

HAVE YOU REDDIT? OBAMA WON! – The 2008 US presidential election campaigns’ coverage on selected social news sites

András Szabó
University of Vaasa
Vaasa, Finland
e-mail: andras.szabo@student.uvasa.fi

ABSTRACT

The aim of this paper is to describe the role of social news websites in the political public sphere. It presents an analysis of the agenda and the sources of selected social news sites, focusing on reporting about the 2008 US presidential (pre-)election campaigns, and using the mainstream media output as a basis for comparison. The analysis uses the Habermasian model of two-track deliberative democracy as a theoretical background. Based on the analysis, this paper claims that social news websites, while promoting a democratic editorial process, are unlikely to become substitutes of the mainstream media – on which they strongly depend. However, social news websites also demonstrate considerable potential for providing citizen journalists – communicators outside the established media – with a real opportunity to compete for attention with the mainstream media.

Keywords: social news website, public sphere, 2008 US elections, citizen journalism

1. INTRODUCTION

This paper, presented at the 2009 International Conference on New Media and Information, studies the role that social news websites play in the operation of the political public sphere. It does so through the content and source analysis of such sites, in the context of the 2008 (pre-)election campaigns in the United States.

1.1. Background – the two-track Habermasian public sphere

Habermas’ public sphere theory has quite a history, full of revisions, changes and re-interpretations (cf. Habermas 1989, 1987, 2004, 2006 and others). As a starting point of my research, I rely on his latest overviews and revisions of the topic (2006), along with the seminal *Theory of Communicative Action* and the – as I understand, often overlooked – *Between Facts and Norms* volume.

The main problem for the public sphere in Habermas’ theory of deliberative democracy is that of the legitimation of power (Habermas 2005: 386). To put it in a nutshell: legitimate political decisions are ones that aim at reaching what is in the best interest of the society; however, this “best interest of the society” cannot be the sum of the interests and wants of its members, or some kind of blind compromise to them; instead, it needs to be developed on a higher level, relying and reflecting on these individual interests. (Habermas 1987, vol 2.: 77 – 82.)

The way to guarantee that such decisions will be made is to have them made in the course of an ideal procedure of deliberation. If decisions are hammered out in reasoned, open, non-forced and honest deliberation among all the participants involved, then these decisions can be considered legitimate. Importantly, this legitimizing potential of deliberation – according to Habermas – stems from the

communicative rationality that is universally inherent in our speech acts. (Habermas 2005: 385.)

The idea of universal communicative rationality has been subject to ample critique (e.g. Dahlgren 2000: 98 – 112). Claiming that people are engaging in rational discourse every single time they discuss something seems to be flying in the face of what empirically could be described as reality. If there is a possibility of ideal speech situation in practice, it hardly ever actually takes place in real life situations. Most of the time, we do not settle our disputes or converse with our friends or workmates in deliberation.

But as I understand it, Habermas is well aware of the practical, empirical limitations of his theory; in fact, I would argue that his understanding is perfectly compatible with a very realist model of what it takes to be a citizen, taking into account various psychological and even physiological factors in defining our capacities and limitations as human beings (cf. Lippmann 1993, 2004).

As he says flat out in *Between Facts and Norms*, deliberation cannot organize the totality of society: “[...] I would like to understand the procedure from which procedurally correct decisions draw their legitimacy [...] as the core structure in a separate, constitutionally organized political system, but not as a model for all social institutions [...]” (Habermas 2004: 305). Instead, he develops a “two-track” approach to society-wide deliberation – a flexible and permissive model which does *not* assume an all-encompassing authority to communicative action in the creation of a public sphere that is supposed to legitimize decisions of various institutions of the state.

In this model, a set of “wild flows of communication” emerges from “weak,” anarchic, certainly not communicatively organized (but very “human”) publics. These spontaneous, wild flows of communication shed light on what the problems are within a society; e.g. what the political system should deliberate about¹. These wild flows are then filtered by channels controlled by “informed, elite discourse” – most notably, organs of the national quality media. These filters distil what the important issues, and the main considerations about them, are. Naturally, the media also mediates information from the political system towards the public. (Habermas 2004: 304 – 308)

So the public sphere is understood as the sphere whereby this filtering – and discussion of problems in the light of the actions of the formal decision making organs of the state – takes place. “*To put it in a nutshell, the deliberative model expects the political public sphere to ensure the formation of a plurality of considered opinions*” (Habermas 2006: 416).

This two-track model of deliberative democracy is not without deficiencies. It can be pointed out that Habermas does not seem to be interested in what actually takes place inside the political decision making core of society; and the model’s unrealistic expectations towards the media system seem to severely compromise its empirical applicability.

By “unrealistic expectations” I would like to point to the following problem. The model expects two steps to be taken in political decision making. First, wild

¹ But – and this is often overlooked, although in my understanding this holds the key to understand how and why *blogs* can be important for the public sphere – this is not their only function; for even in unorganized, unstructured, haptic communication, the act of communicating something means taking one step closer to understanding it; to making it clear for *oneself*. (Cf. Habermas 2005: 389, 2006: 420.)

flows of information report about (some of the) perceived problems in the society, and the political public sphere crystallizes strands of “considered opinion,” by filtering these wild flows of communication and incorporating the contributions of the media, experts, lobbyists and activists. Second, based on these considered opinions, the actual deliberation can take place within the core of the political system.

