What Is News When the Audience is Editor?

Jonathan Stray

JMSC6002, HKU Fall 2009

12/12/2009

Abstract

There are now several web sites which allow users to assemble news content from around the internet by means of voting systems. The result is a new kind of front page that directly reflects what the audience believes to be salient, as opposed to what the editorial staff of a newsroom believes the audience should know. Content analyses of such sites show that they have little overlap with mainstream media agendas (5% in a previous study). In fact, many of the items selected by users are would not traditionally be considered "news" at all. This paper examines the shift from editor to audience agendas in the context of previous theories of news production, discusses existing content analysis work on the subject, and reports on a new 30 day study of Digg.com versus NYTimes.com.

Introduction

No news organization can cover everything. Traditionally, it is ultimately the editor of a news publication who decides what is newsworthy: what stories reporters will follow, and what stories will be published. It has been considered part of the value of a news organization to determine what its audiences need to know about.

It's never been entirely clear how professional journalists decide which events are worth reporting, out of all the events taking place in the world. Neither has it been obvious how editorial choices relate to the audience's personal judgments about what is important, but such questions were largely theoretical before the advent of the web. "I own a newspaper,

you do not" was always the implicit end to discussions about who got to decide what was news.

Today, publishing is near-free and the news package has been disaggregated. An online audience member can select single stories that interest them, without reading or even really being aware of the traditional news package. Alongside this disaggregation we find a new class of online applications that re-aggregate content from multiple sources. Readers vote on pages from across the web, and the top-rated items are displayed on the aggregator's home page.

News consumers are literally tearing the world's newspapers apart and re-assembling them to fit their own agendas, including lots of content not traditionally considered news at all.

This paper examines what we can learn about the online audience's judgment not only of what is important but what is news at all, and how it differs from that of traditional newsrooms. I review previous work on "news values" and "news agenda" in professional journalism, look at measurements of what audiences view online, and report on my own 30 day quantitative study of Digg as compared to the New York Times.

Features of the audience-generated agenda

Online audiences seem to be selecting for themselves a radically different set of stories and topics than that assembled for them by the mainstream media. The most relevant work on this topic is 2007 study by the Pew Research Center's Project for Excellence in Journalism [1]. The PEJ investigation followed Digg, Del.icio.us and Reddit for one week, as well as the more conventionally edited Yahoo News, collecting a total of 644 stories. It compared these to 1,395 stories from the same period published in the PEJ's News Coverage Index, a collection of mainstream news sources in print, online, network TV, cable, and radio. The report's "key findings" are worth excerpting:

- The news agenda of the three user-sites that week was markedly different from that of the mainstream press. Many of the stories users selected did not appear anywhere among the top stories in the mainstream media coverage studied.
  ...

- The sources user news sites draw on are strikingly different from the mainstream media. Seven in ten stories on the user sites come either from blogs or Web sites such as YouTube and WebMd that do not focus mostly on news.

...

- Despite claims that the Web would internationalize consumers' news diets, coverage across the three user-news sites focused more on domestic events and less on news from abroad than the mainstream media that week.

These points suggest the major themes that will recur in this paper. User news judgment is vastly different from editor news judgment. Users do not appear to care whether or not stories sources are produced in traditional journalistic fashion. And "serious" journalism (on, e.g. international topics) is unpopular.

Theories of news agenda

How do professional journalists decide what stories to follow and publish?

Schudson [2] examines research into this question beginning with the "gatekeeper" model. In this framework, those who are in a position to decide what is published control what information may pass from the world into the news. The notion is that the gatekeepers will inject their particular perspectives and biases into the news. And yet, early studies revealed that there is little variation in the wire stories chosen for publication by different local editors. Moreover, the "gatekeeper" approach doesn't answer the question of what comes to the gate and how.

