Explication of Political User-Generated Content and Theorizing about Its Effects on Democracy with a Mix-of-Attributes Approach and Documenting Attribute Presence with a Quantitative Content Analysis

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Ivan B. Dylko

Graduate Program in Communication

The Ohio State University

2011

Dissertation Committee:

Professor William Eveland, Advisor
Professor Kelly Garrett
Professor Michael McCluskey
Abstract

The present study attempts to stimulate a new program of communication-effects investigations needed to catch up with the significant recent technological transformations in the communication environment. Eveland’s (2003) mix-of-attributes (MOA) approach is argued to represent a needed adjustment to how communication field investigates media effects in today’s communication environment. The MOA approach is used to explicate political user-generated content (UGC) and its five technological attributes. In this study, the most popular political UGC and traditional media Web sites are content analyzed to obtain a detailed description of the attribute presence across prominent groups of UGC Web sites, and to compare presence of attributes on UGC versus traditional news Web sites. Cluster analysis is used to develop a theoretically- and empirically-grounded classification of political UGC. Despite relatively low presence of attributes across different UGC Web sites, the study confirms usefulness of the MOA framework. Presence of attributes on traditional news Web sites suggests that the theoretical importance of the attributes might increase over time. This study advances communication-effects theory by: (1) examining the nature and potential effects of political UGC; and by (2) illustrating how MOA approach can be applied, given its strengths and weaknesses. Additional implications of results, study limitations, and directions for future research are discussed.
Dedicated to my parents
Acknowledgments

I want to thank my advisor, William Eveland, for his wise guidance, extensive support, and patience during the work on this project. I want to thank my committee members, Kelly Garrett and Michael McCluskey, for their encouragement and numerous useful recommendations that helped improve this work. I am thankful to the members of the Communication, Opinion and Political Studies group in the School of Communication for providing thorough and always helpful critique of many developing ideas related to this project. Finally, I want to thank several dedicated undergraduate research assistants who helped improve the codebook and helped code part of the data used in this research.
Vita

1999 ....................................................... Gymnasia N2

2005 ........................................................ B.A. Communication, Alcorn State University

2007 ........................................................ M.A. Communication, Ohio State University

2008 to present ........................................ Graduate Teaching Associate, School of Communication, The Ohio State University

Publications


Fields of Study

Major Field: Communication
# Table of Contents

Abstract .................................................................................................................................................. ii

Acknowledgments ................................................................................................................................. iv

Vita......................................................................................................................................................... v

List of Tables .......................................................................................................................................... ix

List of Figures ......................................................................................................................................... x

Chapter 1: Introduction ........................................................................................................................ 1

Chapter 2: Theorizing Challenges in Today’s Communication Environment ...................................... 3

Chapter 3: User-Generated Content ...................................................................................................... 12

Chapter 4: Mix-of-Attributes (MOA) Approach ................................................................................... 24

Chapter 5: Using the MOA to Explicate Political UGC Attributes ..................................................... 31

  Political Participation ......................................................................................................................... 31

    Participatory Democracy Theory ................................................................................................. 32

    Variables and Processes Affecting Political Participation .............................................................. 41

Technology ............................................................................................................................................ 44

Technological Attributes of UGC and Their Hypothesized Influence on Political Participation ........ 46

  Task-Specific Information Retrieval Performance ............................................................................ 47

  Information Environment Customizability ....................................................................................... 53

  Content Manipulability .................................................................................................................... 62

  Direct Participation Facilitation ....................................................................................................... 65
List of Tables

Table 1. Web Sites in the Sample..........................................................80
Table 2. Descriptive Characteristics for Attribute Scores on UGC versus Reference Group Web Sites ..........................................................97
Table 3. Median Scores of Each Web Site Group on Attributes..................105
Table 4. Descriptive Characteristics of Data Outliers Replaced in Cluster Analysis........112
Table 5. Cluster Analysis Procedure Validity Assessment: Comparison of Two Split-Sample Cluster Solutions.......................................................116
Table 6. Descriptive Characteristics for Attribute Scores..................................................121
Table 7. Median Scores of Web Site Clusters on Attributes.................................125
Table 8. Post-Hoc Tests of Difference among Conventional UGC Groups on Attribute Scores.144
List of Figures

Figure 1. Four Groups of Technological Features Representing Information Environment

Customizability..........................................................................................................................55

Figure 2. Frequency Distributions of Attribute Scores..............................................................100

Figure 3. Error Bars Plotting Rescaled Attribute Scores in Different Web Site Groups..........106

Figure 4. Cluster Analysis Dendrograms: Comparison of Two Split-Sample Cluster Solutions.114

Figure 5. Cluster Analysis Dendrogram Displaying Web Sites Grouped According to Their

Scores on Attributes..................................................................................................................117

Figure 6. Rescaled Frequency Distributions of Attribute Scores in Clusters..........................123

Figure 7. Error Bars Plotting Rescaled Attribute Scores by Cluster and Web Site Group ....126

Figure 8. Distribution of Web Site Groups across Three Clusters.........................................129
Chapter 1: Introduction

The present research is an attempt to stimulate a new program of communication-effects investigations that are argued to be superior to the conventional investigations in terms of theoretical nuance and long-term theoretical relevance. The primary impetus for this work is a realization by various communication-effects researchers that the field, which has progressed quite successfully in the past, has recently lost its theoretical vigor, in large part due to significant transformations in the communication environment that were unmatched by the necessary transformations in the communication field’s methodology and theory (Bennett & Iyengar, 2008; Chaffee & Metzger, 2001; Eveland, 2003). The historical context of the field’s development is described, and today’s theoretical challenges hindering the field are detailed.

This work adopts Eveland’s (2003) MOA approach and applies it to explication and theorizing about the political effects of UGC. This investigation is an attempt to illustrate the concrete theoretical and practical benefits of the MOA approach that can minimize two of today’s theoretical challenges that are stalling the advancement of the communication-effects paradigm. Advantages and disadvantages of the MOA approach are described. In addition to reiterating the methodological contribution made by Eveland through his MOA approach, and illustrating how it can be applied, this investigation also advances communication-effects theory by systematizing existing thinking about the phenomenon of political UGC and its political implications. UGC has become increasingly important and pervasive (Pew, 2010); however, scholars have yet to address the phenomenon in a systematic fashion (some exceptions include Shao, 2009 and Van Dijck, 2009). This work attempts to fill that void.
Following methodological prescriptions of the MOA approach, this study identifies and explicates five technological attributes defining the core of political UGC: task-specific information retrieval performance, information environment customizability, content manipulability, direct participation facilitation, and community orientation. The attributes are described in terms of their relationship to already familiar communication constructs, such as interactivity or customizability, and detailed conceptualizations and operationalizations of attributes are provided. Explication of the attributes’ nature facilitates theorizing about their effects on important political outcomes. Several multi-step mediation models, with attributes as independent variables, political participation as the dependent variable, and various political, social, and psychological variables (e.g., depth of information processing, aggregation of social capital) as mediators, are proposed. This study employs content analysis to examine two groups of Web sites—the most popular political UGC and traditional media Web sites—to document the degree to which the technological attributes are present in the context of each group. The content analysis makes it possible to not only obtain needed information to develop an empirically-grounded classification of political UGC (based on the presence of the mix of attributes on each UGC Web site), but also to obtain information needed for formulation of more nuanced hypotheses about political UGC Web sites’ potential effects.
Chapter 2: Theorizing Challenges in Today’s Communication Environment

Nearly a decade ago, Bimber (2003) argued that the modern communication environment is characterized by: (1) an abundance of highly cost-efficient information distribution channels; (2) the facilitation of communication among ordinary citizens; (3) the ability to cheaply distribute and store large quantities of information; (4) the multiplication of cheap information distribution and collection channels. All of these characteristics are even more pervasive today, and all have important implications for how we should view and study the communication environment. For the purposes of the present discussion, two characteristics of the communication environment need to be described in detail. First, today’s communication environment is characterized by a substantially greater diversity of media forms compared to the recent past (Prior, 2007). Over the last twenty years, we have moved from face-to-face, print, radio, and television broadcast forms of communication to numerous “mass-interpersonal” hybrids, as exemplified by blogs, social-networking Web sites, online discussion forums, and so forth. The complexity of the underlying technology appears to drive the diversity of media forms and appears to facilitate more diverse and efficient uses of those media forms. Web portals, social-network Web sites, blogs, content-sharing Web sites, collaborative-authoring Web sites and others enable users to: (1) *efficiently* create (or remix) and share their own content, as well as, promote, classify, vote and comment on content created by others; (2) *efficiently* customize their information environment by subscribing to information feeds from specific sources and getting that information in preferred formats; (3) *efficiently* obtain (news) information while establishing social relations and interacting with other users (Shao, 2009).
The second important aspect of today’s communication environment is its rapid and continual transformation, often resulting in hybridization of existing media that merge to form novel-appearing media forms: Newspaper and television Web sites often incorporate content-sharing functionality and discussion forums, discussion forums often incorporate instant messaging and blogs, blogs often incorporate content-classification and content-sharing capabilities, content-sharing Web sites often incorporate content-classification capabilities and private one-to-one communication features, and so forth.

Such diversity and the rapid pace of transformation and hybridization are all important to the field of communication, especially, its subfield of communication effects (for simplicity, communication-effects field will be used throughout this paper), because social and technological change greatly influence the field’s core models, assumptions, and research directions (Bennett & Iyengar, 2008). In this work, “media or communication effect” refers to any possible short- or long-term influence of exposure, attention, or use of a given media form (independent variable) on the outcome of interest (dependent variable). Such influence includes: direct impact of independent variable on dependent variable, indirect impact of independent variable on dependent variable through mediator variable(s), amplification or dampening of other variable(s) effects on dependent variable (i.e., moderation) (McLeod & Reeves, 1980). The current work focuses on individual-level (or micro) effects, which can be classified into “(1) opinion formation and change; (2) cognitive, (3) perceptual, and (4) behavioral” (McLeod, Kosicki, & McLeod, 2009, p. 229) categories.

These two largely Internet-driven characteristics of today’s communication environment make the present time very different from the time when Paul Lazarsfeld originated the media-effects (communication- and media-effects are used interchangeably) research tradition in the early 1930s (Bennett & Iyengar, 2008; Chaffee & Metzger, 2001; Rogers, 1997). These
characteristics make today not only an interesting and exciting, but also challenging and frustrating time to be a communication-effects scholar. Bennett and Iyengar reflected on this dynamic in the field as follows:

... the field is adrift theoretically, seldom looking back to see where foundational modern theory needs to be adapted and, in some cases, overthrown, in order to keep pace with the orientations of late modern audiences, and new modes of content production and information delivery... (p. 713)

This author does not advocate overthrowing foundational communication-effects theory or the field’s dominant methodologies. However, it seems reasonable that the field can benefit from utilizing novel approaches to supplement existing ones. The frustration with how theorizing is currently practiced in the communication-effects field expressed by Bennett and Iyengar (2008) reflects, in the opinion of this author, the fact that the elements of today’s communication environment (diversity of media forms and constant transformation and hybridization of existing media) create two serious and related theorizing challenges that hinder our field’s advancement. The first challenge relates to theorizing about the effects of using today’s media (existing media challenge), and the second challenge relates to anticipation of effects of using yet-to-be-developed media (future media challenge).

The existing media challenge refers to the difficulty of theorizing about effects of today’s media forms, primarily due to the diversity of exemplars belonging to those media forms. In other words, this challenge refers to a situation where even narrowly defined media forms are characterized by substantial heterogeneity of media outlets within these forms, which hinders theorizing about the effects of those particular media forms as relatively uniform groups.

For example, media-effects scholars studying political news blogs – a rather narrow genre of media – frequently encounter this challenge. Such research often employs survey methodology (Eveland & Dylko, 2007; Johnson & Kaye, 2004), in which participants are asked about their frequency of reading blogs, and other questions tapping various dependent variables
of interest. The correlations of their answers are subsequently assessed. The results of such media-effects investigations might be inconclusive because political news blogs are a relatively heterogeneous group. Unlike traditional news reporters, the plurality of whom attend journalism and communication programs with relatively uniform curricula and share various professional values, such as objectivity, accuracy, need to hold government officials accountable, and others (Gans, 1979; Weaver, Beam, Brownlee, Voakes, & Wilhoit, 2007, p. 44), bloggers do not go through uniform training and, therefore, might not share uniform values and abide by a uniform set of professional standards (Dylko & Kosicki, 2006; Woodly, 2008; however, see Hindman, 2009 who argued that many elite political bloggers have professional journalism background). In addition to values and standards, different blogs are characterized by different economic and operational models: One might be focused on a very narrow issue (e.g., communication of scientists with the public), updated occasionally and maintained largely as a hobby by a university professor attempting to project a certain image within the academic and professional community and to affect public intellectual discussions taking place in a society (e.g., Framingscience). Another blog might also be written by an academic, but cover a vast array of subjects, serve the function of providing supplementary income and be characterized by frequent and short posts, referring readers to online destinations that match the blogger’s ideology (e.g., Instapundit). Other blogs can resemble news organizations more than amateurish individual diaries, where original stories are reported by experienced, professional reporters working full time and supported by advertising revenue (e.g., Talkingpointsmemo, The Huffington Post). Other blogs are fully incorporated with professional news organizations and are hosted by these organizations’ Web sites (e.g., The Corner by The National Review Online, On Media by Politico). Yet, other blogs (e.g., Powerline, Crooksandliars, Dailykos) are written by political
activists trying to alter public opinion on important social issues (Davis, 2009; Ekdale, Namkoong, Fung, & Perlmutter, 2010).

Besides the business and operational models, blogs differ in style of presentation, which might be important from the media-effects perspective. For example, some blogs do not allow readers to comment on posts (e.g., Instapundit), others let only registered users comment (e.g., Hotair, The Huffington Post), whereas others allow anyone to comment (e.g., Framingscience). Some blogs routinely offer videos and photographs (e.g., Hotair, Crooksandliars), whereas others offer mostly textual content (e.g., Instapundit). Some blogs offer a large number of links to other Web sites, whereas others use hypertextuality sparingly (Herring, Scheidt, Kouper, & Wright, 2007). Some blogs actively attempt to recruit their readers to get engaged in the political process (e.g., Blogforamerica, Crooksandliars), whereas others do not explicitly encourage political participation (e.g., Framingscience).

All of these differences in operational models, goals, and style might result in different usage of these blogs. For example, blogs written by partisans might attract politically interested citizens who can use them to find mobilizing information (Kerbel & Bloom, 2005; Lemert, 1984; Trippi, 2005) and opportunities to get politically involved. Blogs written by academics and which are more policy, rather than politics, oriented might be used by readers to develop a thorough and multi-faceted understanding of various issues. Blogs written by professional journalists in the context of professional news organizations might be similarly relied upon by readers for learning about competing perspectives on different issues. Blogs written by partisan activists and other individuals might be used for reinforcement of one’s own views, rather than for encountering counter-attitudinal information. As a result of different usage (and possibly different content), the effects of blog reading should also predictably vary, with some readers developing greater political knowledge, others developing more extreme political attitudes, and, yet others becoming
more politically involved. To sum up, the above discussion makes it obvious that the blogosphere is an extremely heterogeneous environment and that treating all political news blogs as a single phenomenon appears to be theoretically untenable, especially in investigations concerned with the effects of blog usage. Similar theoretical challenges are encountered when theorizing about other types of increasingly diverse media forms, hindering theory building in communication-effects area.

Although treating any particular medium as homogenous is arguably more problematic today than in the past (due to an unprecedented diversity of today’s media), this problem existed and was recognized before. For example, researchers from various disciplines have long recognized that individuals approach media with different motivations that are often anchored in the nature of the medium itself (Kosicki & McLeod, 1990; Rayburn & Palmgreen, 1984; Salomon, 1983, 1982).

The second theorizing challenge – the future media challenge – refers to conventional communication-effects theorizing being rather ill-suited for predicting and explaining effects of media forms that have not yet been developed. To understand this challenge, it is helpful to examine the history of the communication-effects field. When the field originated, in the 1930s (Rogers, 1997), communication technology development was relatively incremental, especially compared to the present era. Media were relatively stable, and whereas it was not completely unimaginable that new media forms would develop, it was justifiably hard to imagine that some of the existing media might disappear, drastically change, or be replaced by new media. Even if such possibilities were entertained, scholars could be assured that such changes would not take effect soon enough to influence their programs of research, or fundamentally alter existing research paradigms.
Since the beginning of the 20th century the rate of the communication technology inventions has been steadily accelerating. It also appears that the communication technology transformation rate has been steadily accelerating. This claim does not appear to be empirically confirmable. However, it is based on several plausible assumptions: (1) recent time has seen a reduction of cost and improvement in quality of communication, which facilitates sharing and development of knowledge (Bimber, 2003); (2) the recent era has seen an increase in computing power, which reduces the “research-and-development” cycle, facilitates technological breakthroughs, and accelerates the rate of introducing new products to consumers; and (3) both assumptions mentioned above speed up the transformation of communication technology. As a result, it appears that in the past, communication scholars had fewer tangible incentives to develop models that could readily incorporate new media or predict new media’s effects. This situation was further exacerbated in the early stages of the media-effects research tradition, where the field was primarily concerned with practical problems (e.g., how to persuade citizens to support America’s involvement in WWII; determining the effects of TV violence; determining the most effective strategies for health communication campaigns), rather than with basic research focused on the creation of theoretical models of communication that would be able to incorporate future media forms, and be not so heavily focused on the media forms that were prominent during different historical periods (McLeod & Reeves, 1980). Although solving applied problems is still a very important function and activity taking place in communication-effects research, our field has become more mature as a science, and a strong argument can be made that the situation is different today than during the early stage of the communication-effects research tradition. Additionally, as previously mentioned, the Internet poses several serious theorizing challenges that need to be addressed.
Recent communication history has seen the development of such Internet-based media forms as web portals, discussion boards, blogs, social-network Web sites, collaborative-authoring and -classification Web sites, content-sharing Web sites, and many others. The introduction of many of these media forms, was followed by researchers asking important communication-effects questions, such as: What effects do these media forms have on bridging digital divides and empowering society’s nonelites (Bruns, 2005; Chadwick, 2006; Dylko, Beam, Landreville, Geidner, 2009; Hindman, 2009); what effects do they have on political attitude polarization (Sunstein, 2002; Wojcieszak & Mutz, 2009); what effects do they have on engaging citizens into political and civic spheres of their society (Eveland & Dylko, 2007; Gueorguieva, 2008; Kerbel & Bloom, 2005; Trippi, 2004; Woodly, 2008); what effects do they have on the nature of public discourse (Maratea, 2008; Munger, 2008; Stromer-Galley, 2002; Woodly, 2008; Wright & Street, 2007); what effects do they have on reinvigorating (Ellison, Steinfield, & Lampe, 2007; Valenzuela, Park, & Kee, 2009; Wellman, Quan-Haase, Witte, & Hampton, 2001) or undermining (Keen, 2008; Sunstein, 2002) community and social capital, and so forth (Tremayne, 2007). These questions are important. However, today’s communication environment (and the field’s conventions rooted in the communication-effects fields’ history) often “seduces” researchers to go about answering them in a rather atheoretical manner. This happens largely because researchers have no uniform theoretical media-effects framework from which to examine these questions. The main reason such research appears to be somewhat atheoretical is because it is often unclear why researchers expect to find different effects when a new media form emerges. For example, why exactly should we expect political blog use to have different effects from use of political discussion boards, or from participation in political networks of popular social-network Web sites? All three appear to represent a way to get political information while socially interacting with fellow Web site users, and all three empower users to create and share political
content. Definitions of these media forms and theoretical distinctions among them are often left without explication. Because a systematic framework for making comparisons among media forms is missing, each new media form appears in research reports to be unprecedented or even revolutionary in terms of its nature and effects. However, often they are not revolutionary, or even particularly different from the earlier media forms. The reason such similarities are hard to detect is because of the lack of an overarching media-effects framework, suited for incorporating novel media forms into existing theoretical models and capable of offering a theory-based description of media forms’ nature and effects.

Our recent communication technology history tells us that the “future media challenge” will continue to hinder theory building because novel media forms are sure to be developed. It is imperative for media-effects researchers to create a framework that permits the systematic comparison of novel media forms to the existing ones. This will facilitate a better understanding of the nature and potential effects of media forms that existed for some time, forms that were only recently introduced, and possible future media forms. Without such a framework, researchers will be inclined to reinvent the theoretical wheel (Eveland, 2003) and the communication effects field will continue to be theoretically adrift (Bennett & Iyengar, 2008).
Chapter 3: User-Generated Content

One recent manifestation of the media transformation has been labeled UGC, which is a theoretically interesting phenomenon because of its novel appearance, relatively novel nature, and potential to produce important effects (Bruns, 2005, 2008; Chadwick & Howard, 2009; Davis, 2009; Pew, 2009). Very briefly, UGC is defined here as information products that are freely available online, which are political in topical orientation, which are shaped to a significant degree by the users’ actions, and where such actions are voluntary (a more thorough and complete discussion about the nature of UGC is provided below). Theoretical interest in UGC has been amplified by its popularity (Harrison & Barthel, 2009). To properly evaluate UGC popularity, several different popularity indicators should be examined: (1) popularity of features characteristic of UGC, (2) participation of individuals in content generation, (3) popularity of media forms representative of UGC, and (4) popularity of Web sites falling under the rubric of UGC. First, in terms of UGC-related functionality, a recent Pew (2010) study reported that in the beginning of 2010, of all online news users, 44% considered it important to be able to “easily share online news content with others, through emails or posting to other Web sites, like Facebook” (p. 42); 37% considered it important to have “opportunities to comment on stories” (p. 42); and 25% considered it important to be able to follow the Web sites through social-network Web sites. Second, in regards to content creation, Pew reported that “37% of Internet users have actively contributed to the creation, commentary, or dissemination of news” (p. 44). More specifically, 25% of Internet users commented on a news story or a blog post, 17% have posted thoughts and links to news sources on social-network Web sites; 11% have engaged in content
classification (i.e., tagging, categorization); 9% have posted their own article, photo, or video online. Third, Pew reported that 57% of online Americans use social-network Web sites, such as Facebook or MySpace, and 28% of online Americans get news via such Web sites. Pew also reported that 19% of online Americans use status-update Web sites, like Twitter. Another Pew (2009) study reported that as of May 2008, 52% of Internet users watched videos on video-sharing Web sites, like YouTube; as of December 2008, 47% used Wikipedia; as of August 2006, 37% uploaded photos to a photo-sharing Web site; as of December 2008, 32% read someone’s online blog; as of April 2009, 31% rated a person, product or service; as of September 2007, 30% posted comments or reviews of products and service; as of December 2006, 28% classified photos, posts, and other online content; and as of December 2007, 22% posted comments to online news groups, blogs, photo Web sites, and so forth. Fourth, if we examine the 10 highest traffic Web sites in the United States, seven of them fall under the rubric of UGC: Facebook (2nd most trafficked), YouTube (4th most trafficked), Wikipedia (6th most trafficked), Twitter (7th most trafficked), Blogger (8th most trafficked), Ebay (9th most trafficked), and Craigslist (10th most trafficked) (Alexa, 2011). The above discussion shows that the UGC phenomenon is real, highly popular, and diverse.

