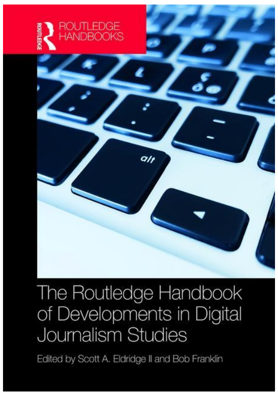

The Routledge Handbook of Developments in Digital Journalism Studies

Scott A. Eldridge, Bob Franklin

An Approach to Assessing the Robustness of Local News Provision

Publication details
https://www.routledgehandbooks.com/doi/10.4324/9781315270449-9
Philip M. Napoli, Matthew Weber, Kathleen McCollough
Published online on: 30 Aug 2018

8

AN APPROACH TO ASSESSING THE ROBUSTNESS OF LOCAL NEWS PROVISION

Philip M. Napoli, Matthew Weber, and Kathleen McCollough

As local journalism evolves in response to challenges posed by the technological changes that have taken place in the media sector (Anderson et al., 2012; Downie and Schudson, 2009; Picard, 2014), one growing concern is that significant differences exist across communities in terms of the extent to which journalism sources are serving people's information needs. Research raises concerns that, in some communities, local journalism is essentially collapsing, with the decline and (in many cases) disappearance of traditional news outlets leaving massive unfilled gaps (what Stites (2011) has termed "news deserts"; see also Ferrier's (2013) analysis of "media deserts"). These gaps may create greater opportunities for political and corporate corruption to flourish in ways that can undermine effective democratic participation (Starr, 2009).

The extent to which local news is collapsing may vary according to the particular characteristics (demographic, economic, political, technological) of individual communities (e.g., Pew Research Center, 2015). A 2015 report noted, for instance, that large U.S. cities such as New York, Washington, DC, and Los Angeles are employing an increasing proportion of the country's professional journalists, with smaller cities experiencing dramatic declines (Tankersley, 2015). These patterns suggest an emerging condition of "journalism haves and have-nots" across the country. As the Knight Commission observed in 2009:

> the nation's vast information needs are met unequally, community by community. Some populations have access to local news and other relevant information through daily newspapers, radio, and television broadcasts, local cable news channels, hyper-local websites, blogs, mobile alerts. . . . Others are woefully underserved.
>
> *(Knight Commission, 2009: 3)*

However, anecdotal evidence like that presented in the Knight Commission report has yet to be accompanied by systematic empirical data on the state of local journalism that allows for analyses across multiple communities.

Other areas of public interest, such as economic development, the environment, political participation, and community engagement, have reasonably well-developed methods and measures for assessing the health of local communities (e.g., Community Health Status Indicators Project

Working Group, 2009; Sustainable Jersey, 2013). The same level of method and measure development has not been the case, however, for local journalism.

This chapter describes ongoing efforts to address this analytical challenge. In it, we present a multi-level methodological framework for assessing local journalism and the extent to which it addresses communities' critical information needs. The development of such a method and the accompanying measures provide a valuable analytical tool for news organizations, funders of journalism initiatives, advocacy groups, citizens, and policymakers.

The first part of this chapter contextualizes this methodological approach within a sequence of calls for methods and measures that could inform public and private efforts to strengthen local journalism. The second section presents the methodological approach and describes the rationales underlying its design. The third section describes the types of metrics that can be constructed and analyses that can be conducted with the data. The concluding section considers further steps in the development and application of this methodology.

Local journalism and the need for assessment tools

Concerns about the state of local journalism have been widespread, as technological changes dramatically affect the economics of the commercial news business and the dynamics of how consumers access, share, and even produce journalism (see Barnett and Townend, 2015; Nielsen, 2015). These concerns have been particularly pronounced in the United States, where there is often a greater reliance on the commercial model of news production than in virtually any other developed nation (Schudson, 2005; see also Pickard, this volume, Chapter 16).

Reflecting the intensity of these concerns in the US, the Knight Commission released a 2009 report identifying access to credible and relevant information as a key requisite for healthy communities (Knight Commission, 2009). As the Knight Commission (2009: 39) emphasized, it was particularly important to assess the quantity and quality of information available to communities: "If activists, policy makers, and the general public had more concrete ways of describing, measuring, and comparing the systems of community news and information flow, it would be much easier to mobilize public interest around community information needs".

The Knight Commission (2009) report led to policy-related inquiries across various branches of the U.S. government (see Napoli and Stonbely, in press; Leibowitz, 2009; U.S. Senate, 2009). Most relevant to this analysis is the attention that the Federal Communications Commission (2009) devoted to the issue. Picking up on the Knight Commission's work, in 2009 the FCC initiated a comprehensive assessment of how community information needs are being met in the broadband era (Waldman, 2011). Among the FCC inquiry's conclusions was a call (echoing the Knight Commission) for a thorough accounting of the journalism provided at the community level.

