

**“What Aggregators Do: Rhetoric, Practice, and Cultures of Digital and Analog Evidence in Web-Era Journalism.”**

**C.W. Anderson, College of Staten Island (CUNY)**

[cwa2103@columbia.edu](mailto:cwa2103@columbia.edu)

Paper Submitted to the 12th International Symposium for Online Journalism  
Austin, Texas | April 2-3 2010

## Abstract

This paper analyzes an increasingly valorized form of newswriting-- “serious, old fashioned,” “boots on the ground reporting”-- through an exploration of its purported occupational opposite, news aggregation. The paper begins with a qualitative content analysis of the March 4, 2010 FCC workshop “The Future of Media and Information Needs of Communities: Serving the Public Interest in the Digital Era,” in which journalists and scholars, using public rhetoric, attempted to draw a sharp, clear boundary between original reporting and aggregation. The paper turns, in its second section, to an exploration of the actual hybridized practices of journalistic aggregation. Connecting these threads is an argument, drawing from research in science and technology studies, that knowledge claims must be examined *both* as pure, line drawing arguments *and* as messy hybrids in which mangled practices need to be constantly purified through the aforementioned rhetorical work. The paper concludes by pointing towards an under-theorized aspect of this process of occupational differentiation-- the role played by actual things (digital objects like hyperlinks, databases, and algorithms) —in the establishment of the boundaries that make up our socio-technical world. It is this analysis of the connection between fields and things that might provide the most fruitful avenue for journalism research in the digital age.

## Introduction: Aggregate This!

The so-called “battle between bloggers and journalists” continues to rage, long past its expiration date, persisting in a journalistic world of increasing occupational overlap and hybrid work practices. In his recent excavation of the psychological roots of the debate, NYU media scholar Jay Rosen called the argument a “psychological thing,” and argued: “there’s something about bloggers versus journalists that permits the display of a preferred (or idealized) self among people in the press whose work lives have been disrupted by the Internet ... Spitting at bloggers,” he concludes, “is closely related to gazing at your own reflection, and falling in love with it all over again” (Rosen 2011). For Rosen, the rhetoric of bloggers versus journalists is a pathology in which self-mythologized communication producers denigrate their “evil opposite” via the creation of their ideal other.

Rosen thus follows in a long line of scholars (Anderson 2009, Carlson 2007, Carey 1997, Lowrey 2006, Singer 2003, Zelizer 1992) who have argued that the tensions between different occupational categories of media producers can be found in the nexus of jurisdictional struggles and rhetorical claims to professional expertise. Drawing on extensive newsroom fieldwork, semi-

structured interviews, and content analysis, this paper argues that the primary “jurisdictional conflict” in journalism today actually lies between *reporting and aggregation* rather than blogging and journalism. The roots of the conflicts animating today’s journalism can this best be understood by examining different visions of what constitutes meaningful empirical evidence in the digital age, analyzing the tensions that surround the definition of “proper” newswork, and even by thinking philosophically about the nature of digital reality itself. The meaningful arguments in journalism today, in short, must be approached as questions of *ontology, not epistemology*.

A recent article by *New York Times* editor Bill Keller perfectly captures the crux of this debate about original reporting and news aggregation. In an opinion piece called “Aggregate This,” Keller made it clear that while “aggregation” can mean smart people sharing their reading lists, plugging one another into the bounty of the information universe ... [it] too often it amounts to taking words written by other people, packaging them on your own Web site and harvesting revenue that might otherwise be directed to the originators of the material” (Keller 2011a, np). Amongst the primary culprits in Keller’s world of barely disguised thievery was the Huffington Post and it’s CEO, Arianna Huffington. In a follow-up column, Keller was even more blunt: “aggregating the work of others is no substitute for boots-on-the-ground journalism” (Keller 2011b, np). In her reply to Keller, Huffington chose not to directly address questions of theft, preferring to emphasize her traditionalist bona fides. “Even before we merged with AOL,” Huffington argued, “HuffPost had 148 full-time editors, writers, and reporters engaged in the serious, old-fashioned work of traditional journalism” (Huffington 2011, np).

This paper analyzes this increasingly rare (and increasingly valorized) “serious, old fashioned,” “boots on the ground journalism” through an exploration of its purported occupational opposite – news aggregation. The paper begins with a qualitative content analysis of the March 4, 2010 FCC workshop<sup>1</sup> “The Future of Media and Information Needs of Communities: Serving the

---

<sup>1</sup> A note on public hearings: the public hearing transcripts discussing the “future of journalism,” primarily carried out in Washington DC in 2009 and 2010 by the Federal Trade Commission (FTC) and the Federal Communications Commission (FCC) have been chosen as the primary corpus through which professional boundary drawing battles will be analyzed. Obviously, many other corpora on this topic are possible: twitter discussions, blog postings, industry trade journals, conference proceedings, journalism school textbooks, and so forth. Public hearings, however, are unique insofar as they mark a primary moment in the U.S. media

Public Interest in the Digital Era,” in which public rhetoric attempted to draw a sharp, clear boundary between original reporting and aggregation. The paper turns, in its second section, to an exploration the actual hybridized practices of journalistic aggregation. Connecting these threads is an argument, drawing from research in science and technology studies, that knowledge claims must be examined *both* as pure, line drawing arguments *and* as messy hybrids in which mangled practices need to be constantly purified through the aforementioned rhetorical work. The paper concludes by pointing towards an under-theorized aspect of this process of occupational differentiation-- the role played by actual things (digital objects like hyperlinks, databases, and algorithms) —in the establishment of the boundaries that make up our socio-technical world. It is this analysis of the connection between fields and things that provides, I argue, the most fruitful avenue for research on journalism in the digital age.

