The Effects of Aggregation on the Economics and Perception of News

Adam Wagner
Ohio University
E. W. Scripps School of Journalism
Journalism 416: Online Journalism Seminar
Contact information: aw333507@ohio.edu
Abstract

Aggregation is an increasingly popular way of people viewing their news, as a significant portion of Internet users utilize services such as Google News or Yahoo to find information or Web sites that they are searching for. The popularity of these sources has caused questions about many issues, including fair use, the economics of the media on the Internet, and the changes to the kind of news people frequent, to percolate. Generally, users of aggregators care about advanced search functionalities, a user-friendly interface, and quality and reputation of sources.
Introduction

Aggregators, or mass collectors of information, have existed as long as computers have, meaning since around the early 1970s. According to Chowdhury and Landoni, “A news aggregator collects news information, and often summarizes it in a pre-defined or user-defined format, from a variety of channels” (Chowdhury & Landoni, 2006). Since that time, three aggregators—Dialog, Factiva, and LexisNexis—have been the world’s preeminent information collectors, charging monthly fees for their content and holding contracts with the companies whose material they use. That changed, however, with the invention of the Internet, as the information that these services were charging for was suddenly available en masse and for little to no money.

Perhaps the best way to think about the relationship between aggregators and traditional media outlets is to think of the media outlets as department stores. They have high quality goods, they have a standard array of goods, the visitor can expect to spend a fair amount of time there, and it is beginning to die off. An article in the Economist said, “the fate of the news package is similar to that of the department store. Some customers have been lured away by discount chains; others have been drawn to boutiques” (“Tossed by a gale,” 2005). Aggregators are the news equivalent of Walmart. They have a wide variety of goods, they are massive, they depend on other sources for cheap content, and they are thriving.

Despite all of that, the threat from aggregators was only truly legitimized, however, in 2002, when search engine giant Google created Google News, a service that combs the Internet for information and posts it in blurb form for users to view, linking it to the original place of publication. This allows people to choose the stories that they want to read and ignore the stories they don’t care about. Furthermore, Google’s status as a search engine allows people to do what
aggregators have spent over thirty years charging for little or no fee, namely type in a term and see a collection of all of the most recent information about that issue.

This threat has caused many issues not only amongst pay-for aggregation services, but, more importantly, amongst the mainstream media. Aggregators have brought about issues of fair use, the economics of paper in an increasingly online world, and the changing nature of information. By forcing entrenched news outlets to face these questions, aggregators are playing a vital part in the continuing business and practical evolution of the news media, aiding in the simplification and mass dissemination of information that seems to be the general direction that news is heading towards.

In a testimony to Congress about the future of journalism, Arianna Huffington, the owner of The Huffington Post, said, “Can anyone seriously argue that this isn’t a magnificent time for readers who can surf the net, use search engines, and go to news aggregators to access the best stories from countless sources around the world—stories that are up to the minute, not rolled out once a day?” (Huffington, 2009).

**Literature Review**

*Aggregators’ use of other source’s material as it pertains to fair use issues and the media’s reaction to it*

Traditional news outlets feel threatened by aggregators, believing that the typical format of posting a headline and the first line or two of a story infringes on their copyright. This issue has recently been brought to both the court of law and the court of public opinion, as there have been lawsuits against aggregators and prominent media figures coming out against the search engines.
In 2006, Copiepresse, one of Belgium’s largest newspaper companies, sued Google in Belgium, claiming that Google’s practice of putting the headline of a story and a few sentences of it online violated copyright laws and the database rights of newspapers. Copiepresse also claimed that users could access content that was no longer meant to be free through a cache located on Google’s servers. The Belgian court found in favor of Copiepresse, ordering the search engine to remove all content of Copiepresse newspapers from its Web site, with a fine of €1,000,000 each day Google left the content up (Laurent, 2007). The court also ordered Google to post the Order of the Belgian version of its site.