However, what exactly is UGC? In this work the name UGC is adopted (Van Dijck, 2009), instead of multitude of potential (and sometimes less inclusive) other terms such as user-generated media (Shao, 2009), produsage (Bruns, 2008), Web 2.0 (Chadwick & Howard, 2009; O’Reilly, 2005), collective intelligence (Sunstein, 2006), citizen journalism (Bruns, 2005), participatory news (Deuze, Bruns, & Neuberger, 2007), social media (Pew, 2010), and others. The term UGC appears to most accurately tap the phenomenon which is examined here, and which is conceptualized below. Having settled on the name, what exactly does it mean? In this study, the focus is on political UGC, with “political” being defined as follows: Politics and
political are the matters that deal with “authoritative allocation of goods, services, and values” in one’s society (Delli Carpini & Keeter, 1996, p. 12). A more thorough elaboration on how this study operationalizes “political” is provided in the Method section.

To formulate a theoretically-grounded definition of UGC, it is necessary to examine existing research on UGC and related topics. It should be noted that although an extensive body of research exists on different incarnations of UGC, such as social-network Web sites or blogs (Boyd & Ellison, 2008; Davis, 2009; Eveland & Dylko, 2007; Herring, Scheidt, Kouper, & Wright, 2007; Hindman, 2009; Valenzuela, Park, & Kee, 2009; Woodly, 2008; Wright & Street, 2007), there are only a few investigations dealing with the broader phenomenon of UGC (Bruns, 2005, 2008; Shao, 2009; Sunstein, 2006; Van Dijck, 2009). The discussion that follows is centered on theoretical insights about the nature of UGC.

Chadwick and Howard (2009) examined political web 2.0, which despite a different name, refers fundamentally to the same phenomenon as political UGC. The researchers described the following web 2.0 characteristics: (1) Greater control over the content by users through technological capabilities for content editing, commenting, rating, and modifying – all of which adds new dimensions and additional value to the content. (2) Collective intelligence, an idea that was thoroughly explicated by Sunstein (2006), suggesting that amateurs are able to surpass experts in terms of the quality of their work, when amateurs are working voluntarily together and experts are paid for their work, but are working alone. An illustration of collective intelligence is when small acts (e.g., writing product reviews on Amazon.com, or rating online videos on YouTube, or leaving comments on blogs) that can be done by almost anybody, when performed by a large enough number of people can produce unexpectedly significant results (e.g., a greatly improved product search and purchasing decision-making process on Amazon, efficient search for videos judged by a community of one’s peers to be of high quality on YouTube, uncovering
hard-to-find evidence in the case of CBS’s “Memogate” [see Dylko & Kosicki, 2006] that enabled conservative political blogs to influence traditional news media agendas in 2004). Keen (2008), however, argued against the collective intelligence idea by suggesting that the quality of volunteer work done by amateurs is inferior to that of experts, and, in fact, UGC products might undermine cultural, legal, and political aspects of the society. (3) “Perpetual beta,” which is a technological term for a product that is partially released to the public who are encouraged to send their feedback to the engineers while the product is still being developed and improved by the manufacturer. The perpetual beta idea was thoroughly examined by Jarvis (2009), who argued that unfinished content does not equate to inferior content. Popular UGC, such as not carefully vetted (and potentially incorrect) blog posts or incomplete Wikipedia articles, not only project authenticity, but also explicitly extend invitation to the users to improve that content. This results in blog posts being corrected because of the comments left on the blog by the readers, and Wikipedia growing more and more comprehensive and accurate because of Wikipedia’s community participation.

Bruns (2008) also theorized about what he called “produsage,” a phenomenon of perpetual and collaborative creation and modification of content in pursuit of content improvement. His theorizing has direct parallels with Chadwick and Howard’s (2009) work on web 2.0, and fundamentally addresses the same phenomenon as UGC. Bruns emphasized several elements essential to produsage: (1) Individuals often perform rather narrow and simple tasks, which by themselves do not represent anything significant, but when combined, form a larger, more valuable product. This element clearly mirrors the idea of collective intelligence described by Chadwick and Howard (2009) and Sunstein (2006). (2) Nobody is excluded from the produsage process, and everyone willing to participate can do so. This characteristic resembles the increased user control over the content described by Chadwick and Howard (2009). (3) Those
who participate in produsage self-nominate and are not directed by any single person or organization. Altruism, interest in the particular topic area, or desire for self-expression stimulate voluntary and massive user participation in produsage; and the ability to work on small parts of the larger product obviates the need for central planning and management of such work. (4) Nobody owns the content created in the context of the produsage process, and, instead, the content is freely shared with everybody.

Harrison and Barthel (2009) examined web 2.0 and considered UGC to be one of its applications. They argued that web 2.0 enables users with little technical skills to create and share content collaboratively. In that regard, the researchers agree with Chadwick and Howard’s (2009) collective intelligence idea and Bruns’ (2008) view of produsage, where small input from a lot of people translates into a valuable final product. Harrison and Barthel argued that Web 2.0 is premised on the idea that the user becomes substantially more active than in the past, and is able to share, modify, and create content, which, again, mirrors the active-user theme in the work of Chadwick and Howard, and Bruns.

OECD (2007), in their examination of “user-created content,” offered several characteristics defining the phenomenon: (1) Users create the content outside of their professional work. This characteristic matches the collective intelligence idea in which non-experts often work voluntarily (Chadwick & Howard, 2009; Sunstein, 2006). (2) The content is made available to everyone over the Internet. This characteristic is similar to Bruns’ (2008) observation that content created as a result of produsage is not owned, but instead shared by all.

Although the above discussion of other researchers’ insights is helpful for developing a conceptual understanding of political UGC, most of this theorizing on UGC provides little guidance in regards to how social scientists should operationalize UGC in their empirical investigations. The current work attempts to overcome this shortcoming by integrating several
common themes that emerged in the above discussion and translating them into empirical language. First and foremost, active-user dimension (increased user control over content and increased interaction with content) appears to be the core common thread running through most of the UGC theorizing. The second theme appears to be a recognition that user activities can come in different forms, ranging from significant (e.g., creating a new Wikipedia entry) to minimal (e.g., “digging” a news story), and manifesting through different forms (e.g., posting comments, editing a wiki, uploading multimedia content, classifying content). The third theme is that the content should possess some authenticity associated with the fact that it was created by fellow users voluntarily and often in collaboration with other users, as opposed to being professionally created by well-funded organizations trying to promote their commercial interests. The fourth and final theme is that the content should be freely available to anyone over the Internet.

Using these four themes as a theoretical foundation, in this work political UGC is conceptually defined as: (a) information products (individual articles, multimedia content, etc.) (b) that are freely and openly available on the Internet (although, sometimes it is necessary to register for free with the Web site and/or create one’s profile on the Web site), (c) which are political in topical orientation, (d) which are shaped to a significant degree by the users’ actions, (e) and where such actions are voluntary. Despite its breadth, this definition excludes the following non-exhaustive list of Web sites: political blogs written by professional journalists as part of their job, articles written by professional journalists that were inspired by a tip from an audience member, articles written by professional journalists on professional news Web sites that allows commenting, and so forth. The definition does include a public affairs news discussion section on a health-oriented online discussion forum.
Admittedly, several valuable characteristics that are present in the above-described theorizing on UGC are missing from this definition. Collaborative creation is one of them and, although very important for some types of UGC (e.g., wikis), it is not present or required for other UGC forms (i.e., content-sharing Web sites, such as YouTube). Also, perpetual improvement is another characteristic that is missing from the definition. Again, although theoretically important for some UGC types (wikis, to lesser degree for blogs and participatory journalism Web sites), is not important for other UGC forms that are relatively static at the level of a given piece of content (e.g., discussion forums, content-sharing Web sites) where such content generally does not undergo perpetual and/or significant modification after it is posted. In the present definition an attempt was made to be broad enough to be inclusive of all forms of political UGC that are important for political outcomes examined in this work.

What particular media forms exemplify political UGC under the formulated definition? To begin answering this question Google, Google Scholar, the ISI database, Wikipedia, and Urban Dictionary snowball search was conducted by starting with the keyword “user-generated content” and progressing to pages marked as “related” or “similar.” The same procedure was used for all other UGC synonyms, until the point of theoretical saturation. This search produced a comprehensive list of all possible names for UGC phenomenon, and it included the following terms: blogs, bloggers, blogging, blogosphere, citizen journalism/media, civic journalism, collaborative filtering/journalism/tagging, collective intelligence, commons-based peer production, community journalism, consumer-created/fortified/generated content, consumer-created/fortified/generated media, content-sharing Web site, crowdsourcing, customer review Web site, deliberative social networks, democratic media, discussion board/forum, enterprise bookmarking, folksonomy, interactive journalism, grassroots journalism/media, mass collaboration, micro-blogs/bloggers/blogging/blogosphere, open publishing, open source
publishing/journalism, participatory journalism/media, prosumer, prouser, public journalism, remixing, smart mob, social bookmarking/media/networking/news/web, swarm intelligence, user-created/fortified/generated content, user-created/fortified/generated media, virtual volunteering, web 2.0, wiki, wiki journalism/news, and wikinomics. This list is a vivid illustration of why classification and explication work is urgently needed in the area of UGC, as it contains many overlapping names (e.g., public journalism, community journalism, open-source journalism). After identifying the terms, their meaning was established by using Google, Google Scholar, ISI, Wikipedia, Urban Dictionary and several other tools.

Using this keyword list, their definitions, and the UGC definition developed above, another effort was made to identify various UGC types. Using both the scholarly and popular literature, as well as various sources mentioned above, a conventional typology of political UGC was created, consisting of: blogs, wikis, participatory journalism Web sites, discussion boards, socially aggregated tagging Web sites (folksonomies), social-network Web sites, and content-sharing Web sites. Although not a comprehensive list of all possible UGC types (e.g., it excludes open-source software), this list appears to include all of the UGC types important to the individual-level political outcomes examined in this work. The core technological aspects of these UGC forms are briefly described below.

A blog (often called a weblog or an online diary) is a Web site in which the blogger (administrator of the Web site) makes posts on the main section of the page, which is structured to display the most recent posts on top of the page. Such structuring leads to an emphasis on recency and brevity of the posts, often at the expense of depth and a thorough vetting of the content (although this is not necessarily the case, as exemplified by Jay Rosen’s influential blog, PressThink, which emphasizes depth and deemphasizes recency). The posts often contain information about the date and time they were made, and tags or labels are given to the post by
the blogger for easier sorting and identification of the posts on the same topics. They also often contain reader comments that can be turned on or off by the blogger. One notable characteristic of blogs is the practice of archiving all posts, which facilitates retrieval of any posts made by the blogger, often directly from the main page of the blog. Another characteristic is the use of permalinks (permanent Internet addresses to each post that can be bookmarked, emailed, etc.) on blogs, which often leads to an online cross-blog discussion focused on specific blog posts (Davis, 2009; Hindman, 2009; OECD, 2007; O’Reilly, 2005).

A wiki, the second UGC type, is fundamentally a way to collaboratively create documents, in which many users can participate independently and without the need to know or interact with one another. A wiki’s main unique characteristic is that it offers users a particularly high level of control over the content, by allowing users to create a new article from scratch, or erase a large portion of an existing article, or otherwise substantially modify it, as opposed to just voting on the article or discussing the article with others (Bruns, 2009; OECD, 2007). Wikipedia, is undeniably the best known wiki Web site. In addition to its voluminous wiki section (over 3,600,000 English-language entries, as of May 2011), Wikipedia also hosts an online bulletin board to enable its users to discuss various encyclopedia’s articles, rules of operation, special projects, or other issues of interest to the Wikipedia community.

Participatory journalism (also called collaborative journalism, citizen journalism, interactive journalism, grassroots journalism, open-source journalism, etc.) Web sites combine the work of professional journalists with input from users, resulting in various Pro-Am journalistic collaborations, from crowdsourcing projects in which readers provide limited information to a journalist who then organizes it in a preferred way, to articles written entirely by users and published with minimal editing by professional news organizations. Deuze, Bruns, and Neuberger (2007), described several characteristics of participatory journalism, two of which
should be mentioned here: (1) content is actively shaped by professional journalists and their audience; and (2) such collaborative work is often decentralized. Today, almost all news outlets qualify under the definition of “participatory journalism” if the degree of required citizen involvement is low. However, under the present UGC definition only those participatory journalism Web sites that are characterized by a “significant” user involvement in content creation qualify as a UGC form.

Chung (2008) and Domingo et al. (2009) documented the following practices by traditional news organizations: many organizations allow their audiences to submit textual or multimedia content for dissemination to the organization’s mass audience; organizations often provide electronic space to host discussion forums for their audience, which often have social-network-style capabilities (such as creation of profiles, establishing connections with fellow users, etc.); organizations often host users’ blogs or dedicate sections on their Web sites to publishing user-generated stories; organizations often solicit story tips from the audience; the audience is given an opportunity to tag (classify), vote on, and comment on the news stories.

A discussion board (also called an online discussion forum or electronic bulletin board) is a Web site often focused on a specific topical domain (e.g., coin collecting, fishing, bicycling, politics, religion) where individuals can obtain information on that topic and/or establish interpersonal contact with other individuals interested in the topic. Technologically, discussion boards resemble a tree: The first level of a tree consists of the board’s main sections (e.g., lobby, news, FAQs, trading); the second level consists of subsections that branch off of main sections (e.g., trading foreign cars, trading domestic cars); the third level branches off of the subsections, and consists of threads started by board users (the threads are essentially initial posts made by board users on the general topic of the subsection where the thread is located); and the final level are posts or replies to the initial thread. Some boards facilitate interaction between their members
by providing private and public ways to communicate (e.g., chat, email, thread replies, blogs). These boards vary widely in publicness (from requiring registration to read content, to being able to start threads without registration) and in the degree of moderation (from no moderation to heavy moderation), rules for which are generally described in the forum’s FAQs section (Himelboim, Gleave, & Smith, 2009; Wright & Street, 2007).

Socially aggregated tagging or classification Web sites (also known as a folksonomies, or enterprise or social bookmarking Web sites) allows users to collaboratively develop a classification system around various textual, photo, or multimedia content (Golder & Huberman, 2006). On folksonomy Web sites, users do not necessarily create the content, but instead classify it (although, some Web sites combine both generation and classification of content). Therefore, the unique characteristic of a folksonomy is not adding new content to the Web site’s database, but rather making the content more organized and structured, which improves the information system’s information retrieval performance. The ability to tag or classify content is sometimes incorporated into content-sharing Web sites (e.g., Flickr), blogs (e.g., hosted by Blogger), and social-network Web sites (e.g., Facebook).

Social-network Web sites are currently the most popular UGC genre. Facebook, Linkedin, and Twitter are in the top 20 most trafficked U.S. Web sites (see Alexa, 2011). These Web sites enable users to connect to other users or organizations through a network of interconnected profile pages that form a defined social structure existing among individuals and groups on the Web site (Boyd & Ellison, 2008; OECD, 2007). The Web sites help maintain existing relationships, as well as develop new ones. These profiles vary in publicness, number of connections, and other characteristics (Boyd & Ellison, 2008). Individuals and groups using the Web sites can communicate with one another through blogs, chat, status updates, “wall” comments, tags, “liking” and commenting on each other’s content.
Finally, the content-sharing Web site category also encompasses various hugely popular Web sites, such as YouTube or Flickr. These Web sites contain a vast diversity and volume of textual and/or multimedia content submitted by users, who often not only submit content, but rate, tag, and comment on it. The ease of uploading content, numerous content search options, ease of content rating, and vast databases are some of the technological aspects making this UGC type attractive to the users.
Chapter 4: Mix-of-Attributes (MOA) Approach

The above description of UGC is a good illustration of today’s communication environment: It is very diverse and media forms that populate it often overlap, making it hard to create rigorous scientific classifications of these media forms. The above description of blogs, social-network Web sites, folksonomies, content-sharing Web sites, participatory journalism Web sites, wikis, and discussion boards shows that most of them overlap to a significant degree. For example blogs, social-network Web sites, participatory journalism Web sites, and content-sharing Web sites often incorporate folksonomy capabilities of tagging or classifying content. Discussion boards and social-network Web sites often incorporate blogs as one of the communication options. Wikis and participatory journalism Web sites often incorporate discussion boards. Discussion boards often incorporate social-network-style features, where users can create profiles and connect those profiles to other board members. Folksonomies often integrate content-sharing capabilities. Such overlap is problematic because if we hypothesize that use of different UGC types might produce different effects on the users, it might be extremely challenging to discern what exactly it is about the UGC Web sites that produces the hypothesized effects. More specifically, it is theoretically problematic if different technological features are expected to produce different effects, and these features overlap across Web sites. In the language of experimental research such situations confound experimental conditions, making it impossible to determine the precise causal process producing the effects. This strongly argues against utilizing frequency of using different UGC types as independent variables in media-effects investigations.
Unfortunately, frequency of UGC use as independent variables is exactly the conventional approach in the communication-effects research tradition. Many conventional communication-effects studies (McLeod, Scheufele, & Moy, 1999; Mutz, 2002; Scheufele, Nisbet, & Brossard, 2003; Shah, Kwak, & Holbert, 2001) rely on survey methods to measure the frequency of respondents’ use of different media (e.g., television, newspaper, radio), which represent independent variables. These studies also measure other constructs (e.g., knowledge, self-reported behavior and attitudes), which represent the dependent variables. Some sort of correlation technique (regression, multilevel or structural equation modeling, etc.) is then used to establish relationships between independent and dependent variables. This approach has worked well in the past because different media were fairly distinct from one another. For example, it was easy to differentiate a newspaper from radio in the past. Their form (visual vs. auditory), their use (more user involvement in reading a newspaper vs. less user involvement in listening to radio), their content (easy to understand, conversational language on radio vs. more technical and formal language in a newspaper), user control (more control in a newspaper, where the user can refer to it at a later time, or read the same material twice vs. less control in radio, where, unless recording equipment is used, the program cannot be listened to more than once), and other dimensions allowed for a fairly easy differentiation among these media forms. However, today’s media are different. The same newspaper today has a Web site which also likely provides users with audio podcasts and other multimedia content that they can use at their convenience. The same radio station also has a Web site that provides textual information, along with podcasts and other features that allow users to consume the content in their preferred way, increasing user control over the content. When it comes to UGC, the above descriptions of UGC types suggest that the situation has become even more problematic. A new methodological approach is needed to reflect the realities of today’s – and tomorrow’s – communication environment.
In situations in which precision in explaining causal mechanisms is important, and where new media forms are the focus, the MOA approach appears to be theoretically appealing (Eveland, 2003). The MOA approach suggests that we should view media as nothing but a combination of different attributes (such as degree of user control, degree of interactivity, type of modality, etc.), where each media form (e.g., print newspaper) can be given scores on each of these attributes (e.g., moderately high on user control, very low on interactivity, textual and pictorial modality). The MOA approach suggests that when we say a “media form” or a “medium” we essentially mean a specific set of attributes and specific degree to which each attribute is present. To oversimplify, when we say a “newspaper” we actually mean a set of paper sheets with text and images on every page (the following forms of information representation are present to a high degree: photographs and text), which has sections organized by topic that are placed in a pre-determined order (degree of user control: moderate), which the user has to leaf through and read with some degree of concentration in order to extract information from it (degree of involvement required from the user: somewhat high). In other words, attributes act as reference points that we can use to compare and differentiate one media form from another (e.g., traditional form of a newspaper is different from radio in terms of information representation).

When using the MOA, the analytical focus shifts from broad media forms closer to what we actually expect to produce specific effects: media forms’ underlying structures, or attributes. In other words, the way we operationalize our independent variables changes from relying on questions tapping frequency of specific media form usage, to using the presence of technological attributes underlying these media forms. Such a shift of focus is not an inconsequential linguistic change - it has deep theoretical and methodological implications. It facilitates the development of a more detailed understanding of the nature of each media form (by thoroughly describing which attributes, and to what degree, are present in any media form), and its potential effects (by
specifying which attributes should be experimentally examined to establish their specific effects). It also aids in systematic classification efforts, since the attributes can serve as classification criteria for creating a typology of media forms, similar to how different anatomical criteria are used for classification of animal species in biological taxonomy (Romesburg, 2004). Finally, the MOA approach facilitates theory building in today’s quickly evolving communication environment since it shifts scholarly attention from distracting, puzzling, and revolutionary-appearing novel media forms to their already familiar and well-researched underlying attributes that can be used for classification, effect prediction, and causal mechanism explanation.

The major benefit of the MOA approach is that it discourages researchers from reinventing the theoretical wheel with the introduction of each novel media form, and instead encourages researchers to place each novel media form into an existing and well-understood matrix of attributes, whose properties are known. Using this framework allows researchers to treat novel media forms as nothing but an alteration in an attribute mix that existed before in the context of another existing media form. Such an alteration can be represented by some existing attribute being absent, or becoming less or more prominent, or (much more rarely) having a new attribute being added to the existing mix. Therefore, as our communication environment and existing media forms will undoubtedly continue to evolve, the MOA approach provides communication-effects researchers with a theoretical structure (i.e., matrix of attributes and their effects) that can serve as a useful, temporally stable and relevant foundation for understanding the effects of any novel media form.

To develop a ready-to-use framework for the communication-effects field, researchers need to follow these three steps: (1) Identify attributes that (a) characterize a certain media form of interest and that (b) are theoretically expected to influence specific dependent variables of interest; (2) use systematic content analysis to document that the degree to which each attribute is
present; and finally (3) use experiments to establish the specific effects that (a) individual attributes and (b) the attribute mixes documented with content analyses produce on the dependent variables of interest. The second step – use of systematic content analysis – is optional. For investigations concerned with the effects of existing media it is necessary to content analyze the media of interest (or experimental stimulus materials that are supposed to represent these media) to document which attributes are present. However, the MOA approach can be used for another purpose: the development of novel media forms designed to facilitate very specific effects. For example, if attribute A, B and C are explicated (step 1) and experimentally shown to produce effects X, Y and Z (step 3), it is not essential to know whether any of the existing Web sites have attributes A, B, and C present on them. If these attributes are deemed important (i.e., they are known to produce desirable effects) these attributes can be technologically implemented in context of different Web sites. Therefore, it appears that there are two broad purposes that MOA can serve: (1) Establishing effects of attributes and designing Web sites that produce very specific effects. For this purpose, only steps 1 and 3 are needed. (2) Documenting which attributes exist on popular Web sites and establishing effects of these attributes. For this purpose, all three steps are needed. It is a position of this researcher that the first purpose is currently more important because (a) there is an urgent need to expand the list of attributes researchers can use, and steps 1 and 3 are less resource and time intensive than steps 1, 2, and 3; (b) media forms will continue rapid transformation, making it likely that content analyses results (step 2) might provide insights that might lose their theoretical and practical relevance quickly. In other words, the first purpose is more achievable and is more future-oriented, whereas the second purpose is relatively more laborious and past-oriented.