### **Public Rhetoric and the Contest Over Journalistic Jurisdiction**

Journalism scholar Wilson Lowrey has advanced a useful intellectual framework by which to examine the conflicts over the boundaries of 21st century journalism. Drawing on Andrew Abbott’s ecological approach outlined *The System of the Professions*, Lowrey argues that journalism should be seen as existing within “an interrelated system and as compet[ing] for jurisdiction over ‘work tasks’, or ‘human problems amenable to expert service.’” (Lowrey 2006, 480). Journalists, Lowrey notes, are engaged in a “battle” for control over their work practices, the positive rhetorical claims they make on behalf of those practices, and the jurisdiction created by this combination of occupational practice and expert claim. Other groups, Lowrey argues, can seize neglected or vulnerable areas of journalism’s professional jurisdiction by claiming it as their own, or by claiming that the positive public outcomes pointed to by original jurisdiction occupants no longer hold within a new technological, social, economic, or political environment. Such

---

system when the values of journalism are specifically addressed by formal bodies who have the power to (occasionally) set practice-defining policy. It might be inferred, then, that it is in these public policy setting that rhetorical arguments will be the most sharply defined. It is against this background that the historical analyses of the 1934 Radio Act (McChesney 1995) the Hutchins Commission (Pickard forthcoming), and the 1927 Radio Act (Sylvain 2010) have been carried out.

structural changes can also open up new jurisdictional spaces to be seized. As Anderson and Schudson put it, also drawing on Abbott:

“Expert” professionals—in this case, journalists—seek, via occupational struggle, to monopolize a form of journalistic expertise, which itself is discursively constructed out of various journalistic practices and narratives, including the claim to professional objectivity (Schudson and Anderson 2009, 96).

Reporters and aggregators, under this model, can thus be expected to battle over journalistic jurisdiction (Gieryn 1983, Zelizer 1992) by making particular rhetorical claims about what they do and why they do it in public settings or in “inter-professional venues such as conventions and trade publications” (Lowrey 2006, 482). Through this lens, Keller’s arguments about the special value of “boots on the ground journalism,” Huffington’s response to Keller, and even Rosen’s analysis of the responses to Huffington’s response, can be seen as aspects of a larger jurisdictional struggle between workers engaged in practice of news aggregation and workers doing original reporting.

Despite increasingly sophisticated theoretical work on this topic, empirical research devoted to the analysis journalism’s jurisdictional struggles remains rare. The first section of this paper undertakes a qualitative content analysis of the March 4, 2010 FCC workshop<sup>2</sup> “The Future of Media and Information Needs of Communities: Serving the Public Interest in the Digital Era”<sup>3</sup> in an effort to add some substance to this discussion of rhetorical boundary-maintenance. This

---

<sup>2</sup> A note on public hearings: the public hearing transcripts discussing the “future of journalism,” primarily carried out in Washington DC in 2009 and 2010 by the Federal Trade Commission (FTC) and the Federal Communications Commission (FCC) have been chosen as the primary corpus through which professional boundary drawing battles will be analyzed. Obviously, many other corpora on this topic are possible: twitter discussions, blog postings, industry trade journals, conference proceedings, journalism school textbooks, and so forth. Public hearings, however, are unique insofar as they mark a primary moment in the U.S. media system when the values of journalism are specifically addressed by formal bodies who have the power to (occasionally) set practice-defining policy. It might be inferred, then, that it is in these public policy setting that rhetorical arguments will be the most sharply defined. It is against this background that the historical analyses of the 193x Radio Act (McChesney 1995) and the Hutchins Commission (Pickard forthcoming, Sylvain 2010) have been carried out.

<sup>3</sup> The analyses here is exploratory in two senses: first, the completed paper will contain more thorough analyses of *other* FTC and FCC hearings that took place in 2009 and 2010 in Washington DC, in addition to the one discussed here; second, the content analysis itself is preliminary insofar as it has been carried out by single researcher. There is thus no discussion of inter-coder reliability in this section of the paper, and its arguments should be taken as highly provisional.

section of the paper utilizes qualitative content analysis (Hsieh and Shannon, 2005), to understand the process by which FCC commissioners and witnesses discussed journalistic products (“news reports,” “news stories”) as either *original* or *aggregated*. What were the tacit conceptions of original reporting and aggregation that commissioners drew upon during the March 4 hearing? And how did commissioners and witnesses articulate their understanding of the relationship between original and aggregated journalistic content within the larger digital news ecosystem?