> The term "gatekeeper" is still in use and provides a handy, if not altogether appropriate, metaphor for the relation of news organizations to news products. A problem with the metaphor is that it leaves "information" sociologically untouched, a pristine material that comes to the gate already prepared; the journalist as "gatekeeper" simply decides which pieces of prefabricated news will be allowed through the gate.

In pursuit of a more sophisticated model of how the news content is decided, Schudson identifies three broad schools of theory.

In political economy theories of news production, structural constraints determine what it is possible to publish, regardless of the intentions of individual journalists. Chomsky's "propaganda model" is the archetypal example. According to Chomsky the news is strictly limited to and complicit in reporting only what is favorable to maintaining the (unjust) status quo. Issues of political interest, commercial pressures and the capitalist structures of society are cogently discussed in this branch of theory.

"Organizational" theories are also structural, but don't necessarily see collusion between newsrooms and elites. Instead, the focus is on how the individual journalist within these structures ends up having little choice in how they operate. Noting that the great majority of news stories are the result of official reports from government agencies, Schudson says that the reporter sees the world as "bureaucratically organized":

> One study after another comes up with essentially the same observation, and it matters not whether the study is at the national, state, or local level -- the story of journalism, on a day-to-day basis, is the story of the interaction of reporters and officials. Some claim officials generally have the upper hand. Some media critics, including many government officials, say reporters do. But there is little doubt that the center of news generation is the link between reporter and official, the interaction of the representatives of the news bureaucracies and the government bureaucracies. This is clear especially when one examines the actual daily practices of journalists.

Constructivist theories that investigate how meaning is produced for journalists and audiences, taking into account the cultures in which they both live. It differs from the organizational perspective in that it examines the symbols and ideas available to journalists. This body of theory is the one best suited to deal with framing choices, and questions of what is and isn't surprising (and therefore newsworthy) within a particular culture.

Each of these theories provides interesting analytical tools, yet there is no overall theory that reliably answers our basic question: what makes the news? Lacking a clear explanation of this point, it's difficult to know how well the public is being served by traditional news sources.

What is newsworthy?

If we ask journalists how they decide what beats to follow, what leads to investigate, and what stories to produce, we typically get answers involving the "newsworthiness" of various events. Yet journalists are at a loss to explain what this actually means. One veteran editor described news judgment to me as "tribal," i.e. publication dependent and essentially arbitrary -- which is of course at odds with theories of "objective" reporting. Hall writes about the difficulty of defining "newsworthy" in a 1973 essay on photojournalism [3]:

> "News values" are one of the most opaque structures of meaning in modern society. All "true journalists" are supposed to possess it: few can or are willing to identify and define it. Journalists speak of "the news" as if events select themselves. Further, they speak as if which is the "most significant" news story, and which "news angles" are the most salient are divinely inspired. Yet of the millions of events which occur every day in the world, only a tiny portion ever become visible as "potential news stories"; and of this proportion, only a small fraction are actually produced as the day's news in the news media. We appear to be dealing, then, with a "deep structure" whose function as a selective device is un-transparent event to those who professionally most know how to operate it.

Perhaps our most comprehensive understanding of what "newsworthiness" actually means comes from Shoemaker [4] and colleagues. They performed very diverse studies of what people consider to newsworthy and found something surprising: the news agenda doesn't reflect *anyone's* personal judgment!

> In our study of news in ten countries, Akiba Cohen and I (2006) discovered a disconnect between what people think is newsworthy and how prominently newspapers display the stories. People in four types of focus groups -- journalists, public relations practitioners, low socio-economic audience, and high socio-economic audience -- were asked to rank ten headlines according to their newsworthiness, each set being taken from their local newspapers several months earlier. The stories ranged (in percentiles) from the most prominent as displayed in the newspaper to the least prominent.

> As expected, people within each focus group ranked the stories in much the same way, but we also found that journalists agreed with public relations practitioners, high SES audience members agreed with low SES audience members, journalists with audience members, and so on -- no matter what their station in life, people agreed on how newsworthy the events were. This was true in each of the ten countries we studied.