What is missing is what happens at elections and referendums – where decisions are not made by an informed deliberative group of politicians following constitutionally established practices of ideal deliberation, but by the “weak publics” whose communicative power, however weak, is going to be transformed into the hardest of political power through voting. In these cases, the media is supposed to create and guarantee the ideal conditions of deliberation all across society.

This is a utopian expectation.

However, the model might still be empirically relevant, if we understand it as an ideal standard of measure for the situation that can empirically be observed. Hence, Habermas (2006) formulates a criterion to the media system; stipulating that in an ideal scenario it needs to be *reflective, open, undistorted* by business interests, and *independent* of the political system.

Much talk nowadays about the media, even without such a theoretical background as I outlined here, is indeed concerned with its supposedly open, reflexive, undistorted and independent nature. Community media or “grassroots,” citizen journalism, being, by definition, open, democratic and undistorted by economic or political interests, is sometimes seen as something that could restore and safeguard these qualities to the media. (Bruns 2008, Dahlgren 2005, Gillmor 2006, Sirkkunen & Kotilainen 2004.)

In this context, *social news sites* are in a peculiar position.

1.2. The concept and potential of social news sites

What I call social news sites are websites which aggregate and filter previously published online contents, according to the preferences of their respective user communities. In other words, such sites are news sites, where the community acts as both the gatherer and the editor of news. Users of the site submit material (in fact, they submit hyperlinks to various material) and vote and/or comment on other users’ submissions. Those submitted items that prove the most popular get published on a more visible part of the site – what I refer to as their front page. This editorial process is supposedly highly democratic (in that each member of the site has only one vote to cast on each submitted item).

Social news sites are “social” because they encourage and facilitate interaction between members of their communities, e.g. through on-site-messaging and the possibility to develop and maintain networks of contacts (i.e. to have “circles of friends”) on these sites themselves. In this respect, they are similar to social network sites such as Facebook (cf. Hogan 2008: 151).

These principles above are the general principles that such sites follow; but of course no two of these sites are identical – and through their peculiar features they might to some extent deviate from these guidelines in their operation. Social news

sites include Digg (www.digg.com), Reddit (www.reddit.com), Newsvine (www.newsvine.com) and Propeller (www.propeller.com), to name but a few².

A number of features suggest that such sites are in an important position in the political public sphere, understood in the context of deliberative democracy outlined above. Namely:

(a) Social news sites treat all news sources equally, as regards their presentation. When a hyperlink gets submitted to such a site, it is visually presented in a very generic fashion. An article from the *New York Times*, when submitted, might be put next to the scribbled online notes of a solitary political activist and a video clip from *The Today Show*. Ideally, the sole factor deciding whether or not an item gets popular is its quality. This principle of equality thus promotes openness, combating exclusion of particular groups from the media sphere on political or business-related grounds.

Naturally, the principle of openness does not claim that every single contribution to a social news site would be worthy of attention. And those who subscribe to a normative theory of society – such as Habermas himself – might question the capacities of a random online community to appropriately judge the social / journalistic merits of various articles. Nevertheless, the idea that on such sites articles from all sources (bar illegal ones) can be subjected to reasoned debate suggests that these sites could indeed increase the reflexivity of the media system³.

(b) Social news sites are consequential – they are examples of “strong” discussion spaces (Janssen & Kies 2005) – and I hypothesize this consequentiality to counter trends of fragmentation of discussion, often regarded as a plague of online communication (e.g. Sunstein 2002: 48-49, 54-62 and others). What I mean can best be described through a comparison of social news sites and blogs.

On any given blog, responding to an argument in a post by commenting on it entails few tangible (visible) consequences. Responding to a counterargument, the author of the original post might update it so as to incorporate competing views or to refine the original claims of the post – but this cannot be guaranteed, no matter how compelling the counterargument or how accurate the contributing remark is. The host

² I would like to stress the difference, often overlooked, between social news sites and social bookmarking sites (such as StumbleUpon or Delicious). Although these two groups of websites are similar in the technology and principles they use (whatever gets the approval of the community will be moved to a more visible area of the site), their purposes are different, for the focus of news sites is on – well: – news, while the point of using social bookmarking sites is to bookmark web contents that are judged to be important, useful or interesting sources of information that are worth to return to later. Individual news items are typically not like this – there is nothing less relevant than yesterday’s news. To phrase it yet another way: from the user’s point of view, it makes sense to submit an individual news item (such as a particular blogpost in one of Guardian’s blogs) with considerable actual news value to a social news website, where it will get voted to the front page and disappear in due course; but it makes sense to bookmark on a social bookmarking site the home page where the news item was published (in the example, the home page of the blog itself) – so that the site itself could be revisited later, even when the particular post that originally attracted the user’s attention is woefully out-of-date.

³ In the case of Newsvine, this principle does not fully stand – this particular news outlet is owned by the MSNBC news enterprise (Newsvine 2009), and follows in its architecture and layout a more traditional approach, with syndicated professional content on its front page. However, users to the site are still encouraged to contribute, both by submitting previously published articles (“seeding”) and by producing their own articles in their respective columns; and voting still affects the position of all the items that get submitted to the site.

of a blog and a visitor to a blog are not equal in their conversation (*if* there is some kind of conversation at all).

Commenters starting their *own* blogs is a logical “solution” to this problem. If there is no point in arguing in vain on a blog, one can just as easily launch one's own blog to pontificate. It is easier to establish a new place for discussion than to contribute to another from an unequal position, and with the knowledge that the best argument cannot be guaranteed to win – and even if it does, this will result in no noticeable change. Hence the connection to the fragmentation of online communication flows.