Of course, “original reporting” and “aggregation” were not the only, or even necessarily the primary, topics discussed at the FCC on March 4, 2010. The nuances of broadcast regulation made up a lengthy portion of the hearing, as did more macro-level arguments about the economics of newswork in the digital age. The uncertain line between original and aggregated content comprised up a substantive portion of this larger economic debate, however, a fact that provides further evidence for my earlier argument that the jurisdictional conflict within journalism is very often centered around notions of “ontological originality.”



that still originate the overwhelming amount of the news we get,” noted FCC commissioner Michael Copps, “on the order of three-quarters or more, and that number is going to go down only slowly.” (FCC 2010, 13) In a second representative comment, media historian Paul Starr noted:

Paid circulation is also in long-term decline, yet newspapers have financed most of the original reporting at the state and local level, and as both their advertising revenues and circulation dropped, they are cutting back for original reporting more rapidly than new resources are developing on line. (ibid, 31)

Here and elsewhere, speakers draw a connection between reporting, positive democratic outcomes, and the importance of newspapers to the journalistic process. Newspapers are described as the “originators” of most original content; indeed, hearing chairman Steven Waldman joked that the Project for Excellence in Journalism “should get a royalty every time [the committee]” mentioned the group’s study finding that 95% of news stories with new information came from traditional media, primarily print journalism (ibid, 304). The taken-for-granted importance of original reporting articulated in the first third of the hearing did not include much discussion about the meaning of “original reporting,” or the manner in which technological changes were affecting that reporting. Indeed, much of the discussion in the early part of the hearing could have been presented at any time in the late 20<sup>th</sup> or early 21<sup>st</sup> century; the sense of “crisis” within the news industry expressed by participants could have been applied to earlier periods of concern about shrinking news coverage. Indeed, one speaker made note of the fact that one reason for the historical success of FM radio was a 1965 regulatory decision that these stations had to produce a certain amount of original programming in exchange for their licenses. The analysis, here and elsewhere, bore an uncanny resemblance to earlier declinist arguments about the disappearance of a journalistic “golden age” under impact of forces of commercialization, corporatization, or technology. The actual mechanics of that decline seemed less important than the general lament.

In the final third of the hearing, on the other hand, the tenor and content of the conversation shifted as witnesses directly debated the relationship between digital technology, search engines, and the economic impact of these technologies on original reporting practices.

The rhetoric during this final portion of the March 2010 hearing crystallized around a debate between Associated Press General Counsel Srinandan Kasi and CUNY journalism school professor Jeff Jarvis. Kasi threw down the gauntlet, arguing, in his words, that “effectively what we see happening [online] ... is that those who are in a position to fulfill demand -- consumer demand online for news are, therefore, controlling modernization -- have literally no cost of news gathering.” (ibid, 250). In other words, the vast majority of digital revenue on the internet went to online search engines like Google that filtered, sorted, and highlighted digital content. Meanwhile, argued Kasi, the originators of that content were left out in the cold:

There's a real cost in news gathering and that goes to what I call the cost of the first copy. But the technology allows you to have secondary copies at no cost. So what happens when a number of these secondary copies get into the same ecosystem and effectively compete for opportunity -- modernization opportunity -- with the first copy? (ibid, 254)

In defending Google and other search engines, Jarvis advanced a different argument, one that emphasized Google’s ability to create audiences and deliver them to publishers. “Google gives you value,” he told Kasi and the FCC “There are two creations of value today -- the creation of the content and the creation of the audience for that content. Each bring value ... It's up to [publishers] to decide whether you can create a relationship and value out of that.” (ibid, 282). But if Jarvis saw Google as an indexer of content and a driver audiences, the executives working for traditional media companies perceived it as a fragmentizing original content substitute, one which occupied an unfair position in the digital value chain. The name they gave to this process of fragmentation, excerption, and indexing was *aggregation*. If you analyze web traffic, argued Kasi, “what you'll see is that the *aggregation sites* actually enjoy the benefit of the traffic flow. Over half the surveyed audience got the news from aggregation sites” rather than the original content producers Kasi concluded (ibid, 253).

In this debate, we encounter a de-emphasis on the positive democratic outcomes outlined in the first debate and a re-emphasis on the economic injustice of news aggregation. The assumed public importance of original reporting operates in the background of this conversation, however. Also at work here is a criticism-- not of “traditional” news aggregation-- but rather the

massive algorithmically powered process of indexing and search. The definition of what, exactly, original reporting is, why it is important, and how it compares to the aggregation of information is addressed only obliquely. Also almost entirely ignored in the March 24 hearing are the websites many people think of as classic news aggregators—blogs and hybrid bog / news websites like *Gawker*, *Media Gazer* and the *Huffinton Post*. Nevertheless, the animus to a variety of technological and digital information processes on the part of news executives comes though clearly in these hearings. The enemy is named, and the enemy is aggregation.

Only at the conclusion of the hearing does Chairman Steve Waldman attempt probe more deeply the question of what original reporting might actually mean, and why it might be important.

Turning to Jarvis, he inquires about whether the he is worried by the recent Pew Report on Baltimore which concludes that the vast majority of original news content comes from traditional print media. “I don't agree with the premise,” Jarvis responded

because I think there's a definition of news. There's a top-down definition of news. When I ran community sites and we had a site for ballerinas under the news tab it said the leotards are in. Well, to them that's news. There's many different definitions of news. And I think that the flaw in the Pew study was its definition of news and media and distribution were very limited (ibid, 305).