In 2013, the FCC followed its report by commissioning a study that would provide a "Multi-Market Study of Critical Information Needs" (Social Solutions International, 2013). The proposed study included content analysis of journalism from a range of media, including television, radio, newspaper, and internet, analyzed in terms of the extent to which it covered "critical information needs". However, the proposed research produced a firestorm of controversy, on the basis of concerns that the research represented government intrusion into newsrooms (see Pai, 2014). Congressional hearings ensued, as well as threatened legislation to kill the research (Eggerton, 2014). As a result of these pressures, the FCC first scaled back and then ultimately canceled the entire study (Flint, 2014; for a more detailed discussion, see Napoli and Friedland, 2016).

More recently, the Pew Research Center (2015) has produced a thorough analysis of the local news ecosystems in three U.S. communities of different sizes. Extending upon earlier ecosystem research on the flow of news in Baltimore (Pew Research Center, 2010), this study included inventories of local media outlets, surveys of news consumers, and analysis of social media data. While incredibly useful, the scope and depth of this analytical approach comes with costs that prohibit scaling up to analyze a much larger sample of communities, which would enable more extensive comparisons and, potentially, more generalizable findings.

Indeed, case studies examining the state of local journalism in a single community, or in a very limited number of communities, are more easily found (Durkin and Glaisyer, 2011; Durkin and Hadge, 2010; Gloria and Hadge, 2010; Morgan, 2011, 2013; Pew Research, 2010; Ramos et al., 2013; Ryfe et al., 2012). The lack of studies on a larger scale are most likely a reflection of the challenges associated with developing a methodology that could be applied to a larger sample of communities at a manageable cost. In research in which larger samples of communities are analyzed, the analytical framework is typically limited in terms of focusing on a single platform and/or focusing on a particular issue or type of news (see Becker and Yanich, 2015; Fico et al., 2013; Holt and Karlsson, 2015; Karlsson and Holt, 2014; St. John et al., 2014; Williams et al., 2015). The narrow focus often occurs because the primary unit of analysis in such research is typically the individual media outlet or platform type or the individual news story rather than the community as a whole.

As this review indicates, there remains a gap in terms of a robust but scalable analytical approach to broadly assessing local journalism across communities, platforms, and issues that could be utilized by foundations, policymakers, researchers, and industry professionals. The methodological approach presented here represents an effort to fill this gap.

Assessing local journalism infrastructure, output, and performance

In order to outline a new analytical approach, the assessment of local journalism has been broken down into three connected conceptual dimensions: (1) the journalistic *infrastructure*; (2) journalistic *output;* and (3) journalistic *performance*, each with an associated methodological component. This analytical approach borrows from, and modifies, the well-known *structure – conduct – performance* theoretical framework from industrial organizational economics (see Caves, 1992). This framework presumes causal relationships between the structure of markets (as represented by the number, size, and characteristics of market participants), which in turn affects the conduct of firms in these markets, which in turn affects firms' market performance.

For this analysis, the notion of infrastructure refers to the number, size, and characteristics of media outlets within a community. Output refers to the volume of journalistic output produced by these media outlets. Performance refers to some fundamental qualitative dimensions of this journalistic output (see later). The goal here, though, is less about exploring relationships between the three conceptual dimensions of local journalism ecosystems than it is about offering a reasonably comprehensive empirical framework for assessing the state of local journalism in ways that can facilitate comparative analyses across communities and/or over time, and for (ultimately) facilitating research that identifies those community characteristics that are useful for explaining the state of local journalism in individual communities.

Assessing journalistic infrastructure

A key dimension of any local journalism ecosystem is the basic footprint of local journalism, in terms of the number of outlets capable of producing/disseminating local news and information within a given community. A fundamental premise of democratic theory-based approaches to journalism is the notion that the democratic process and an informed citizenry are enhanced via the presence of a diversity or plurality of sources of news and information (Baker, 2002). This perspective can encompass very basic indicators, such as the number of sources serving a community, or, it can delve deeper into the nature of these sources, including criteria such as market shares or ownership characteristics (see Napoli, 2001). Reflecting this theoretical perspective, this methodological approach begins by examining the journalistic *infrastructure* in a community.

A starting point for assessing a local journalism infrastructure involves identifying each source of journalism within a particular community (see Lin and Song, 2006). Such an activity has become more complicated, as keeping pace with the profile of any local journalism ecosystem is much more challenging in this time in which various journalistic initiatives are rapidly started and stopped.

Given the inadequacy of available data sources (such as the directories of local media outlets that existed in the pre-internet age; see Barnett and Townend, 2015), any effort to create an inventory of the sources of local journalism serving a community is, to some extent, an *ad hoc* endeavor. When one looks at previous research in this vein, relatively little concrete methodological detail is provided as to exactly how the inventory of local journalism sources was constructed, which is a reflection of the somewhat improvisational, somewhat impressionistic nature of the process.

However, for this methodological approach, we established – and outline in detail here – a concrete, multi-stage data-gathering protocol in order to provide as much clarity and transparency about the process as possible. The process draws from – and to some extent combines – approaches employed in previous research. It involves consultation with the most authoritative relevant directories available and supplements these consultations with a systematic search and discovery process that involves both online searching and engagement with members of the communities being studied.