Words and actions *on social news sites*, in contrast, lead to easily perceptible changes, as the very point of these sites is to publish material according to the will of their users, expressed in votes and in comments. If you don't like an article on a social news website, you can downvote it, and explain your reasons for doing so as a comment. (Often on social news sites, comments themselves can also be rated by other users. Thereby you can expect that if you make a compelling argument, it will be rated positively, meaning it will be more visible, meaning it will be more influential than the other comments.)

In other words, contributions are rewarded on social news sites, unlike on blogs and on conventional online newspapers; thus arguing is not “in vain,” which provides an incentive to use such sites for all users, regardless of their views about a given subject matter. Even if it might be utopian to expect total meritocracy to develop on social news sites, the arrangement of these sites promotes discussion among parties who represent conflicting views, countering the trend of polarization and fragmentation of communication flows.

(c) In addition, since social news sites are free of charge, and they support their operation through advertising revenues, it is in their interest to attract the largest possible number of visitors and contributors. Meaning that instead of being partisan in their bias, arguably the best course of action they can take – from a business point of view – is to stay independent and to encourage debate in such a way that the largest possible number of parties gets involved.

In sum, an underlying assumption of my study is that social news websites encourage deliberation – through various structural features such as the generic presentation of submitted material or the possibility to introduce some kind of meritocracy for members of their user communities. If this is so, that would mean that social news sites represent some kind of *external, independent, democratically organized* and *reflexive* scrutiny on the mainstream media.

Whether or not deliberation *actually* takes places on these sites, and whether or not this deliberation has a meaningful connection to what articles end up published on the front page, are empirical questions. However, in order to assess the role of social news sites in the political public sphere, we might also approach them from the side of their contents and sources – in looking at how they actually contribute to the output of the media system, and how they affect the flow of information within it. In this paper, I try to offer some tentative answers to this second pair of questions.

I believe that **the analysis of the contents and sources of social news websites provides an important clue in understanding their role in the political public sphere in deliberative democracy.**

2. RESEARCH QUESTIONS

The distinguishing feature of social news sites – when they are viewed as media organs – is that they follow markedly different editorial methods than any given outlet of the established mainstream media. Since the editorial methods of social news sites are, by definition, democratic, and supposedly free from the systemic pressures that conventional media is subject to, one might expect that these editorial methods will result in normatively *better, higher quality* media output – at least insofar as the agenda of social news sites would be a “truer” representation of public opinion.

This is the promise. On the other hand, those who are pessimistic about the capacities of random online communities to recognize what topics are relevant might expect that the output of social news sites will be normatively *worse* than that of the mainstream media (in the case of politics, less substantial, more “tabloid-like,” more focused on hype and entertainment than on real issues).

Without dealing with the normative side of this question, it is reasonable to suppose that the agenda of social news sites, because of the underlying difference in the editorial process, will be *different* to that of the mainstream media (Xenos 2008: 490 – 491). The first research question of this paper aims at ascertaining this assumption, posing the question in the context of the 2008 US presidential elections (I explain why I chose this topic to focus my research on in the following part of this paper).

RQ1. To what extent did the agenda of selected social news sites differ from the agenda of mainstream media in the coverage of the 2008 US presidential election campaigns?

This first question is closely linked to the second one, which deals with the sources that social news sites refer to (i.e. the sources that produce the content which ends up re-published on the pages of social news sites).

One argument that questions the revolutionary potential of grassroots, citizen journalism and the “we media”-concept is tied to the resource-intensive nature of media enterprises (cf. Szabó 2008). It takes considerable human, financial, technological and informational resources to run a media organ, and especially one involved in the regular and reliable *production* of up-to-date, relevant and accurate news content. Wannabe journalists and activists are reasonably expected to lack such resources, as a result of which they face serious barriers when trying to compete with established media for attention.

However, social news sites, as mentioned before, merely aggregate and filter previously published content, and they treat this content in a uniform manner, regardless of this content having been produced by the professional media, or by citizen journalists.

Social news sites are thus expected to take advantage of a large number of varied sources, and thus it is logical to understand the number and variety of their sources as a proxy measuring their *openness* and *inclusive* nature – in other words, measuring to what extent they provide alternative sources of information, and to what extent they rely on the mainstream media in the production of their contents. Hence, the second question of this paper is:

RQ2: How many and how diverse sources did selected social news sites use – in other words, to what extent did these social news sites rely on mainstream media organs in the production of their contents – in their coverage of the 2008 US presidential election campaigns?

When answered, these questions will only provide one piece of the proverbial puzzle. In order to fully understand the role and place of social news websites in the political public sphere, they should be analysed from a number of other, different perspectives too.

Understood as media organs, the quality of their representation of news should also be analysed – i.e. whether or not they provide ideologically coherent narratives, and whether their representation of issues and persons involved is fair, objective and accurate. In addition, as I alluded to it above, social news sites could also be understood not as media outlets, but as spaces of discussion; and from this perspective, the quality of deliberation (deliberation?) should also be subject to research. Further, in order to make sense of the findings presented here, the analysis of *users* of such sites is also recommended – e.g. following the uses and gratifications theory, in order to understand how social news site users' media consumption habits influence their behaviour on such sites, or to ascertain which demographic groups use social news websites, and how this might influence the differences in the such sites' agenda⁴.