We are thus left with tantalizing questions about whether or not certain categories of original reporting are undervalued, along with the even more the metaphysical question of what is original reporting even is. Although these questions were not addressed in the March 24 hearing, several larger themes are prominent. They include: the assumed connection between original reporting and a healthy democracy, the negative economic consequences of the digital economy for original content producers, and the confusion about what aggregators are and how they are different from search engines. The most important theme, however, is simply the transparently negative animus directed against the idea of aggregation by many traditional reporters and news executives, along with the idea that it is radically distinct from original reporting. It is in this fairly simplified fashion that public hearings like the ones held by the FCC and FTC throughout 2009 and 2010 may have their most impact on jurisdictional contests.

## From Purification to Complexity and Back Again

In the early pages of *Science in Action*, anthropologist of science Bruno Latour draws a distinction between what he calls “ready made science” and “science in the making.” Accompanied by the image of the “two faced Janus” (with the wise, bearded face representing ready-made science and the youthful, insouciant face standing in for science in the making), Latour argues (Latour 1987) that all settled controversies represent a series of unopened black-boxes, and that the outcome of rhetorical and praxis-based controversy is to either open or close these boxes.. In Latour’s analysis, ready-made science represents the default state of affairs, a situation in which distinctions are clear and simplicity is valued. As Graham Harman puts it in his overview of the concept of the black box:

In a sense, all human activity aims to create black boxes. Boeing engineers labor to create a new model of jet, which will never reach the market if its various parts break down during test flights. In forming a friendship, settling a marriage, or composing a manuscript, our hope is to establish something durable that does not constantly fray or break down. A job in which our roles are reassigned each week, or with the constant danger of being sacked by an emotionally unstable superior, is more of a headache than anyone can endure. Earning a doctoral degree would not be worth the trouble if our transcript and thesis were scrutinized monthly by a panel of experts for the rest of our lives, or if long-time professors had to retake their comprehensive exams every summer. In everyday language we now refer to certain cars and people with the wonderful phrase ‘high-maintenance’. By definition, a black box is low-maintenance. It is something we rely on as a given in order to take further steps, never worrying about how it came into being. The reason it can be either so refreshing or so annoying to speak of one’s work with outside amateurs is that they lack awareness of the black boxes widely recognized in our respective professions (Harman 2009, 37).

For our purposes, the utility of Latour’s argument is that it gestures at the fairly common-sense notion that all stable definitions of originality and aggregation disguise their incredibly complex histories. They also bracket off the tangled, halting practices of actual journalistic work. The differences between an “aggregator” and an “original reporter” are never as clear in actual practice as they are during testimony in front of a public commission. Indeed, even during the already examined FCC hearings, the lines between aggregator and original reporter were not entirely clear. Most noticeable was the oft-fudged distinction between aggregation as the

compilation and contextualization of previously reported stories and the notion of aggregation as an algorithmic sorting of articles carried out on a massive scale by powerful algorithms. Once we shift our analytical lens from the domain of rhetoric to the domain of practice, the complexity of the distinction between aggregation and original reporting becomes even more tangled. In this second section, I want to analyze the manner by which seemingly-solid occupational boundary lines are actually comprised of a myriad of complex, uncertain, unstable practices. I argue that the boundary line between aggregation and original reporting only becomes stable rhetorically or retroactively, once the controversy over the definition of aggregation is settled. Studying these micro-practices qualitatively can do more than simply add complexity to a previously clear situation; the purpose of this second strand of analysis is not to simply ethnographic masochism. Rather, the analysis of the manner by which rhetoric is complicated in practice, and by which practices become retroactively pure, help point us toward new analysis of the evidentiary status of news objects, along with a deeper understanding of the work of journalism in the digital age.

### **Aggregation in Action**

Between 2008 and 2010, site visits to the offices of *Philly.com* (the website aggregating and repackaging the content of both the *Philadelphia Inquirer* and the *Philadelphia Daily News*), the New York City-based media news and opinion blog *JorunoBlab*<sup>4</sup>, the offices of the *Newark Star-Ledger*, and the offices of the Washington Post allowed me to formulate a schematic overview of aggregation as its own unique form of newswork. An important complement to these site visits were a series of in-depth, semi-structured interviews with former and current news aggregators, primarily in Washington D.C. and New York City. All in all, I conducted eleven semi-structured interviews with news aggregators and engaged in more than 100 hours of aggregation-specific fieldwork (much of it occurring during an extensive period of research in Philadelphia). As with the content analysis presented above, the qualitative research into news aggregation

---

<sup>4</sup> A pseudonym

discussed here should be considered highly provisional, with more research (and hopefully more site visits) to come. It was an odd aspect of this research that my attempts to gain access to supposedly “informal” news aggregation organizations like the *Huffington Post*, *Gawker*, and *Mediaite* were considerably more fraught than my earlier participant observations at more formal news companies like *Philly.com* and *The Newark-Star Ledger*.

### **A Schematic Overview of Aggregation as Journalistic Work**

At the time of my initial research, between 2008 and 2009, *Philly.com* was located in its own non-descript downtown offices, separate from the newspapers they aggregated. As in many of the newsrooms I visited, employee desks occupied the central space of the main room, with executive offices ringing the outside walls. Surprisingly, the web production team took up only a fraction of the space inside of *Philly.com*— in terms of space allocated and the number of employees, the production team was far outnumbered by the marketing and advertising departments. The *Washington Post*, on the other hand, had by the summer of 2010 completed a long and somewhat tortuous process of “digital integration,” with the formerly separate employees of *Washingtonpost.com* now scattered across the larger *Washington Post* newsroom.<sup>5</sup> At the *Post*, the “continuous news desk” stood at the center of the main newsroom, surrounded by technical gadgetry and desks for “user-interactivity” and public-relations/social media. The more traditional newspaper desks (local, politics, business, obituaries) formed an outer ring around the continuous news desk, sections immediately distinguished by their employee workspaces covered with paper and coffee mugs rather than mp3 cables. The offices of *JournoBlab*, finally, were uncomfortably wedged into a SoHo office building above a very hip record store; there was less a sense of conscious desk organization here than simply a feeling of far too many people in far too small a space (indeed, by the time I visited them in 2010, *JournoBlab* was preparing to move into larger offices).

