Chapter Three: Researching Community Information Needs

Katherine Ognyanova

katya@ognyanova.net

Rutgers University

The complex processes of information packaging, dissemination, and consumption in a contemporary communication ecology have been examined in a variety of disciplines employing a wide range of theoretical and methodological frameworks. This chapter provides a brief overview of major theory-driven approaches contributing to the evaluation of community information needs (CIN) and the extent to which those needs are met in a given context. The frameworks outlined here are grounded in (1) communication ecology and multilevel approaches, (2) economics research, market and audience analyses, (3) information inequality and digital exclusion, (4) mass communication and content analysis, and (5) computational social science and network analysis. While this list is by no means exhaustive, much of the relevant academic, industry, and policy research falls into one or more of those five categories.

Communication ecologies and community research

Understanding a community is a critical step in identifying and evaluating individual and group information needs. Research in that area builds on a long tradition of studies exploring the social fabric of neighborhoods and its implications for civic engagement (Putnam 2000) – a theoretical focus that has retained its relevance in the age of digital communication.

A defining feature of studies in this tradition is the multilevel approach examining nested structures – from individuals, through families, to neighborhoods and cities, combined with a high-level view on the forces that shape an urban environment. Those investigations often use multifaceted mixed-method designs combining quantitative and qualitative analysis.
Mixed-method studies of metropolitan, neighborhood, and interpersonal communication processes are in the core of major ecological projects conducted in a wide range of communities. Research efforts in the U.S. have explored communication and media ecologies in large metropolitan areas like Los Angeles (Chen et al. 2013), Philadelphia (Schaffer 2010), Seattle (Friedland 2013) and Baltimore (Pew Research 2010); as well as mid-sized and smaller cities like Madison, Wisconsin (Friedland et al. 2007) Macon, Georgia (Mitchell et al. 2015), New Brunswick, Newark, and Morristown in New Jersey (Napoli et al. 2015).

Sampson’s book *Great American City* (2012) grounded in the Chicago School tradition of urban sociology provides one of the most comprehensive recent examples of ecological community research. The analyses reported in the volume are based on eight years of cohort studies, community surveys, systematic social observation, experiments, and multi-wave network analysis seeking to identify community leaders.

Another set of studies explicitly adopting the multilevel ecological approach to urban neighborhoods is grounded in *communication infrastructure theory* (Ball-Rokeach and Jung 2004, Chen et al. 2013). This framework examines community information needs in the context of larger social systems incorporating demographic and institutional processes. The media ecology of residents is seen as a subset of their total connections to communication resources, both interpersonal and organizational. A central concept in that research is the *neighborhood storytelling network*: a system encompassing residents, local media, and organizations, as well as the connections within and among them (Kim, Jung, and Ball-Rokeach 2006). A strong storytelling network was found to enhance civic outcomes including engagement, collective efficacy, and neighborhood belonging.

Focusing on the interplay between social relations and space in an urban environment, Friedland (2001) advances the *communicatively integrated community* framework which encompasses power relations and communicative action at a global, regional, metropolitan, and local level. His follow-up works in this ecological line of research examine the interplay between community and media over time (Friedland et al. 2007), as well as the interactions between civic and communication ecologies (Friedland 2013).

Some recent efforts to explore the information needs of communities have focused on an ecological understanding of journalism. One example comes from *The Media + the Public Interest*
Initiative, exploring three cities in the state of New Jersey (Napoli et al. 2015). The parameters examined in that research include the local journalism infrastructure, encompassing news sources and their social media presence; as well as the journalistic output and performance, measuring the number of news stories and posts, and their ability to address the critical information needs of the community.

The efforts to examine community information needs through ecological studies expand beyond the realm of academia. In a high-profile report, the Knight Commission urged researchers, policy makers, and organizations to pursue three key objectives: (1) making relevant and credible information available to all, (2) strengthening the capacity of individuals to engage with information by providing access to needed tools and skills, and (3) promoting engagement with both local information and the public life in communities, leading to a more effective self-governance. Following that call, the FCC commissioned two comprehensive reports exploring the U.S. media landscape and the critical information needs of Americans (Waldman 2011, Friedland et al. 2012).