Even something as simple as defining and identifying a source of local journalism is a more complex and challenging process than it once was, and any efforts to define and identify sources can be critiqued as being too narrow or too expansive. The definitional approach we employed leaned toward being expansive. First, we did not employ any criteria based on the presence of minimum levels of journalistic content; at this very local level, we wanted to include both dedicated and what we might term *tangential* or *potential* sources of journalism. Thus, for instance, local radio stations would be included regardless of their format as part of the journalistic infrastructure of a community; any journalistic distinction between a news and music station occurs in the analysis of journalistic output and performance dimensions (see later). And so, all community-based media outlets (defined as television stations, radio stations, print/online newspapers, and community and hyperlocal news websites) were included in the infrastructure level of analysis. Our search protocol then sought to identify other potential online journalism sources (such as blogs) left out of more the formal organizational criteria of a media *outlet*. However, in an effort to impose some boundaries on the abundance of content available online, such an online source needs to exhibit evidence on its homepage of addressing one or more of the critical information needs described later. Community news and information-focused blogs have failed to materialize in a very significant way in applications of our search protocol thus far, perhaps due to limitations in our protocol (and the associated databases) or to the dearth of such community-centric blogs within the communities of our analysis.

The focus of this methodological approach is *local* journalism. We define local in terms of the geographic boundaries of the communities being studied. Thus, this approach focuses on the journalistic sources that reside within and serve the selected communities (Lin and Song, 2006). Operationalizing these parameters means excluding larger regional, county, or state-level outlets that may cover the local community (e.g. if analyzing Newark, NJ, the *New York Times* would be excluded) but that are not geographically based in the community and primarily focused on the community.

The search process for identifying relevant journalistic sources is conducted in three stages:

Stage 1: Consult relevant media directories

A number of print and online data sources are available to identify media outlets at the local level. Because research has shown that many such directories (including those offered by commercial providers or government agencies) tend to be incomplete, multiple directories were consulted. These directories are listed in Table 8.1. This process involves locating those media outlets associated with the set of communities being analyzed. It is important to note that our design and application of this methodological approach have thus far focused on individual municipalities as the unit of analysis but can be modified to accommodate other units of analysis such as individual designated market areas (DMAs) or individual states.

Stage 2: Supplement directory data with manual search

In order to supplement the data gathered from the directories, the second stage of data-gathering involves a manual search for relevant journalistic sources. Following the approach employed by Ramos et al. (2013), this process involves keyword searches via search engine and then visiting those sites produced by the search queries to identify links to other relevant sources. Furthermore, those sites that were linked to by the original site are subsequently examined to determine whether they contain links to any additional relevant sites. Keyword searches employ the name of the town, county, and region, along with associated media terms such as 'news', 'blog', 'radio', and 'television'. In addition, when a community is known by a particular nickname (e.g., 'Brick City' for Newark), that terminology is also employed in the search process. If the analysis is expanded to include a wider geographic domain (e.g., DMA), search parameters for identifying sources simply need to be expanded according to the definition of the local level.

Stage 3: Targeted interviews with community members

In order to identify potential sources not identified by Stages 1 and 2, a final step involves conducting targeted interviews with community members in positions to be well-informed about the journalistic sources serving the local community (see Morgan, 2011). Specifically, for each

Table 8.1 Media directories/databases consulted

Medium	*Directory*
Television	Community Media Database
	Association of Public Television Stations' Station Directory
	FCC Broadcast Television License Database
Radio	NPR Labs Mapping and Population System
	FCC AM and FM Broadcast License Database
Print	Library of Congress Directory of Newspapers
	Editor & Publisher International Data Book
Online	Knight Foundation's Directory of Community News Sites
	Columbia Journalism Review's Guide to Online News Startups
	Online Newspaper Directory for the World
Multiplatform	National Directory of Ethnic Media
	Cision Media Database

community being analyzed, 3–5 interviews are conducted with individuals in the following categories: (1) local government (2) local news media; (3) activist organizations; and (4) ethnic community organizations.

Assessing journalistic output

The logical question that arises from the *infrastructure* assessment described earlier is how much journalistic output does the infrastructure generate? Thus, the *output* component assesses the aggregate journalistic output within a selected community within a specified period of time. The concern is one of quantity (the qualitative dimension is addressed in the *performance* dimension), because a reasonable indicator of the health of a local journalism ecosystem is the amount of journalism produced for and within the community. Ultimately, though, this assessment of journalistic output serves as a vital intermediate step toward conducting the more substantive assessment of journalistic performance (see later).

This stage in the process involves gathering and analyzing a sample of news content from all of the sources identified in each of the communities being analyzed. For our pilot test of this methodology (see Napoli et al., 2017), we analyzed three communities in New Jersey (US) through a sample week of content. For this small-scale effort, it was relatively easy to analyze all of the relevant content on a daily basis for each day of analysis. However, to conduct this research on a larger scale, content archiving is necessary. So, for instance, we are currently content analyzing a constructed week's worth of local news output for 100 randomly sampled U.S. communities. Sampled communities had a maximum population size of 300,000 and a minimum population size of 20,000.