3. CONTENTS AND AGENDAS

3.1. Method and sampling

The comparison of the news output of social news sites to that of “the mainstream media (as such)” seems a daunting task. In a similar study, Xenos compared the contents of blog posts reporting about a given topic to articles published in *The New York Times*; arguing that data gathered this way was “broadly representative of discourse in [... the] traditional news media,” as the “*Times* coverage provides a useful proxy for the general tenor and content of broader national news media” (Xenos 2008: 492). Although no evidence is offered for this claim, it is true that the *Times* circulation figures (The New York Times 2009) already suggest a considerable weight in setting the media agenda in the US.

However, I chose not to extrapolate a picture drawn from the data of one publication to the whole of the mass media, but to use, as a starting point, the description of it provided by the Project for Excellence in Journalism (PEJ) research project. The project, run by the Pew Research Center – a spinoff from the Columbia University (Pew 2009a) – provides weekly content reports of the “mainstream or establishment daily news media in the United States,” including: network TV news, newspapers, online news sites (but not blogs or social news sites), cable news and radio news. The PEJ reports identify key topics for the given week, and key persons who the news have been talking about (Pew 2009b). For example, such a report might state that “the Iraq war” as a topic was featured in 40% of the newsworld (the combined and weighed output of mainstream media) during a given week.

⁴ E.g. if an item of the mainstream media agenda is conspicuously missing from that of social news sites, is that so because users are already familiar with it (through conventional media organs), and think that there is nothing new to say about it, or because they would not consider that particular topic relevant in the first place?

I focused not on the newsworld as such, rather, on the broad topic of the presidential (pre-)election campaigns. This decision was motivated by a number of reasons. First, last year's US elections provided a case where the Habermasian model could be analysed on a large scale and in great detail. It was a case of the public sphere exercising a crucially important function of supplying people with "considered opinion" based on which they, through voting, eventually provided their party (and candidates) of choice with political power. This is at the heart of the political public sphere.

Second, the elections were expected to be *considered important*, both by media organs and political activists alike. This was to lead to large quantities of media output about the elections.

Third, and in connection with the second point, it was also expected that much of this media output would be openly or covertly biased, and representing conflicting viewpoints. These conflicting viewpoints were expected to generate debate and discussion (this is a conceptual necessity of the public sphere); and in turn these discussions were expected to demonstrate how (if?) the public sphere "works" in practice. (And these discussions, naturally, should not be limited to discussions conducted between private citizens, but also between media organs, or between readers/viewers/listeners and media organs.)

And fourth, the campaigns provided a set of events that had their own agenda – and thus, the concrete, real-life events of the campaign opened up the possibility to meaningful comparison of how, and to what extent, these real-life events became part of the *media* agenda.

Hence I was drawn to sample the *political* contents of social news sites, to identify the most salient topics, and to check to what extent these corresponded to the most salient topics of the mainstream media, identified by the PEJ reports.

From February to November 2008, I was collecting data (with the help of automated downloading software) from social news websites. For reasons of feasibility, I randomly chose two one-week periods as the sampling period for this study, and I picked three social news sites to analyse. Thus the final sample for the content analysis consists of the most popular, front-page-voted, election-related articles of *Digg*, *Reddit* and *Newsvine*, in the periods of March 31st – April 6th, and October 6th – October 12th.

I included these three sites in the analysis because of the – supposedly meaningful – differences between them. While the principle of the democratic filtering of pre-existing content applies for all three sites, they are different in size, popularity, nuances of their working mechanism, and in their ties to conventional media. *Digg*, the largest (Alexa 2009a) or most popular of these sites is operated by an independent company (Digg 2009), while the somewhat smaller (Alexa 2009b) *Reddit* belongs to the publisher Condé Nast (Condé Nast 2009), and *Newsvine* to the MSNBC news network (Newsvine 2009). In further stages of the research I plan to examine these differences and their effects on the output (and discussion) going on in these sites. For this content analysis, I merely would like to point out that these differences do exist – which means that talking about all these sites together, in an undifferentiated manner, under the same general term as social news sites might be misleading.

In any case, I included in the sample those four campaign-related articles from each of these three sites, which received the most votes (or “diggs”) from the community, and so was considered to be the most popular, for every given day of the sample periods. (The sampling, based on the list of most popular articles in the past 24 hours, was carried out at midnight Eastern Time. If there were fewer than four campaign-related articles voted to the respective front pages of these sites, I naturally included fewer of them in the sample, too.) This produced the total sample of 149 items, from which 4 had to be dropped because these items were removed from the web by the time I got to archiving them.

Table 1.
Sample size for content analysis

	Digg	Reddit	Newsvine
Week 1 (March 31 st – April 6 th 2008)	26	28	14
Week 2 (October 6 th – 12 th 2008)	27	22	28
Total number of sampled items	53	50	42

What I refer to as an “item” is an “online article,” which incorporates textual articles of various sizes, photographs, pictures, videos (typically streamed from an on-demand video service such as YouTube) and audio recordings, or any given combination of these. Once again, it is important to note that whatever the actual content of these items, their presentation on the social news sites is standardized along the lines of a headline-plus-lead format.

The items in the sample were retrieved in their entirety and randomly distributed to coders. The coders were asked to decide whether the main topic of each item corresponded to any of the topic categories drawn from the weekly PEJ reports, or not. A total of 9 coders participated in the project, and each item in the sample was coded by two different coders, so as to ensure the reliability of its findings⁵.