The need to follow these three steps naturally leads to a discussion of the MOA downside: As a research strategy, it is substantially more time and labor intensive than the
conventional media-effects research approach. If a researcher is interested in determining the effects of a specific media form on specific outcomes of interest, utilizing the MOA approach turns this interest into a full-fledged program of research, rather than a single study that can be done quickly and on a small budget. This occurs because each of the three steps requires a substantial intellectual undertaking, especially when the MOA approach is just beginning to be implemented in the communication-effects domain, characterized by a void of needed work on attribute explication. The first step, which is undertaken in the present work, might involve synthesizing theory from diverse literatures, since it requires connecting technological attributes of a media form to psychological or social mechanisms that the technological attributes are expected to trigger, to some ultimate outcome in which the researcher is interested. Each of these three components is likely to encompass very diverse literatures, as this project will illustrate. The second step, which is also undertaken in the present work, might be challenging when one is focusing on a media form that is rather heterogeneous in terms of outlets representing it (such as UGC). Potentially the most challenging step is the third one. For example, even in relatively simple situations (if one has just three attributes, each attribute has three levels [low, medium, and high] of presence in the context of a medium, and each attribute is expected to affect three dependent variables of interest), to experimentally establish the effect of each attribute, at each level, on each dependent variable of interest would require several dozen experimental conditions. A solution to this problem is discussed in the Discussion section.

However, as the MOA approach becomes more widely implemented, the exact presence of many attributes across media forms becomes known, and the effects of many attributes become well understood, which substantially reduces the amount of work that needs to be done in every subsequent MOA investigation. Despite the resource requirement, the MOA approach is still a highly appealing research strategy due to the framework becoming easier to utilize as it is
becoming adopted in more investigations; due to its explanatory and predictive nuance that greatly minimizes the two challenges to communication-effects theorizing described in the beginning of this manuscript; and due to the absence of any adequate alternative research strategy suited for today’s quickly-evolving communication environment that is characterized by numerous hybrid media forms. Because of these considerations, the MOA approach is utilized in this project.
Chapter 5: Using the MOA to Explicate Political UGC Attributes

Before discussing the attributes, it is important to describe the theoretical boundaries surrounding this project that determined which attributes were selected, and how. The scope of the present work is narrowed by two criteria. The first criterion is the focus on political participation as an outcome. The second criterion is the focus on communication technology.

**Political Participation**

Chosen primarily because of its importance to political and communication theory, political participation is conceptually defined as actions by citizens aimed at affecting public policy (Verba et al., 1995). Political participation has been historically regarded as an essential (or at least a desirable) element of a healthy democracy, from the era of the Greek city states (Dahl, 1989), to the founding of the United States of America (Sheldon, 1991), to the present day (Chadwick, 2006; Knight Commission, 2009; Mutz, 2006; Verba et al., 1995). The reason political participation is well regarded by both theorists and practitioners alike (Knight Commission, 2009; Mutz, 2006), is because equally distributed and relatively high levels of political participation among citizens produces various desirable effects, such as: (1) allowing citizens to make politicians more accountable for their actions, which then leads to (2) an improvement in the government’s efficiency and effectiveness, which increases the (3) sense of the government’s legitimacy (Pateman, 1970); and (4) bolstering of citizens’ belief that they can have a meaningful influence on their political system (Finkel, 1985; McLeod et al., 1999). Political participation is often studied as a dependent variable, and the focus is on what increases or decreases it. In many such investigations, researchers examined how different media use and
interpersonal communication behaviors could facilitate or hinder participation (Bimber, 2003; Eveland & Hively, 2009; Eveland & Scheufele, 2000; Garrett, 2006; McLeod, et al., 1999; Mutz, 2002, 2006; Prior, 2007; Scheufele, 2000; Scheufele, Nisbet, & Brossard, 2003; Scheufele, Nisbet, Brossard, & Nisbet, 2004; Stein, 2009). This present study continues this tradition by focusing on potential effects of political UGC on political participation.

*Participatory democracy theory.* As mentioned above, political participation is deemed to be one of the most important constructs in political science. One way to understand the role and implications of political participation for democracy is by examining various prominent theoretical models of democracy (particularly participatory democracy). Another way is to weigh the desirable and undesirable output of political participation. Both approaches are employed below.

Participation of citizens in political life of their society has been extensively examined by many political theorists, each of which provided their own take on what “democracy” is and what role political participation of citizens should play in it. This section does not attempt to provide a comprehensive review of all the arguments surrounding the role of political participation, and much less review all models of democracy (e.g., direct, representative, deliberative, thick, thin); however, this section does address the major arguments and visions of *participatory* democracy and participatory aspects of different prominent models of democracy. Addressing the following questions structures the discussion below: How important is political participation for the democratic theory and to the properly functioning modern democracy? What levels of political participation, by whom, and under what conditions appear to be the most optimal? How does reality measure up to the ideal? Answering these questions is particularly relevant in today’s era characterized by plentiful opportunities to get involved in public affairs that are created by new information communication technologies.
The first model examined here was proposed by Pateman (1970) who contrasted participatory democracy model with conventional democracy model and argued that the former was more desirable. Pateman (1970) argued that the conventional thinking about what democracy should be (what she called “contemporary theory of democracy”) unjustifiably and mistakenly diminished the role of political participation by lay persons. According to the conventional thinking, Pateman argued, citizens should be primarily responsible for voting and choosing among the competing elites during periodic and fair elections. Universal suffrage was considered to be the main mechanism ensuring political equality. Additionally, the conventional thinking regarded stability of a democratic system to be crucial, and lack of political interest and lower participation rates among low SES (socioeconomic status) individuals to be key to such stability, because low SES individuals were supposedly characterized by “non-democratic attitudes” (p. 14), such as high authoritarianism.

As an alternative to this vision, Pateman proposed “participatory theory of democracy” (p. 42), which was quite different. According to this model, formulated by Pateman from the writings of Rousseau, John Stuart Mill, and G. D. H. Cole, universal suffrage was not a sufficient element of a democratic system. Political participation of a vastly greater number of citizens was suggested to be highly preferable, because such participation nurtures the democratic qualities, attitudes, and political sophistication among citizenry making them valuable, effective, and motivated participants in political life of their country. These consequences of the widespread participation, and not the limited participation of the low SES individuals, was the key to democracy’s stability, because the democratic system would be “self-sustaining through the educative impact of the participatory process” (p. 42). Additionally, participation would facilitate system’s stability through an increase in government legitimacy since participation “aids the acceptance of collective decisions” (p. 43). According to Pateman, an important task for political
theorists and practitioners was to understand why actual participation rates were low among various segments of population, and then find ways to facilitate the widespread political participation among those segments. The present situation, characterized by relative political apathy (Verba et al., 1995) clearly falls short of the bar set by Pateman.

The second view of political participation was articulated by Verba et al. (1995) in what may be argued to be the most comprehensive empirical work on the nature and antecedents of political participation in the United States. The first sentence of their influential volume was: “Citizen participation is at the heart of democracy” (p. 1). Although the statement clearly indicates how important political participation is deemed to be by the researchers, and although it implies some important effects of political participation on democracy, the authors acknowledged that they did not focus on the consequences of political participation, but rather on its antecedents and nature.

Despite the fact that the authors did not argue for any single coherent model of (participatory) democracy, they did suggest that several elements are indicative of a healthy democracy: equality of representation and “goodness” (i.e., sophistication and tolerance) of citizens engaged in political life of their country. Equally distributed levels of participation across all social strata were theorized to positively affect the first element; whereas, ironically, greater levels of participation among some social groups (and not others) would facilitate the second element. The researchers found that some social groups (e.g., disadvantaged low SES) participated less; however, the groups that did participate were more politically knowledgeable and tolerant. The authors also argued that the system where representation inequality results from a free choice of certain groups not to participate is much more democratic than the system where representation inequality results from systematically different availability of resources. The authors found evidence that a lack of representativeness in American political system (e.g., rates
of making donations to political causes) sometimes stemmed from lack of resources (e.g., family income and education). Therefore, the authors found flaws as well as strengths in today’s American democracy.

The third model of democracy examined here was articulated by Dahl (1989), who only indirectly dealt with participatory democracy. He observed that the growth of political unit’s size (from an ancient democratically-run city state to a modern-day democratic country of even the smallest size), inevitably reduced the number and diversity of available political participation options for an average citizen. However, in describing modern polyarchies (the term used to describe today’s most democratic regimes and the most feasible and democratic forms of political organization available to a large country), Dahl suggested seven criteria that have to be met for a system to be classified as a polyarchy. Of those seven, four clearly reference political participation: free and fair elections, inclusive suffrage (“practically all adults have the right to vote in election of officials” (p. 221)), right to run for office given to “practically all adults” (p. 221), and association autonomy (a freedom to form political parties and special interest groups). This theorist suggested excluding some parts of society from participation where the criteria for exclusion is democratically determined (e.g., children, mentally ill).

Dahl also suggested an important role for the communication technology in terms of ensuring stability of modern polyarchies by narrowing the “growing gap that separates policy elites from the demos” (p. 338). He suggested that today’s communication technology should be tapped in order to make accurate information about policy issues on the national agenda easily and universally available to all citizens, which would inform and prepare citizens for meaningful participation in political discussions and help them influence policy agenda.

However, Dahl saw a much smaller role for an average citizen in terms of partaking in politics then did Pateman (1970). Dahl suggested that technical difficulties that are inherent in
many of today’s social issues prevented every citizen from learning about every issue, making it practically impossible for citizens to actively participate in all areas of political sphere. Instead, some groups of citizens might become expert on some issues, and other groups on other issues, creating a system of the issue publics, which can then develop sufficient expertise enabling these issue publics to meaningfully affect national policy on the issues falling under their expertise. Dahl’s vision appears to be fairly close to the political realities in today’s most democratic nations.

The last model discussed here is Zimmerman’s (1986) theorizing about participatory democracy, who, like Pateman (1970), saw a greater role for the public’s political participation. However, his justification of that role was rather different. This theorist’s position can be summarized by the following quote:

We are convinced, however, that active and widespread citizen participation is essential for the best functioning of the various units of government in the United States and that intermediary bodies do not always represent accurately the views of the citizenry. (p. 2)

Zimmerman’s position is rather pragmatic in that he acknowledges that some governmental programs (especially the ones that are uncontroversial and are running smoothly) require little participation from the public, whereas other programs (especially those involving large sums of money and affecting many citizens) require a widespread participation. The latter type of programs, Zimmerman argued, could greatly benefit from the expertise that resides in the public in terms of the program planning, implementation, and evaluation.

Additionally, Zimmerman suggested that representative system of government, although appealing in theory, can facilitate corruption and manipulation of the public, unless representatives are guided by the noblest motives and highest ethics standards. This scholar suggested structuring the governmental system in a way that makes direct interventions by the public (e.g., voting, petition referendum, recall) easier to undertake.
Zimmerman suggested that citizens could best participate if they learned and exchanged their views on the relevant issues. Otherwise, their participation would be less effective and more expensive for the government than it should be. For this model of government to work, Zimmerman placed a heavy burden on media to accurately inform the public, and suggested that representatives and bureaucrats should actively engage citizens in decision making during all stages of major public projects.

Having examined several theoretical models of democracy, it is useful to perform a more focused cost-benefit analysis of political participation to better understand its implications. Potential advantages of participation can often be placed on a continuum from the most elevated and abstract to the most practical and mundane. On the most abstract end of the continuum is Thomas Jefferson’s vision of direct political participation as an essential means for personal fulfillment and development of an individual. Sheldon (1991) found similarity between Jefferson’s political thought and Lockean liberalism by virtue of… [Jefferson’s] vision of the polity as deliberatively nurturing citizens’ social faculties, through direct participation in community life and governance, and providing the economic, educational, and political prerequisites of that independent democratic citizenry. (p. 54)

For this founder, who adapted classical democratic theory to the unique American landscape, active and regular participation in the political life of one’s community was a virtue in itself, which not only contributed to the development of an individual’s mind, but also helped realize human social nature and prevented a state from collapsing into a tyranny (Sheldon, 1991). Such vision is clearly present in Pateman’s (1970) theorizing on participatory democracy.

Another, somewhat less abstract positive outcome of a widespread participation is equality of representation – in a practical, as opposed to theoretical sense (Zimmerman, 1986). As Verba et al. (1995) aptly argued in their influential volume: “Equality in political rights does not, however, create substantive equality in their effective use” (p. 11), meaning that having some
rights is not tantamount to deriving benefits by actively exercising those rights. Verba and colleagues explained that politicians, just like most other people, are only exposed to a limited amount of information that comes from a limited number of sources. These politicians base their decisions on such information, and if the active citizens, whom politicians most frequently hear from, are not representative of the larger population, many citizens may be systematically misrepresented or underrepresented. Such misrepresented and underrepresented citizens do not exercise the rights they are given, and do not enjoy political equality that they are entitled to. Therefore, increasing participation rates among the underrepresented would increase the overall political equality. Other researchers agree and argue that aggregate increases in participation levels may mask an increase in a “participation gap.” Similar to the diffusion of information process and original formulation of knowledge gap hypothesis, only targeting the most needy and vulnerable groups will result in greater equality. This happens because when the entire population is targeted with some monolithic programs (to increase voter turnout, educate citizens about political issues, or facilitate adoption of some new agricultural technique), people who were better off initially will derive significantly greater benefits from those programs than people who were worse off (McLeod, Eveland & Horowitz, 1998; Rogers, 2003; Tichenor, Donohue & Olien, 1980).

An even less abstract, but no less important positive outcome of participation is a greater legitimacy of government actions in the eyes of the populace (Pateman, 1970). Citizens who partake in political processes more, are likely to understand the hurdles and opportunities inherent in those processes better, and thus realize the limitations and difficulties faced by the political leaders better, and appreciate their work more. Additionally, citizens who have contributed to formulation of governmental decisions should develop a sense of ownership of those various governmental policies and programs (Zimmerman, 1986).
Even more practical and, perhaps, the most important outcome of the widespread political participation is an opportunity to hold public officials accountable (Zimmerman, 1986). The most obvious way to keep “good” public servants in office and force “bad” ones out or to promote a desired policy and fight against a detrimental policy is by voting, campaigning, and protesting (Dahl, 1989; Pateman, 1970; Verba et al., 1995).

The least abstract benefit of the widespread participation is enlisting public’s “detailed knowledge of local conditions, needs, and desires” (Zimmerman, 1986, p. 3) that can help better design and implement important social programs. Zimmerman argued that particularly complex and expensive social programs could benefit from the public’s input.

There are several other plausible positive outcomes of political participation. The first one is an improvement of factual political knowledge. This improvement might occur when an individual learns more about the political process by working inside the political system – a sort of learning-by-doing or an “active learning” (Prince, 2004) process. Participation may also bolster a sense of political efficacy, because individuals may learn more about the political system through their involvement, and as a result see where they can apply themselves more effectively (Finkel, 1985; McLeod et al., 1999). After becoming more knowledgeable and efficacious, individuals’ interest in politics may also increase. This situation is similar to being interested in some game only after figuring out the rules of the game and becoming good at it. All of these three variables: political knowledge, efficacy, and political interest seem to be involved in a spiral of reciprocal causality with participation, because all three have been empirically demonstrated to predict or highly correlate with participation (Eveland & Scheufele, 2000; Scheufele, Nisbet, Brossard & Nisbet, 2004; Zimmerman, 1989), as will be discussed more thoroughly below.

However, a widespread political participation of citizens also has its costs. Zimmerman (1986) identified several dangers associated with the efforts to increase overall citizen
participation. One major danger is that involvement of too many incompetent “lay citizens” (p. 4) in technical plans, programs, or issues may lead to inefficiencies, either by resulting in more bad decisions or resulting in more delays caused by the governmental efforts to educate the populace about complex and subtle details of those plans, programs, or issues. In other words, when there is a popular rule, or a truly widespread participation, there is a danger that “the wisest and ablest will be overshadowed by the lack of knowledge, skill and experience of the majority” (Held, 1987, p. 93). Such a scenario is not too far-fetched, because citizens are known to have alarmingly low political knowledge (Delli Carpini & Keeter, 1996).

The second danger of a widespread popular participation in public affairs is the cost that government may run into trying to get citizens engaged and informed (Zimmerman, 1986). Expensive recruitment efforts and turnout drives by government or different special-interest groups may not be worth the outcomes obtained from the widespread involvement of public in politics, especially when public has to undergo an extensive education on the relevant issues. The third danger is valuing interests of one’s small locale more than the interests of one’s larger community (Zimmermann, 1986).

In conclusion, it appears that the abstract notion of political participation is vital to much of theorizing on what constitutes a healthy democracy. Additionally, it appears that the advantages of participation on the aggregate outweigh the disadvantages, even when large and representative segments of a society actively participate in politics. Zimmerman argued that these “costs are a small price to pay for the benefits of a healthy political system” (p. 4). However, there is still some disagreement among the scholars about the extent, type, and context of the most optimal incorporation of citizens’ political participation into the normative democratic models.
Variables and processes affecting political participation. A large body of work on political participation suggests that numerous variable and processes can hinder or facilitate citizens’ individual-level political participation, and several variables and processes of the greatest interest to political communication-effects research will be examined here. The rationale for focusing on these particular variables will become evident after the discussion of the UGC attributes that is offered below.

The first variable of interest to this study is political knowledge, most commonly defined as knowledge of discrete political facts (e.g., the name of the country’s vice president, percent of the senators needed to override a presidential veto, see Delli Carpini & Keeter, 1996). There is substantial evidence in political science and communication literatures showing that political knowledge is a strong and positive predictor of political participation (Cho & McLeod; 2007; McLeod et al., 1999; Scheufele, 2000; Scheufele, Nisbet, & Brossard, 2003; Scheufele, Nisbet, Brossard, & Nisbet, 2003). It can be hypothesized that having extensive insight into politics, the political system, and political issues might help citizens better appreciate the importance of their active participation, because they can clearly see and understand the blemishes of their political system and more clearly perceive a need for their action directed at correcting them. Additionally, knowing more about politics might make one more interested (in intellectual sense) in politics. Political interest is a very strong and positive predictor of political participation (Bimber, 2001; Eveland & Scheufele, 2000; Kim, Wyatt & Katz, 1999; Verba et al., 1995). Finally, political knowledge was found to be a positive and direct predictor of political participation, even after controlling for efficacy (McLeod et al., 1999; Scheufele et al., 2003; Scheufele et al., 2004).

The second variable of interest to this study is political efficacy. Political efficacy is generally viewed as consisting of two dimensions: internal and external. Internal political efficacy refers to a perception of oneself being sufficiently competent in political matters to be able to
exert a meaningful influence on political system and government (Acock, Clarke, & Stewart, 1985; Finkel, 1985). External political efficacy, which is very much related to internal efficacy, is generally defined as a perception of the government, political leaders, and the overall political system being responsive to citizen input and influence (Acock et al., 1985; Finkel, 1985).

Political knowledge, discussed above, appears to affect both dimensions. The more one knows about government, the more one feels qualified to be able to partake in their society’s political life. Similarly, the more one understands his or her government, the better one understands how to apply one’s efforts, and knows the contexts where political system and their players can be most responsive. Political efficacy, predictably, exerts a strong positive impact on individual-level political participation (Finkel, 1985; McLeod et al., 1999; Scheufele et al., 2003).

The third variable is political attitude extremity (or polarization). One way to conceptualize this construct is by differentiating two types of polarization: polarization as a process versus polarization as a state. The former can be defined as an extent of opinion opposition that is increasing over time between two or more social groups. The latter, which seems to be a less intuitively appealing view of the concept, is defined as “the extent to which opinions on an issue are opposed in relation to some theoretical maximum” (DiMaggio, Evans, & Bryson, 1996, p. 693). Political attitude extremity (or polarization) has several negative effects (e.g., increased social conflict, incivility, inability to find common ground, reduced political trust), but it is also known to decrease attitudinal ambivalence, and thus increase political participation (DiMaggio et al., 1996; Hetherington, 2008; Mutz, 2002; Mutz, 2006; Mutz & Reeves, 2005; Sunstein, 2002).

The fourth variable is social capital. Social capital is a resource or a property of a social network (i.e., without the network, there is no social capital) where some networks have more, others have less of this property (Coleman, 1988). Social capital is a multidimensional construct,
consisting of two parts: bonding social capital and bridging social capital (Putnam, 2000).

Bonding social capital refers to the provision of emotional and psychological support, increase of the sense of belonging, and other elements making a given social network more cohesive and reinforcing one’s group’s exclusive identity (Putnam, 2000; Skoric, Ying, & Ng, 2009). Bridging social capital refers to coming in contact with individuals from outside of one’s social network and being able to find common ground with them and overcome social cleavages (Norris, 2002; Putnam, 1995). Social capital is commonly represented by strong norms of reciprocity (willingness to do things for one another, existing among the members of the social network), high interpersonal trust, high commitment to the social network, and willingness to work on common problems (Shah et al., 2001; Skoric, Ying, & Ng, 2009; Wellman, Quan-Haase, Witte, & Hampton, 2001). Social capital has been found to be associated with individual-level political participation (Putnam, 1995; Skoric et al., 2009; Valenzuela et al., 2009).

The fifth variable is “interpersonal recruitment” or “requests to participate” in politics. Conceptually, interpersonal recruitment is nothing more than being asked by another individual to participate in politics. The notion of interpersonal recruitment into politics was most systematically studied by Verba et al. (1995) in context of their Civic Volunteering Model. The authors suggested that often

…political activity arises more or less spontaneously from individuals when they become excited about issues, connect politics to their basic interests, or get involved out of a sense of civic duty. Frequently, however, they become active because someone asked. (Verba et al., 1995, p. 133)

The authors also pointed out that a substantial portion of requests to engage in politics comes from individuals known to the person (e.g., a co-worker, a neighbor, friend), and such requests were the most effective due to their “peer pressure” dimension. However, there is evidence of requests emanating even from strangers being effective, because such requests are often successful at increasing the perception of personal benefit one can derive from political
participation (Wielhouwer & Lockerbie, 1994). Additionally, and somewhat unsurprisingly, what
makes a request to participate more effective is whether or not it is directed at a target that is
likely to say “yes” to the request. Such targets are politically interested individuals, who are more
inclined to participate in politics than the general public (Lim, 2010; Verba et al., 1995).
Therefore, based on large volume of research on interpersonal recruitment into political action,
one of the major forces facilitating political participation is being asked to get involved.

Technology

The second narrowing criterion is that the UGC’s attributes had to be technological in
nature. Technology is important to how a given media form is used, and what effects such usage
can produce. For example, a computer, a cell phone, or a television set are all meant to be used a
certain way. Generally, they are used in that way (e.g., a television set is not used to make phone
calls, and cell phone is not traditionally used for processor- and graphics-intensive computing
applications requiring a large monitor and a powerful processor), primarily due to their
technological features and design. However, individuals are capable of finding unforeseen uses of
technology that often go counter to how engineers meant for the technology to be used. A
computer is now used to watch television shows, a television set is used to browse the Internet,
and a cell phone is used for watching TV shows and for relatively resource-intensive applications
that previously could only be carried out on a computer (although, an argument can be made that
development of the Internet and an improvement in information processing technology facilitated
those unforeseen uses). Indeed, although technology is important, it by no means is determinant
of uses and effects of that technology, as adaptive structuration theory suggests (DeSanctis &
Poole, 1994; Fulk & Boyd, 1991). Often technology is not used as intended, or is not used at all,
which reduces the importance of technology in producing various effects (Bachen, Raphael,
Lynn, McKee, & Philippi, 2008).
However, technology is deemed important because it allows users to (a) either use it as intended; (b) use it not as intended; or (c) not use it at all. Without a specific technology, only option “c” is a viable option. Technology often facilitates what individuals already want to do, or increases individuals’ options for behaving one way or the other. Such a perspective on technology parallels Bennett and Iyengar’s (2008) position that technology is important and should be more thoroughly considered for inclusion in theoretical communication-effects models. In some sense technology generally, or the technological attributes examined here specifically, act as a strong force in terms of setting in motion various processes that ultimately produce strong and important effects.