For this analysis, we partnered with the Internet Archive's Archive-IT platform to conduct a custom web crawl and archive for future analysis. Even an ambitious Internet archiving project, such as the Internet Archive, does not comprehensively archive local news sources associated with all communities (particularly smaller communities). Rather, much of the crawling is random in nature. Thus, for this custom archive, we provided the Internet Archive with the URLs for the nearly 700 media outlets associated with the 100 sampled communities. We also provided them with seven randomly sampled days for our constructed week – constructed from the four upcoming months. The Internet Archive then proceeded to crawl and archive all of the homepages – as well as any pages linked by one step (one click) from those homepages – for each media outlet. This process has produced roughly 2.7 terabytes of data and included approximately 12 million archived documents (web pages, images, movies, etc.).

Each media outlet's homepage serves as its indicator of local news provision. Using website homepages as representative of an outlet's news output builds on the tradition of sampling a newspaper's front page, which is at once the most likely to be seen by readers and also represents the news outlet's judgment as to the most important news it has to offer (e.g. Benson, 2013). Homepage stories are counted/coded for each sample day only if they were posted on the sample day. Thus, for instance, if the homepage on the analyzed day of July 27, 2016, contains stories dated earlier than July 27, those stories are not included in the analysis. The goal here is to aggregate and analyze the news output produced on each selected day.

The methodological approach employed for the output dimension – and the performance dimension that follows – relies on the journalistic content available online, regardless of the outlet's "native" platform. Thus, the journalistic outputs of daily and weekly newspapers, magazines, radio stations, television stations, and local cable channels all were assessed via their online content offerings, in the same way that the outputs of online news sources such as community journalism sites were assessed.

This approach (which is a reflection of the effort to create a realistically scalable methodological approach) runs counter to the common assertion that certain types of legacy media (e.g., local weekly print publications, ethnic media outlets) remain slow to utilize the internet as a means of disseminating their content. We believe that we are at a point in the evolution of legacy media and their place within the broader media ecosystem that this generalization likely no longer holds true. The economic and strategic pressures and incentives to have an online presence, combined with the inherent economic imperative to distribute content production costs across as broad an audience base as possible (Hamilton, 2004), suggest that the content available online can serve as a reliable *indicator of the relative journalistic output* across individual outlets, regardless of their "native" platform. The key term here is *indicator*, as we are not seeking to produce a comprehensive inventory of journalistic output, only a set of indicators that are conceptually and methodologically robust; that can be employed in comparative analyses across communities or over time; and that can be scaled to multiple communities at reasonable cost. We have found that, within the 100 communities currently being analyzed, incidences of a local media outlet lacking an online presence are rare.

Assessing journalistic performance

At the *performance* level, the goal is to provide measures of the extent to which the output serves the information needs of local communities. Thus, we focused on three criteria: (1) whether the content is original; (2) whether it is about the local community; and (3) whether it addresses communities' critical information needs (CINs).

These criteria are clearly rough and superficial indicators of the complex notion of the "quality" (Lacy and Rosenstiel, 2015) of the journalism being produced. However, these criteria provide a relatively simple, economical, straightforward, and replicable set of indicators of journalistic performance. They also address the fundamental concern about whether journalistic sources are addressing each communities' critical information needs.

The notion of "critical information needs" has been central to the ongoing discourse about the performance of local journalism (Knight Commission, 2009; Waldman, 2011). The approach employed here draws upon this discourse and the research it has inspired (e.g., Becker and Yanich, 2015; Friedland et al., 2012). Specifically, each story is content analyzed to determine whether it fits into one or more of the critical information needs categories identified in Friedland et al.'s (2012) comprehensive literature review prepared for the FCC. Friedland et al. (2012) provide eight categories of community critical information needs. These categories are as follows (see Friedland et al., 2012, for detailed descriptions that provided the basis for each category's coding criteria):

1 Emergencies and risks
2 Health
3 Education
4 Transportation systems
5 Environment and planning
6 Economic development
7 Civic information
8 Political life

These categories provide a comprehensive and relatively straightforward schema for content analyzing local news stories/posts in a way that could certainly be used to explore differences in critical information needs across communities and differences in the extent to which local journalism sources are addressing specific critical information needs. These are intended as universal categories and do not reflect the fact that communities certainly differ in terms of the relative importance of each of these categories. For the purposes of our analysis, we are primarily concerned with the basic binary of whether a story addresses a critical information need, in order to compute metrics related to the proportion of journalistic output that addresses critical information needs (see later).

Each story is also coded for whether it is *original* (i.e., produced by the journalism outlet rather than reprinted or shared from elsewhere) and whether it is *about the local community*. The emphasis on original stories serves to separate aggregation, linking, sharing, retweeting, and re-publication activities, in an effort to determine the amount of original journalism output being produced. This measure is intended to tap at the robustness of local journalism sources by determining how active they are in *producing* news stories. A website story is considered 'original' if it has a byline by an outlet's reporter or if it lacks any indication that it is a reposting of content originally produced elsewhere.