---

<sup>5</sup> I should note that, by the time of my follow-up research there in the winter of 2011, *Philly.com* had itself been integrated back into the same building as the Philadelphia newspapers, though in a far more haphazard fashion than the *Washington Post*.

The lengthy amount of time I spent observing *Philly.com* web producers helped me formulate a schematic definition of news aggregation, a generalized description that can be used as a reference point for analyzing other, more unique forms of aggregation work as well as in formulating broader empirical themes. At *Philly.com* the primary role of a web producer was to determine where a news story (usually written by a reporter at the *Inquirer* or the *Daily News*, almost never by an employee of *Philly.com* itself) belonged on the *Philly.com* website. This was mostly an issue of news judgment: was a piece of news worthy of being a “biggie” (the name for the top story on the site), should it be downgraded one notch and function as a “spotlight” story, or should it not be promoted at all? Producers at *Philly.com* rewrote what they saw as web-unfriendly ledes and headlines, procured art, and decided on ways to “build out” stories with links to other, related stories on *Philly.com*. To move into a lead position at *Philly.com*, a story needed to contain art, as well as an additional piece of “user-generated content” (often a comment box, or a poll). A story almost certainly needed a collection of related links if it was going to ascend the ladder of importance at *Philly.com*.

For the purposes of this paper, then, I define news aggregators as *hierarchizers, inter-linkers, bundlers, rewriters, and illustrators of web content*. News aggregation is particularly common in journalistic networks where journalists at the ends of an organizational chain produce pieces of content in an uncoordinated or quasi-coordinated fashion. In many cases, these end-network producers are not formal members of the news institution that is doing the news aggregation. A news aggregator coordinates amongst a series of quasi-institutionalized (or entirely independent) content producers. The primary task of this news aggregator is, then, to build links between independently produced news stories, and to rank these bundled news stories according to a rapidly shifting criteria of importance, popularity, and newsworthiness.

It is interesting to note that many of the individual news aggregators I interviewed told me that they had no real workflow to speak of. “When something comes across the internet, we either grab it or we don’t,” one of the workers at *JournoBlab* told me. “There’s not much more to it than that.” (CITE). Indeed, many of the aggregators I spoke to seemed skeptical that anyone could find anything interesting or meaningful to say about aggregation as a form of newswork.

Nevertheless, after talking to a number of individuals engaged in aggregation work at a number of different organizations, larger themes and structures in the routines of aggregation quickly became apparent.

### *Daily Routines*

For most of the professional content aggregators I spoke to, the news day begins early, and it begins with an immersion in the tsunami of digital content that continually flows across the internet. “I subscribe to about 150 RSS feeds on my Google reader,” a former editor at the Washington DC based website DCist told me. “For a while, I just sit at my laptop skimming head[lines], clicking on ones that seemed relevant, and opening browser tabs. My early morning goal was always to put together something called the “morning roundup,” which was a long-ish summary of three or four top stories, plus descriptions of and links to five or six minor stories.” The DCist editor added “the goal was to have the morning roundup up by 9am, but sometimes it wouldn’t get up until 9:30 or 9:45. There are other aggregators out there more hard-core than me.”

One of these “hard core aggregators” was John Winter, who worked for several months as the local politics and entertainment editor for one of the leading online news aggregators in the country, the *Wellington Gazette.com*<sup>6</sup>. After a short career in weekly journalism and a stint in graduate school, Winter describe jointing the *Wellington Gazette* after school and instantly entering an intense work environment where days began and ended with intense media consumption.

My alarm would go off at 6:55, and I’d jump out of bed and immediately turn on [MSNBC]. I’d watch it while I was combing the online newsires. If I thought that nothing was absolutely breaking right away, I would get some granola. While I was eating, I would look through the websites of the [Wall Street] Journal, and the New York Times blogs. I’d also check the AP and check Reuters.

---

<sup>6</sup> Ex-staffers with the *Wellington Gazette* are subject to a fairly stringent non-disclosure policy, making accurate representations of work there difficult to come by. This ex-aggregator spoke to me under the condition of anonymity.

From this stream of news Winter would craft his aggregated content. Between 7am and 8am, Winter described what he called a “really frenzied process” of refreshing the local politics page and “clearing it out.” “Clearing it out” meant removing an old story and replacing it with a newer one. “A good half the page would be cleared out, would be new, by 8am,” Winter told me. He also noted that he had two content management “buckets” on the backend of the local politics page, one that would manage the content on the news and aggregation portion of the page, and a second that controlled the blog posts submitted by the Gazette’s volunteer blogging community. Most of the blog content was simply arranged in chronological order, while the display of the news involved some prioritization. “There is one story that runs all the way across the top, in huge blaring letters. Then there was space for three stories below that, and below that, there are secondary and tertiary stories.” After clearing out as much of the page as he could, Winter would backtrack and check the television news, the internet, and his feeds again, “and if it was clear, I’d hop in the shower.” But there was often more news, “and there were times when I wouldn’t head for the train until 9 and get into the office until 10 or 10:30. From then on, and for the rest of the day, the goal was is to refresh the page until 6 or 7pm.”