But when we compared the peoples' newsworthiness rankings to how prominently their local newspapers had displayed the stories, agreement was much lower. In some countries there was actually a negative relationship between how newsworthy people thought an event was and how prominently it was covered by the newspaper. In most countries, the relationship was positive, but much weaker than the relationships between the various groups of people.

The newsworthiness of an event is only one of many factors that determines how prominently the story will be covered. We cannot assume that the most prominently covered stories are the ones that people (whether editors, reporters, PR practitioners, doctors, mechanics, or teachers) think are most newsworthy, and we cannot reasonably expect people's mental judgments about what is newsworthy to correlate highly with what actually becomes the social artifact news.

In other words, it does not appear that mainstream news agenda is representative of what even journalists think is newsworthy, let alone the audience.

This finding demands an explanation. Does the news as produced not reflect individual judgment due to structural issues such as, say, political pressures? Might this be an example of groupthink, where each person produces what they imagine everyone else in their culture wants? It is also possible that the people in Shoemaker's focus groups were mis-reporting their judgments in some way, whether through selective perception and cognitive bias, Hawthorne and other experimental effects, or social pressures.

But if we assume for the moment that Shoemaker's finding is believable, we do not yet have any good explanations of why it could be that the "social artifact" of news represents the newsworthiness judgment of no one in particular. But we do get a prediction: given a free choice, audiences would construct a news agenda that is dramatically different from the mainstream status quo.

What are Audiences Actually Reading?

Consumption is one form of audience judgment. Newspaper publishers have long known or assumed that their readers don't have the same priorities as the newsroom. Stereotypically, it is sports and celebrity gossip that draw the most readers. This wasn't necessarily an economic concern in the newsprint era given that the customer could not buy less than a

whole paper. In any case, reader story choice was anyway not immediately measureable. In some studies readers have been asked what they read, but this technique is subject to deep problems related to memory and perception. In another design, readers have been directly observed as they read the paper, but this is an artificial situation and subject to well-known distortions such the Hawthorne effect (where the subject tries to please the researcher).

In contrast, online measurements can be completely unobtrusive. In fact the raw data is already routinely recorded in web server logs, and collected by companies such as Nielsen Ratings. Tewksbury [5] analyzed visits to 13 pre-selected news sites by 9,209 randomly-selected Americans in the months of March and May 2000. From [5] table 2:

Story Category	% views
Sports	26.0
Business and money	13.4
Arts and entertainment	10.9
Features	10.7
U.S. national	10.2
Technology and science	7.0
Interactive elements	7.7
World	6.1
Politics	5.4
Weather	3.6
Health	1.5
Opinion and editorial	1.4
State and local	1.2
Obituary	.1
Other news	2.5
Advertising, index page, etc.	19.1

True to stereotype, sports lead with 26% of views, followed by business, entertainment, features, and national news. The political and international news considered so important by professional journalists together comprise just 11.5% of total page views.

Academic work on news agenda has taken a similarly narrow focus, in that it has not really come to grips with the implications of a readership who doesn't much care for the news that journalists think is important. Media effects researchers have for decades used studies where subjects are asked "what is the most important problem facing the nation?" [6] Within this research paradigm, by definition only "problems facing the nation" can generate news. Where does sports and lifestyle reporting fit into this? Even before the era of user news sites, it seems that journalists and scholars alike had a skewed conception of how their audiences interacted with the news media.

The Audience Agenda-Generation Process

Audience-driven content aggregation sites such as Digg, Reddit, etc. all work on similar principles. Users can submit arbitrary URLs into the system, either directly on the aggregation site or via submission buttons that content publishers make available to readers on their site in the hope that their content will be promoted, e.g. "Digg this" buttons. Similar voting buttons are provided on the aggregation site for each item displayed. Votes collected from all locations are tallied for each item, and the ranked results constitute the user-generated "front page."