The Pew Research Center is another organization evaluating the local news ecologies in a diverse set of U.S. communities (Rosenstiel et al. 2011, Mitchell et al. 2015). Their most recent project focusing on local news in three metro areas includes six studies that (1) identified local news providers; (2) surveyed local residents about their news consumption; (3) analyzed and coded news stories produced in the target areas; (4) conducted interviews with residents, journalists, local officials and businesses; (5) explored the news and information carried by social media platforms, and (6) examined city-level variables including geography, population size, demographics, and broadband penetration.

One important aspect of community research that deserves to be mentioned separately involves the role of technology. Online and mobile platforms, social media and participatory digital spaces have become a crucial part of the communication infrastructure of American communities. Scholars have developed a variety of conceptual frameworks explaining the profound changes in social networks and information flows in a digital age (Hampton 2015, Wellman and Rainie 2012, Castells 1996). Reflecting the critical importance of digital communication patterns, recent works evaluating the health of community news and information systems almost without exception incorporate investigations of the availability and access to digital content.
Economic research: market, and audience analyses

Works in this tradition explore the demand and supply dynamics that influence the availability of news and information in local communities. Economic studies of the media system have examined the effects of audience preferences, advertising demands and competition (Dimmick 2003, McManus 1994). Media outlets vary in size, location, projected identity, social context, target audiences, political orientation, production technologies, available resources and ties to other organizations. All of those characteristics – and more – affect content production (Allern 2002). This has clear implications for informed citizenry as mainstream media still have a leading role in producing current affairs news and shaping public opinion (Shehata and Stromback 2013).

The economic approach has been used to examine the factors that cause underproduction of local, political, and public affairs coverage. In a comprehensive investigation of the market forces underpinning the media business, Hamilton (2004) describes the economic characteristics of news products as information goods. His book discusses the complicated realities of commercial media models that need to satisfy both advertiser demands and consumer preferences. Story coverage is tailored to maximize its appeal to key demographic groups: those who are most likely to buy the advertised products. When target audiences place low value on hard news, media outlets have an incentive to reduce current affairs and political reporting in favor of entertainment and sports coverage.

Works in that line of research have explored the impact of economic factors like ownership and market structure on a variety of outcomes including individual news consumption (Althaus, Cizmar, and Gimpel 2009), political behavior and election turnout (Althaus and Trautman 2008), the availability of local news (George and Waldfogel 2006), the quality of political coverage (Dunaway 2008), and the political slant of news sources (Gentzkow and Shapiro 2010).

One central theme in economic studies of the media system is the relationship between markets and media diversity. Diversity is a major regulatory concern, as maintaining a pluralism of voices in the media is essential for a healthy democracy. The diversity principle can be seen as comprising three separate measurable components: source, content, and exposure diversity (Napoli 1999).

The source dimension refers to the ownership and workforce diversity of media organizations and program producers. The content component examines the range of programing available to audience members, and its diversity in terms of type, target audience, and represented viewpoints.
The *exposure* dimension shifts the focus from the content that is made available to the content that people actually consume. Metrics of that type evaluate the number and type of outlets and programs selected by the public, as well as the range of viewpoints presented by those outlets.

The Federal Communications Commission (FCC) has launched a number of efforts to evaluate media diversity. One such effort involved the development of a Diversity Index for local media markets meant to serve as an evaluation instrument in media ownership regulation. The index sparked controversy and was eventually challenged in court and its use was suspended (Lloyd and Napoli 2007). The Commission’s further efforts to examine the critical information needs of U.S. communities were put on indefinite hold after coming under strong criticism from several media outlets and members of Congress.