Despite losing that case, however, Google still seems fairly safe on American soil, as “fair dealing,” the European counterpart to “fair use,” is much more relaxed (Hogge, 2006). In 2005, Agence France-Presse sued Google for $17.5 million on in America, a case that was eventually settled by a licensing agreement between the two companies.

Part of the reason for Google’s success on American soil is that four specific factors are taken into consideration when questions of fair use arise (“Measuring Fair Use,” n.d.). The first is the transformative factor, whether the work has been used to create something new or whether it has simply been copied. If it has been used to create something new, then fair use has not been violated. Google can’t claim this defense, however, because they are undoubtedly just copying the information. The second step is the nature of the copyrighted work. According to Stanford University, “Because the dissemination of facts or information benefits the public, you have more leeway to copy from factual works such as biographies than you do from fictional works such as plays or novels” (“Measuring Fair Use,” n.d.). Google could claim this defense, as they are simply helping relay information to a wider audience than it would otherwise receive. The third step is the amount and substantiality of the portion taken, which, in Google’s case, is
always very small. The search engine takes little more than the headline and a sentence or two of text—just enough to give the reader a vague idea of what the story is about, but not enough to ruin the story for him. The final factor taken into consideration, and perhaps the most important for Google, is the effect of the use upon the potential market. Basically, if the way someone uses an article prevents the creator from gaining increased profits from it, fair use has been violated. Google, however, could argue that it is actually helping bring readers to the Web site, thereby helping increase the profit and helping the organization. By fulfilling three of the four factors of fair use, Google has established that it will be very safe in the event of another fair use lawsuit.

These legal protections have not prevented media magnates from being very outspoken about their feelings towards Google, though. Rupert Murdoch, owner of media conglomerate the News Corporation, which owns Fox News and the *Wall Street Journal* used the Cable Show, an event in Washington, D.C., to say, “Should we be allowing Google to steal all of our copyrights? The answer should be thanks, but no thanks” (Smillie, 2009). The News Corporation’s anger at Google continued, as Robert Thomson, the managing editor of the *Wall Street Journal*, lashed out at Google, saying, “The whole Google sensibility is inimical to traditional brand loyalty. Google encourages promiscuity . . . and therefore a significant proportion of their users don’t necessarily associate that content with the creator. Therefore revenue that should be associated with the creator is not garnered” (Schulze, 2009). These news organizations are primarily worried about how Google is affecting their profit margins, although some members of the media see Google as a useful tool. Jim Brady, the executive editor of the *Washington Post*, said, “Clearly, some people use Google News as a primary news source, which makes them a competitor. But they are a driver of a significant amount of traffic, which we appreciate” (Helft,
2008). Brady raises one of the other major issues that established news outlets have with Google and other aggregators—that they are intrinsically and irrevocably changing the face of news.

The Changing Face of the News and the Role That Aggregators Are Playing

Part of the reason that the established media outlets are concerned about aggregators is that some users are beginning to prefer the headline and two to three sentence snippet that aggregators offer to the full version that newspapers and other media organizations offer. According to the New York Times’ Miguel Helft, they are just as concerned that when users do click a link, “instead of entering a preferred news site through its front page, users are being routed to a single article, perhaps deep inside the site, and when they are done reading it, they move on” (Helft, 2008).

These habits seem to be obeying trends across media, however, as readers are beginning to pursue the information that they want, regardless of the source. Furthermore, experts on online media believe that shutting out aggregators is actually damaging to a site as a whole. Nick Denton, the CEO of news aggregator Moreover, said, “Criticism [of aggregators] is typical of applying traditional media to the Internet, which is an entirely different model. It’s the walled garden sites that tend to do badly” (Martin, 2001). Part of the reason for this is that information has become an entity that people expect to be free and, furthermore, they do not care where it comes from. “To today’s Internet user, all content is equal,” (Plosker, 2004) and that sentiment has only become more common since that article was printed.