The resulting rankings are not equivalent to a poll of readers. To begin with, votes must be counted in a time-limited way, or such sites would rank the most popular content of all time, as opposed to popular *recent* content. The number of votes for an item also depends greatly on the number of people exposed to that item, and this has much to do with factors that influence the extent of "viral" transmission of an item through social media, including emotional response (see e.g. [7]) and the social network topology around the people who have an interest in the topic. Further, readers voting on the site are far more likely to vote for items that are more prominently displayed -- and thus already more popular.

Nonetheless, this voting process produces some sort of snapshot of aggregate audience interest. It's a relatively opaque sort of snapshot, and doesn't clearly represent anything in particular. It will favor already popular items and items with emotional content -- but so does pop culture. It's not obvious that this user-generated news agenda is any "better" or "worse" than the agenda of a newsroom.

What we can say is that type of audience-generated agenda clearly draws on a wider array of sources than a traditional news publication. Any web page can be voted upon, not just content from "news" sites. Videos are relevant, as are blog postings on arbitrary topics. Crucially, the audience agenda generation process seems to involve only very weak preconceptions about what is potentially newsworthy -- "anything on the web" -- as compared to journalists and academics. Because users vote individually, mostly in private, not for pay, and effectively anonymously, we might also expect user agendas to be free from the structural and sociological constraints acting on a newsroom. Is it possible that user-generated agenda are simply a more honest reflection of what all of us actually consider newsworthy?

Audience-Generated Agendas in Detail

The PEJ study [1] examined user-edited sites in several ways. One of the most revealing is the (lack of) overlap between these sites and the stories in the Pew's news coverage index of mainstream media outlets.

> In the user-generated sites, [mainstream media] stories were barely visible. Overall, just 5% of the stories captured on these three sites overlapped with the ten most widely-covered stories in the Index (13% for Reddit, 4% for Digg, and 0% for Del.icio.us).

Again we see that there is very little overlap between what the mainstream media considers "news", and the stories that users choose for themselves. Even accounting for the demographic skew of these sites -- which are arguably still "early adopter" and over-represent tech geeks -- the lack of agreement is astounding.

They further examine the content in terms of the top five subject categories for each site.

News Index (Mainstream Media)

Topic	Story %
International (non-US)	15
Disasters/accidents	11
US Foreign Affairs	10

Immigration	8
Government	7

Digg

Topic	Story %
Technology and science	40
Lifestyle	11
International (non-US)	6
Business	6
Government	6
Celebrities	6

Reddit

Topic	Story %
Technology and science	22
Lifestyle	15
Government	13
International (non-US)	7
Crime	6

Del.icio.us

Topic	Story %
Technology and science	41
Lifestyle	20
International (non-US)	16
Business	6
Miscellaneous	6

The difference of focus between editor- and audience-generated agenda stands out here. The audience sites had heavy coverage of technology and science issues, which again suggests demographic differences. Lifestyle was also much more popular. As in the online readership survey, what a traditional journalist would call "hard news" barely registers.

My own study of user vs. mainstream media agenda was a 30 day comparison of Digg and the New York Times, from 4 October 2009 to 4 November 2009. Each evening shortly after midnight, I took a screen capture of the front pages of both sites. With the browser window set to a full screen height of 1024 pixels, and ignoring the smallest sized links on the NYT page, both screen shots averaged about nine stories per day. I categorized the stories on each site and also tracked, for each day, how many stories appeared on both sites. In total, I collected 238 stories from Digg and 227 stories from the New York Times this way.

Rather than the default Digg page as used in the PEJ study, I used the "24 hour news" view in order to capture a list that was a little more comparable in time scale to a daily newspaper, and contained the "news" label which may tell content voters that "newsworthiness" is being asked for.