3.2. Results...

Since the topic categories varied depending on the sample period (as, naturally, different topics featured the two different sample weeks), reliability was separately calculated for the two periods, using Krippendorff’s alpha for inter-coder agreement. For the first week (March – April), the reliability quotient was found to be 0.68, while for the October period, Krippendorff’s alpha was found to be 0.67. These are acceptable, but barely acceptable figures, warning that the findings of this study will require further analysis – however, I shall feel safe in relying on them when drawing tentative conclusions. (One suspected reason for the relatively low inter-coder agreement is the fact that the PEJ reports offered only limited information as to

⁵ Intuition, and the considerable amount of time spent browsing the social news sites in question, lead me to believe that the way each item is presented (framed) on these sites affects the way the items are perceived by their readers. This is because, upon submitting a particular piece, users are required to give it a title and to provide a lead-in paragraph, or description, to it. The title and the lead might or might not be the same as in the original publication; and this frame might well be quite different in its accuracy, bias or even content to the original article.

From this content analysis, I excluded the analysis of framing. Coders did not see how the articles were presented on the social news sites; they did not see what their new title or description sounded like. Such an analysis, probing into the effects of framing on the popularity of news items on social news sites, shall be conducted subsequently.

the definition of the topic categories themselves.) With such reserves in mind, the comparison of items lead to the following findings.

Table 2.
Comparison of main topics of news items

	Digg (n=53)	Reddit (n=50)	Newsvine (n=42)
Share of items dealing with topics salient in the mainstream media , %	58.49	46.00	59.52
Share of items dealing with other topics , %	41.51	54.00	40.48
Total	100	100	100

Topics salient in the mainstream media, as identified on the basis of the PEJ report (Pew 2009c), for week 1 of the sample period: discussion whether Hillary Clinton should quit the race for the presidential nomination; the financial status of the candidates' campaigns; celebrity and superdelegate endorsements for the candidates; Barack Obama bowling; the Rev. Jeremiah Wright and Obama's ties to him; Hillary Clinton "escaping sniper fire" in Bosnia; John McCain's past and his tour of revisiting important places in his life and career; the Clinton's tax returns; and the campaign strategy of Barack Obama.

Other salient topics of week 1 featuring the analysed social news sites: opinion polls; Obama's personality and his race as an issue; Hillary Clinton's alleged wrongdoings in her career as a lawyer; results from the Democrat's primary election in Texas; debunking a story employed by the Clinton campaign; the candidates' views on Iraq; and Obama's views on abortion.

Topics salient in the mainstream media, as identified on the basis of the PEJ report (Pew 2009d), for week 2 of the sample period: harsh attacks by the candidates (Obama's ties to William Ayers, McCain's involvement in the Keating 5 case, anger at Republican campaign events and the campaigns' other personal attacks on the rival candidate); the televised debate of McCain and Obama on October 7th: the candidates' reactions to the financial crisis and their views on the economy as an issue; voting issues (voting practicalities, irregularities tied to the voting process); opinion polls.

Other salient topics of week 2, identified on social news websites: John McCain as a person; the candidates' views on Iraq; (Republican vice-presidential candidate) Sarah Palin's alleged abuse of power as governor; Obama's race as campaign issue; and Palin's ties to the Alaskan Independence Party.

Since these numbers summed up in Table 2. are not so far from each other (SD=6.14), it can reasonably be argued that this study suggests that during the 2008 presidential election and pre-election campaigns, on average, 54.67% of the popular items on social news sites dealt with topics highly salient in the other arms of the US mainstream media, while 45.33% of the front-page worthy items addressed issues that were not salient in other arms of the US mainstream media.

3.3. ...and interpretation

A more extreme result would be easier to interpret. A total lack of similarity between the agendas would suggest that social news sites are decidedly an alternative to the mainstream media, while a near-perfect agreement between the agendas would suggest that these sites are just another channel for the mass media to disseminate

contents. Still I would venture to interpret this in-between result as suggesting two, somewhat contradictory ideas.

First, that different, supposedly democratic, meritocratic and open editorial processes of the social news sites do result in a different output of news – perhaps because they incorporate many different (and non-professional) concepts of “news.” These results point to a non-negligible potential of social news sites to deviate from the mainstream-set media agenda.

Second, however, that mainstream media organs are still very important in their agenda-setting function for social news sites. It should not be forgotten that, naturally, all media organs have their idiosyncratic agendas, and that the agenda of “the mainstream media” was conceptualized in this study as the totality of the most important common points in these various agendas – without taking into consideration the variance of topics between individual media outlets. In other words, just because a particular topic was not found salient in the mainstream media as such in a given week, this does not mean that this topic has not been dealt with in *some particular* mainstream media organs. Which, in turn, suggests that social news might have relied upon the mainstream media *more often* than in about 55% of the time.

4. THE ANALYSIS OF SOURCES

4.1. The concept of “source”

Sources in online communication have been conceptualized in numerous ways (for an excursus on the theoretical background of source conceptualization, cf. Sundar & Nass 2001). In my understanding, sources refer to various media organs: online or print versions of newspapers and magazines, TV and radio channels, online news sites, blogs and discussion forums. I use the term “source” in reference to the organ that can be seen as having *produced* the item (article) in question (which is why YouTube or similar sites are *not* considered sources in my research, as they merely provide a channel for their users to disseminate content).