The importance of these technological attributes is structurally similar to James Fishkin’s (1997) conception of the importance of “institutional design” in the context of the deliberative democracy. Fishkin views institutional design as a tool to overcome human beings’ cognitive and psychological shortcomings, and he considers institutional design as highly important facilitator of the “right” type of deliberation among citizens. Fishkin recognizes that such institutional design, to be the most effective, should incorporate a thorough understanding of the social and psychological nature of individuals. However, the design (e.g., moderation of deliberative discussions and presence of opposing expert opinions) can channel or encourage various social and psychological processes (e.g., promote equality in deliberative discussions and maximize exposure to and understanding of opposing arguments on an issue), while minimizing other social and psychological processes (e.g., reduce domination of discussion by high SES individuals and reduce selective exposure to only attitude-congruent information). Similarly, in the present work, technological attributes constitute a certain design that can hinder or facilitate various social and psychological processes, producing indirect and sometimes direct effects on important political outcomes.
Despite a strong emphasis on technology in this work, this author acknowledges that many human or social factors (e.g., usage of technology, contexts of usage, users of technology, content of media form, perception of the media form, motivations of the users, and others) are all extremely important to understanding and predicting communication effects. A thorough discussion on this topic is present in the Discussion section. However, the argument made in this work is that research on human and social factors would be most productive if it is based on a solid understanding of the “main effects” of technology on various political outcomes of interest.

This author disagrees with the technological determinist position of considering technology the only important source of insights for understanding media effects on individuals. The main reason only technological attributes are being examined in the present work is practical: Investigating all important social and human factors, as well as technological factors, in a single study is practically impossible. As in a properly-designed experimental investigation, the researcher cannot include every single important factor as an experimental manipulation. The present work only begins to examine the political effects of political UGC use by focusing on technological attributes. This line of research should be carried further by examining how various social and human factors interact or mediate the effects of technology, as well as produce effects independently, in order to obtain an intellectually satisfying answer to whether and how use of political UGC produces political effects.

*Technological Attributes of UGC and Their Hypothesized Influence on Political Participation*

In the present work the focus is on five technological attributes, each of which arguably (a) most explicitly characterizes political UGC and (b) appears to have the clearest relationship with political participation. This investigation does not attempt to either examine every single technological attribute of political UGC, or to examine every political UGC attribute that has political implications. Such a task would be well beyond any single research project. However,
this work covers most of the technological attributes that appear to characterize the essence of political UGC and that have strong implications for political participation. The attributes examined here, namely task-specific information retrieval performance, information environment customizability, content manipulability, direct participation facilitation, and community orientation, were selected through a process involving recurrent inductive and deductive methods. An exploratory content analysis of various political UGC Web sites was conducted. Subsequently, existing research on different related constructs, such as interactivity, social capital, political participation, or effects of different media types falling under the present political UGC definition was reviewed. Several iterations of these steps were carried out. Several of the attributes are similar to constructs that have long and rich research traditions, and which had to be adjusted to make them appropriate for the political UGC context (i.e., task-specific information retrieval performance, information environment customizability). Other attributes, although similar to some familiar concepts, had to be formulated without a coherent and ready-to-use theoretical foundation behind them (i.e., content manipulability, direct participation facilitation, and community orientation).

*Task-specific information retrieval performance.* This attribute enables users to quickly and easily locate a desired piece of information in the context of an information system (i.e., Web site) improving the system’s ability to efficiently and effectively locate information the user is looking for. Conceptually, this attribute makes it more likely that a user would find what (s)he wants when (s)he has a specific informational need or goal. This attribute helps conserve users’ cognitive, time, and effort resources as users are performing information search tasks. Task-specific information retrieval performance can operationally be represented by the latitude that users have in adjusting their search parameters during each search task. Such operationalization is deemed appropriate because it facilitates (a) an effective removal of irrelevant content; and (b)
Several types of features are expected to improve the task-specific information retrieval performance of an information system. The first group of features, which can broadly be labeled “content type sorting,” consists of the ability of users to sort (as in “ordering” all the data according to some criteria) or filter (as in removing some data) search results by content creation date, content source, content modality, and so forth. This group of features allows narrowing down results returned to the user according to various criteria that the user specifies. Some information retrieval tasks require the user to locate the most recently created content, whereas others might require filtering out all multimedia content, and other tasks might require the user to examine only textual content. Having various sorting options would increase the performance of the information system during those particular information retrieval episodes by making a given information system an efficient tool that could incorporate those various task-specific search parameters (Escudeiro & Escudeiro, 2009; Xiong & Wang, 2008).

An example illustrating how having more versus less information retrieval options might make it more likely that a user’s information needs would be satisfied might help clarify the proposed theoretical mechanism for this attribute. Imagine a user interested in finding a full-length speech on race given by Barack Obama during the 2008 presidential campaign, and recorded in high definition by some professional news organization. The user can use such keywords: “Barack Obama race speech.” Hypothetically, millions of search results might be returned, whereas only a small fraction of them represent what the user is looking for. If the user can search (or filter results) for video content versus other types of content, the number of
returned results will decrease, and the likelihood of the user finding what (s)he wants will increase. If the user is also given an option to search (or filter results) for high versus low definition video, the number of returned results will decrease, and the likelihood of the user finding what (s)he wants will increase further. Additionally, if an option to search for content produced by news organizations versus other organizations is offered, the number of returned results will decrease even more, and the likelihood of the user finding what (s)he wants will increase even more. Also, searching for speeches of certain length (e.g., 30 minutes long to 60 minutes long) will further reduce the number of returned results, and increase the likelihood of the user finding what (s)he wants. Finally, if searching for videos posted or created during 2008 versus other years is offered, the number of returned results will decrease, and likelihood of the user finding what (s)he wants will increase yet again.

The second group of retrieval performance indicators, which can be termed “social popularity indicators,” most directly represents the essence of UGC because it is based on “digital artifacts” left by the users or that reflect users’ interaction with the content. This group of features is structurally similar to the above group of “content type sorting” features; however, the main difference lies in the substantially greater role of the users in the development of the “social popularity indicators” group of features. This group can be represented by the technological ability of an information system to sort through content based on the number of user-generated comments or replies to the specific piece of content, content ratings generated by users, number of content views or clicks, number of times content was shared, and so forth. Some features in this group (e.g., ratings) provide a reliable indicator of the content’s quality (as will be described more thoroughly below). Other features (e.g., number of times the content was shared or viewed) might be an indicator of the content’s overall quality, as well as of some specific facet of the content (e.g., sensationalism, creativity, entertainment value, production value). In all cases, this
group of features helps identify content of interest to users during particular information retrieval tasks, such as locating the most entertaining or creative or high-quality content (Xiong & Wang, 2008). There is evidence that some of the features from this group (i.e., user tags) improve performance of the general information retrieval tasks (Sinclair & Cardew-Hall, 2008).

Once again, it might be helpful to offer an example to illustrate the appropriateness of using this group of features to tap the information retrieval performance attribute. If we use the same search scenario from the previous group of information retrieval features, a user might want to find not just any speech on race by Barack Obama, but his most famous speech (because Obama might have plausibly given many speeches on the issue). To do so, user can sort the search results by “view count” or the overall rating of the video. Such sorting capability will make it easier for the user to locate the specific content (s)he wants, and thus satisfy his or her informational need more efficiently.

Because the “social popularity indicators” group is the most directly relevant to UGC, this group of features deserves a more detailed elaboration. Social popularity indicators exemplify the increased role and increased control of users over the online information environment. When users contribute to the creation of and use the system of social indicators of popularity, they perform what Bruns (2005) calls “gatewatching,” which is contrasted with “gatekeeping.” In gatewatching, users observe information that flows from various different sources and they choose whether or not to recommend it, how to vote on it, how to tag it, and how to otherwise interact with it, making it easier for other users to determine if this content is relevant and/or worth consuming. Voting, recommending, or tagging some content is a behavior indicating that some particular content stands out in some fashion. The number and positive valence of votes and tags, as well as the number of recommendations, clearly taps the “social popularity” concept. The popularity of such Web sites as Reddit (ranked by Alexa, 2011, as the 50th most trafficked Web
site in the U.S.), SlashDot (ranked by Alexa, 2011, as the 708th most trafficked Web site in the U.S.), and Digg (ranked by Alexa, 2011, as the 99th most trafficked Web site in the U.S.) shows that the social indicators of popularity system is perceived to be useful by large audiences. Each of these Web sites gives a substantial amount of control to their users in determining which content on the Web site should be more or less prominently displayed. According to Sunstein (2006), massive groups of online users can act as a sophisticated arbiter of content quality. If one person notices some content and decides to recommend it, or leave a positive comment on it, or repost it to their Facebook page, chances are there is something in that content making it worthy of consumption (Jarvis, 2008). If, however, thousands or millions of people notice some content and collectively agree that the content is worth recommending, then it is an almost absolute certainty that the content possesses valuable characteristics that peers of this massive group of users would find equally worthy of their attention. This parallels the logic of collective intelligence, discussed by Chadwick and Howard (2009) and Sunstein (2006), where amateurs working collaboratively and voluntarily can outperform experts working alone and for pay. The success of the social indicators of popularity system and of gatewatching might suggest that massive numbers of online users working collaboratively and independently can often make equally good (or even superior) judgments about the quality of content compared to professional editors of traditional news organizations, which makes reliance on social indicators of popularity an efficient information search strategy.

How can improving task-specific information retrieval performance influence political participation? In short, information retrieval performance is expected to increase political participation; however, such effect should be conceptualized as a multi-step mediation model. Being able to quickly and easily locate the desired political information can be expected to reduce the cognitive or mental load on the users. Reducing demands in one part of the cognitive system,
leaves more resources that can be used in other parts of the cognitive system, and the important task becomes to minimize or redirect cognitive demands from non-essential cognitive tasks (looking for information) to the essential ones (deeply processing this information and learning) (Mayer & Moreno, 2003). The reduction of cognitive load, associated with reduced time and effort needed to identify relevant content, should leave more cognitive resources that can be used to deeply process the retrieved information (Eveland & Dunwoody, 2000). Deeper processing is associated with an improvement in learning (Anderson & Reder, 1979; Craik & Tulving, 1975; Salomon, 1983, 1984). Additionally, the relevance of the retrieved information in itself is known to increase one’s depth of information processing (Frymier & Nadler, 2007; Graber, 1988; Lang, 2000; Wise, Bolls, & Schaefer, 2008), and such depth is directly and positively related to learning. Greater learning from exposure to political UGC is likely to facilitate political participation. As was discussed above, there is substantial evidence that political knowledge is a strong and positive predictor of political participation (Cho & McLeod; 2007; McLeod et al., 1999; Scheufele, 2000; Scheufele, Nisbet, & Brossard, 2003; Scheufele, Nisbet, Brossard, & Nisbet, 2003).

However, having “too many” choices for structuring one’s information retrieval query, which is identical to having very high levels of this attribute presence, may become ineffective in terms of increasing political participation. Very low and very high levels of this attribute presence might unnecessarily increase cognitive load. Research in educational psychology shows that users benefit from encountering interactive features in a multimedia learning system, but the users can be overwhelmed by too many interactive features because it takes effort to analyze and respond to each feature (Mayer & Moreno, 2003). The general logic of such a finding is applicable to the information retrieval domain, where having too many information retrieval options might increase the cognitive load because it would take the user more time and effort to decide which
options to use during any given information retrieval task. If this expectation is accurate, then the information retrieval performance attribute offers “diminishing returns” in terms of facilitating political participation: The effect can be positive up to a certain point, where it disappears. Whether such curvilinearity exists, and at which point the curve becomes flat or even bends down, are interesting empirical questions that need to be investigated.

Another important consideration is the possibility that some social-popularity indicators can be rigged. For example, it is quite plausible that some overzealous marketing firm would anonymously leave positive ratings for the products of its clients. In such a case, use of ratings might increase the user’s frustration with the Web site, instead of reducing cognitive load. Such rigging attempts can be countered with technology: Web site can create a reputation system, where ratings of credible, regular, and expert users can count more than ratings of anonymous “drive-by” users whose reputation is unknown. However, the recent redesign of the Web site Digg suggests that even such measures are far from perfect (Wasserman, 2010).

*Information environment customizability.* This technological attribute enables users to modify their personal information environment by systematically and automatically excluding disliked sources of information and disliked content topics, and by selecting the preferred sources of information and preferred content topics that are then automatically and consistently displayed to users in a variety of forms (e.g., RSS, email summaries, embedded news feeds). *Systematic* and *automatic* exclusion or inclusion of certain content is important to the conceptualization of this attribute, because this attribute taps how the user screens out certain information due to relatively long-term content preferences, which can be contrasted with the previous attribute in which task-specific or short-term content preferences determine which information is screened out and which one is included during each information retrieval episode. Information environment customizability refers to the user being technologically empowered to alter the default
information flow inside information systems, often by making it match their predefined content preferences (Kalyanaraman & Sundar, 2006). An important point is that only content available to that particular user is modified, while content available to other users remains the same.

Technologically, information environment customizability can be explicit, where users take specific steps towards adjusting their information environment. But, also can be implicit, where the information system analyzes users’ online behavior and customizes the information environment based on that behavior, without any purposeful effort from the users. Explicit customizability can be conceptualized as a technological “tool” designed to help user customize his or her information environment and offered to the user by the information system, whereas implicit customizability can be conceptualized as a system that is using such a tool, instead of simply offering it to the user and expecting the user to employ it. Although the two customizability forms are different in terms of intentionality and effort on the part of the users, both are expected to produce similar effects, which are discussed below. The similarity in effect is expected because both types of customizability enable user to automatically and systematically screen out some information and include other information, and where such content discrimination occurs over a relatively long term. In other words, these two types of customizability provide slightly different mechanisms for achieving the same result. Therefore, both of these customizability dimensions are present in the information environment customizability attribute.

The information environment customizability attribute can be operationally defined by presence of several groups of technological features, displayed in Figure 1. The first group of features, “explicit people customizability” refers to the user being technologically empowered to select individuals that (s)he wants to be friends with, or communicate and share information with. Some specific technological features representing this group are: (a) the ability to pick “friends”
or social network members, and receive information from them; (b) the ability to adjust (block, reduce) the level of the amount of incoming communication from particular user(s) on the Web site, one’s “friends” or social network members; and (c) the ability to subscribe to channels of users, or to “follow” certain users or groups. This group of features is expected to validly operationalize the information environment customizability attribute because individuals obtain large quantities of information from interpersonal sources (one’s social network members), and who those interpersonal sources are should be important to the type of information that ends up surrounding the user in context of a given information system.

<table>
<thead>
<tr>
<th>Explicit</th>
<th>Implicit</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>People</strong></td>
<td>Ability to pick social network members; ability to adjust the amount of incoming information from particular user(s)</td>
</tr>
<tr>
<td><strong>Content</strong></td>
<td>Ability to adjust topical domains/sources to be displayed to the user when (s)he visits the Web site</td>
</tr>
</tbody>
</table>

Figure 1. *Four Groups of Technological Features Representing Information Environment Customizability.*

The second group of features, “implicit people customizability,” refers to features whereby the information system automatically suggests social relationships based on preferences of one’s social network members, or automatically suggests “friends” to follow or users to
subscribe to based on the proximity of those individuals to one’s profile information. This group of features also appears to validly tap the customizability attribute because it is very similar to the “explicit people customizability” group discussed above. The only difference from the previous group is that the system encourages the user even more aggressively to select some social network members and not others.

The third group of features, “explicit content customizability,” refers to user being given technological tools to alter his or her information environment by purposefully choosing the sources, or types of content, he or she is consistently receiving. This group can be represented by such features as: (a) the ability to subscribe to specific sources, channels in context of a Web site; (b) the ability to adjust the default settings allowing one to choose only some topical domains to be displayed to the user when (s)he visits the Web site’s homepage or one’s profile page. This group of features accurately represents the customizability attribute because the user’s information environment can be significantly modified, and such modifications persist over time.

The fourth group of features, “implicit content customizability,” refers to an information system automatically offering users certain content based on different parameters, such as their current and past searches, their content preferences, and their profile data. This group of features also appears to be a valid operationalization of information environment customizability because the system actively customizes the user’s information environment based on the user’s past actions or profile characteristics.

Some similarity of the task-specific information retrieval performance and information environment customizability should be acknowledged, since both help search for and filter out certain type of information (Belkin & Croft, 1992). Both help individuals reduce effort when interacting with vast amounts of information, which conserves cognitive resources that can be applied to processing information deeply. However, there are several important distinctions. First,
high information environment customizability could encourage selective exposure to attitude-congruent content, whereas high task-specific information retrieval performance does not encourage it. Second, highly customized systems (that have low scores on other attributes) can contain vast amounts of relevant information that are never made available to the users. Because users adjust their default information intake, they screen out substantial amount of information on a consistent long-term basis, which they have determined to avoid, but which, depending on the search task, might be highly relevant. The result is a greatly reduced possibility of serendipitous exposure to relevant information during various information search tasks. The possibility of such serendipitous exposure is greater in information systems with high information retrieval performance (and low on other attributes), because no information is screened out by default or over the long term. Instead all information is analyzed before the system determines whether it is relevant or not and presents users with results to their search queries. When some information is screened out, it occurs on a case-by-case (or search-by-search) basis. In other words, all or most of the content delivered to the user in such an information system is expected to be relevant, whereas the content in a highly customizable system might not be particularly relevant, and some of the (serendipitously) relevant content might be inadvertently screened out in a highly customizable system. However, it should be acknowledged that, at least in theory, customizability or personalization is meant to deliver highly relevant information. For example, Negroponte (1995) suggested that future computers will know so much about each individual user that they will be able to discern our needs with the same degree of subtlety as other human beings. Despite prominent recent examples of Amazon providing useful customized product recommendations or Facebook serving ads that match one’s psychological and behavioral profile, this vision appears to remain rather futuristic. In the future such customizations might very well become extremely effective in delivering highly relevant content (and avoiding irrelevant content). As of today, the
above suggestion that information systems with high task retrieval and low customizability will deliver more relevant content to users than information systems with high customizability and low task retrieval appears to be very plausible.

High information environment customizability is expected to increase political participation, again through a multi-step mediation process. High customizability allows individuals to submerge themselves in a homogenous and attitude-congruent information environment, where sources of information (organizations or individuals) offer content with which users agree. Cases in which customizability might not lead to selective exposure will be discussed below. Most individuals generally, and particularly individuals with high prior knowledge in the relevant topical domain and with strong attitudes on the relevant issues, are expected to consume more and process more deeply such information that fundamentally agrees with their prior viewpoints, while avoiding and processing less deeply information that disagrees with their prior viewpoints (Iyengar & Hahn, 2009; Knobloch-Westerwick & Meng, 2009; Taber & Lodge, 2006; however see Garrett, 2009 for evidence of individuals not screening out attitude-incongruent information as much as conventionally thought). In the political communication domain, the information environment customizability attribute can contribute to a spiral-like process. When users are consistently exposed to primarily attitude-congruent information, and process such information at deep levels, they become more likely to seek out the already well-understood attitude-consistent information and become more capable of deeply processing such information. This occurs because each instance of exposure and processing of familiar attitude-consistent information makes an existing mental schema, containing this information, denser (Graber, 1988). Mental schemas are dense when links (connections that can be quickly activated) between the nodes (concepts, cognitions, etc.) are frequently activated and become strong (one node quickly activates adjacent nodes) and redundant (one node can activate adjacent nodes
through a greater number of links). Each instance of exposure and processing of familiar attitude-consistent information also expands an existing mental schema, in the sense of connecting more and more nodes to it. All of this facilitates deep processing and subsequent integration of the attitude-consistent information into one’s mental schema in the future. This perpetuates the tendency to selectively consume and deeply process attitude-congruent information, while screening out and processing shallowly attitude-incongruent information. Such a spiraling process strengthens one’s existing views (e.g., increasing one’s attitude extremity) and reduces attitudinal ambivalence (Sunstein, 2002). Strong political attitudes and lack of political attitude ambivalence are known to be strong positive predictors of political participation (Mutz, 2002).

The above discussion of attitude-congruent information can help further crystallize the difference between task-specific information retrieval performance and information environment customizability. It is proposed that attitude-congruent information will be much more frequently made available to individuals in highly customizable systems (with low scores on other attributes), and less so in systems with high task-specific information retrieval performance (with low scores on other attributes). Admittedly, attitude-congruent information can be obtained in systems that have high task-specific information retrieval performance. However, the individual would have to make a conscious effort to seek out such information, during every single information retrieval episode. In the context of highly customizable systems, the attitude-congruity preference kicks in automatically, as suggested by numerous selective-exposure investigations (Iyengar & Hahn, 2009; Knobloch-Westerwick & Meng, 2009; Taber & Lodge, 2006), which produces a situation in which primarily attitude-congruent information is delivered to the individual without any conscious effort on the part of individual to seek out that particular type of information during every information retrieval episode. Additionally, such information may or may not be particularly relevant to the individual’s specific goals during search tasks.
Effects of this attribute might be limited by how information system administrators choose to implement several groups of customizability features. For example, if the software code used to recommend content to the users is based on determining similarity of user’s political ideology and the political ideology represented in the content, and then choosing to recommend the content that is most similar to the user’s ideology, the selective exposure effect will result. However, if the system is designed to estimate how topically similar a given piece of content is to one’s interests (as expressed in one’s profile, or through prior content exposure behavior), or how similar other users are in terms of age, geographical location, or hobbies, then the system’s recommendations of content or “friends” will not necessarily increase selective exposure in terms of attitude congruent v. incongruent information. Therefore, exactly how these implicit customizability features are implemented appears to be important to determining the ultimate effect of the information environment customizability attribute.

However, it can be argued that one’s hobbies, geographic location, and other psychographic characteristics are over- or underrepresented in people of various political ideologies (e.g., there is a correlation between income [which often restricts the hobbies one can choose among] and being conservative [Gerber, Huber, Doherty, Dowling, & Ha, 2010; Zipp & Fenwick, 2006]; and there is a correlation between political ideology and different psychological characteristics [see Gerber et al., 2010; Tetlock, 2005]). Additionally, it can be argued that even when the information system suggests certain pieces of content, or certain users as potential “friends,” individuals might choose not to follow the system’s recommendations. That is, an individual seeking out politically like-minded social network members may ignore the system’s recommendations when the recommended “friends” do not match the user’s ideological preference. If individuals are motivated to expose themselves to primarily attitude-congruent
information, selective exposure might still result even from using information systems not explicitly designed to promote such selective exposure.