Another key theme here is audience fragmentation (Napoli 2011): the idea that as information gets increasingly personalized, mass audiences may dissolve into small isolated groups. Scholars have predicted a coming era of *cyberbalkanization* (Sunstein 2007) and filter bubbles (Pariser 2011). Empirical research, however, has found no conclusive evidence to support those predictions. In a set of studies unpacking audience fragmentation across traditional and online news sources, Webster (2014) finds high levels of duplication across media outlets and no evidence of isolation in like-minded consumption groups.

**Information inequality and digital exclusion**

Evaluating information inequalities is crucial for any effort seeking to identify the complex social, economic, and technological factors that come together to produce informed communities. Contemporary debates of inequality in this context incorporate three major aspects: (1) predictors and patterns of disparity in the production of information relevant to different social groups; (2) inequalities in the quantity and quality of access to information and participation; and (3) differences in training, skills, and digital literacy levels. Some of the major parameters defining the information divides in the U.S. today include geography, language, age, income, education, race and ethnicity, immigration status, disability, gender, and sexual orientation.

Works evaluating the inequality in content production systems have often focused on the role of female and minority ownership and employment in media organizations. Policy-relevant research has demonstrated a link between the quality and quantity of content directed at different demographic groups and their presence in the workforce and management of news companies.
(Bachen, Hammond, and Sandoval 2007). Nonetheless, women and ethnic minorities remain underrepresented in many areas of the media industry, especially at the higher levels (Hunt 2014, American Society of News Editors 2015, Papper 2015). Discussing the costs of exclusion, Costanza-Chock and Wilson (2012) examine disparities in ownership in print, broadcast, and online media. While they find more diversity of ownership on the Web compared to offline outlets, their results suggest that even online sources are disproportionally infrequently owned by people of color.

Early research on information inequality focused on the digital divide – individual and community-level disparities in access to information and communication technologies. Today, the gap in access to devices and Internet connection is narrower, though recent reports still show lower penetration in rural communities, among the elderly, people with disabilities, and for those in the lowest education and income brackets (Rainie 2015). Even larger disparities remain with regard to the quality of access and technology used across groups. A recent report by the White House (Council of Economic Advisers 2015) describes a broadband gap in the U.S.: high-speed Internet services still have much lower availability, market competitiveness, and end-user penetration in rural and low-income communities.

Perhaps more importantly, inequalities remain in the areas of digital literacy, skills, and types of use. Different online activities have different outcomes, some more beneficial than others. Internet use can focus on entertainment, or it can provide a chance for advancing one’s education, career, and financial status by accruing economic, social, and cultural capital (van Deursen and van Dijk 2014).

Examining differences in digital participation, Wei (2012) found that people with lower income and education had a narrower scope of online activities and used the Internet primarily for entertainment and socializing. Women and senior citizens also took part in fewer activities on the Web. Büchi, Just, and Latzer (2015) similarly found systematic differences in types of Internet use across gender, age, and socioeconomic groups.

These digital inequalities are also associated with participation gaps. Studies examining the general impact of Internet use on political and civic engagement have reported mixed results (Boulianne 2009). We do know, however, that specific online activities and goals are linked to higher political
participation (Dimitrova et al. 2014) and efficacy (Ognyanova and Ball-Rokeach 2015). This makes the demographic disparities in online activity patterns particularly consequential.

**Mass communication and content analysis**

Compared to ecological, economic, and inequality studies, classic mass communication research has been less prominent in the policy conversations around community information needs. Works in that tradition, however, can make substantive contribution to the debate by illuminating important aspects of media production and content. This section provides a few select examples of relevant theoretical frameworks, and discusses the use of media content analysis across research traditions.

While economic studies investigate the market forces behind news production, the mass communication literature explains how media content is shaped by journalistic standards and practices. *Gatekeeping theory*, for instance, explores how the vast number of potential news stories gets selectively narrowed down to the coverage actually carried by news media. A major contribution of gatekeeping research is the identification of multiple critical points where news selection happens. Shoemaker and Vos (2009) describe five levels where gatekeeping processes may occur. The *individual level* refers to effects coming from the demographic and personal characteristics of news workers. The *routine level* deals with the prevailing practices and standards of journalistic work. The *organizational level* captures properties of media companies, including their ownership, structure and size. The *social institutions level* looks into external factors relevant to the media industry – audiences, advertisers, political institutions, and interest groups. Studies at the *social system* level explore gatekeeping controls imposed by a country’s economic, political, or cultural system.