Another role that aggregators play is bringing content together in a way that is a necessity in the modern world. The Internet allows the average person to have access to more news today than at any other point in human history, leading to a seemingly constant state of overload. By bringing information together in shortened form, however, aggregators allow people to scan
news for stories that they are particularly interested while also gaining some understanding of what is happening in the world. Today’s free aggregators accomplish this better than any other source, as a case study from January, 2006, to August, 2007, and published in the *Journalism & Mass Communication Quarterly* demonstrated that Google News averaged roughly 31 stories a week on the topic in comparison to LexisNexis’ seven (Chowdhury & Landoni, 2006). When duplicate stories were removed from the occasion, as Google takes everything from its sources and doesn’t check to see whether it already has a story or not, Google still averaged 20 stories a week.

Google and its fellow aggregators cast a wider net over information than ever before, but do so in a way that allows the typical person to stay informed without being overwhelmed, shattering long-standing brand loyalties, and therefore annoying established media outlets, along the way. Despite being annoyed by aggregators, however, the media establishment could learn about its intended audience and how news should be presented by looking at how aggregators work.

*What Makes a Good Aggregator Successful and What the Established Media Can Learn From Them*

In a media climate where 29 percent of Internet users, or 43 million people, use the Web solely to check the news three more times a week, aggregators play the vital role of boiling information down to the bare minimum and making it easy to find (Dayton, 2005). Some aggregators, however, handle the business of gathering information better than others.

Google News, for instance, supposedly draws from over 4,500 sources to find the news that it puts on its site. It does not, however, publish the sources it uses anywhere, thereby creating a feeling of secrecy and discomfort amongst the rest of the media, who have called for
Google to become more transparent (Wiggins, 2006). Other ways that Google could improve include listing the criteria for a sources inclusion, noting when sources are added or removed from the list, and giving users the option to filter certain sources.

Another issue with Google is that it relies solely on its algorithm to decide which stories are printed, claiming that it is trying to avoid bias and print as many points of as possible. This algorithm, however, has proven to ignore breaking news sometimes, as was the case when Tim Russert died and the story didn’t reach Google’s front page for over an hour even though it was on MSNBC’s and Yahoo’s front pages (Helft, 2008). The algorithm works well most of the time, but Google should retain an editorial team or partner with a wire service to better handle breaking news.

Perhaps most disconcerting is the fact that Google can be tricked. Google has had to pull a story after finding out that it was a fraud several times throughout its short history. Richard Wiggins wrote about how he used a PR firm to trick Google into running a press release that he wrote about himself very high under any result for “Tim Berners-Lee,” the man widely recognized as the inventor of the Internet (Wiggins, 2006). “Any news aggregator will always carry content that is controversial; sometimes it will carry plain propaganda. One can not expect Google News to offer ‘the truth’ or to satisfy all audiences. But one should expect Google News to represent what say media say—not “just random text that anyone types into a faux PR firms’ submission window,” (Wiggins, 2006). In order to fix this, Google could take the same approach that aggregators such as Totalsnews, Infoseek, and Northern Light have taken and cross-source news before printing it, ensuring that a story is carried by at least two reputable sources and is not completely made up, like Wiggins.
One of Google’s major strengths is that because of its status as a search engine and media giant, it has access to massive libraries of both images and video, allowing it to constantly provide compelling multimedia. Yahoo and MSNBC have taken similar approaches, offering short video clips as well as printed stories. Aggregators such as LexisNexis, however, have been unable to take advantage of this opportunity yet, while traditional media outlets are just becoming familiar with it.

In a 2006 study, two professors asked 45 people what they valued most in an aggregator. The most common answer was advanced search functionalities, while a user-friendly interface and high quality reputable sources also ranked very highly (Chowdhury & Landoni, 2006). The two authors went on to list the seven most important factors of a news aggregator, which are advanced search functionalities, user-friendly interface, high quality and reputable sources, facilities for searching back in time for related information items, browsing functionalities, showing information items in chronological order, and personalization (Chowdhury & Landoni, 2006).

Google News was found to have the best ability to retrieve information out of the five aggregators studied, with the authors concluding that the search service is, “. . . quite effective and efficient, with some improvements needed on the definition of categories . . .” (Chowdhury & Landoni, 2006). Since the study was printed, Google has better defined its categories and continued to improve the personalization and news alerts that the authors of the study found especially useful.