All the aggregators I interviewed described their mornings in similar terms, as the immediate immersion in the flood daily news and an attempt to pick out those stories that would most directly interest their website’s audience, draw traffic, or both. The remained of the day usually involved sitting in front of a computer and repeating the same basic tasks, at a less frenzied place. The importance of an early start and intense start to the day, while apparently trivial, actually contrasts in some significant and symbolic ways with the habits of traditional journalists; during my earlier Philadelphia fieldwork, professional reporters working for newspapers spoke happily about their late-morning starts and their late evenings. But beyond simply “getting up” and “reading the internet,” is there anything more to say about the work of news aggregation? There is. To understand, on a deeper level, the set of skills that go into aggregation work, I want to turn discussion of the relationship between aggregation and news judgment.

## *The Skills of Journalistic Aggregation*

Sitting in a darkened midtown bar that has long been one of the favorite haunts of journalists working for the New York City tabloid *New York Post*, Hasaani Gittens was emphatic. “There are skills to news aggregation,” he told me. As a self-described “former “old school rewrite man” for the *New York Post* now working as a web editor at the local NBC website, Gittens is the kind of journalist who often gets ignored in the classic studies of both digital and analog journalism, and his choice of the location for our interview was an interesting one. Our bar was what might be described as a “*New York Post* bar,” distinct from the “*New York Daily News* bar” on the other side of Times Square, and Gittens drew on his journalistic experience, his rewrite experience, and his aggregation experience when outlining the qualities of a good news aggregator. For Gittens, there was a distinct typology of skills needed to report, to rewrite, and to aggregate:

As an aggregator, you really need to know what's a good story. It's heavy on news sense. And you need to know what's a good story for your audience. But it's not about reporting skills. To be a reporter, on the other hand, you need to know certain neighborhoods. Most of all, you have to have the guts to ask questions. You need to know how to talk to people to get them to talk to you. You need to know how to take notes. You need to know how to develop and cultivate sources. As an aggregator, you're, number one, going to need to know how to have a sense of story.

When I pressed him about how aggregation differed from his days as a rewrite man, Gittens had a ready response: “You do a little journalism as a rewrite man,” he told me. “There's a skill involved in crafting a witty turn of phrase out of a boring story, and a lot of the stories you end up rewriting for a tabloid are boring stories. So you can really have three kinds of skills. You can have reporting. You can have rewriting. You can have aggregation.” But the dominant quality of a good aggregator was “news judgment. Good reporters don't always have the greatest news judgment in world, but good editors always do, and so do aggregators.” You have to know how to write a good headline, Gittens went on. You also have to know how to write a good summary of a story, and how to incorporate visual graphics and photos. “You need to know how to write a hed that gets you traffic. But most of all you need to know what's news.” The former editor of DCist

described a similar panoply of skills, emphasizing news judgment above all. “You need to know what stories readers will be interested in. You need to keep track of your sources. You need to be quick. And you need to add value to the story you’ve aggregated. But news judgment is the most important.”

John Winter, the former aggregator with the Wellington Gazette, also contrasted the relationship between aggregation and reporting, though he phrased the distinction in almost entirely cultural terms. “No one I worked with at the Wellington Gazette *wanted* to call themselves a journalist,” he told me. “Because they didn’t like it. They *looked down* on it. They thought journalism was dying, and they didn’t want to be part of a dying thing. And they actually looked down on journalists, for cultural reasons. The journalists they knew were old, and none of them had gone to the same prep schools they had.” Winter said that he found the experience incredibly depressing. “I’d gone to journalism school,” he told me. “And here I was with these people who totally looked down on what I did, or at least what I’d wanted to do when I was in school.” At the same time, Winter noted that, in the months since he’d left the Gazette, he’d retrospectively begun to see what the work he did there as journalism.” Aggregation was journalism, he told me, because it was no longer just the work of a few fringe bloggers and digital cranks. “Now traditional news organization have started aggregating and applying their ethics to it,” he said.

The fact that news organizations are themselves now engaged in aggregation work is another example of the hybridized complexity I referred to at the conclusion of the previous section. A further complexity stems from the fact that news organizations *often deliberately seek to be aggregated*: “we are they aggregator and the aggregates,” Gittens laughed, describing the process of deliberately seeking out links from high-traffic web sites like the Drudge Report and Yahoo News. Gittens spoke of spreadsheet kept by the aggregators on staff at NBC News that had contact information for various web editors at major internet sites, as well as potential hooks that might encourage them to link back to a story. It was considered a major success, particularly in financial terms, when other aggregators linked to NBC.