The emphasis on locality seeks to analyze the extent to which the output of local journalism sources focuses on the local community. This measure taps into the extent to which local journalism is truly local, in the sense that stories provide community members with news and information about, and directly relevant to, their communities. The narrow measure of local content reflects the long-standing *localism* principle, which has featured prominently in democratic theory perspectives on media and in media policymaking (see Napoli, 2001). From this perspective, the extent to which citizens are engaged with and capable of informed democratic participation in their communities is a function of the availability of local news and information about their communities (see George and Waldfogel, 2006). For this variable, we utilize a strict definition of community, where we identified an item as about the community only if the subject was an issue/event oriented around the specific town.

Together, these three variables of focus reflect some of the primary concerns about the state of local journalism today: (1) that the economic pressures on local journalism create overwhelming incentives to aggregate and repurpose existing content rather than engage in original reporting (see Anderson et al., 2012; Doctor, 2010); (2) that the changing technological dynamics for news distribution and consumption are exacerbating the extent to which large-market or out-of-market news can infiltrate local communities (George and Waldfogel, 2006); and (3) that the increasing challenges associated with attracting and retaining an audience for news are compelling news outlets away from substantive topics in favor of an emphasis on "soft" news, celebrity, and sensationalism (see Patterson, 2013). For these reasons, these three variables represent a useful set of top-level indicators of how well local journalism facilitates informed participation and engagement in local community affairs.

Trained coders conduct the content analysis with the goal of achieving high levels of intercoder reliability. Google Translate facilitates coding of foreign language content. Subscriptions were obtained for any sites whose current content is behind a paywall making them inaccessible to the Internet Archive's web crawler.

Metrics and analysis

The data gathered here can be employed to facilitate multivariate analysis and/or to facilitate more basic comparative analyses across communities. From a multivariate standpoint, we are primarily interested in whether it is possible to identify those characteristics of individual communities that are significantly related to the robustness of local journalism. Essentially, we are interested in identifying the characteristics of 'at-risk' communities. For this analytical approach, the raw counts of outlets, story output, and number/proportion of stories that are original, are local, and that address a critical information need serve as the relevant dependent variables.

Another indicator of the robustness of local journalism that we have constructed with these data is a concentration index that indicates the extent to which the journalistic output in a community is emanating from many or few sources. This measure is derived from the well-known Herfindahl-Hirschman Index (HHI). The traditional HHI is calculated by summing the squared shares of each firm in a market to produce a measure of concentration. It is expressed as follows: $H = \sum_{i=1}^{N} s_i^2$. In the case of this analysis, shares of total journalism output within a community are used in place of market shares. This HHI can also be calculated to focus exclusively on specific categories of journalistic output, such as original news stories, local news stories, or stories addressing a critical information need. These categories can be analyzed individually or in combination (e.g., how concentrated is the output of news stories that are original and local and that address a critical information need?).

Such output concentration measures reflect the historical significance of source diversity in the journalistic sphere as well as the frequent use of the HHI as a tool for assessing diversity in media (see Napoli, 2001). The extent to which journalistic output is concentrated within few sources would seem to be a relevant indicator of the health of local journalism ecosystems. The utility of the HHI in this context is purely comparative. While in economics variations in the HHI have been found to be associated with variations in the behavior of firms (thus, an HHI of 1,300, for instance, has a specific meaning), here the HHI is being used as a comparative metric with the individual values, having no inherent interpretation.

We are particularly interested in determining if, or to what extent, the robustness of local journalism is a function of community characteristics such as average income, ethnic diversity, and proximity to a large media market. Based on indications from the pilot research (see Napoli et al., 2017), we are curious as to whether these factors may be negatively related to the robustness of local journalism. In conducting such an analysis, it is also important to take into account relevant controls such as population size and density as well the geographic size of the community. We have gathered such data for each community being analyzed.

For more descriptive, comparative analyses across communities, obviously comparing raw numbers in terms of infrastructure, output, and performance is problematic, given the large degree of variance that is likely to exist across key variables and the challenges associated with measuring those variables. We are currently conducting multivariate analyses in order to identify the best approach for applying controls across communities in a way that facilitates meaningful apples-to-apples comparisons.

In addition, social network analysis provides a means to measure the robustness and resilience of a local news ecosystem. Prior work has, for instance, shown that network analysis can be used to effectively map the connections between sources within a given locality (Weber and Monge, 2011; Anderson, 2013). Moreover, connections between local news sources provides a means for understanding connectivity and for mapping the sharing of information (Lowrey and Erzikova, 2016). In addition, the sharing of content and the use of hyperlinking to connect online local news sources indicates a level of competition and robustness that is generally not present when websites are disconnected from one another. Indeed, some have suggested that reciprocity is a hallmark of local news, and local news sources are reliant on the flow of information from a given community (Harte et al., 2017); it is plausible to expect the same principle to extend to connectivity amongst websites, which is key to building robustness in online communities.

The aggregated web data from each community provides a basis for creating and measuring social networks within local news communities. A social network can be created by capturing the hyperlinks that exist between websites. This can be done manually or automatically using computational methods (see Lin et al., 2014). In addition, variables capturing specifics of a given local news outlet can be added to enhance the analysis (Weber, 2012). It is important when using social network analysis to consider what counts as a "connection" between two websites; often, it is useful to look beyond the existence of a single hyperlink and to consider more stringent definitions of connection. Subsequently, a broad number of measures can be used to assess the strength of a local news community – for example, considering the issue of robustness, the network concept of embeddedness, which captures the degree to which an entity is embedded within its broader community. In network terms, this can be measured in a number of ways, including density of ties or clustering. Resilience is the degree to which a network of entities is able to withstand disruptions and is closely connected to embeddedness. Resilience can further be captured by higher level measures such as triads, or the degree to which entities are connected to more than one other website.