It should also be acknowledged, however, that there might be some groups of individuals and some contexts in which a greater number of customizability options will not necessarily lead to greater selective exposure to attitude-congruent information. For example, if individuals are experts in some domain, they might not want to limit their information exposure to just attitude-congruent information. These individuals might feel comfortable encountering attitude-incongruent information because they can easily counter argue it (Taber & Lodge, 2006). Another group of individuals, the “news junkies,” who consume vast amounts of news content and do not discriminate among attitude-congruent and attitude-incongruent information. For this group of individuals, informational value of diverse information is of prime importance. Yet, another group of individuals are those experiencing an emotional state of anxiety and for whom informational utility is extremely important (Valentino et al., 2009). For this group, the value of information “diversity” might outweigh the desire to seek out attitude-congruent information. Finally, some individuals might experience less of a “psychological pleasure” from encountering attitude-congruent information, which might stem from the fact that these individuals are characterized by personality characteristics such as as open-mindedness or experience-seeking, or if these individuals score very low on such traits as dogmatism, authoritarianism or intolerance of ambiguity (Lavine, Lodge, & Freitas, 2005; Ozer & Benet-Martinez, 2006). Another plausible reason for experiencing little psychological pleasure from encountering attitude-congruent information is that some individuals simply do not hold any particular position in the debates or divides that are generally present on most socio-political issues. Such individuals might be interested in an issue, but perhaps they have not yet formulated their own position on an issue (e.g., they are uninformed), or perhaps all positions appear equally
valid to them (e.g., they are cross-pressured or deeply ambivalent). For the above groups of individuals, or under the above conditions, having more customizability options in an information system might not lead to a greater selective exposure, and might actually result in a reduced selective exposure. The above described groups of individuals and contexts, however, arguably represent a rather limited number of cases. It is plausible that in the aggregate (across all available cases), having more customizability features should be expected to increase selective exposure – although for some groups it might be a small increase, for others groups it might be a large increase, and yet for other groups there may be no increase or even a decrease in selective exposure.

*Content manipulability.* This attribute allows users to modify information available to everybody on a Web site. Concepts like "technological interface" interactivity (Kiousis, 2002; Sundar, 2004) and "adaptive interactivity" (Deuze, 2003) are somewhat similar to the idea of content manipulability – all tapping the degree to which content found on the information system can be modified as a direct result of the users’ actions during their normal use of the Web site.

Content manipulability can be operationally defined by the presence of several groups of features. The first group, "main content manipulability," refers to technological features that enable the user to substantially alter the content itself. This group is represented by such features as: (a) the ability to directly upload content (make posts, upload video, audio, graphics, etc.) or edit a wiki; (b) the ability to start a blog on the site. This group validly represents content manipulability attribute because, by definition, the user can contribute to and modify the actual content on a Web site.

The second group of features, "meta-data manipulability," refers to technological features enabling user to create or modify “secondary” or “non-essential” content on a Web site. The following features represent this group: (a) the ability to tag or classify content; (b) the ability to
vote on or rate content; and others. The terms “secondary” or “non-essential” do not mean “unimportant.” These terms simply refer to the idea that without the main content (e.g., an article about Obama) users will not be able to use these features (e.g., one cannot rate an article about Obama, if this article does not exist); while users can use the main content without these “meta-data” manipulability features. These particular features facilitate the development of content that serves to enhance the primary content on a Web site. Such “non-essential” content as content tags or votes represents a unique type of content because it describes or adds new and meaningful dimensions to the primary content. In some information systems, such as Amazon or Digg, this supplementary content actually serves a very important function and adds a substantial value (in case of Amazon, product reviews and ratings might improve the quality of purchasing decisions; in case of Digg, the number and type of votes on news stories directly determines how prominently they are displayed and can play a crucial role in determining whether any given piece of content is viewed or not) to the primary content.

Some of the same technology is tapped in both task-specific information retrieval performance and content manipulability attributes. However, the technology is implemented in a different fashion. For example, the ability to see a total number of comments a given piece of content receives taps task-specific information retrieval performance, whereas ability to comment on a particular piece of information taps content manipulability. These two features are based on the same idea of commenting. However, these two features are different. First, information system often allows one to leave a comment on a particular piece of content, whereas it does not allow one to sort the content by the total number of comments. In this case, one part of the commenting functionality is enabled, whereas the other is disabled. Second, although based on the basic idea of commenting, one implementation taps the usefulness of the total number of comments as a heuristic for judging the content quality, whereas the other implementation taps
the usefulness of allowing readers to comment to increase (perceptions of) interactivity available on the Web site. Therefore, it is argued that both task-specific information retrieval performance and content manipulability are two different technological attributes.

Also, the difference between content manipulability and information environment customizability needs to be delineated. Both generally tap the degree to which information available on a Web site can be modified. In the case of customizability, filtering of the incoming information that is available to oneself is the focus. In the case of content manipulability, active production of information that is then available to every Web site user is the focus. For example, one user can submit a video on YouTube (which taps content manipulability of YouTube), whereas another user can choose to subscribe to the first user after having seen the video (which taps customizability of YouTube). The resulting effects of the two are also expected to diverge, as discussed next.

Content manipulability is expected to increase political participation indirectly – through political efficacy and political knowledge. When users are given the opportunity to have an impact on their information environment by creating or modifying political UGC, they see how their political actions (being defined as posting a political video, or editing a political article on a wiki page, etc.) produce tangible results. Seeing such results should create a sense of empowerment, which is internal political efficacy represent. Such theorizing has been partially confirmed by Leung (2009) who found a weak, but statistically significant, connection between generating online content and political efficacy and psychological empowerment. Political efficacy, as previously mentioned, facilitates political participation (Finkel, 1985; McLeod et al., 1999; Scheufele et al., 2003).

Besides efficacy, content manipulability should improve political learning. Being an active content producer and consumer, which is facilitated by content manipulability, leads to the
following process: When users of a Web site are allowed to create or modify the Web site’s content, they may process the content more deeply, even without actually creating or modifying any content. Eveland (2004) described a phenomenon that he called “anticipatory elaboration” and Pingree (2007) examined a similar phenomenon that he called “expectation of expression”, both of which refer to the effects of one’s expecting to engage in political conversation. When an individual anticipates that she will be discussing politics with politically-interested co-workers, friends, or family members, the individual tends to pay more attention to political news, and elaborate (which is almost synonymous with “deeply process”) more on that information. Both attention and elaboration increase political knowledge, since both help individuals retain more political facts in memory. Therefore, even the potential of political conversation (which is similar to a potential of creating or modifying political content on Web sites with high content manipulability) facilitates deep processing and political learning. Although an interpersonal conversation with members of one’s social circle and modification of content on a Web site are different behaviors (the former has a dimension of social peer pressure encouraging a person to learn about the topics her social circle is interested in, whereas the latter may not have that dimension), they are both characterized by an opportunity to express oneself. In the case of content manipulability, such an opportunity is coupled with a sense of political efficacy, which can produce a strong motivation to learn and think more about politics. Greater political knowledge, as mentioned earlier, is a strong, positive direct and indirect (through efficacy) predictor of political participation (Finkel, 1985; McLeod et al., 1999; Scheufele et al., 2003).

Direct participation facilitation. This attribute directly helps users get involved in politics by reducing the costs associated with political participation (e.g., costs in terms of effort, time, money, and other resources). Research on the role of technology in promoting collective action, facilitating social movements, and encouraging political participation (Bimber, 2003; Garrett,
The direct participation facilitation attribute is operationally defined by presence of “political hyperlinks,” such as: (a) a hyperlink for sending email messages to public officials or government agencies that can be used in support or opposition to some political cause; (b) a hyperlink to a letter form that is either blank, partially filled, or pre-filled, which user can complete and send to a pre-filled email address (of a public official, agency, etc.); (c) a hyperlink to a Web site where a monetary donation can be made; (d) a hyperlink to an online “townhall” or space where citizens can meet and interact with public officials or political candidates. These features validly tap the direct participation facilitation attribute because they reduce the costs of participation. Clicking on the hyperlink conveniently brings users to a destination where they can engage in actual political participation acts (e.g., donating money, contacting politicians or media, attending “virtual meetings”).

This attribute combines the Web site’s content and technological structure. The technological structure is represented by such elements as use of hyperlinks. The content element is represented in the destination of the hyperlink. In other words, the presence of the hyperlink capability in itself is necessary, but not sufficient for properly representing this attribute. The nature or content of the link is equally important.

The direct participation facilitation attribute should be differentiated from Lemert’s (1984) mobilizing information. Mobilizing information is information such as an address of a government agency, phone numbers, or tactical suggestions on how citizens can organize a protest or a grassroots calling campaign. In essence, mobilizing information educates individuals
about how to get politically engaged, and it stops there. However, there is some evidence that exposure to mobilizing information increases intent to participate in politics (Bybee, 1982). The direct participation facilitation attribute picks up where mobilizing information leaves off. Direct participation facilitation does not educate the users, but instead it provides practical and immediate tools that help users to get politically engaged. Both mobilizing information and the direct participation facilitation attribute should produce the same result – directly increase political participation – but their methods are different.

The effect of this attribute is very straightforward: by making participation more efficient (as a result of reducing costs associated with participation), political participation should increase. The reduction of costs is a powerful economic force in the political context (Downs, 1957). Downs (1957) and Wielhouwer and Lockerbie (1994) suggested that citizen participation in politics seems almost irrational, given the very low probability that their individual actions would have any impact, and given the costs associated with such participation. Decreasing the costs, or increasing the (perception of) impact one’s participation can have on political outcomes are known to increase political participation. The direct participation facilitation attribute examined here decreases the costs of participation.

However, Bimber (2003) offered anecdotal evidence of political participation acts that were easy to engage in (e.g., emailing public officials, posting comments on blogs) being less effective at getting politicians’ attention compared to the acts that are harder to engage in (e.g., visiting the politician’s office, making a personal phone call, writing a hand-written letter, hiring a lobbyist). If the “cheap” political communication and participation activities are less effective, it can discourage citizens from engaging in such acts, because citizens’ perception of how much their actions impact on the political system will be negatively affected. However, the comparative
influence of such acts on the political system and the popularity of easy-to-engage-in acts among citizens have yet to be empirically established.

Additionally, when political participation becomes an easy-to-engage-in activity, it can actually undermine the experience of political participation. If political participation has traditionally been an activity among politically interested and like-minded individuals, reduced costs of participation can attract individuals less interested in finding a like-minded community and potentially even attract one’s political opponents willing to subvert political efforts of some group. For example, an opponent of some political candidate can easily click on a hyperlink bringing the individual to the politician’s “online public meeting” Web site. The individual can then proceed to flame, troll, or subvert the discussion, or otherwise attempt to alter the dynamic of the meeting. Such practices can spoil the experience of the politician’s supporters, who might become less willing to engage in these sorts of political participation acts in the future. Thus, while facilitating political participation among like-minded people, this attribute can also undermine such participation by making it easier for intruders to spoil the participation experience.

Community orientation. This attribute encourages users to become members of a greater number of online communities and encourages more frequent communication among these users. An online community is defined as a group of individuals who interact, share information, or congregate in the same online space because they share some common interest. The interest they share can be anything, from hobbies (e.g., athletic activities), to political preferences, to diseases, to geographic proximity, to familial ties (Wellman et al., 2003).

The community orientation attribute is operationally defined by the presence of two groups of different technological features. The first group is focused on the presence of trust-facilitating “reputation information” (Resnick, 2002, p. 9; Dellarocas, 2010) about individual
These features include: (a) publicly displayed information about number and types of fellow user’s “friends,” “fans,” “followers;” (b) seniority or length of being a Web site’s member; (c) “rank” in virtual hierarchy within a particular online environment; (d) cumulative ratings or other feedback from other users. This group of features validly represents the community orientation attribute because these features bring out information useful for making a decision about whether to become a group member or not, and features that encourage users to be good “members” of their online community. More details about how this group of features facilitates community development are described below.

The second group of features allows the user to identify communities existing on a Web site, and identify reputational information of these communities. This group is represented by such features as: (a) ability to locate groups existing on a Web site; (b) ability to see the number and types of group members, followers, other organizations affiliated with online groups; information about how long the group existed; (c) ability to sort the existing groups by various criteria (e.g., topic, size, seniority). This group of features represents the community orientation attribute because it helps users make efficient and accurate evaluations of various groups available in the context of a Web site, and select the ones that match the user’s preferences, or decide that the groups do not fit well with the users’ preferences.

The community orientation attribute is expected to facilitate political participation, through several multi-step mediation mechanisms. Specifically, the impact of this attribute on political participation is expected to occur as a result of putting in motion three social and psychological processes: (1) connection of individuals to more people and more social networks; (2) increase in per capita social capital; and (3) increase in political discussion frequency.

The first process, connecting of individuals to a greater number of people and social networks, is the direct outcome of individuals being able and willing to join more online
communities (Resnick, 2002). The provision of reputational information and the ability to join or start an online community should increase the ability and willingness of individuals to join online communities that fit their needs. It also makes it likely that these individuals continue to be loyal and “good” members of the communities, which might attract an even greater number of members (Dellarocas, 2010). When one is connected to a larger number of individuals and social networks, theory suggests that they are more likely to participate in politics because if an individual is connected to more people, the probability of being asked to participate is greater (Verba et al., 1995). Research on participation in voluntary civic organizations and social movements finds the same effectiveness of the interpersonal recruitment mechanism for getting people involved (Garrett, 2006; Klofstad, 2007; Stein, 2009).

The second process, accumulation of per capita social capital (unit of social capital per person), is the outcome of a greater number of individuals connected to a greater number of networks. Such an expectation is purely probabilistic. Social capital is generally regarded to be a property of social network. Some social networks have more of this property, others less (Coleman, 1988). The community orientation attribute facilitates the formation and expansion of social networks, which creates a potential that networks with at least some social capital will be formed. The more social networks with the social capital property exist, the more per capita social capital should be expected to exist (given that the number of networks with social capital property increases, while the number of individuals in the general population remains constant, the “per capita” social capital must increase). Reputational information is considered to be particularly important for social networks possessing high degree of social capital to be formed. Reputational information enables individuals to evaluate the available social networks they can choose from, and then commit to the network that they feel suits their preferences. As mentioned above, reputational information should also encourage individuals to be “good” members of the
social network, facilitating trust and healthy interpersonal dynamic among the members
(Dellarocas, 2010). Social capital has been found to increase individual-level political
participation (Putnam, 1995; Skoric et al., 2009; Valenzuela et al., 2009).

Finally, being connected to more people and more social networks makes it likely that an
individual will have more frequent interpersonal discussions. As the frequency of interpersonal
communication (both online and offline) increases, the frequency of political and civic discussion
should also increase (Eveland & Kleinman, 2010, November). Political discussion by itself is a
strong influence on one’s political knowledge since individuals have more opportunities to
elaborate on and make sense of the political content they have heard or seen elsewhere (Eveland,
2004). Additionally, even the potential of interpersonal political discussion might encourage
individuals to more deeply process political information, which further improves political
learning (Eveland, 2004; Pingree, 2007). Political knowledge is a positive predictor of political
participation. Finally, increased frequency of political discussion is also expected to improve
one’s odds of being asked to participate and ultimately be recruited into political action (Verba et
al., 1995), and being asked to participate in politics is a strong positive predictor of actual
participation (Garrett, 2006; Klofstad, 2007; Stein, 2009; Verba et al., 1995).
Chapter 6: Hypotheses and Research Question

The major goal of the empirical part of this study is determining how much each technological attribute is present on the most popular Web sites belonging to these UGC categories: blogs, wikis, content-sharing Web sites, participatory journalism Web sites, and discussion forums. Social-network Web sites were removed due to practical impossibility to sample the needed content on those Web sites. As described above, exploratory content analysis and prior research suggest that there is a strong reason to expect the five technological attributes to be present in context of the political UGC Web sites (Boyd & Ellison, 2008; Bruns, 2005; Domingo et al., 2009; Jarvis, 2009; O’Reilly, 2005; Sunstein, 2006; Valenzuela, et al., 2009; Wright & Street, 2007).

To reiterate, several characteristics of UGC were identified above: (1) Greater control over the information (Bruns, 2008; Harrison & Barthel, 2009; Sunstein, 2006). Such control might take the form of offering users technological capabilities for content editing, commenting, rating, and modifying. This characteristic is partially a result of the existing technology, and partially a result of the “produsage” culture discussed above (see Bruns, 2008). This characteristic strongly overlaps with the conceptualization of the manipulability attribute. Customizability attribute also resulted from the increased user control over the information. However, such control takes the form of adjusting what information to let into one’s communication environment, and what information to filter out. Therefore, the increased control characteristic also overlaps with the customizability attribute. (2) Collective intelligence, or accumulation of the meta-data, is another characteristics which might result from the first characteristic (Bruns, 2008;
Jarvis, 2009; Sunstein, 2006). Creation of “content about content” is theorized to be an important technological affordance in the present work. This characteristic partially overlaps with the idea behind information retrieval attribute. (3) Users share content, while developing social relations with one another (Harrison & Barthel, 2009). Such practice over time might have led to formation of structured communication spaces, such as social-network sites and discussion forums. Development of reputational information is a logical outgrowth of sophisticated and elegant forms of online communities, helping users assess each other and various online communities. Reputational information is a substantial conceptual part of the community orientation attribute. Finally, there is evidence that novel user-driven media forms, such as political blogs are characterized by a different set of norms, values, goals, and operational procedures (Dylko & Kosicki, 2006; Woodly, 2008). Specifically, political news blogs are expected to take a more activist role and be less limited by a desire to appear objective, than traditional news organizations (Davis, 2009; Ekdale, Namkoong, Fung, & Perlmutter, 2010). Such theorizing underlies the expectation that political UGC, which in many ways is similar to political blogs, should be characterized by presence of participation facilitation attribute.

Based on prior research (Bruns, 2005; Domingo et al., 2009; Jarvis, 2009; Sunstein, 2006), and based on the present explication work, it is expected that attributes should have a greater presence on political UGC than on traditional news Web sites. To explicate UGC through its attributes, it is necessary to demonstrate that the scores of UGC Web sites on these attributes are statistically different than those of non-UGC Web sites. To do so, it is important to select an appropriate reference (or contrast) group whose scores on attributes would be recorded and then compared against those of UGC. Using Web sites of traditional journalistic organizations seems to be appropriate.
Traditional journalistic organizations’ Web sites (the most popular Web sites of newspaper, TV or cable, and magazine news organizations) are relatively similar to UGC in terms of the broad environment in which they operate (i.e., Internet), the end product (i.e., information available to the public), and topical orientation (i.e., political content). Because of such similarity, showing that UGC Web sites differ from non-UGC Web sites is more challenging, but would be more theoretically interesting. Using a reference group drastically different from UGC (such as printed political books, or news programs shown on television) would make it theoretically less interesting because results would be very unsurprising.

Traditional journalistic Web sites have been recently shown to adopt many of the UGC innovations, such as commenting on stories, hosting discussion boards, and so forth (Domingo et al., 2009; Thurman, 2008) in order to prevent further erosion of audience and to better engage the audience in the news production process to improve the quality of the news products. However, due to (1) their operational modes, characterized by frequent necessity to appeal to a large audience, having large bureaucracies with numerous employees, dependence on advertising and being relatively slow and risk averse in terms of adopting technological innovations (Jarvis, 2009; Sigal, 1973; Shoemaker & Reese, 1991), and (2) values, exemplified most vividly by adherence to the notions of accuracy, objectivity, and displaying a sense of unique authority or entitlement over news production and delivery, and disdain towards citizen journalists or alternative ways to report and deliver news (Dylko & Kosicki, 2006; Gans, 1979; O’Sullivan & Heinonen, 2008; Trippi, 2004; Woodly, 2008), these media organizations on many levels are far from the pure UGC forms examined here. Therefore, contrasting UGC against traditional news organizations’ Web sites appears theoretically interesting and appropriate for the goals of this investigation.

Therefore, it is hypothesized that:
H1: The five attributes should have a greater presence (number of technological features) in political UGC (most popular blogs, wikis, content-sharing Web sites, participatory journalism Web sites, and discussion forums) than in traditional news Web sites (most popular newspaper Web sites, TV or cable Web sites, and magazine Web sites).

The five attributes are theorized to be important to defining political UGC and explaining its effects on political participation. However, the existing classification of political UGC does not appear to be associated with the relative presence of these five attributes (or with any other systematic classification principles).

Additionally, exploratory content analysis of several exemplars of UGC indicated that presence of attributes should vary across different UGC groups. For example, a nonsystematic examination of different Web sites across various UGC types suggests that information retrieval might have only low presence on blogs and wikis, moderate presence on participatory journalism sites, and high degree of presence on content-sharing sites and discussion forums.

In regards to customizability attribute, a nonsystematic examination of UGC Web sites suggests that there might be a relationship between the number or diversity of sources that contribute content to an information system, and the number of customizability features. Herring, Scheidt, Kouper, and Wright (2007) found that blogs are most often written by one individual (e.g., Instapundit), or by several individuals who often share the same ideology or interest (e.g., PowerLine). Because such structure produces relatively homogenous content (e.g., there is lack of oppositional content), it appears that customizability options might be relatively less useful. However, on content-sharing sites or discussion forums, a great diversity of existing sources and content makes customizability options more useful, enabling users to screen out unwanted sources or content in a systematic and efficient fashion.
Similarly, exploratory examination of UGC Web sites also suggests that we should expect to find variance in content manipulability. The presence of this attribute might depend on the amount of control over the content that Web site owners want to retain, which in turn depends on what amount of control is the most optimal for a successful operation of the Web site. For example, the idea of a wiki or content-sharing Web site will simply not work, if users are not given substantial opportunities to create and share content. On the other hand, blogs that do not allow user comments, or participatory journalism Web sites that minimally utilize amateurs might still have viable business models. Similar expectation of variance is justified for participation facilitation and community orientation attributes.

However, due to the fact that all five attributes are being currently explicated in the present work, and the fact that they have not been extensively examined in prior research, it is impossible to formulate precise predictions about exact attribute distributions across UGC forms. To develop a classification of political UGC based on the five attributes, and to determine how attributes are distributed across UGC forms, the first research question asks:

RQ1: What is the precise pattern of the five attributes’ presence across the following political UGC types: blogs, wikis, content-sharing Web sites, participatory journalism Web sites, and discussion forums?

Finally, two pairs of attributes are expected to correlate: information retrieval and customizability on the one hand, and participation facilitation and community orientation on the other hand. The first pair are expected to correlate together because the volume and heterogeneity of information on a Web site appears to be correlated with both attributes. The greater the amount and diversity of information in a database, the more useful different information retrieval options can be expected to be. If information is scant, there appears to be little need to implement elaborate search functionality. Likewise, if content is homogenous, customizability functionality
will be unnecessary because customization will likely result in highly similar content being filtered in and filtered out, reducing the usefulness of the attribute. Because the amount and diversity of information are possibly correlated with both customizability and information retrieval attributes, the latter two attributes might be also correlated in terms of their presence across Web sites. Therefore, the third hypothesis is:

H2: The presence of task-specific information retrieval performance and information environment customizability will be positively correlated across all UGC Web sites.