I established one special type of source, namely that of “perceivably independent private individual.” The term refers to sources – most likely: private individuals – who can be seen as creators of content working outside of the institutional, organizational framework of a media organ. “Citizen journalists” would certainly fall into this category. However, I used the phrase “*perceivably* independent” as a reminder that the real identities of online actors always remain doubtful (e.g. it is certainly possible that someone acting in the guise of a citizen activist is in fact a professional journalist in the employment of a large national daily).

In my research, I distinguished between “submitters” of items to social news sites, “primary” sources and “secondary sources.” The term “submitter” denotes a person that submits an item (by providing a hyperlink to it) to a social news website. A “secondary source” is a source to which this hyperlink points (in other words, what is considered “secondary source” is the source that was submitted to the social news site). A “primary source” in my understanding is a source that is referred to (either by hyperlinks or by other, textual or visual means) by a secondary source, *and* provides significant pieces of information, which the secondary article relies on.

It is best to illustrate this through an example. User “Biff” submits to Digg an article, published in the online newspaper *The Huffington Post*. In this case, Biff would be the submitter and *The Huffington Post* the secondary source. However, it is found that the *Post’s* article consists of direct quotes from, commentary on, and hyperlinks to another article, originally published in *The New York Times’* online edition – which latter is then coded as the primary source of the article.

Thus, sources are categorized in the function of their place in the news production chain, and in this conceptualization I draw a line between the first-hand production, and the secondary treatment of news (this latter including commentary, explanation or illustration of previously published information). Primary sources are seen as producers of new content an original reporting, while secondary sources are seen as providing little new information, and contributing mostly by offering commentary, analysis or summary of material produced by primary sources.

This is not to downplay the importance of secondary sources. The secondary treatment of news items often results in articles of high quality; and surely one of the most important functions of this secondary processing of news is to put individual news items into context, offer interpretation to news, question or argue about their relevance, validity and objectivity etc. Much of this activity is in fact the very task of the political public sphere. But secondary sources, no matter how sophisticated or influential, **depend upon** primary sources. This is the key difference in their conceptualization.

One limitation of this research is that sources are only analysed in two steps. Returning to the example above, if *The New York Times* – considered being the primary source – also did not produce the article in question, but merely borrowed it from another source, this original source remains unseen in my analysis.

It follows from the concepts above that articles may have multiple primary sources. It also follows that a source in a given case might be considered both as secondary and as a primary source.

4.2. Sampling and data

In analysing the sources of social news sites, I used a broader sample than in the analysis of the agenda of these sites. While I followed the same sampling principles as in the content analysis (sampling up to four popular-voted articles from the political section of social news sites every day), the sampling period spanned two months: from the 3rd of March to the 3rd of May 2008. Table 3. offers an overview of the three sites in question, and their respective samples.

Table 3.
Overview of sampled news items and their sources

<i>Social news site</i>	<i>Number of items sampled (n)</i>	<i>Number of distinct primary sources</i>	<i>Number of distinct secondary sources</i>	<i>Avg. number of primary sources per news item</i>	<i>Avg. number of news items per (distinct) primary source</i>	<i>Avg. number of news items per (distinct) secondary source</i>
Digg	235	151	82	1.74	1.55	2.87
Reddit	212	158	98	1.97	1.34	2.16
Newsvine	141	108	60	1.91	1.31	2.35
<i>Total / avg.</i>	588	275*	165*	1.87	2.14**	3.56**

*: total number of distinct sources across the three observed websites; because of overlaps, this is a smaller figure than the sum of the number of distinct sources on each of the sites

** : total number of articles divided by the number of distinct sources registered across all three social news sites

The reliability of these findings was tested on a random subset of the whole sample, consisting of 73 items. Agreement between the original test and the re-test for the identification of secondary sources was found to be, unsurprisingly, 100%. Agreement for the identification of primary sources was found to be 82.81%. (Given the large and varying number of sources, and the fact that chance agreement between the test and the subsequent retest is bordering impossible insofar as the coding involves not only the assignment of certain values to a variable, but also the identification of the *possible* values the variable could be assigned to⁶, I felt confident in using percentage agreement as a measure of reliability.)

Table 3 testifies of a high degree of openness of the observed social news sites: the number of both the primary and the secondary sources suggests that articles that end up on the front pages of such sites come from a large pool of (secondary) sources; and that these sources refer on average to 1.87 other, primary sources. If you imagine social news sites as an analogy to conventional newspapers, these figures mean that if the paper contains 588 articles, these were written by 165 different editorial teams, who referred to 275 different primary sources). But to shade the picture, let's take a look at the most popular sources.

⁶ I.e. it is highly unlikely that a coder would make up a source and assign it randomly to an article.

Table 4.

The most frequently referenced *secondary* sources

Digg (n=235)		Reddit (n=212)		Newsvine (n=141)	
Secondary source	Share in articles (%)	Secondary source	Share in articles (%)	Secondary source	Share in articles (%)
The Huffington Post	18.30	p.i.*	5.66	p.i.*	20.57
p.i.* , The Daily Kos	5.53 each	The Huffington Post	5.19	AP	17.02
The New York Times	4.68	The Daily Kos	4.72	The Washington Post	4.26
Barack Obama's campaign**	3.83	The Seminal	4.24	American Thinker, The Chicago Tribune, The New York Times, The Philadelphia Inquirer	2.84 each
CNN, The Raw Story, The Washington Post	3.40 each	AlterNet, CNN	2.83 each	AlterNet, The Huffington Post, Yahoo! News	2.12 each

*: *perceivably independent private individual (e.g. a private blogger, Flickr or YouTube user)*
 **: *including his website and press releases directly issued by the Obama campaign*

Table 5.

