Mapping Iran's online public: Politics and culture in the Persian blogosphere*. Berkman Center Research Publications, no. 2008-1. Cambridge, MA: Harvard University.
- Kenix, L. J. (2009). Blogs as alternative. *Journal of Computer-Mediated Communication*, 4, 790-822.
- Kerbel, M. R. (2009). *Netroots: Online progressives and the transformation of American politics*. London, UK: Paradigm Publishers.
- Li, D., & Walejko, G. (2008). Splogs and abandoned blogs: The perils of sampling bloggers and their blogs. *Information, Communication & Society*, 2, 279-296.
- Lomborg, S. (2009). Navigating the blogosphere: Towards a genre-based typology of weblogs. *First Monday*, 5.
- Lund, J. (2010). Blogging in the face of social media. The maturing of a media group. *Digital View: Life on the Danish Internet*, February 17, 2010.
- MedieNorge/SSB. (2009). Internettbruk en gjennomsnittsdag. Retrieved March 10, 2010, from <http://www.medienorge.uib.no/?cat=statistikk&medium=it&aspekt=&queryID=315>
- Moe, H. (2010). Everyone a pamphleteer? Reconsidering comparisons of mediated public participation in the print age and the digital era. *Media, Culture & Society*, 4, 691-700.
- Morrissey, B. (2005). “One in five blogs is spam.” *AllBusiness*. Retrieved March 10, 2010, from <http://www.allbusiness.com/marketing-advertising/4160767-1.html>

- Øvrebø, O. A. (2006). Under medienes overflate. Et forskningsprosjekt om blogging og journalistikk [A Research Project on Blogging and Journalism]. Retrieved March 10, 2010, from <http://www.oov.no>
- Park, D. W. (2009). Blogging with authority: Strategic positioning in political blogs. *International Journal of Communication*, 3, 250-273.
- Park, H. W., & Kluver, R. (2009). Affiliation in political blogs in South Korea: Comparing online and offline social networks. In G. Goggin & M. McLelland (Eds.), *Internationalizing Internet studies* (pp. 252-265). New York, NY: Routledge.
- Park, H. W., & Thelwall, M. (2006). Hyperlink analyses of the World Wide Web: A review. *Journal of Computer-Mediated Communication*, 4.
- Rettberg, J. W. (2008). *Bloggning*. Cambridge, UK: Polity Press.
- Rogers, R. (2009a). *The end of the virtual. Digital methods*. Amsterdam, The Netherlands: University of Amsterdam Press.
- Rogers, R. (2009b). Mapping public web space with the IssueCrawler. In C. Brossard & B. Reber (Eds.), *Digital cognitive technologies: Epistemology and knowledge society* (pp. 115-126). London, UK: John Wiley.
- Saxvik, J. (2009). Hvor mange registrerte blogger er det i Norge? [How many registered blogs are there in Norway?]. Retrieved March 10, 2010, from <http://www.halogen.no/haloblogg/hvor-mange-registrerte-blogger-er-det-i-norge>
- SSB. (2009). IKT i husholdningene, 2. kvartal 2009 [ICT in the households, 2 quarter 2009]. Retrieved March 10, 2010 from <http://www.ssb.no/ikthus/>
- Sunstein, C. R. (2007). *Republic.com 2.0*. Princeton, NJ: Princeton University Press.
- TNS Gallup. (2010). Topplisten [The top list]. Retrieved March 10, 2010, from <http://rapp.tns-gallup.no/Default.aspx?aid=9072261>
- Tremayne, M. (Ed.). (2007). *Bloggning, citizenship and the future of media*. London, England: Routledge.
- Tveit, M. S. (2009). "300.000 bloggere i Norge" [300 000 bloggers in Norway]. *Fædrelandsvennen*. Retrieved March 10, 2010, from <http://www.fvn.no/tema/teknologi/article655750.ece>
- Wasserman, S., & Faust, K. (1994). *Social network analysis. Methods and applications*. Cambridge, UK: Cambridge University Press.
- Wei, L. (2009). Filter blogs vs. Personal journals: Understanding the knowledge production gap on the Internet. *Journal of Computer-Mediated Communication*, 4, 532-558.
- Wojcieszak, M. (2010). 'Don't talk to me': Effects of ideologically homogeneous online groups and politically dissimilar offline ties on extremism. *New Media & Society*, 4, 637-655.

Bio

Hallvard Moe (hallvard.moe@uib.no) is an associate professor at the Department of Information Science and Media Studies, University of Bergen.