The second pair (i.e., participation facilitation and community orientation) are expected to be correlated because direct participation facilitation and community orientation are attributes whose role in promoting political participation is intuitive even for nonexperts. Web site administrators willing to promote political participation among their users might extensively utilize both of these attributes simultaneously. The key factor determining presence of participation facilitation attribute appears to be the motivation of the content creator or poster. If the motivation is to mobilize other Web site users, then this attribute will be more extensively present. For example, many political blogs (e.g., Crooks and Liars, Instapundit) often circulate petitions, or post requests to donate money or contact their representatives and senators. However, not all political blogs engage in such active recruitment (e.g., Framing Science), plausibly because motivations of political bloggers are very diverse (Ekdale, Namkoong, Fung, & Perlmutter, 2010).

One might also argue that because presence of many attributes depends on the motivations or goals of the Web site administrator or users, technological attributes that are “commonly known” to achieve certain goals should also correlate. Participation facilitation is theorized to have a direct impact on participation, whereas community orientation attribute is expected to trigger the greatest number of processes leading to increases in political participation.
Therefore, Web site administrators willing to promote political participation among their users might extensively utilize both of these attributes: They might simultaneously attempt to foster a sense of community (Eaton, 2010) and offer Web site members tools for direct participation in politics (Kerbel & Bloom, 2005; Trippi, 2004), all of which is likely to create a correlation between these two attributes. Therefore, the last hypothesis is:

H3: The presence of direct participation facilitation and of community orientation will be positively correlated across all UGC Web sites.
Chapter 7: Method

Sampling

The study employed a nonprobability purposive stratified sampling design. The study examined the most popular (according to the traffic or alternative authority or popularity metrics) Web sites. There were an equal number (30) of Web sites in each of the five political UGC groups (blogs, wikis, content-sharing Web sites, participatory journalism Web sites, and discussion forums) and one reference group (traditional elite news Web sites: newspaper Web sites, TV or cable Web sites, and magazine Web sites), a total of 180 Web sites (see Table 1 for the full list of the Web sites).
<table>
<thead>
<tr>
<th>Address</th>
<th>Alexa Worldwide Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Blogs</strong></td>
<td></td>
</tr>
<tr>
<td>hotair.com</td>
<td>3694</td>
</tr>
<tr>
<td>dailykos.com</td>
<td>4803</td>
</tr>
<tr>
<td>zerohedge.com</td>
<td>5649</td>
</tr>
<tr>
<td>breitbart.tv</td>
<td>6645</td>
</tr>
<tr>
<td>Biggovernment.com</td>
<td>9088</td>
</tr>
<tr>
<td>michellemalkin.com</td>
<td>9730</td>
</tr>
<tr>
<td>redstate.com</td>
<td>11289</td>
</tr>
<tr>
<td>crooksandliars.com</td>
<td>11371</td>
</tr>
<tr>
<td>americablog.com</td>
<td>12468</td>
</tr>
<tr>
<td>powerlineblog.com</td>
<td>15748</td>
</tr>
<tr>
<td>commondreams.org</td>
<td>16473</td>
</tr>
<tr>
<td>wonkette.com</td>
<td>17621</td>
</tr>
<tr>
<td>jihadwatch.org</td>
<td>18317</td>
</tr>
<tr>
<td>bigjournalism.com</td>
<td>22050</td>
</tr>
<tr>
<td>volokh.com</td>
<td>33948</td>
</tr>
<tr>
<td>althouse.blogspot.com</td>
<td>45873</td>
</tr>
<tr>
<td>order-order.com</td>
<td>54985</td>
</tr>
<tr>
<td>digbysblog.blogspot.com</td>
<td>61359</td>
</tr>
<tr>
<td>eschatonblog.com</td>
<td>72477</td>
</tr>
<tr>
<td>balloon-juice.com</td>
<td>74789</td>
</tr>
<tr>
<td>sayanythingblog.com</td>
<td>77316</td>
</tr>
<tr>
<td>outsidethebeltway.com</td>
<td>94470</td>
</tr>
<tr>
<td>talkleft.com</td>
<td>94714</td>
</tr>
<tr>
<td>coloradopols.com</td>
<td>118007</td>
</tr>
<tr>
<td>legalinsurrection.blogspot.com</td>
<td>118063</td>
</tr>
<tr>
<td>hurryupharry.org</td>
<td>124773</td>
</tr>
<tr>
<td>newshounds.us</td>
<td>125013</td>
</tr>
<tr>
<td>mydd.com</td>
<td>131231</td>
</tr>
<tr>
<td>cafehayek.com</td>
<td>134779</td>
</tr>
<tr>
<td>nicedeb.wordpress.com</td>
<td>155417</td>
</tr>
<tr>
<td><strong>Wikis</strong></td>
<td></td>
</tr>
<tr>
<td>wikipedia.org</td>
<td>7</td>
</tr>
<tr>
<td>wikia.com</td>
<td>204</td>
</tr>
<tr>
<td>dmoz.org</td>
<td>535</td>
</tr>
</tbody>
</table>

Notes: 30 Web sites from each group included in the analysis. The Web sites are sorted from the most to least trafficked within each group.

Table 1. Web Sites in the Sample
<table>
<thead>
<tr>
<th>Address</th>
<th>Alexa Worldwide Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>wikihow.com</td>
<td>656</td>
</tr>
<tr>
<td>wiktionary.org</td>
<td>881</td>
</tr>
<tr>
<td>aboutus.org</td>
<td>1558</td>
</tr>
<tr>
<td>wikibooks.org</td>
<td>2665</td>
</tr>
<tr>
<td>wikiquote.org</td>
<td>2710</td>
</tr>
<tr>
<td>wikitravel.org</td>
<td>3315</td>
</tr>
<tr>
<td>encyclopediadramatica.com</td>
<td>3627</td>
</tr>
<tr>
<td>wikileaks.org</td>
<td>4692</td>
</tr>
<tr>
<td>wikisource.org</td>
<td>5356</td>
</tr>
<tr>
<td>openstreetmap.org</td>
<td>10092</td>
</tr>
<tr>
<td>geonames.org</td>
<td>12950</td>
</tr>
<tr>
<td>everything2.com</td>
<td>13867</td>
</tr>
<tr>
<td>strategywiki.org</td>
<td>14291</td>
</tr>
<tr>
<td>wikinews.org</td>
<td>16179</td>
</tr>
<tr>
<td>productwiki.com</td>
<td>17269</td>
</tr>
<tr>
<td>world66.com</td>
<td>19050</td>
</tr>
<tr>
<td>wikiversity.org</td>
<td>19737</td>
</tr>
<tr>
<td>wikidweb.com</td>
<td>20970</td>
</tr>
<tr>
<td>foodista.com</td>
<td>32450</td>
</tr>
<tr>
<td>mariowiki.com</td>
<td>33465</td>
</tr>
<tr>
<td>owasp.org</td>
<td>35557</td>
</tr>
<tr>
<td>sourcewatch.org</td>
<td>47660</td>
</tr>
<tr>
<td>citizendium.org</td>
<td>51022</td>
</tr>
<tr>
<td>gasbuddy.com</td>
<td>51781</td>
</tr>
<tr>
<td>conservapedia.com</td>
<td>63838</td>
</tr>
<tr>
<td>gentoo-wiki.com</td>
<td>67127</td>
</tr>
<tr>
<td>ballotpedia.org</td>
<td>76788</td>
</tr>
</tbody>
</table>

**Content-Sharing Web sites**

<table>
<thead>
<tr>
<th>Address</th>
<th>Alexa Worldwide Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>youtube.com</td>
<td>3</td>
</tr>
<tr>
<td>flickr.com</td>
<td>33</td>
</tr>
<tr>
<td>photobucket.com</td>
<td>67</td>
</tr>
<tr>
<td>4shared.com</td>
<td>72</td>
</tr>
<tr>
<td>imageshack.us</td>
<td>75</td>
</tr>
<tr>
<td>twitpic.com</td>
<td>105</td>
</tr>
<tr>
<td>dailymotion.com</td>
<td>109</td>
</tr>
<tr>
<td>deviantart.com</td>
<td>114</td>
</tr>
<tr>
<td>ehow.com</td>
<td>128</td>
</tr>
<tr>
<td>tumblr.com</td>
<td>130</td>
</tr>
<tr>
<td>vimeo.com</td>
<td>177</td>
</tr>
<tr>
<td>hubpages.com</td>
<td>205</td>
</tr>
<tr>
<td>metacafe.com</td>
<td>209</td>
</tr>
<tr>
<td>squidoo.com</td>
<td>232</td>
</tr>
<tr>
<td>scribd.com</td>
<td>253</td>
</tr>
<tr>
<td>tweetphoto.com</td>
<td>279</td>
</tr>
<tr>
<td>Address</td>
<td>Alexa Worldwide Rank</td>
</tr>
<tr>
<td>-------------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>yourfilehost.com</td>
<td>450</td>
</tr>
<tr>
<td>break.com</td>
<td>462</td>
</tr>
<tr>
<td>yfrog.com</td>
<td>475</td>
</tr>
<tr>
<td>fotolog.com</td>
<td>483</td>
</tr>
<tr>
<td>veoh.com</td>
<td>484</td>
</tr>
<tr>
<td>webshots.com</td>
<td>796</td>
</tr>
<tr>
<td>collegehumor.com</td>
<td>993</td>
</tr>
<tr>
<td>pandora.tv</td>
<td>1511</td>
</tr>
<tr>
<td>esnips.com</td>
<td>1921</td>
</tr>
<tr>
<td>phbase.com</td>
<td>2170</td>
</tr>
<tr>
<td>photo.net</td>
<td>2190</td>
</tr>
<tr>
<td>fotki.com</td>
<td>4163</td>
</tr>
<tr>
<td>flurl.com</td>
<td>5778</td>
</tr>
<tr>
<td>picturetrail.com</td>
<td>13281</td>
</tr>
<tr>
<td>associatedcontent.com</td>
<td>550</td>
</tr>
<tr>
<td>merinews.com</td>
<td>5883</td>
</tr>
<tr>
<td>nowpublic.com</td>
<td>6554</td>
</tr>
<tr>
<td>english.ohmynews.com</td>
<td>10229</td>
</tr>
<tr>
<td>globalvoicesonline.org</td>
<td>19385</td>
</tr>
<tr>
<td>groundreport.com</td>
<td>72805</td>
</tr>
<tr>
<td>kuro5hin.org</td>
<td>80747</td>
</tr>
<tr>
<td>universalhub.com</td>
<td>83359</td>
</tr>
<tr>
<td>sportingo.com</td>
<td>109407</td>
</tr>
<tr>
<td>babblestorm.co.uk</td>
<td>137215</td>
</tr>
<tr>
<td>yourhub.com</td>
<td>176566</td>
</tr>
<tr>
<td>blufftontoday.com</td>
<td>187254</td>
</tr>
<tr>
<td>newwest.net</td>
<td>214476</td>
</tr>
<tr>
<td>broowaha.com</td>
<td>229290</td>
</tr>
<tr>
<td>orato.com</td>
<td>262808</td>
</tr>
<tr>
<td>mytowncolorado.com</td>
<td>369780</td>
</tr>
<tr>
<td>phillyfuture.org</td>
<td>399048</td>
</tr>
<tr>
<td>urbanvancouver.com</td>
<td>906151</td>
</tr>
<tr>
<td>coastside.com</td>
<td>1150801</td>
</tr>
<tr>
<td>salidaditizen.com</td>
<td>1178091</td>
</tr>
<tr>
<td>citizenwausau.com</td>
<td>1215032</td>
</tr>
<tr>
<td>therapidian.org</td>
<td>1225826</td>
</tr>
<tr>
<td>ibrattleboro.com</td>
<td>1231983</td>
</tr>
<tr>
<td>citizenjoe.org</td>
<td>1334958</td>
</tr>
<tr>
<td>freshare.net</td>
<td>1795632</td>
</tr>
<tr>
<td>mymissourian.com</td>
<td>4220888</td>
</tr>
<tr>
<td>inewss.com</td>
<td>6478957</td>
</tr>
<tr>
<td>winnipegfirst.ca</td>
<td>6937010</td>
</tr>
<tr>
<td>forumhome.org</td>
<td>13880181</td>
</tr>
<tr>
<td>sierrabear.com</td>
<td>n/a</td>
</tr>
</tbody>
</table>

Participatory journalism Web sites

<table>
<thead>
<tr>
<th>Address</th>
<th>Alexa Worldwide Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>associatedcontent.com</td>
<td>550</td>
</tr>
<tr>
<td>merinews.com</td>
<td>5883</td>
</tr>
<tr>
<td>nowpublic.com</td>
<td>6554</td>
</tr>
<tr>
<td>english.ohmynews.com</td>
<td>10229</td>
</tr>
<tr>
<td>globalvoicesonline.org</td>
<td>19385</td>
</tr>
<tr>
<td>groundreport.com</td>
<td>72805</td>
</tr>
<tr>
<td>kuro5hin.org</td>
<td>80747</td>
</tr>
<tr>
<td>universalhub.com</td>
<td>83359</td>
</tr>
<tr>
<td>sportingo.com</td>
<td>109407</td>
</tr>
<tr>
<td>babblestorm.co.uk</td>
<td>137215</td>
</tr>
<tr>
<td>yourhub.com</td>
<td>176566</td>
</tr>
<tr>
<td>blufftontoday.com</td>
<td>187254</td>
</tr>
<tr>
<td>newwest.net</td>
<td>214476</td>
</tr>
<tr>
<td>broowaha.com</td>
<td>229290</td>
</tr>
<tr>
<td>orato.com</td>
<td>262808</td>
</tr>
<tr>
<td>mytowncolorado.com</td>
<td>369780</td>
</tr>
<tr>
<td>phillyfuture.org</td>
<td>399048</td>
</tr>
<tr>
<td>urbanvancouver.com</td>
<td>906151</td>
</tr>
<tr>
<td>coastside.com</td>
<td>1150801</td>
</tr>
<tr>
<td>salidaditizen.com</td>
<td>1178091</td>
</tr>
<tr>
<td>citizenwausau.com</td>
<td>1215032</td>
</tr>
<tr>
<td>therapidian.org</td>
<td>1225826</td>
</tr>
<tr>
<td>ibrattleboro.com</td>
<td>1231983</td>
</tr>
<tr>
<td>citizenjoe.org</td>
<td>1334958</td>
</tr>
<tr>
<td>freshare.net</td>
<td>1795632</td>
</tr>
<tr>
<td>mymissourian.com</td>
<td>4220888</td>
</tr>
<tr>
<td>inewss.com</td>
<td>6478957</td>
</tr>
<tr>
<td>winnipegfirst.ca</td>
<td>6937010</td>
</tr>
<tr>
<td>forumhome.org</td>
<td>13880181</td>
</tr>
<tr>
<td>sierrabear.com</td>
<td>n/a</td>
</tr>
</tbody>
</table>
Table 1 continued

<table>
<thead>
<tr>
<th>Address</th>
<th>Alexa Worldwide Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Discussion Forums</strong></td>
<td></td>
</tr>
<tr>
<td>imdb.com/boards</td>
<td>35</td>
</tr>
<tr>
<td>forums.digitalpoint.com</td>
<td>104</td>
</tr>
<tr>
<td>forums.cnet.com</td>
<td>108</td>
</tr>
<tr>
<td>gamespot.com/forums/index.html</td>
<td>163</td>
</tr>
<tr>
<td>boards.ign.com</td>
<td>176</td>
</tr>
<tr>
<td>vnboards.ign.com</td>
<td>176</td>
</tr>
<tr>
<td>if.invisionfree.com/index</td>
<td>176</td>
</tr>
<tr>
<td>sitepoint.com/forums</td>
<td>252</td>
</tr>
<tr>
<td>forums.filefront.com</td>
<td>376</td>
</tr>
<tr>
<td>forums.adobe.com</td>
<td>386</td>
</tr>
<tr>
<td>forums.sun.com</td>
<td>457</td>
</tr>
<tr>
<td>tambayan.1.forumer.com</td>
<td>538</td>
</tr>
<tr>
<td>dpreview.com/forums</td>
<td>556</td>
</tr>
<tr>
<td>forum.wordreference.com</td>
<td>572</td>
</tr>
<tr>
<td>fatwallet.com/forums</td>
<td>583</td>
</tr>
<tr>
<td>forums.winamp.com</td>
<td>618</td>
</tr>
<tr>
<td>slickdeals.net/forums</td>
<td>632</td>
</tr>
<tr>
<td>newgrounds.com/bbs</td>
<td>688</td>
</tr>
<tr>
<td>forums.xbox.com</td>
<td>723</td>
</tr>
<tr>
<td>masalaboard.com</td>
<td>744</td>
</tr>
<tr>
<td>rotten tomatoes.com/vine/index.php</td>
<td>758</td>
</tr>
<tr>
<td>my.opera.com/community/forums/home.dml</td>
<td>799</td>
</tr>
<tr>
<td>forum.ebaumsworld.com</td>
<td>808</td>
</tr>
<tr>
<td>boards.harrypotter.warnerbros.com</td>
<td>811</td>
</tr>
<tr>
<td>ultimate-guitar.com/forum</td>
<td>870</td>
</tr>
<tr>
<td>archive.gamespy.com/forums</td>
<td>941</td>
</tr>
<tr>
<td>webhostingtalk.com</td>
<td>952</td>
</tr>
<tr>
<td>forums.oscommerce.com</td>
<td>973</td>
</tr>
<tr>
<td>phpbb.com/community</td>
<td>1040</td>
</tr>
<tr>
<td>talktalk.co.uk/forums</td>
<td>1060</td>
</tr>
<tr>
<td><strong>Reference Group</strong></td>
<td></td>
</tr>
<tr>
<td>bbc.co.uk</td>
<td>42</td>
</tr>
<tr>
<td>cnn.com</td>
<td>58</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>91</td>
</tr>
<tr>
<td>timesofindia.indiatimes.com</td>
<td>173</td>
</tr>
<tr>
<td>guardian.co.uk</td>
<td>209</td>
</tr>
<tr>
<td>foxnews.com</td>
<td>213</td>
</tr>
<tr>
<td>washingtonpost.com</td>
<td>319</td>
</tr>
<tr>
<td>latimes.com</td>
<td>372</td>
</tr>
<tr>
<td>usatoday.com</td>
<td>407</td>
</tr>
<tr>
<td>bloomberg.com</td>
<td>454</td>
</tr>
<tr>
<td>abcnews.go.com</td>
<td>481</td>
</tr>
</tbody>
</table>
In choosing the number of Web sites for each UGC type and for reference group two considerations were important: (1) arriving at a number that represented the actual population of interest to the study and (2) arriving at a number that was manageable in terms of effort and time needed to conduct the study.

First, this study defined its population of interest as the most popular UGC Web sites in each of the five UGC groups and the most popular traditional journalistic organizations’ Web sites. In this study, the goal was to examine the Web sites that can potentially affect the greatest number of users. Therefore, the number 30 was considered a good approximation for what constitutes the theoretical population of interest to this study. Second, (a) because this study examined 180 Web sites, (b) because this study attempted to document the presence of five technological attributes (with a total of 35 technological features), and (c) because this study

<table>
<thead>
<tr>
<th>Address</th>
<th>Alexa Worldwide Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>examiner.com</td>
<td>482</td>
</tr>
<tr>
<td>forbes.com</td>
<td>546</td>
</tr>
<tr>
<td>businessweek.com</td>
<td>556</td>
</tr>
<tr>
<td>time.com</td>
<td>572</td>
</tr>
<tr>
<td>sfgate.com</td>
<td>661</td>
</tr>
<tr>
<td>cbsnews.com</td>
<td>719</td>
</tr>
<tr>
<td>smh.com.au</td>
<td>721</td>
</tr>
<tr>
<td>cnbc.com</td>
<td>754</td>
</tr>
<tr>
<td>nypost.com</td>
<td>758</td>
</tr>
<tr>
<td>chicagotribune.com</td>
<td>1247</td>
</tr>
<tr>
<td>theage.com.au</td>
<td>1448</td>
</tr>
<tr>
<td>theglobeandmail.com</td>
<td>1612</td>
</tr>
<tr>
<td>hindustantimes.com</td>
<td>1721</td>
</tr>
<tr>
<td>newsweek.com</td>
<td>1757</td>
</tr>
<tr>
<td>suntimes.com</td>
<td>1767</td>
</tr>
<tr>
<td>csmonitor.com</td>
<td>1882</td>
</tr>
<tr>
<td>economist.com</td>
<td>2058</td>
</tr>
<tr>
<td>ajc.com</td>
<td>2077</td>
</tr>
<tr>
<td>azcentral.com</td>
<td>2106</td>
</tr>
</tbody>
</table>
utilized a thorough examination of a large portion of each Web site (where each Web site was a rather sophisticated and vast information system), it was important to arrive at a manageable number of the Web sites for inclusion into the content analysis sample. Due to this consideration, the number 30, again, appeared to be appropriate.

A common problem in content analyses of online content is the great difficulty of obtaining an exhaustive list of Web sites that constitute one’s population of interest – a necessity for selecting a random sample of units for analysis (McMillan, 2000). To quote Riffe, Lacy and Fico (1998): “… the Internet is like a city without a telephone book or map… New houses are being built all the time, and old houses are being deserted with no listing of the changes” (p. 118). This situation generally prevents researchers from obtaining an acceptable “probability” sample of online content. To use probability sampling in this study, an exhaustive list of all Web sites in the five political UGC categories, and in the reference group, would be needed. Such exhaustive, or even quasi-exhaustive lists, do not appear to exist.

In the present study, there were five strata of UGC Web sites and one strata of reference group Web sites. Equal strata sizes were used in part to optimize the robustness of the cluster analytic procedure utilized in this study (Romesburg, 2004). The stratified sampling technique also allows for the comparison of differences among the strata, as well as allows one to combine results for all strata to provide a cumulative description of the entire sample (Frankel, 1983).

To identify the list of the most popular Web sites in each political UGC strata, an effort was made to find the most reliable and complete directories of these Web sites that provided popularity or traffic information for the Web sites. Identification of such directories took place in Summer and Fall 2010. Ideally, three different directories for each stratum would be identified. Then, the thirty most popular Web sites from each directory would be identified and then merged into a single list. Finally, the thirty most popular Web sites (using Alexa rankings) from the
merged list would be used for analyses. However, the number of available and, more importantly, authoritative, directories for each political UGC strata varied from one to three. Web sites that had not been updated within past year were excluded.

**Blogs.** It was necessary to remove a number of blogs because they did not fit the definition of political UGC presented above. Some of the blogs that were removed were: the professionally-produced blogs (e.g., *Talking Points Memo*); blogs that presented themselves as professional organizations (with paid staff, formal organizational structure - such as *The Huffington Post*); blogs that were formally produced or affiliated with think tanks, news organizations, and other elite organizations (e.g., blogs hosted by *The Washington Post, Times* magazine. Blogs written by individuals with professional journalism background were included only if the individuals or blogs were not explicitly and formally affiliated with any of the organizations mentioned above. Blog descriptions in directories and blog “About Us” pages on the actual blog were examined at the time of the blog list construction. Blogs not focused on politics (which did not mention “politics” or “public affairs” or “political news” as one of their core focus areas) were excluded. Blogs whose address was moved or changed or which were inaccessible were excluded. The following three directories were used: (1) Technorati; (2) Truthlaidbear; and (3) Google directory > Society > Politics > News and Media > Weblogs > 5 subdirectories:

- Top 6 from “Progressive and Left” subdirectory
- Top 6 from “Conservatism” subdirectory
- Top 6 from “Libertarianism” subdirectory
- Top 6 from “United Kingdom” subdirectory
- Top 6 from “United States” subdirectory
Wikis. These three directories were used: (1) wiki1001.com/wikidirectory; (2) wikindex.com; and (3) meta.wikimedia.org/wiki/List_of_largest_wikis. Those wikis whose address was moved or changed or which were inaccessible were excluded (this rule also applies to all Web sites in the sample).