Network analysis provides another lens through which local news can be examined and assessed. Focusing on connections between local news sources provides a means for understanding the degree to which a community of information exists. Moreover, by loosening the definition of local, it is also possible to examine the degree to which information is restricted to a single locale or flowing across local boundaries.

Conclusion

The methodology outlined here is intended to facilitate both explanatory analyses of the factors related to the robustness of local journalism and descriptive, longitudinal analyses that facilitate the monitoring of trends over time. In terms of the former, expansion to a larger, more diverse sample of communities (incorporating, for instance, communities larger than 300,000 residents and smaller than 20,000 residents) would be beneficial in terms of enhancing generalizability. In terms of the latter, expansion in terms of systematic, annual data-gathering would be essential for longitudinal analysis.

Despite the fairly superficial nature of the content analysis, data-gathering at the scale presented here is resource intensive and time consuming. Going forward, one key question is whether the content analysis process could be automated. Automating the process would certainly make expansion and/or longitudinal application of the methodology more feasible. Given the fairly superficial nature of the content variables being analyzed and advances in machine learning and textual analysis, this would seem to be an avenue worth pursuing. The data gathered thus far could serve as the basis for developing, refining, and testing such an approach.

This research initiative has resonated with content archivists at the Internet Archive and the Library of Congress, who recognize the importance of local journalism and understanding the factors that affect its evolution, and who also recognize that local journalism, at the level of granularity being analyzed here, largely slips through the cracks of most web archiving efforts. For this reason, the Internet Archive has taken it upon itself, independently of our research project, to archive the first day of every month for all of the URLs associated with our 100 sampled communities. The Library of Congress is also exploring initiating a more systematic effort into archiving local journalism.

The content archive upon which our research is based could support a wide range of additional research inquiries. Indeed, our intention is to make this archive publicly available to facilitate additional research. For instance, one researcher interested in crime coverage is exploring the possibility of using the data to explore the relationship between local crime statistics and local crime coverage. Such research, along with our ongoing research, focuses on the community as the unit of analysis. Of course, these data could facilitate more granular units of analyses including the individual media outlet or story. From such perspectives, it would be possible, for example, to explore the relationship between outlet characteristics (size, medium, ownership, etc.) and journalistic output as a way of enhancing our understanding of contemporary news organizations. These are just some of the possible ways that this archive of local journalistic output could be utilized in future research. And, of course, were this archive to become longitudinal in nature, the range of research possibilities may be enhanced further.

Further reading

Researchers are devoting increasing attention to local journalism. A key conceptual framework for this research is the notion of media deserts, which has its origins in the work of Michele Ferrier (2013; see also Ferrier et al., 2016). For an international perspective, Rasmus Kleis Nielsen's (2015) edited volume *Local Journalism: The Decline of Newspapers and the Rise of Digital Media* is a wide-ranging and valuable collection. For a focus on strategic approaches to preserving the economic viability of local journalism, Penelope Abernathy's (2014) book *Saving Community Journalism: The Path to Profitability* is a valuable resource. For those with an interest in local news audiences and how they go about navigating their local news environments, the News Measures Research Project has recently published a qualitative analysis based upon focus group research conducted in a number of New Jersey communities (McCollough et al., 2017).

References

Abernathy, P. M. (2014) *Saving Community Journalism: The Path to Profitability*. Chapel Hill, NC: University of North Carolina Press.

Anderson, C. W. (2013) *Rebuilding the News: Metropolitan Journalism in the Digital Age*. Philadelphia, PA: Temple University Press.

Anderson, C. W., Bell, E. and Shirky, C. (2012) *Post-Industrial Journalism: Adapting to the Present*. Retrieved from http://towcenter.org/research/post-industrial-journalism/

Baker, C. E. (2002) *Media, Markets, and Democracy*. New York, NY: Cambridge University Press.

Barnett, S. and Townend, J. (2015) "Plurality, policy and the local: Can hyperlocals fill the gap?" *Journalism Practice*, 9(3), 332–349.

Becker, A. and Yanich, D. (2015) *Citizens' Critical Information Needs and Local TV News*. University of Delaware Center for Community Research and Service Working Paper. Retrieved from www.ccrs.udel.edu/sites/ccrs.udel.edu/files/BeckerYanich%20Broadcast%20News%20CIN%20Model%2030%20July%202015_0.pdf

Benson, R. (2013) *Shaping Immigration News: A French-American Comparison*. Cambridge, UK: Cambridge University Press.

Caves, R. E. (1992) *American Industry: Structure, Conduct, Performance*. New York, NY: Prentice Hall.