Among other themes, that line of research has explored media coverage patterns under different economic conditions. Soroka (2012), for instance, uses the framework to examine the relation between the state of the economy, inflation and interest rates, unemployment, and their coverage in mainstream media. Recent gatekeeping work has also examined the content diversity and systematic bias in reporting on cable and online news (Gonzalez-Bailon et al. 2014).

*Agenda-setting* research is similarly relevant as it explores the information priorities of individuals, social groups, and media sources. In one representative study coming from that tradition, Tan and Weaver (2013) investigate the diversity of issues that received public and media attention over
time. Other works have explored the influence of online content and social media on journalism and news reporting, especially in the context of political information (Conway, Kenski, and Wang 2015, Quandt 2008).

Agenda-setting studies have traditionally relied on content analysis to measure the priority of different themes in news coverage, and on opinion polls to evaluate the public interest in a variety of topics. With the increasing importance of digital platforms, scholars have started exploring new ways of assessing audience priorities. Bastos (2015) examines the diffusion of stories through social media to identify the types of content favored by online consumers. Lee, Kim, and Scheufele (2015) propose using search engine requests as a proxy for individual interest in news topics. Their study finds the volume of Google searches for economic information over time is associated with the salience of economic issues measured through opinion polls.

A content analysis of media stories is often used to determine their topic, valance, geographic focus, political slant, and other key parameters. This classic analytical strategy of mass communication is now routinely employed by other research traditions discussed here. Scholars have used content analysis to evaluate, for instance, the relationship between ownership structure and local news content (Yanich 2010), as well as the extent to which broadcast news serve the public interest (Kaplan and Hale 2010).

While the bulk of studies using content analysis have focused on newspaper and broadcast stories, researchers increasingly use this analytical strategy to process online content, often comparing it to traditional media coverage (Carpenter 2010). The Pew research Center has been at the forefront of these efforts, with a number of projects collecting newspaper, broadcast, blog, and social media stories over time and coding them into thematic categories (Mitchell et al. 2015, Pew Project for Excellence in Journalism 2010).

Comprehensive content analysis, whether at the local or national level, has been difficult to maintain and scale as it required a large number of human coders. New methods discussed in the next section of this chapter have vastly expanded the viability of that option in large longitudinal projects.

**Computational social science and network analysis**

Advances in computational social science provide sophisticated analytical techniques that can illuminate the social, technological, and economic processes underpinning the information systems
of communities. New methodological tools facilitate the examination of large-scale digital trace data. One feature of particular relevance here is the capacity to conduct automated text, image, audio, and video analysis. While still inferior to human coding in terms of precision and flexibility, automated content analysis can work at a very large scale, producing consistent results across a variety of unstructured content. One obvious application is a thematic categorization of news stories and social media posts, facilitating the analysis of content diversity for multiple outlets over time.

In one recent example, scholars analyzed 30 thousand news stories produced over the span of 30 years to determine how media coverage and public perceptions about the economy influenced each other (Soroka, Stecula, and Wlezien 2015). Another study used a computational approach to examine a million news articles and over 5 million tweets, unpacking different aspects of the public conversation around mass shootings (Guggenheim et al. 2015). In a high-profile paper assessing the way people share and view political news on social media, Bakshy, Messing, and Adamic (2015) examined the content streams of over 10 million Facebook users. The analysis found three major factors reducing the political diversity of news consumed by Facebook users: the individual preference for friends who share your political views; the individual propensity to click on links that align with your political views; and Facebook’s news feed algorithm, which selectively filters and orders the posts and stories that users see.