Google and other aggregators will continue to improve and innovate, even as traditional media outlets gradually begin to gain a better handle on the Internet and how to best utilize its varied resources.
Future Plans

How Aggregators and News Outlets Can Benefit From Each Other

The relationship between aggregators and various news outlets does not need to be nearly as hostile as it has become in recent months. Unfortunately for the traditional media, aggregators have been able to better exploit the short attention spans of the modern person, relying on short blurbs and headlines instead of full-length stories, infuriating the media who are treating most information the same way they did 15 years ago. When Thompson made his comments about aggregators, Michael Wolff, the founder of Newser, wrote:

The value of information is going down because technology has made it so much more plentiful; the Wall Street Journal no longer has a monopoly on data, access, diligence, research, understanding, consistency, or reliability. The WSJ and the New York Times and every other newspaper can fold and we here at Newser will still be sorting through vast quantities of news, opinion, and analysis. The mountain of information only gets bigger (Wolff, 2009).

This “mountain of information” means that existing media organizations must begin to adapt to the times in order to survive.

The Associated Press has begun to take just such actions, investigating working with other content creators to create a new aggregator, trying to “build an online destination where it hopes Web users can easily find and read its news stories and those of other content creators” (MacMillan, 2009). This site would be much different than traditional aggregators, relying on “landing pages” to host articles. All of the participating news outlets would share the revenue from these sources, which would include both local and national media outlets. Art Howe, a Pulitzer-winning journalist at the Philadelphia Inquirer, said, "What the AP is doing is trying to create a mechanism to reestablish the balance of power between search engine companies and content creators. They are taking the first steps to creating a new model that benefits the publishers — the people who are paying writers and editors and photographers” (MacMillan,
2009). Media experts are skeptical of whether such a plan will work, however, as 19 percent of the users going to news sites in March, 2009, went through Google, while 13 percent used Yahoo as a portal to news (MacMillan, 2009).

Jeff Jarvis, director of the interactive journalism program at the City University of New York's Graduate School of Journalism, said that instead of battling Google, news organizations should “try to make Google News better” (MacMillan, 2009). There are several ways that this could be accomplished, one of which would be utilizing the AP’s ideas about local ideas, of which AP Chief Executive Tim Curley said, “There are some pretty exciting ways to showcase some of the core features of local newspapers” (MacMillan 2009). One of these ways could be to create an aggregated page for each individual paper and make it easily accessible on the paper’s site, using the Google algorithm but only one paper’s appearance and content. If proven effective, Google could provide the papers a small fee, thereby gaining a large profit due to the sheer number of papers that they would be working with, and the papers would gain advertising revenue. Furthermore, this hyper-local aggregated page would appeal to both the online generation, which likes to have all of its news in one place, and the newspaper generation, which has strong loyalties to a particular paper or place.

The media is at an awkward place in its history, lying on a nexus between the past and the present and not yet understanding where it is all headed. This awkwardness, however, is a great opportunity to change the face of news for the better, finding a way to allow people who have never had access to news before to gain it and allowing those who did to peruse more information than ever before.

**Conclusion**
Aggregation represents a new way of finding and accessing news, and only by recognizing that it is likely one of the next frontiers and adapting to change instead of resisting it will traditional media outlets be able to benefit from it the most. Aggregators are dependent on news outlets for content, and news outlets are reliant on aggregators for sources, so the best result for both parties would be to work together to help shape the economic and practical world of journalism in the near future.

By working with large companies like Google and Yahoo and having access to their resources, small organizations would have the ability to enhance their content with much more multimedia content, as well as being able to link and be linked to other news organizations. By taking these actions, the media outlets could support themselves, in a way creating their own aggregator with the established company at its center.

No matter what happens next, though, news has been permanently altered by the Internet, and learning how to most effectively to adapt to technology while maintaining high profit margins and quality content is the next great frontier of the media, whether aggregation plays a part in that or not.
References


Effects of Aggregation