The most frequently referenced (5% or larger share) *primary* sources

Digg (n=235)		Reddit (n=212)		Newsvine (n=141)	
Primary source	Share in articles (%)*	Primary source	Share in articles (%)*	Primary source	Share in articles (%)*
CNN	11.49	The New York Times, The Washington Post	11.32 each	AP	17.73
The New York Times	10.21	ABC	10.85	p.i.**	14.89
MSNBC	8.08	CNN	8.96	ABC	12.05
ABC, AP	7.66 each	AP	8.49	The New York Times	11.35
Fox News, The Huffington Post	6.38 each	MSNBC	6.13	MSNBC	7.80
Barack Obama's campaign***, The Washington Post	5.96 each	The Huffington Post	5.66	CNN	7.09
		Fox News	5.19	The Washington Post	6.38
				The Huffington Post	5.67

*: *e.g. a given source was referred to in this percentage of the articles; since articles might have multiple primary sources, these figures are not cumulative*
 **: *perceivably independent private individual*
 ***: *including his website and press releases directly issued by the Obama campaign*

These tables point to an open, rather even distribution of sources referenced on the observed social news sites, where no single dominant source emerges. However, a more influential core of sources can indeed be detected. Drawing a line between

“light” sources (referenced 1 or 2 times in the observed 2-month period), and “heavy” ones, we can see the following distribution of weight:

Table 6.
“Heavy” secondary sources in the total news output

Digg		Reddit		Newsvine	
Share of “heavy” sources in the total number of secondary sources (%); n=82	Percentage of items accounted for by “heavy” sources (%); n=235	Share of “heavy” sources in the total number of secondary sources (%); n=98	Percentage of items accounted for by “heavy” sources (%); n=212	Share of “heavy” sources in the total number of secondary sources (%); n=60	Percentage of items accounted for by “heavy” sources (%); n=141
28.05	79.71	27.55	61.32	16.67	59.57

Thus e.g. on Newsvine, 16.67% of the cited secondary sources accounted for almost 60% of the total number of popular articles, while on Reddit, about one-third of the distinct secondary sources were referenced at least 3 times (“heavy” sources), and these sources accounted for 61.32% of the site’s total output. A similar comparison could not be made using the numbers of primary sources (since a single article might refer to multiple primary sources), but so much can be established, that from among the primary sources cited, 80.79% was only referenced in 1 or 2 sampled article on Digg. The corresponding figures are 76.58% for Reddit and 78.70% for Newsvine.

4.3. Results explained

Thus there emerges an oligarchy of most important sources for social news websites, without a single, decisively dominant source, but with a more important core of “heavy” sources and a highly fragmented periphery of “light,” seldom referenced sources.

Looking at the table of most often referenced primary sources, it can be seen that **ABC, AP, CNN, The Huffington Post, MSNBC, The New York Times** and **The Washington Post** had, across the board, a defining role in making up the “power core” of primary sources for the campaign-related contents of the analysed social news websites (with **Fox News** also being highly visible on two of the three analysed websites).

With the exception of The Huffington Post, these media could all well be categorized under the “mainstream” tag, and Arianna Huffington’s online journal has also become a professional media undertaking, with a considerably large and professional team of editors and journalists. (The Huffington Post 2009.)

Thus, the analysis of sources corroborates with findings from the analysis of the contents of social news websites, as regards the importance of traditional, established, and/or mainstream media in providing the contents for social news websites.

However, we can also observe how the supposedly independent, private, activist “citizen journalists” act as important secondary, and – which is perhaps surprising – also primary sources.

The fact that P.I.s could become so influential on Newsvine might be at least partially explained by the way in which the site functions. Unlike on Digg and on

Reddit, users on Newsvine are encouraged to contribute to the site in their own respective columns. In addition, Newsvine is much smaller than both Digg and Reddit (Alexa 2009c), in terms of the number of users and visitors to the site. The community there is not only smaller, but it also closer-knit than on the other two sites; and the stronger relationship between community members is supposed to contribute to a relatively high popularity of user-created content (as “Newsvine-friends” read and vote on each other’s articles). Finally, as I mentioned before, Newsvine visually presents itself in a way more similar to conventional newspapers, and in this arrangement syndicated articles from its main sources (AP and MSNBC) are published at the start on the front page, consequently, users might perceive that there is no point in voting on these professionally written and edited articles, to the detriment of articles of fellow citizen journalists.

But the output of all three sites suggests that **independent internet users and citizen journalists** acting outside the organizational and institutional framework of any kind of media organ **do have a real and significant chance of producing material that gets voted to the front page** – in this sense, competing with the established mainstream media.

5. CONCLUSIONS

Both the analysis of the contents, and the analysis of sources of the three sampled social news sites suggest that these websites rely heavily on traditional, established media organs – and thus it is seen unlikely that social news sites could *replace* or substitute these latter. At the same time, both prongs of the analysis also indicate a non-negligible potential of these sites to channel the attention of their viewers to alternative sources – including independently acting private individuals and activists (“citizen journalists”) –, often taking up issues missing from the general agenda of mainstream media.