Community Health Status Indicators Project Working Group. (2009) *Data Sources, Definitions, and Notes for CHSI2009*. Department of Health and Human Services, Washington, DC. Retrieved from http://communityhealth.hhs.gov

Doctor, K. (2010) *Newsonomics: Twelve Trends that Will Shape the News You Get*. New York, NY: St. Martin's Press.

Downie, Jr., L. and Schudson, M. (2009, November/December) "The reconstruction of American journalism." *Columbia Journalism Review*. Retrieved from www.cjr.org/reconstruction/the_reconstruction_of_american.php

Durkin, J. and Glaisyer, T. (2011) *An Information Case Study: Scranton*. Retrieved from http://oti.newamerica.net/publications/policy/Scranton

Durkin, J., Glaisyer, T. and Hadge, K. (2010) *An Information Case Study: Seattle*. Retrieved from http://oti.newamerica.net/publications/policy/seattle

Eggerton, J. (2014, February 25) "House to hold hearing on FCC study." *Multichannel News*. Retrieved from www.multichannel.com/news/policy/house-hold-hearing-fcc-study/307382

Ferrier, M. (2013) "Interview with Debra Galant." *NJ News Commons*. Retrieved from www.youtube.com/watch?v=nK9OqINpyys

Ferrier, M., Sinha, G. and Ourich, M. (2016) "Media deserts: Monitoring the changing media ecosystem." In M. Lloyd and L. A. Friedland (eds.), *The Communication Crisis in America, and How to Fix It*. New York, NY: Palgrave Macmillan (pp. 215–233).

Fico, F., Lacy, S., Baldwin, T., Bergan, D., Wildman, S. S. and Zube, P. (2013) "Newspapers devote far less coverage to county government than to city governance." *Newspaper Research Journal*, 34(1), 104–111.

Flint, J. (2014, February 28) "FCC cancels study after backlash from lawmakers." *LA Times*. Retrieved from www.latimes.com/entertainment/envelope/cotown/la-et-ct-fcc-cin-study-20140228-story.html

Friedland, L., Napoli, P. M., Ognyanova, K., Weil, C. and Wilson, E. J. (2012) *Review of the Literature Regarding Critical Information Needs of the American Public*. Retrieved from http://transition.fcc.gov/bureaus/ocbo/Final_Literature_Review.pdf

George, L. and Waldfogel, J. (2006) "The *New York Times* and the market for local newspapers." *American Economic Review*, 96(1), 435–447.

Gloria, K. and Hadge, K. (2010) *An Information Community Case Study: Washington, D.C.* Retrieved from http://oti.newamerica.net/publications/policy/an_information_community_case_study_washington_dc

Hamilton, J. T. (2004) *All the News That's Fit to Sell: How the Market Transforms Information Into News*. Princeton, NJ: Princeton University Press.

Harte, D., Williams, A. and Turner, J. (2017) "Reciprocity and the hyperlocal journalist." *Journalism Practice*, 11(2–3), 160–176.

Holt, K. and Karlsson, M. (2015) "'Random acts of journalism?': How citizen journalists tell the news in Sweden." *New Media & Society*, 17(11), 1795–1810.

Karlsson, M. and Holt, K. (2014) "Is anyone out there? Assessing Swedish citizen- generated community journalism." *Journalism Practice*, 8(2), 164–180.

Knight Commission. (2009) *Informing Communities: Sustaining Democracy in the Digital Age*. Report of the Knight Commission on the Information Needs of Communities in a Democracy. Retrieved from www.knightcomm.org/read-the-report-and-comment/

Lacy, S. and Rosenstiel, T. (2015) *Defining and Measuring Quality Journalism*. News Measures Research Project Working Paper. Retrieved from http://mpii.rutgers.edu/wp-content/uploads/sites/129/2015/04/Defining-and-Measuring-Quality-Journalism.pdf

Leibowitz, J. (2009) *'Creative Destruction' or Just 'Destruction': How Will Journalism Survive the Internet Age?* Opening Remarks for FTC News Media Workshop. Retrieved from www.ftc.gov/sites/default/files/documents/public_statements/creative-destruction-or-just-destruction-how-will-journalism-survive-internet-age/091201newsmedia.pdf

Lin, J., Gholami, M. and Rao, J. (2014, April) "Infrastructure for supporting exploration and discovery in web archives." *Proceedings of the 23rd International Conference on World Wide Web*, ACM (pp. 851–856).

Lin, W. and Song, H. (2006) "Geo-ethnic storytelling: An examination of ethnic media content in contemporary immigrant communities." *Journalism*, 7(3), 362–388.

Lowrey, W. and Erzikova, E. (2016) "News media ecosystems and population dynamics: A cross-cultural analysis." In *Advancing Media Production Research*. Basingstoke, UK: Palgrave Macmillan (pp. 79–94).

McCollough, K., Crowell, J. K. and Napoli, P. M. (2017) "Portrait of the online local news audience." *Digital Journalism*, 5(1), 100–118.

Morgan, F. (2011) *The Research Triangle, North Carolina: A Region of Locally Owned Media Outlets and Entrepreneurs on the Verge*. New America Foundation. Retrieved from http://oti.newamerica.net/publications/policy/the_research_triangle_north_carolina

Morgan, F. (2013) "The stories not told: A case study of the information needs of Siler City, North Carolina." *I/S: A Journal of Law and Policy for the Information Society*, 8(3), 481–529.