Content-sharing Web sites. These were primarily Web sites allowing users to upload their content to a centralized database of a single Web site, and where this content could later be searched and used by other Web site users. Peer-to-peer sites and software, bit-torrent sites, and social link-sharing sites were excluded in order to simplify the sampling procedure. A search for a “content-sharing site” list or directory showed that there were no up-to-date authoritative lists. Therefore, Alexa’s list of the world’s top 500 Web sites was used to find the most popular content-sharing Web sites. Despite its authoritativeness, the list is not organized by Web site type (i.e., all 500 Web sites were presented, not just content-sharing sites). The author went through the entire list and manually identified the content-sharing Web sites. Web site names and short descriptions were consulted to determine if they were in fact content-sharing sites. If the name or description were not informative enough for a particular Web site, that Web site was examined by clicking on its URL. This procedure generated a list of 21 most popular content-sharing sites. To identify the remaining nine Web sites, a directory of popular image-sharing Web sites (www.en.wikipedia.org/wiki/List_of_photo_sharing_Web_sites) and a directory of popular video-sharing Web sites (www.reelseo.com/list-video-sharing-Web_sites/) were utilized. A list of the top 10 Web sites from each directory (a total of 20 for both directories) was created and sorted by Alexa’s worldwide ranking. The nine most popular Web sites (which were not already on the original list) were added to the original list of 21 Web sites. Thus, these three directories were used: (1) Alexa – top 21 Web sites; (2) Wikipeda and (3) Reelseo – top nine Web sites.
**Participatory journalism Web sites.** These Web sites were characterized by a substantial contribution of the users (non-professional journalists) to content creation. More specifically, to be included in the list, each Web site had to be characterized by a high degree of collaboration between professional journalists and amateur readers, such as either: (a) an amateur writes the story, then a journalist edits it and posts it online, or (b) an amateur writes the story, then a journalist reviews it and decides if it would be posted online. It was not sufficient for a Web site to indicate that any posts or articles can be removed from the Web site if they contained certain content (e.g., pornography, advertising). An examination of the Web site’s “About Us” page, “FAQ” page, and various pieces of content (i.e., news articles) was conducted to determine if it met the above criteria. Such strict criteria coupled with a lack of up-to-date comprehensive and authoritative directories of prominent participatory journalism Web sites ordered by popularity, created challenges in identifying the needed number of Web sites. Therefore, the following procedure was used. Several directories of participatory journalism Web sites were identified, including: (1) www.google.com/Top/News/Media/Participatory; (2) http://www.cyberjournalist.net/news/002226.php; (3) http://www.sourcewatch.org/index.php?title=List_of_citizen_journalism_Web sites. Of the three, only Google directory provided Web site traffic information. Therefore, Alexa’s world ranking was used to order the Web sites in each of the three directories from most to least trafficked. Top 30 Web sites that met the inclusion criteria were included in the sample.

**Discussion forums.** Top English-language forums from bigboards.com were used, after being sorted by “Alexa Ranking.” Alexa’s traffic score indicated by bigboards was recorded because discussion boards were often incorporated into other Web sites and their URL addresses often consisted of several orders of domain names (several words separated by the dot symbol), which made it impossible to check the traffic score on Alexa.com.
Reference group. Traditional elite news Web sites included: newspaper Web sites, TV/cable Web sites, and magazine Web sites. The top Web sites belonging to the traditional news category were selected, without ensuring a proportional representation of Web sites belonging to the three traditional news subcategories. Alexa’s list of the most popular news Web sites (http://www.alexaw.com/topsites/category/Top/News) was used.

Variables and Codebook

Each attribute was represented by an additive index, constructed as follows: Each feature representing an attribute was given two points if it was present on the Web site’s homepage, one point if it was present elsewhere on the Web site, and zero if it was absent from the Web site. All feature scores were added into a cumulative score for the attribute. The codebook included in Appendix A provides detailed information about how the features representing the five attributes were coded.

Improving Comparability across Web Sites

The author manually saved relevant sections of each Web site before beginning to assess intercoder reliability. Procedures for identifying and saving content are detailed below.

General procedures for identifying and saving content. First, the author created an account for each Web site to sign in and to have access to all available technological features. Second, when sections of a Web site had a URL unrelated to the main Web site, such sections were excluded. For instance, DailyKos had a main section on its homepage, called “DKosopedia,” which was located at: dkosopedia.com/wiki/Main_Page and several Web sites had a “photo” section hosted on Flickr. Third, “international,” “national,” or other types of “editions” of the Web site were not examined separately. Only if the homepage of the Web site was in language other than English was an alternative appropriate other section consulted. Fourth, if, in order to register, some Web site required the submission of an application to be evaluated at
a later time, then registration was avoided for practical and theoretical reasons and the Web site was examined without researcher being its member. Fifth, all Web site sections were saved using the FireFox “Print Screen” option. FireFox (version 3.6.9) extension ScreenGrab 0.96.3 was utilized. This extension allowed saving the entire length of a page as a single file, which facilitated a more accurate reproduction of the actual pages. Sixth, when saving these sections, FireFox browser extension, NoSquint 2.0.3 was used to display all links in red color and underlined, without altering how other Web site elements were displayed. This was necessary to enable identification of the hyperlinks for the participation facilitation attribute. In the absence of these measures, the ability to identify all hyperlinks would be lost because the link functionality is disabled by the PrintScreen procedure. Additionally, FireFox extension Adblock Plus 1.2.2 was used to remove ads from the Web sites to simplify the page layout and reduce coder fatigue. A thorough examination of this extension’s reviews (located at: https://addons.mozilla.org/en-US/firefox/addon/1865/reviews/) and lengthy pretesting of the extension showed no evidence that any non-advertising content was removed.

Specific procedures for identifying/saving content. The following sections on each Web site were identified and saved:

1. Registration page (saved only when this page provided any information about the Web site, such as “terms of service,” “about us,” etc.)

2. Homepage

3. Front page of each technologically distinct section, including:

   A. Blogs
   B. Discussion forum
   C. Wiki
   D. Videos or Photos (excluding podcasts)
4. For homepage, section front pages, and profile pages only: “A” below must be persistent (i.e., such customizations had to remain even after the browser window was closed and the user navigated away from the page to other Web sites).

A. When elements on the homepage could be moved around the page (e.g., user could move “national news” section from the top to the bottom) screen was captured when the elements were being moved. Or, when a page element could be hidden or shrunk, the page was captured while making it apparent that such elements could be hidden or shrunk (e.g., by putting a cursor on the “hide button”, to make the “hide” message pop up). This step was necessary to enable coders to identify different customizability features.

5. Search options before conducting the search (on each technologically distinct section, if search options differed).

6. Search results sorting options after conducting the search (on each technologically distinct section, if sorting options differed). If the user was taken to the Google homepage (google.com), the functionality of Google was not analyzed or included in the sample.

7. Web site section or search option for groups or organizations present on the Web site (e.g., “community” section) was saved. When groups were present, profile pages of the two most popular, or in the absence of popularity indicators, randomly selected groups were saved. Random numbers were generated by using “Random Number Generator” Android application (version 1.01) developed by Alex Sierro. The application was configured to generate two random numbers from 1 to 10. If the number “4” was given, the fourth piece of content, counting from the top of the page, was included into the sample.

8. “Browsing” sections or sections where the Web site’s content was grouped by topic, authors, date, or some other criteria.
9. After logging into the user profile area, the following sections were saved:
   A. Main user profile page
   B. Subscriptions settings page
   C. User profile section indicating: (1) Join date; (2) groups he or she is affiliated with on the Web site; (3) friends or connections of the user

10. FAQs section. Only questions on:
   A. Customization, user behavior tracking, search
   B. How users could contribute content

11. Privacy policy section
   A. Main Web site section
   B. User profile section

12. Section detailing user’s (potential) involvement with the site, including:
   A. About Us page
   B. “Become a contributor” or “for contributors” or other similar page

13. Content identification on blogs v. other sites. Sites belonging to different political UGC groups and to the reference group were very different from one another. In particular, these Web sites fell into two categories: (1) Web sites that were primarily political in overall orientation (i.e., blogs), and (2) Web sites that were not explicitly political, but which had sizable political content sections (all the remaining categories of Web sites). This difference created a need for two approaches to identifying political content in those Web sites. For the blogs, two blog posts were saved. For the second group of Web sites, political content on the Web sites was identified with a keyword search. Keyword “Obama” was utilized because: (1) it was expected to help retrieve content from across ideological spectrum; (2) because President Obama, by virtue of his office, was the most
prominent politician in the country at the time of the study; and (3) because there was some evidence that the plurality of the popular political UGC content (as represented by YouTube most viewed political news videos during the 2008 presidential election campaign) was focused on Obama (see Dylko et al., 2009). To insure comparability across all types of Web sites, the same sections were saved on all Web sites:

A. On blogs, the two most popular or, in the absence of popularity sorting option, randomly chosen pieces of content on each of the Web site’s technologically distinct sections were saved. A piece of content is a blog post, an article, a discussion thread, and so forth. Whenever available, first comments (located on the first page of comments section) left on the piece of content were saved. The actual page from which the piece of content was meant to be read or viewed was captured.

B. On all other sites, when identification of political content was performed through a keyword search, the two most popular results were saved. If results could not be sorted by popularity, the two pieces of content were chosen randomly. Whenever available, first comments (located on the first page of comments section) left on the piece of content were saved. In cases where multimedia (e.g., videos and photos) had to be saved, screenshots of page from which each piece of content should be viewed was taken. These pieces of content were not downloaded, because downloading strips the content of hyperlinks and other embedded content. These pieces of content were viewed and whenever any “clickable” links appeared, a screenshot of the page was taken.
C. If a Web site had several technologically distinct sections, and if search options and content on each section differed, two pieces of content from each section were saved.

14. For the saved pieces of content, usernames of a maximum of two content posters (or commenters, wiki editors, etc.) were identified. This researcher clicked on the usernames, and the resulting page with user information was captured for a maximum of four users per Web site. When no users associated with the two pieces of content could be identified, at least two users from elsewhere on the Web site were identified randomly, whenever possible.

A. Reputational information (e.g., stars by the username) was saved by putting a cursor on the visual representation of the user reputation in order to activate a pop-up message describing the user’s reputation.

Coders and Reliability Assessment

The coding of the Web sites proceeded in several steps. First, coders were recruited from a pool of OSU undergraduate students to work as undergraduate research assistants as part of an independent study class. Second, the recruited coders were trained to use the codebook, which entailed group discussion of the codebook, group application of the coding instructions, and independent application of the coding instruction. The training took about 35 hours in the beginning of Fall 2010 during which several recruited students left, showed poor progress during training, utilized excessively deep (and therefore inappropriate) level of analysis while coding content, or were deemed insufficiently proficient in the use of communication technology to be used as reliability assessment coders. Several students, therefore, were assigned work on the project that did not involve coding content or intercoder reliability assessment. The training process led to minor-to-moderate modifications to the codebook (the final version is available in
the Appendix). A total of two coders, in addition to this researcher, were utilized. These two coders have successfully completed training, proved to be very dependable, and demonstrated a very high degree of familiarity with the UGC Web sites (or developed such familiarity during training). The coders earned a course credit and a moderate financial compensation for their work. The first coder, Daniel Case, was specifically trained to code customizability and participation facilitation attributes. The second coder, Ujala Rizwan Abbasi, was trained to code all five attributes as a precaution against unexpected circumstances. However, only the work on information retrieval, manipulability, and community orientation produced by Ujala was used. This researcher coded all five attributes during reliability assessment. For each attribute, reliability was calculated between two coders only (one of which was always the author and the other either Daniel [for customizability and participation facilitation] or Ujala [for information retrieval, manipulability, and community orientation]) who coded 10% of the actual content (Riffe, et al., 1998).

The intercoder reliability for each attribute index was assessed by Krippendorff's alpha. Krippendorff's alpha coefficients for the attributes were: 0.80 for information retrieval; 0.85 for customizability; 0.88 for community orientation; 0.81 for participation facilitation; and 0.69 for content manipulability. Reliabilities were deemed satisfactory, due to coefficients greater than 0.80 being generally recommended for manifest variables (see Riffe, Lacy, & Fico, 1998), whereas coefficients greater than 0.667 being acceptable for “drawing tentative conclusions” about manifest variables (see Riffe, et al., 1998, p. 151). After concluding the reliability assessment, the remaining content was coded by the researcher and two coders. The cases in which disagreement between coders was present in the 10% of the content used for reliability assessment, a coin toss determined the coder whose interpretation was adopted in each individual case of disagreement.
Chapter 8: Results

Before describing results obtained to test the hypotheses and answer the research question, several descriptive characteristics of the data need to be addressed. Tables 2A and 2B show that the attribute scores were relatively low for both UGC and non-UGC Web sites. The starkest contrast is for participation facilitation attribute, for which the theoretically possible maximum scores was 14, whereas the mean was 0.13 ($SD = 0.49$) for UGC and 0.00 ($SD = 0.00$) for the reference group (Table 2B).
<table>
<thead>
<tr>
<th></th>
<th>Information Retrieval</th>
<th>Customizability</th>
<th>Manipulability</th>
<th>Participation Facilitation</th>
<th>Community Orientation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>78.57</td>
<td>100.00</td>
<td>100.00</td>
<td>62.50</td>
<td>100.00</td>
</tr>
<tr>
<td></td>
<td>36.90</td>
<td>65.23</td>
<td>19.83</td>
<td>23.33</td>
<td>33.04</td>
</tr>
<tr>
<td></td>
<td>35.71</td>
<td>64.28</td>
<td>12.50</td>
<td>25.00</td>
<td>28.57</td>
</tr>
<tr>
<td></td>
<td>28.57</td>
<td>64.29</td>
<td>12.50</td>
<td>25.00</td>
<td>28.57</td>
</tr>
<tr>
<td></td>
<td>15.06</td>
<td>15.20</td>
<td>18.09</td>
<td>14.94</td>
<td>16.64</td>
</tr>
<tr>
<td></td>
<td>1.23</td>
<td>2.77</td>
<td>1.47</td>
<td>2.72</td>
<td>1.35</td>
</tr>
</tbody>
</table>

Notes:
Comparison is made between five types of UGC Web sites whose scores were combined versus reference group made up of traditional media Web sites. The scores were rescaled to simplify interpretation of results. The following rescaling procedure was used: Each score within a given attribute was divided by the largest score on that attribute (among all Web sites) and then multiplied by 100. The resulting rescaled scores range from zero (if the lowest score on that specific attribute was zero) to 100 (the highest actual score on that specific attribute).

A minimum scores were not reported because they were “0” for vast majority of attributes except search efficiency (for which it was “7.14” for UGC and “35.71” for the reference group).

Table 2. Descriptive Characteristics for Attribute Scores on UGC versus Reference Group Web Sites.
Table 2 continued

<table>
<thead>
<tr>
<th></th>
<th>Information Retrieval</th>
<th>Customizability</th>
<th>Manipulability</th>
<th>Participation Facilitation</th>
<th>Community Orientation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum theoretically possible score&lt;sup&gt;a&lt;/sup&gt;</td>
<td>22</td>
<td>12</td>
<td>10</td>
<td>14</td>
<td>12</td>
</tr>
<tr>
<td>Maximum score&lt;sup&gt;b&lt;/sup&gt;</td>
<td>UGC</td>
<td>Ref. Gr.</td>
<td>UGC</td>
<td>Ref. Gr.</td>
<td>UGC</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>14</td>
<td>8</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>Mean median mode</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.17,</td>
<td>9.13,</td>
<td>1.59,</td>
<td>1.87,</td>
<td>2.31,</td>
<td>2.03,</td>
</tr>
<tr>
<td>5.00,</td>
<td>9.00,</td>
<td>1.00,</td>
<td>2.00,</td>
<td>2.00,</td>
<td>2.00,</td>
</tr>
<tr>
<td>4.00</td>
<td>9.00</td>
<td>1.00</td>
<td>2.00</td>
<td>2.00</td>
<td>2.00</td>
</tr>
<tr>
<td>SD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.11</td>
<td>2.13</td>
<td>1.45</td>
<td>1.20</td>
<td>1.17</td>
<td>1.10</td>
</tr>
<tr>
<td>Standard error of the mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.17</td>
<td>0.38</td>
<td>0.11</td>
<td>0.21</td>
<td>0.09</td>
<td>0.20</td>
</tr>
</tbody>
</table>

Notes:

a. Total number of features multiplied by 2 (the maximum possible value a feature could receive if it was located on the homepage of a Web site).

b. Minimum scores were not reported because they were “0” for vast majority of attributes except search efficiency (for which it was “1” for UGC and “5” for the reference group).
An examination of histograms for each attribute, pictured in Figure 2A, 2B, and 2C, also suggests that the data generally has a positive skew, with the bulk of Web sites having low scores on the attributes (especially for the participation facilitation). When UGC and reference group sites were analyzed as a single group, Kolmogorov-Smirnov and Shapiro-Wilk tests showed the distributions of all attributes as significantly non-normal. Due to consistently similar results, and due to K-S test being more commonly utilized, only K-S test results are reported next. For information retrieval attribute, \( D(180) = 0.15, p < .001 \); for customizability, \( D(180) = 0.20, p < .001 \); for manipulability, \( D(180) = 0.19, p < .001 \); for participation facilitation, \( D(180) = 0.52, p < .001 \); for community orientation, \( D(180) = 0.27, p < .001 \). Substantively identical results were obtained when analyses were conducted on UGC sites and reference group sites, separately: \( D \) statistic was significant at the \( p < .01 \) level for all attributes.
Notes: The Y-axis represents percent of Web sites within the cell. The X-axis represents the rescaled score of Web sites. Each score within a given attribute was divided by the largest score on that attribute (among all Web sites) and then multiplied by 100. The resulting rescaled scores range from zero (if the lowest score on that specific attribute was zero) to 100 (the highest actual score on that specific attribute).

Figure 2: Frequency Distributions of Attribute Scores.
Figure 2 continued

<table>
<thead>
<tr>
<th>All UGC Web Sites</th>
<th>Reference Group Web Sites</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Information Retrieval</strong></td>
<td><strong>Information Retrieval</strong></td>
</tr>
<tr>
<td><img src="image1" alt="Graph" /></td>
<td><img src="image2" alt="Graph" /></td>
</tr>
<tr>
<td><strong>Customizability</strong></td>
<td><strong>Customizability</strong></td>
</tr>
<tr>
<td><img src="image3" alt="Graph" /></td>
<td><img src="image4" alt="Graph" /></td>
</tr>
<tr>
<td><strong>Manipulability</strong></td>
<td><strong>Manipulability</strong></td>
</tr>
<tr>
<td><img src="image5" alt="Graph" /></td>
<td><img src="image6" alt="Graph" /></td>
</tr>
<tr>
<td><strong>Participation Facilitation</strong></td>
<td><strong>Participation Facilitation</strong></td>
</tr>
<tr>
<td><img src="image7" alt="Graph" /></td>
<td><img src="image8" alt="Graph" /></td>
</tr>
<tr>
<td><strong>Community Orientation</strong></td>
<td><strong>Community Orientation</strong></td>
</tr>
<tr>
<td><img src="image9" alt="Graph" /></td>
<td><img src="image10" alt="Graph" /></td>
</tr>
</tbody>
</table>

All scores = zero
<table>
<thead>
<tr>
<th>Blogs</th>
<th>Wikis</th>
<th>Content-sharing sites</th>
<th>Participatory journalism sites</th>
<th>Discussion forums</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Inform. Retrieval</strong></td>
<td><strong>Inform. Retrieval</strong></td>
<td><strong>Inform. Retrieval</strong></td>
<td><strong>Inform. Retrieval</strong></td>
<td><strong>Inform. Retrieval</strong></td>
</tr>
<tr>
<td><img src="image1" alt="Histogram" /></td>
<td><img src="image2" alt="Histogram" /></td>
<td><img src="image3" alt="Histogram" /></td>
<td><img src="image4" alt="Histogram" /></td>
<td><img src="image5" alt="Histogram" /></td>
</tr>
<tr>
<td><strong>Customizability</strong></td>
<td><strong>Customizability</strong></td>
<td><strong>Customizability</strong></td>
<td><strong>Customizability</strong></td>
<td><strong>Customizability</strong></td>
</tr>
<tr>
<td><img src="image6" alt="Histogram" /></td>
<td><img src="image7" alt="Histogram" /></td>
<td><img src="image8" alt="Histogram" /></td>
<td><img src="image9" alt="Histogram" /></td>
<td><img src="image10" alt="Histogram" /></td>
</tr>
<tr>
<td><strong>Manipulability</strong></td>
<td><strong>Manipulability</strong></td>
<td><strong>Manipulability</strong></td>
<td><strong>Manipulability</strong></td>
<td><strong>Manipulability</strong></td>
</tr>
<tr>
<td><img src="image11" alt="Histogram" /></td>
<td><img src="image12" alt="Histogram" /></td>
<td><img src="image13" alt="Histogram" /></td>
<td><img src="image14" alt="Histogram" /></td>
<td><img src="image15" alt="Histogram" /></td>
</tr>
<tr>
<td><strong>Partic. Facilit.</strong></td>
<td><strong>Partic. Facilit.</strong></td>
<td><strong>Partic. Facilit.</strong></td>
<td><strong>Partic. Facilit.</strong></td>
<td><strong>All scores = zero</strong></td>
</tr>
<tr>
<td><img src="image16" alt="Histogram" /></td>
<td><img src="image17" alt="Histogram" /></td>
<td><img src="image18" alt="Histogram" /></td>
<td><img src="image19" alt="Histogram" /></td>
<td></td>
</tr>
<tr>
<td><strong>Comm. Orient.</strong></td>
<td><strong>Comm. Orient.</strong></td>
<td><strong>Comm. Orient.</strong></td>
<td><strong>Comm. Orient.</strong></td>
<td><strong>Comm. Orient.</strong></td>
</tr>
<tr>
<td><img src="image20" alt="Histogram" /></td>
<td><img src="image21" alt="Histogram" /></td>
<td><img src="image22" alt="Histogram" /></td>
<td><img src="image23" alt="Histogram" /></td>
<td><img src="image24" alt="Histogram" /></td>
</tr>
</tbody>
</table>

C
Due to such consistent and statistically significant deviation from normal distribution, an attempt was made to transform the data by using (1) square root, logarithmic, and negative reciprocal transformations. All of the above transformation failed to normalize the data. Regardless of the form of transformation, the $D$ statistic remained significant at the $p < .001$ level for all attributes when all sites were treated as a single group, and with $D$ statistic remained significant at the $p < .05$ level for all attributes when UGC and reference group sites were analyzed separately.

In another attempt to normalize distributions, outliers (cases outside three standard deviations within each attribute) were replaced with scores equal to attribute mean plus two standard deviations (Field, 2005). Once again, the $D$ statistic remained significant at the $p < .001$ level for all attributes when all sites were treated as a single group, and $D$ at $p < .05$ level for all attributes when UGC and reference group sites were analyzed separately.