*“Keeping an Ear to the Internet”*

Journalists engaged in the work of news aggregation consume massive amounts of digital content, and also need to have good news judgment. But what is it that they actually produce? The question is complicated by the rhetoric of content “theft” that permeated the FCC hearing analyzed earlier, as well as the hope of many traditional news organizations to both “aggregagate” and “be aggregated.” But a common description of what news aggregators did was summarized by the ex-editor of DCist: “our job is to make it possible for someone to have a one-stop location for the news they need about DC ... if you only read DCist, we wanted you to know what was going on in town. And we also want to *add value to content as well*. As an aggregator, you should always be adding a little something: a poll, a different take on an issue, our own two cents, bring two related ideas together in one space, a better photo. That kind of stuff.” This idea of being a “one-stop shop” was expressed by every aggregator I spoke to, in almost identical language. “What did Wellington Gazette pay me for?” asked Winter. “They paid me to try to be the one stop shop for news about local politics and entertainment.”

The primary reference point of the one-stop-shop news aggregator was the internet itself. It was from the digital bounty of the online universe that aggregators drew the content they would parse, rebuild, and contextualize. “What we do as aggregators isn’t about journalism,” one high-level news executive told me. “It’s about making sense of the internet.” Comparing his tenure at a major daily newspaper with his time as the manager of a prominent and wealthy news aggregator, the executive noted that traditional journalism had “always been about making sense of the public, and about your local community, but with the internet, we aggregators need to make sense of this other world. It’s why the Huffington Post, for example, has the guts to call an ‘internet newspaper.’ It’s about the internet. It isn’t about journalism, at least the way we’ve always thought of journalism up until now.”

It is interesting to compare the aggregation practices I observed at *Philly.com* and practices at the other news aggregators I spoke to in the months after my fieldwork. Because of the combination of *Philly.com* linking practices (discussed in the final section, below) and the

company's organizational setup, the web producers at Philly.com found them aggregating content from a very narrow set of sources; almost all the content on Philly.com comes from one of the two local newspapers. So while there was a constant stream of digital news being ingested and regurgitated by Philly.com aggregators, the scope from which that content could be drawn was fairly narrow and presented in a fairly rational fashion. Reporters at the *Daily News* and the *Inquirer* were company employees, subject to the same upper management as workers at Philly.com and operating under fairly established deadlines and news routines. Aggregators at Philly.com did *not*, in other words, aggregate the internet, or even the local internet in Philadelphia. True aggregators, in contrast, "keep their ear glued to the internet." Their informational commons, and often their journalistic mandates, are far broader. And while the speed of work at Philly.com was fast-paced, and the content demands were high, these demands and routines paled in comparison to the expectations at more traditional web aggregators.

### *Creativity and Speed*

As part of the discussion of his daily routine at the Wellington Gazette, John Winter made an interesting point about the amount of time and effort he put into his various aggregated pieces of content on a typical day. An outside observer might assume that every aggregated story is more or less the same and requires a more or less identical amount of time and effort. But this is not the case. "There's a degree of whorishness to the fact that we spent so much time at the Wellington Gazette trying to come up with SEO-optimized headlines," Winter told me. "But that was the name of the game, and it really got interesting when we tried figure out these sexy headlines for those really important stories."

You do put up some garbage stories. You have to put it on the page. But I'm not going to spend any mental energy on those stories. I have to put them up because I have to feed the beast, but I'm going to put it up there with just the most straight, boring headline ever, and spend my time figuring out headlines for the stories I really care about or that are really important.

Winter makes the point that, with a goal of refreshing the local politics three or four times a day, producing quality aggregated content—content that, in the words of the editor of DCist, always contains added value—was difficult if not impossible. But, he adds, the regularized production of mediocre content could coexist with the occasional production of higher-quality stories.

There is little doubt, of course, that the demands upon digital journalists are increasing. At the end of our interview, I asked the former editor of DCist if she had any thoughts on recent journalistic complaints by journalists that the increasing speed of their work routines were diminishing creativity. Was increased speed or output demand damaging original reporting? “I don’t really see the point in putting original reporting on a pedestal,” she told me. “It’s important, sure. But it’s only one thing among many other things. I don’t have a lot of respect for these old-time journalists, to be honest. I think writing four stories a month is lazy. I think it’s lame.” In discussing a reporter at the *Washington Post* who often spends a great deal of time on long-form stories, a Post digital editor complains to me that the reporter spends a lot of “time flying down to Florida to talk to fishermen, or to someplace else to stare at horse. During the amount of time he spent doing that he could have written 20 stories. But he has the best journalism job. No journalists have that job anymore.”

DCist publishes between 18-20 stories a day, with upper-level company executives hoping for publication metrics on the 3--story per day range. Gittens estimates that the local NBC website asks web producers to compose 10 stories per day, leading to a total of between 40 and 50 new stories per day for the entire website. Are ten aggregated stories per day a great number? There are few standards by which to answer these questions. While the DCist editor expressed contempt for “traditional journalism whiners,” the aggregator at the *Wellington Gazette* spoke eloquently about his own burnout and the “deep levels of exploitation” he felt while working at the high-octane company. In the end, even the creative workarounds Winter designed to highlight the stories he really cared about while spending next to no time on the rest were not enough to keep him from going back to traditional journalism. “The culture there wasn’t right for me and I did want to have a life. And people there didn’t have a life. Or at least they didn’t want to hang out with me.

So maybe it was me. But I didn't want to work with people who made me feel guilty for having fun. I worked," he concludes. "for people who were joyless."