While on the surface these trends might appear to contradict each other, they both fit into the concept of social news sites acting *reflexively* on the mainstream media, taking advantage of its resources of news production, but monitoring, checking and overwriting its narratives when judged necessary.

As I mentioned on several occasions, this study has important limitations, mostly in that it examines social news sites from only one perspective, without taking into account the analysis of other factors such as the composition of such sites’ audiences (and their media consumption habits), the quality of discussion and argumentation taking place on these sites etc.

Nevertheless, this study suggests that social news sites could in reality fulfil a crucial role in the political public sphere, cooperating with the established media in a way that helps the creation of strands of “considered opinion,” and contributing to society-wide deliberation in scenarios when this is needed – i.e. when, according to the Habermasian model of two-track deliberative democracy, it is the media system that is supposed to guarantee the deliberative formation of public opinion across the whole of the society.

Works cited

- Alexa (2009a). "Website traffic details – digg.com", website. http://www.alexa.com/data/details/traffic_details/digg.com, cited 30.3.2009.
- Alexa (2009b). "Website traffic details – reddit.com", website. http://www.alexa.com/data/details/traffic_details/reddit.com, cited 30.3.2009.
- Alexa (2009c). "Website traffic details – newsvine.com", website. http://www.alexa.com/data/details/traffic_details/newsvine.com, cited 30.3.2009.
- Bruns, Axel (2008). "Life beyond the Public Sphere: Towards a Networked Model for Political Deliberation", *Information Polity* 13.1-2: 65 – 79.
- Condé Nast (2009). "Condé Nast Digital and Condé Nast Magazines", website. <http://www.condenet.com/index.html>, cited 30.3.2009.
- Dahlgren, Peter (2000 [1995]). *Television and the Public Sphere – Citizenship, Democracy and the Media*. London: Sage.
- Dahlgren, Peter (2005). "The Internet, Public Spheres, and Political Communication: Dispersion and Deliberation", *Political Communication* 22: 147 – 162.
- Digg (2009). "Digg – Overview", website. <http://digg.com/about/>, cited 30.3.2009.
- Gillmor, Dan (2006 [2004]). *We the Media – Grassroots Journalism By the People, For the People*. New York: O'Reilly.
- Habermas, Jürgen (1987). *The Theory of Communicative Action*. Boston: Beacon Press.
- Habermas, Jürgen (1989). *Structural Transformation of the Public Sphere*. Cambridge: Polity Press.
- Habermas, Jürgen (2004 [1996]). *Between Facts and Norms – Contributions to a Discourse Theory of Law and Democracy*. Cambridge: Polity Press.
- Habermas, Jürgen (2005). "Concluding Comments on Empirical Approaches to Deliberative Politics", *Acta Politica*, 40: 384 – 392.
- Habermas, Jürgen (2006). "Political Communication in Media Society – Does democracy still enjoy an epistemic dimension?", *Communication Theory*, 16: 411 – 426.
- Hogan, Bernie (2008). "Analyzing Social Networks via the Internet", in: Fielding, Lee & Blank (eds.): *The SAGE Handbook of Online Research Methods*. London: Sage.
- The Huffington Post (2009). "The Huffington Post – About Us", website. <http://www.huffingtonpost.com/p/huffington-post.html>, cited 30.3.2009.
- Janssen, Davy and Raphaël Kies (2005): "Online Forums and Deliberative Democracy" in *Acta Politica*, 40: 317 – 335.
- Lippmann, Walter (1993 [1927]). *The Phantom Public*. New Brunswick: Transaction.
- Lippmann, Walter (2004 [1922]). *Public Opinion*. New York: Dover.
- The New York Times (2009). "Media Kit," website. <http://tinyurl.com/cz66os>, cited 30.3.2009.
- Newsvine (2009). "About Newsvine", website. http://www.newsvine.com/_cms/info/companyinfo, cited 30.3.2009.
- Pew (2009a). "Pew Research Center's Project for Excellence in Journalism – Overview", website. http://www.journalism.org/about_pej/about_us, cited 30.3.2009.
- Pew (2009b). "Pew Research Center – The News Coverage Index Methodology", website. http://www.journalism.org/about_news_index/methodology, cited 30.3.2009.

- Pew (2009c). "Pew Research Center – PEJ Campaign Coverage Index: March 31 – April 6 2008", website. <http://www.journalism.org/node/10514>, cited 30.3.2009.
- Pew (2009d). "Pew Research Center – PEJ Campaign Coverage Index: October 6 – 12 2008", website. <http://www.journalism.org/node/13202>, cited 30.3.2009.
- Sirkkunen, Esa & Sirkku Kotilainen (eds.) (2004). *Towards Active Citizenship on the Net*. University of Tampere, Journalism Research & Development Centre.
- Sundar, S. Shyam & Clifford Nass (2001). "Conceptualizing Sources in Online News", *Journal of Communication*, 51 (1): 52 – 72.
- Sunstein, Cass R. (2002 [2001]). *Republic.com*. Princeton University Press.
- Szabó, András (2008). "A reply to "Life beyond the Public Sphere". Unpublished, available at <http://raatali.wordpress.com/papers/>
- Xenos, Michael (2008). "New Mediated Deliberation: Blog and Press Coverage of the Alito Nomination", *Journal of Computer-Mediated Communication* 13: 485 – 503.

Biographical notes:

András Szabó is a doctoral student at the University of Vaasa, Finland, researching social news websites and their connections to the Habermasian model of public sphere. His blog is at <http://raatali.wordpress.com>.