This brings up another point of critical importance for research exploring patterns of information seeking and consumption on the Internet. Today, the content we see online is rarely curated by humans. Instead, it is selected and organized by sophisticated computer algorithms. They filter and order search results, social media posts, and news stories. We are presented with personalized information based on our location, device, demographics, and past behavior. This is useful and necessary as it helps us navigate the vast oceans of online content. As Bakshy et al (2015) demonstrate, however, personalization can also reduce the diversity of the information we see, decreasing our exposure to a variety of viewpoints. Algorithms can also perpetuate inequalities, as they learn, for instance, to show ads for lower-paying jobs to women compared to men (Datta, Tschantz, and Datta 2014). The lack of transparency in those systems presents a key challenge to our understanding of individual and group interactions with online information.

One important set of methods and theoretical constructs within computational social science comes from network analysis. Network thinking enables us to study the complex interactions between
media and social systems (Ognyanova and Monge 2013). Network strategies have been used to examine the structure of the media industry: a sector increasingly characterized by trends towards consolidation, collaborations, local and global partnerships (Arsenault and Castells 2008). Researchers have also explored the interplay between social ties and information consumption. Friemel (2015), for instance, uses actor-based models to examine the networks of high-school students and the influence of social contacts on individual preferences for TV programs.

Studies taking a network approach have also explored the diversity of media content and audiences. Ognyanova (2013) measures the levels of media fragmentation in a network of mainstream U.S. news outlets. Her work finds an increase in media content homogeneity over time. Webster (2014) examines how individuals connect to broadcast and online sources and finds relatively low levels of audience fragmentation.

Network methods also allow us to track the complex patterns of message diffusion through multiple channels (Aral, Muchnik, & Sundararajan, 2009). Research has examined, for instance, the spread of political and civic information over social media, and the factors predicting a user’s ability to distribute messages to a large audience (González-Bailón, Borge-Holthoefer, and Moreno 2013). Understanding the patterns of content flow across platforms is a key step in the process of evaluating the information sources and distribution channels that individuals and communities rely on.

Computational approaches are particularly useful in large-scale efforts to examine how a variety of outlets serve the information needs of the American public. Tools and techniques of this kind provide a feasible way to track how thousands of sources disseminate news stories among millions of audience members. Computational analyses enable researchers to thematically categorize news stories, evaluate their local relevance and diversity, as well as assess information spread and consumption patterns at a massive scale.

Conclusion

Each of the frameworks discussed in this chapter has well-understood advantages and drawbacks. Ecological projects provide rich information about the focal community, often highlighting key mechanisms and processes that may generalize beyond the local case. Unfortunately, multilevel mixed-method studies are also very resource-intensive and difficult to do at a large scale.
Relevant economic research has had a fairly narrow focus on ownership and market structure, overlooking key social and cultural processes. It has been, however, particularly useful in the context of policy, as its main predictor variables are most amenable to regulation.

Digital exclusion works give us a much needed look at the problematic areas and information gaps that we need to address. Yet the field is still facing serious challenges when it comes to finding consistent and relevant metrics and reliable sources of information.

Traditional mass communication frameworks throw light on important aspects of news production and content, though the discipline is still struggling to redefine itself in a digital age.

Computational methods are scalable and allow us to address complex questions – though it is not always clear whether and how answers obtained through digital trace data map onto offline concepts and activities. A computational approach, furthermore, tends to work best in combination with deep qualitative domain understanding.

Many of the interesting research efforts evaluating community information needs combine usefully multiple approaches. Economic works have occasionally taken an ecological perspective (Dimmick 2003), mass communication authors have considered ownership and market structure (Shoemaker and Vos 2009), and all frameworks include some studies using content analysis, computational, and network methods. As we seek to understand content production, dissemination, and consumption in a new information environment, combining existing frameworks with new analytical tools provides one promising direction for exploration.
References


Lloyd, M., and P.N. Napoli. 2007. Local media diversity matters: Measure media diversity according to democratic values, not market values. Washington, DC: Center for American Progress.