Napoli, P. M. (2001) *Foundations of Communications Policy: Principles and Process in the Regulation of Electronic Media*. Cresskill, NJ: Hampton Press.

Napoli, P. M. and Friedland, L. (2016) "US Communications Policy Research and the Integration of the Administrative and Critical Communication Research Traditions," *Journal of Information Policy*, 6, 41–65.

Napoli, P. M. and Stonbely, S. (in press) "Policy issues surrounding journalism." In J. Nussbaum (ed.), *The Oxford Research Encyclopedia of Communication*. New York, NY: Oxford University Press.

Napoli, P. M., Stonbely, S., McCollough, K. and Renninger, B. (2017) "Local news and the information needs of local communities." *Journalism Practice*, 11(4), 373–395.

Nielsen, R. K. (ed.). (2015) *Local Journalism: The Decline of Newspapers and the Rise of Digital Media*. London: I.B. Tauris & Co.

Pai, A. (2014, February 10) "The FCC wades into the newsroom." *Wall Street Journal*. Retrieved from www.wsj.com/articles/SB10001424052702304680904579366903828260732

Patterson, T. E. (2013) *Informing the News: The Need for Knowledge-Based Journalism*. New York, NY: Vintage.
Pew Research Center. (2010) *How News Happens: A Study of the News Ecosystem of One American City*. Retrieved from www.journalism.org/2010/01/11/how-news-happens/
Pew Research Center. (2015) *Local News in a Digital Age*. Retrieved from www.journalism.org/files/2015/03/PJ_MediaEcology_030515_pdf.pdf
Picard, R. (2014) "Twilight or new dawn of journalism? Evidence from the changing news ecosystem." *Digital Journalism*, 15(5), 500–510.
Ramos, D., Gunes, M. H., Mensing, D. and Ryfe, D. (2013) "Mapping emerging news networks: A case study of the San Francisco Bay Area." *Studies in Computational Intelligence*, 424, 237–244.
Ryfe, D., Mensing, D., Ceker, H. and Gunes, M. (2012, April) "Popularity is not the same thing as influence: A study of the Bay Area news system." Paper presented at the *International Symposium on Online Journalism*, Austin, TX. Retrieved from https://online.journalism.utexas.edu/2012/papers/Mensing.pdf
Schudson, M. (2005) "The US model of journalism: Exception or exemplar." In H. de Burgh (ed.), *Making Journalists: Diverse Models, Global Issues*. New York, NY: Routledge.
Social Solutions International. (2013, April). "Research design for the multi-market study of critical information needs." *Prepared for the Federal Communications Commission*. Retrieved from www.fcc.gov/encyclopedia/research-design-multi-market-study-critical-information-needs
Starr, P. (2009, March 4) "Goodbye to the age of newspapers (hello to a new era of corruption)." *The New Republic*. Retrieved from www.newrepublic.com/article/goodbye-the-age-newspapers-hello-new-era-corruption
Stites, T. (2011, December 8) "Layoffs and cutbacks lead to a new world of news deserts." *Nieman Lab*. Retrieved from www.niemanlab.org/2011/12/tom-stites-layoffs-and-cutbacks-lead-to-a-new-world-of-news-deserts/
St. John III, B., Johnson, K. and Nah, S. (2014) "Patch.com: The challenge of connective community journalism in the digital sphere." *Journalism Practice*, 8(2), 197–212.
Sustainable Jersey. (2013) *A Better Tomorrow, One Community at a Time*. Retrieved fromwww.sustainablejersey.com/fileadmin/media/Grants_and_Resources/Green_Team_Resources/Program_Overview/Sustainable_Jersey_Overview_Brochure.pdf
Tankersley, J. (2015, April 23) "Why the PR industry is sucking up Pulitzer winners." *The Washington Post*. Retrieved from www.washingtonpost.com/blogs/wonkblog/wp/2015/04/23/why-the-pr-industry-is-sucking-up-pulitzer-winners/
U.S. Senate. (2009) "The future of journalism." Hearing before the *U.S. Senate Committee, on Commerce, Science, and Transportation, Subcommittee on Communications, Technology, and the Internet*. 111th Congress, first session, May 6.
Waldman, S. (2011) *The Information Needs of Communities: The Changing Media Landscape in a Broadband Age*. Federal Communications Commission, Washington, DC. Retrieved from www.fcc.gov/infoneedsreport
Weber, M. S. (2012) "Newspapers and the long-term implications of hyperlinking." *Journal of Computer-Mediated Communication*, 17(2), 187–201.
Weber, M. S. and Monge, P. (2011) "The flow of digital news in a network of sources, authorities, and bubs." *Journal of Communication*, 61(6), 1062–1081.
Williams, A., Harte, D. and Turner, J. (2015) "Filling the news hole? UK community news and the crisis in local journalism." In R. K. Nielsen (ed.), *Local Journalism: The Decline of Newspapers in the Rise of Digital Media*. London: I.B. Tauris (pp. 203–223).