### **Digital and Analog Evidence in web Era Journalism**

The preceding overview of aggregation "in action" has helpfully complicated the rhetorically purified notion of aggregation often expressed in public hearings, trade journals, and online polemics. Obviously, the actual *work* of aggregation is a complicated exercise, one in which the line between "original reporting" and "aggregated content" is not entirely clear. Traditional media organizations aggregate, and thus blur the line between what exactly is meant by original reporting and digital content. These organizations also play the role of both "aggregator" and "aggregatee." In short, the work of aggregation is a complex affair.

Perhaps the most intriguing idea expressed by the aggregators interviewed above is the notion of aggregation as "keeping your ear to the internet." A digital editor at the Washington Post told me that "while the thing about aggregation is that its constantly changing as a practice, as new techniques come along, and so on. But it's always the same insofar as what people are trying to do is to tap into the human psyche -- and by that I mean the human psyche as it gets put on the internet ... so aggregation isn't about journalism, at least the way we've always thought of journalism up until now. It's about understanding the web. "

As defined processually the line between aggregation and original reporting is not entirely clear, despite rhetorical attempts at category purification and boundary-drawing. Aggregators and journalists both compile shards of facts, quotes, documents, and links together in order to create narrative-driven news stories. Could it be that the real conflict between aggregation and journalism lies not in the work of either occupation, and not in the way that each defines the other as a sort of pathological doppelganger, but in the type of objects of which they build their stories and that they take as their criteria of evidence? Is it possible that the great conflict over journalism may be centered around the *things* of journalism, rather than the *work* of journalism or the definition of journalism?

Studying the things of journalism would mark a move away from journalistic epistemology and towards a sort of ontology of reporting. This scholarship would look at the meaning and status of hyperlinks, not simply as ways in which the news is framed, but also as the objects out of which journalism crafts stories (Coddington, forthcoming). It might chronicle historical divergence between journalism, the “social survey movement,” and academic sociology in the early 20th century. It could combine the growing number of cross-cultural histories of the interview with research on rise of the document in journalism “and the invention of direct observation. This research could analyze Philip Myer’s work in order to trace the genealogy of hacks and hackers and the rise of variable-based journalism. Ultimately this research would draw upon philosophies of digitization and quantification. Perhaps it is in the realm of the digital—the creation of a world and a set of practices in which the aggregation of purely digital objects is possible, necessary, and socially useful—that the true transition of journalism into something new and different might be glimpsed.

## Secondary Works Cited

- Anderson CW (2009) *Breaking Journalism Down: Work, Authority, and Networking Local News, 1997—2009*. Unpublished doctoral dissertation.
- Carlson M (2007) Blogs and journalistic authority: the role of blogs in US Election Day 2004 coverage. *Journalism Studies*, 8(2): 264-279.
- Carey J (1997) The communications revolution and the rise of the professional communicator. In: Munson E and Warren C (eds) *James Carey: A Critical Reader*. Minneapolis: University of Minnesota Press, 128-142
- Gieryn T (1983). Boundary-work and the demarcation of science from non-science: Strains and interests in professional ideologies of scientists. *American Sociological Review*, 48 (6): 781-795.
- Harman G (2009) *Prince of Networks: Bruno Latour and Metaphysics*. Victoria: re: press
- Hsieh H and Shannon S (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9):1277-1288
- Huffington A (2011) Bill Keller accuses me of “aggregating” an idea he actually “aggregated” from me. Available at [http://www.huffingtonpost.com/arianna-huffington/bill-keller-accuses-me-of\\_b\\_834289.html](http://www.huffingtonpost.com/arianna-huffington/bill-keller-accuses-me-of_b_834289.html)
- Keller B (2011a) All the aggregation that’s fit to aggregate. Available at <http://www.nytimes.com/2011/03/13/magazine/mag-13lede-t.html>
- Keller B (2011b) Aggregation agro. Available at <http://6thfloor.blogs.nytimes.com/2011/03/13/postscript-aggregation-aggro>
- Latour B (1987) *Science in Action: How to Follow Scientists and Engineers Through Society*. Cambridge: Cambridge University Press
- Lowrey W (2006) Mapping the journalist-blogger relationship. *Journalism: Theory, Practice, Criticism*, 7(4) 477-500.
- McChesney R (1995) *Telecommunications, Mass Media, and Democracy: The Battle For Control of US Broadcasting, 1928-1935*. Oxford: Oxford University Press
- Pickard V (forthcoming) Whether the giants should be slain or persuaded to be good: Revisiting the Hutchins Commission and the role of media in a democratic society. *Critical Studies in Media Communication* 27 (4): 391-411.
- Rosen J (2011) The twisted psychology of bloggers vs. journalists. Presented at *South by Southwest*. Austin, TX
- Schudson M and Anderson CW (2009). Objectivity, professionalism, and truth-seeking in journalism. In: Wahl-Jorgensen K and Hanitzsch, T (eds) *The Handbook of Journalism Studies*. New York: Routledge.
- Singer J (2003) Who are these guys? the online challenge to the notion of journalistic professionalism. *Journalism: Theory, Practice, Criticism*, 4(2): 139-163
- Sylvain O (2010) *Domesticating “the Great, Throbbing, Common Pulse of America”: A Study of the Ideological Origins of the Radio Act of 1927*. Unpublished doctoral dissertation.

Zelizer B (1992) *Covering the Body: The Kennedy Assassination, the Media, and the Shaping of Collective Memory*. Chicago: University of Chicago Press.