Results by top 5 story categories

Story Category -- Digg

Topic	Story %
Politics and international	31
Technology and science	27
Lifestyle	13
Arts and entertainment	10
Miscellaneous	8

Story Category -- NYT

Topic	Story %
Politics and international	51
Lifestyle	12
Business	11

Arts and entertainment	10
Sports	9

In these tables, audience and editor agendas don't seem that different. But consider this:

Digg's overlap with...	Story %
All mainstream media	29
The New York Times	3

Although the selection of topics was similar, the actual *stories* almost never overlapped between Digg and the NYT. Only a few very prominent stories were covered in both agendas, such as Obama winning the Nobel Peace Prize. In total, stories in the Digg "news" category came from mainstream media reports only 29% of the time. This is high compared to the PEJ study which averaged 5%. The difference may be attributable to my choice of the "news" category on Digg. The fact that this category exists, and the why and how of different user behavior when using this category, is an area ripe for future research.

Conclusion

Lacking a clear understanding of how a traditional newsroom selects its stories, we can say very little about how audiences might theoretically choose differently. Even journalists cannot explain how story selection really works. However, both online news readership surveys and user-aggregated news sites show very different agendas than the mainstream media. This is most visible when one looks at the overlap between audience and editor-generated agendas -- 5% in the PEJ study and 29% for the Digg "news" category.

Is this because the media never really represented public tastes for information, due to structural constraints or value differences between journalists and their audiences? Structural and organizational reasons seem a more promising explanation then personal judgments, given that Shoemaker's surveys suggest that not even journalists agree with the agendas of their publications.

But we must also contend with the deeper problem of "what is news?" What can thinkably be on a news agenda? In traditional media effects research, only "problems facing the nation" can be on the agenda. On the web, the answer is most naturally "any web page." Most of the web is nothing like news from a traditional journalism perspective, but this doesn't stop audiences from voting it on to the agenda.

And this paper has only scratched the surface. We've discussed only the web, not real-time messaging services such as Twitter and Facebook, and this leads us to a key underlying assumption of all of the work discussed in this paper: the "news" is the same for every member of the audience. This is a very "broadcast" mentality, and the web is not a broadcast medium. In the future, news will be personalized and it will also be personal: the machinations of my social network may not be newsworthy to a journalist, but it's certainly newsworthy to me. The entire concept of "news" is undergoing a transformation.

So what is news in the age of the audience-as-editor? In 2007 entrepreneur Adrian Holovaty founded a site called EveryBlock.com that aggregates local police reports, blog posts, and other web content with determinable location. Users from all over the US can see what is happening literally in their neighborhood. It is a valuable source of news, yet EveryBlock employs no "journalists". The success of this product sparked a lively debate around the question "is data journalism?" Holovaty's answer was [8],

> I no longer see the point in debating the definition of journalism. I'm interested in building products that improve people's lives via information. Whether somebody calls that "journalism" is utterly uninteresting.

References

1. "The Latest News Headlines -- Your Vote Counts." Project for Excellence in Journalism, September 12, 2007. http://www.journalism.org/node/7493

2. Michael Schudson, "The Sociology of News Production", *Media Culture Society*, 1989; 11; 263.

3. Stuart Hall, "The Determination of News Photographs", pp 176-190 in S. Cohen and J. Young (eds), *The Manufacture of News: A Reader*, Beverly Hills: Sage, 1973

4. Pamela Shoemaker, "News and Newsworthiness: A Commentary." *Communications*, 31(1):105-111, 2006

5. David Tewksbury, "What do Americans Really Want to Know?" *Journal of Communication Research*, December 2003

6. Maxwell McCombs, "Building Consensus: The News Media's Agenda Setting Roles." *Political Communication*, 14:433-443, 1997.

7. Jonah Berger and Katy Milkman, "Social Transmission and Viral Culture." Wharton School, University of Pennsylvania, 2009. http://marketing.wharton.upenn.edu/documents/research/Virality.pdf

8. http://www.holovaty.com/writing/data-is-journalism/