Finally, three of the above transformations were performed on data after outliers were replaced. Once again, deviation from normality could not be corrected.

As a result of the statistically significant deviation from normal distribution, non-parametric data analytic methods were used while retaining the raw measures without transformation or alteration of outliers. Non-parametric methods generally rely much less heavily on any data-distribution assumptions (Wasserman, 2006), and have a relatively high degree of statistical power when used on non-normally distributed data (Field, 2005; Sheskin, 2004). Results of all analyses, with the exception of the cluster analysis, are based on data that were not transformed in any way.

Hypothesis 1. The first hypothesis proposed that: “The five attributes should have a greater presence (number of technological features) in political UGC (most popular blogs, wikis, content-sharing Web sites, participatory journalism Web sites, and discussion forums) than in
traditional news Web sites (most popular newspaper Web sites, TV or cable Web sites, and magazine Web sites).” To test this hypothesis, the non-parametric counterpart to the one-way independent samples ANOVA – the Kruskal-Wallis test (Sheskin, 2004; Siegel, 1956) – was utilized to conduct an omnibus test of difference between the groups. This test is recommended for data that are not normally distributed (Field, 2005; Sheskin, 2004). If the Kruskal-Wallis test was significant for a particular attribute (indicating that there was a statistically significant difference between medians in at least two groups on the attribute score), a post-hoc probe was conducted. A non-parametric counterpart to the independent samples T-test, the Mann-Whitney test (Sheskin, 2004; Siegel, 1956) combined with a Bonferroni correction was used to determine which specific UGC groups differed from the reference group to a statistically significant degree. The Bonferroni correction method adjusts the critical value for significance for each test to avoid inflation of the Type I error rate (Field, 2005). Therefore, the \( p \) value in post-hoc analyses reported below was adjusted from 0.05 to 0.01 (0.05 divided by the number of post-hoc tests \([n=5]\) for each attribute). Despite being considered a rather conservative procedure, in the current analyses the Bonferroni correction resulted in only one of 20 pairs (reference v. each UGC form) of Web site groups being regarded as statistically indistinguishable, whereas the difference between the two groups was significant at the \( p < 0.05 \). Table 3 compares each group of UGC Web sites to the reference group on all attributes. Figure 3 illustrates the results visually. A detailed description of data presented in Table 3 is offered next.
### Table 3. Median Scores of Each Web Site Group on Attributes

<table>
<thead>
<tr>
<th>Comparison (reference) group – traditional news sites</th>
<th>Information Retrieval</th>
<th>Customizab.</th>
<th>Manipulability</th>
<th>Community Orientation</th>
</tr>
</thead>
<tbody>
<tr>
<td>UGC – Blogs</td>
<td>4.00***</td>
<td>1.00***</td>
<td>2.00</td>
<td>0.00</td>
</tr>
<tr>
<td>UGC – Wikis</td>
<td>4.00***</td>
<td>1.00</td>
<td>2.00</td>
<td>1.00**</td>
</tr>
<tr>
<td>UGC – Content-sharing sites</td>
<td>5.50***</td>
<td>2.00</td>
<td>3.00</td>
<td>2.00***</td>
</tr>
<tr>
<td>UGC – Participatory journalism sites</td>
<td>6.00***</td>
<td>1.00***</td>
<td>3.00**</td>
<td>1.00***</td>
</tr>
<tr>
<td>UGC – Discussion forums</td>
<td>6.00***</td>
<td>2.00</td>
<td>2.00</td>
<td>3.00***</td>
</tr>
</tbody>
</table>

**Notes:** Entries represent medium scores. Significance tests were based on post-hoc Mann-Whitney non-parametric test, which compared score of each UGC group with comparison (reference) group score for each attribute separately. Bonferroni correction method was utilized to avoid inflation of Type I error (two-tailed *p* value in post-hoc analyses was set to 0.01). Participation facilitation attribute was not included in the table because there was no statistically significant difference in medians of different groups of Web sites on participation facilitation attribute score, as evidenced by Kruskal-Wallis non-parametric test. Therefore, no post-hoc probing was done on participation facilitation attribute.  
** < .01; *** < .001.
Notes:
The bars represent 95% confidence interval for the mean. The Y-axis represents the rescaled score of Web sites. Each score within a given attribute was divided by the largest score on that attribute (among all Web sites) and then multiplied by 100. The resulting rescaled scores range from zero (if the lowest score on that specific attribute was zero) to 100 (the highest actual score on that specific attribute).

Figure 3: Error Bars Plotting Rescaled Attribute Scores in Different Web Site Groups.
There was a statistically significant difference between different groups of Web sites on information retrieval, Kruskal-Wallis $H(5) = 69.30, p < .001$. Post-hoc tests indicated that compared to the reference group ($\textit{Mdn} = 9.00$), blogs ($\textit{Mdn} = 4.00$) were lower on information retrieval, Mann-Whitney $U = 49.50, p < .001$; wikis ($\textit{Mdn} = 4.00$) were lower, Mann-Whitney $U = 16.00, p < .001$; content-sharing sites ($\textit{Mdn} = 5.50$) were lower, Mann-Whitney $U = 122.00, p < .001$; participatory journalism sites ($\textit{Mdn} = 6.00$) were lower, Mann-Whitney $U = 142.50, p < .001$; and discussion forums ($\textit{Mdn} = 6.00$) were also lower, Mann-Whitney $U = 106.50, p < .001$. These results directly contradicted hypothesis 1.

There was a statistically significant difference between different groups of Web sites on customizability, Kruskal-Wallis $H(5) = 47.38, p < .001$. Post-hoc tests indicated that compared to the reference group ($\textit{Mdn} = 2.00$), blogs ($\textit{Mdn} = 1.00$) were lower on customizability, Mann-Whitney $U = 218.50, p < .001$; wikis ($\textit{Mdn} = 1.00$) did not statistically differ, Mann-Whitney $U = 330.50, p = .06$; content-sharing sites ($\textit{Mdn} = 2.00$) did not statistically differ, Mann-Whitney $U = 361.00, p = .18$; participatory journalism sites ($\textit{Mdn} = 1.00$) were lower, Mann-Whitney $U = 210.00, p < .001$; and discussion forums ($\textit{Mdn} = 2.00$) did not statistically differ, Mann-Whitney $U = 356.50, p = .16$. These results did not support hypothesis 1.

There was a statistically significant difference between different groups of Web sites on manipulability, Kruskal-Wallis $H(5) = 20.36, p < .01$. Post-hoc tests indicated that compared to the reference group ($\textit{Mdn} = 2.00$), blogs ($\textit{Mdn} = 2.00$) did not statistically differ, Mann-Whitney $U = 392.50, p = .38$; wikis ($\textit{Mdn} = 2.00$) did not statistically differ, Mann-Whitney $U = 401.50, p = .46$; content-sharing sites ($\textit{Mdn} = 3.00$) did not statistically differ, Mann-Whitney $U = 313.50, p = .04$; participatory journalism sites ($\textit{Mdn} = 3.00$) were higher, Mann-Whitney $U = 280.50, p < .01$; and discussion forums ($\textit{Mdn} = 2.00$) did not statistically differ, Mann-Whitney $U = 381.00, p = .29$. These results provided very limited support for Hypothesis 1.
There was no statistically significant difference between different groups of Web sites on participation facilitation, Kruskal-Wallis $H(5) = 9.02, p = .10$. Therefore, no post-hoc probing was conducted. However, because the scores of reference group were a constant (equal to zero), and scores of various UGC groups were greater than zero, these results provided support for hypothesis 1.

There was a statistically significant difference between different groups of Web sites on community orientation, Kruskal-Wallis $H(5) = 59.68, p < .001$. Post-hoc tests indicated that compared to the reference group of traditional news Web sites ($Mdn = 0.00$), blogs ($Mdn = 0.00$) did not statistically differ, Mann-Whitney $U = 351.00, p = .07$; wikis ($Mdn = 1.00$) were higher, Mann-Whitney $U = 267.00, p < .01$; content-sharing sites ($Mdn = 2.00$) were higher, Mann-Whitney $U = 146.00, p < .001$; participatory journalism sites ($Mdn = 1.00$) were higher, Mann-Whitney $U = 243.00, p < .001$; and discussion forums ($Mdn = 3.00$) were higher, Mann-Whitney $U = 61.00, p < .001$. These results supported hypothesis 1.

Based on the above results, Hypothesis 1 received only very limited support. The data analyses provided evidence that several attributes (i.e., community orientation and manipulability), to smaller or larger degree, tended to be more extensively present on UGC sites. The participation facilitation attribute was entirely absent from the traditional news sites, whereas it was present on UGC sites. Finally, several other attributes (i.e., information retrieval and customizability) tended to be more extensively present on the traditional news sites.

Research question 1. The research question asked: “What is the precise pattern of the five attributes’ presence across the following political UGC types: blogs, wikis, content-sharing Web sites, participatory journalism Web sites, and discussion forums?” A hierarchical cluster analysis was the most appropriate data analytic technique because the method generally enables testing and refining of existing classifications, as well as developing new general-purpose and specific-
purpose classifications (Aldenderfer & Blashfield, 1984). Developing a specific-purpose classification of UGC Web sites based on their attributes is one of the goals of this study.

Cluster analysis is generally considered to be an exploratory procedure that does not have any statistical significance tests associated with it. Cluster analysis helps researchers find similar and dissimilar cases in the data. This method is particularly useful in the beginning stages of developing a formal classification system for some particular phenomenon or population. The method is considered mainly exploratory because of a disagreement about what “similarity” and, therefore, what a “cluster” is (Aldenderfer & Blashfield, 1984; Everitt, 1993). Nonetheless, cluster analysis has been utilized in various scientific disciplines to formulate and even test hypotheses (Romesburg, 2004). Specifically, there are several sequential steps recommended to be followed when performing cluster analysis (Aldenderfer & Blashfield, 1984; George & Mallery, 2009; Romesburg, 2004). The section below proceeds according to these steps.

The first step of the cluster analysis was to standardize the data. Before performing cluster analysis, scores of all the Web sites on attribute presence were standardized into Z-scores. Conversion into Z-scores is one of the most common forms of data standardization in cluster analyses (Aldenderfer & Blashfield, 1984; Romesburg, 2004). Standardization was important because in the current research attributes were measured on different scales (different attributes were operationally comprised of different number of features). Without standardization, attributes that had scales with greater range would disproportionally drive the clustering (Romesburg, 2004). Standardization made all attributes equal in their contribution to the final classification (George & Mallery, 2009). Standardization is also generally considered to be important when distance measures of similarity are used, as is the case here (Aldenderfer & Blashfield, 1984). However, despite clear advantages, data standardization has some disadvantages, such as making variables equal in their contribution to development of a clustering solution while such
equalization might not be theoretically desirable (see Everitt, 1993). This does not present a problem in the present study, where all attributes are viewed as having equal conceptual importance and should have an equal role in determining clusters.

The second step of the cluster analysis was to select a similarity or dissimilarity measure. In this project, the squared Euclidean distance dissimilarity measure was used. Euclidean distance is fundamentally a measure of how far apart (in terms of actual geometric distance) each pair of Web sites are located in relation to one another in a multi-dimensional space. Operationally, squared Euclidean distance is the sum of the squared differences for each variable for each case (George & Mallery, 2009, p. 264). This measure was chosen due to its several desirable properties: (1) it is sensitive to size displacement differences (e.g., differences on each attribute’s score among political UGC types); (2) it is sensitive to mirror image translations (e.g., profiles of attribute scores of two political UGC types are “mirror images” when they are direct opposites of one another); (3) it is the only measure that can be used with the clustering method used in this work and described below (Romesburg, 2004). Additionally, squared Euclidean distance has been widely utilized in existing research (Aldenderfer & Blashfield, 1984; Romesburg, 2004).

The third step of the cluster analysis was to select a clustering method. The cluster analysis performed here utilized Ward’s clustering method (which belongs to hierarchical agglomerative methods family, see Aldenderfer & Blashfield, 1984) because it: (1) outperformed other methods in various Monte Carlo studies when clusters were located close to one another or overlapped – a potential for which exists in this study; (2) outperformed other methods in various Monte Carlo studies when few outliers were present in the data (Aldenderfer & Blashfield, 1984; Punj & Stewart, 1983), which is applicable to the present data (see discussion on replacement of outliers below); (3) was recommended for situations where complete coverage (placement of
every single Web site in clusters) was required, and (4) was extensively used in the social sciences and beyond (Aldenderfer & Blashfield, 1984).

The hierarchical agglomerative clustering methods group all Web sites into progressively larger clusters until an overarching cluster combining all Web sites is formed (George & Mallery, 2009). Ward’s clustering method initially treats each Web site as its own cluster. The method works by joining Web sites (or previously formed clusters of Web sites) in a manner producing the minimum increase in the error sum of squares (ESS) (Aldenderfer & Blashfield, 1984, p. 43). “At each step in the analysis, union of every possible pair of clusters is considered and the two clusters whose fusion results in the minimum increase [ESS] are combined” (Everitt, 1993, p. 65).

In this work standardized scores on each attribute outside of three standard deviations were considered outliers (Romesburg, 2004) and were replaced by standardized scores on that attribute equal to the mean score on the attribute plus two standard deviations. This preserved the shape of the original distribution, in which the outlying scores were close to the extreme of the distribution, but the scores became less extreme (Field, 2005). Replacement of outliers was important for preservation of Ward’s clustering method’s statistical robustness (Aldenderfer & Blashfield, 1984; Punj & Stewart, 1983). For the customizability attribute, two outliers were replaced. For the manipulability attribute, one outlier was replaced. For the participation facilitation attribute, four outliers were replaced. For the community orientation attribute, two outliers were replaced. Before being replaced, each outlier was individually examined, and the results of the examination are presented in Table 4. It appears that the presence of outliers (replaced in cluster analysis, but unmodified in other analysis) was neither due to (a) data entry errors, nor due to (b) presence of important substrata of UGC Web sites that were undersampled (Aldenderfer & Blashfield, 1984).
<table>
<thead>
<tr>
<th>Outlier Attribute</th>
<th>Outlier URL</th>
<th>Outlier Alexa traffic score</th>
<th>Outlier raw score</th>
<th>Outlier Z score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customizability</td>
<td>youtube.com</td>
<td>3</td>
<td>8</td>
<td>4.512</td>
</tr>
<tr>
<td></td>
<td>scribd.com</td>
<td>253</td>
<td>7</td>
<td>3.806</td>
</tr>
<tr>
<td>Manipulability</td>
<td>merinews.com</td>
<td>5883</td>
<td>7</td>
<td>4.095</td>
</tr>
<tr>
<td></td>
<td>dailykos.com</td>
<td>4803</td>
<td>4</td>
<td>8.699</td>
</tr>
<tr>
<td>Participation</td>
<td>nicedeb.wordpress.com</td>
<td>155417</td>
<td>2</td>
<td>4.225</td>
</tr>
<tr>
<td>facilitation</td>
<td>commondreams.org</td>
<td>16473</td>
<td>2</td>
<td>4.225</td>
</tr>
<tr>
<td></td>
<td>youtube.com</td>
<td>3</td>
<td>2</td>
<td>4.225</td>
</tr>
<tr>
<td>Community orient.</td>
<td>gamespot.com/forums/index.html</td>
<td>483</td>
<td>7</td>
<td>3.182</td>
</tr>
<tr>
<td></td>
<td>fotolog.com</td>
<td>163</td>
<td>7</td>
<td>3.182</td>
</tr>
</tbody>
</table>

Notes: Scores outside of three standard deviations were considered outliers. Outlying scores were replaced with standardized scores on that attribute equal to the mean score plus two standard deviations only in cluster analysis.

Table 4. Descriptive Characteristics of Data Outliers Replaced in Cluster Analysis.

To determine how optimal the previously discussed cluster analysis procedures were for the current data, a validity assessment was performed. Step four of cluster analysis, assessment of validity, was carried out via a split-sample technique (Aldenderfer & Blashfield, 1984; Everitt, 1993; Romesburg, 2004). Web sites within each of the five UGC groups were ordered from the most to least trafficked, according to Alexa (2010). Each Web site was assigned a number from one to 150. The sample was split in half, with the first half containing all odd-numbered and the
second half containing all even-numbered Web sites from each of the five UGC groups. This sampling division procedure ensured that the resulting split samples were equivalent in terms of the number of Web sites in each UGC group and in terms of average popularity of Web sites. The cluster solution obtained with the above-described procedure for the first half was compared with the cluster solution for the second half. Dendrograms for both samples are presented in Figure 4.
Notes: The Y axis is a rescaled distance (dissimilarity) score. The X axis is the number of Web sites in each cluster. Dotted line indicates where the dendrograms were cut, resulting in 3 clusters in each sample.

Figure 4: Cluster Analysis Dendrograms: Comparison of Two Split-Sample Cluster Solutions.
The substantive results and evaluation of the two cluster solutions are offered in Table 5. As can be seen in Table 5, cluster solutions in both halves of the sample were rather consistent with each other. In both samples, similar clusters were identified (a detailed description of how clusters were identified is offered below). For example, cluster 1 was dominated by participatory journalism sites, according to the cluster-solutions for both halves. Similarly, cluster 2 was dominated by blogs and wikis, and cluster 3 was dominated by discussion forums and content-sharing sites. Due to split-sample consistency, the validity of the procedure was deemed acceptable. After the validity assessment, the data from both halves of the sample were combined, and the remaining analyses were based on the combined dataset including all 150 UGC Web sites.
<table>
<thead>
<tr>
<th>Cluster</th>
<th>Odd (n=75)</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Blogs: 21%</td>
<td>Wikis: 12%</td>
<td>Content-sharing sites: 18%</td>
<td>Particip. journalism: 37%</td>
<td>Discussion forums: 9%</td>
</tr>
<tr>
<td></td>
<td>(n=75)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Even: 15%</td>
<td>Wikis: 23%</td>
<td>Content-sharing sites: 15%</td>
<td>Particip. journalism: 35%</td>
<td>Discussion forums: 12%</td>
</tr>
<tr>
<td></td>
<td>(n=75)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cluster 2</td>
<td>Odd (n=75)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Blogs: 38%</td>
<td>Wikis: 38%</td>
<td>Content-sharing sites: 4%</td>
<td>Particip. journalism: 9%</td>
<td>Discussion forums: 9%</td>
</tr>
<tr>
<td></td>
<td>(n=75)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Even: 29%</td>
<td>Wikis: 26%</td>
<td>Content-sharing sites: 18%</td>
<td>Particip. journalism: 12%</td>
<td>Discussion forums: 15%</td>
</tr>
<tr>
<td></td>
<td>(n=75)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cluster 3</td>
<td>Odd (n=75)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Blogs: 0%</td>
<td>Wikis: 14%</td>
<td>Content-sharing sites: 36%</td>
<td>Particip. journalism: 5%</td>
<td>Discussion forums: 45%</td>
</tr>
<tr>
<td></td>
<td>(n=75)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Even: 7%</td>
<td>Wikis: 0%</td>
<td>Content-sharing sites: 33%</td>
<td>Particip. journalism: 13%</td>
<td>Discussion forums: 47%</td>
</tr>
<tr>
<td></td>
<td>(n=75)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Notes: Cells represent presence of different Web site types (e.g., blogs, wikis) in each cluster. The number indicates percent of all Web sites within a particular cluster.*

*Table 5. Cluster Analysis Procedure Validity Assessment: Comparison of Two Split-Sample Cluster Solutions.*
The fifth step of the cluster analysis was to cut the obtained dendrogram for the entire sample in order to identify several specific clusters of political UGC Web sites based on their attribute scores. Figure 5 shows where exactly the dendrogram was cut (dotted horizontal line).

Notes: The figure represents dendrogram obtained via cluster analysis. Cluster 1 consists of 63 Web sites, cluster 2 of 46 Web sites, and cluster 3 of 41 Web sites. A more detailed dendrogram describing each Web site is available at: http://goo.gl/V7jOV.

Figure 5: Cluster Analysis Dendrogram Displaying Web Sites Grouped According to Their Scores on Attributes.
It was important to identify the proper area where the dendrogram should be cut. There are several numerical guidelines in regards to locating the proper area: (1) The dendrogram needs to be cut within a large range of the resemblance coefficient for which the number of clusters remain constant (Romesburg, 2004). In other words, researcher needs to examine the value of the fusion points (points on the dendrogram’s Y-axis where the clusters where merged) to “discover a significant ‘jump’ in the value of the coefficient (jump from low value to high value on the Y-axis). A jump implies that two relatively dissimilar clusters have been merged; thus the number of clusters prior to the merger is the most probable solution” (Aldenderfer & Blashfield, 1984, p. 57; also see Everitt, 1993, p. 73 for a similar prescription). It appears that the portion of the dendrogram where cluster 2 and cluster 1 were merged represents the first big “jump” (see Figure 5), with the distance of 7 points from cluster 1 to the fusion point, and with the distance of 5 points from cluster 2 to the fusion point. The fusion point for cluster 1 and cluster 2 (value 13 on the Y-axis) also appears to be sufficiently removed from point 8 (point of Cluster 2 formation) and point 6 (point of cluster 1 formation), unlike the fusion point for two sub-clusters making up cluster 2 (value 8), unlike the fusion point for two sub-clusters making up cluster 1 (value 6), and unlike the fusion point for three sub-clusters making up cluster 3 (value 5).

There is also a non-numerical consideration about where to cut the dendrogram:

“… the concept of similarity exists on a variable scale resolution. Each of us, according to our needs and circumstances, sets the right level of resolution. We are always caught between the opposite extremes of wanting complete detail about nature and yet wanting economy of detail too… The extremes of full detail [i.e., each Web site is its own cluster] and no detail [i.e., all Web sites are combined into a single cluster] have decided disadvantages. The optimal compromise lies somewhere between....” (Romesburg, 2004, p. 22-23).

In other words, having too many or too few clusters appears to be detrimental to understanding the phenomenon under consideration. Based on this (admittedly subjective) non-
numerical consideration, and based on the above-described numerical considerations, cutting the
dendrogram below the point of fusion for cluster 1 and 2 represents an acceptable decision. It
should be acknowledged that a two-cluster solution might also appear plausible, in which case
clusters 1 and 2 will become one. However, the three-cluster solution was deemed more
theoretically informative (i.e., providing more detailed information) and justified by the
appropriate considerations described above.

Next, a second (optional) validity check was performed on the obtained cluster-solution.
This validity assessment was carried out by comparing clusters on an external variable not used
for obtaining the cluster solution. Such procedure was recommended because of its ability to
demonstrate theoretical usefulness of the cluster solution (Aldenderfer & Blashfield, 1984). The
external variable for the validity check was the worldwide Web site traffic rank offered by Alexa
(2010) – the only variable conveniently available in the dataset. The rank represents how popular
a given Web site is, or how big its traffic is (with lower values representing greater popularity or
traffic). The rank scores ranged from 3 (for YouTube.com) to 13,880,181 (for ForumHome.org).
The mean rank was 310,656.03 ($SD = 1,423,749.27$) and the median was 5,778.

Formal tests of deviations from normal distribution (Kolmogorov-Smirnov and Shapiro-
Wilk tests) showed that distributions of Alexa ranks in each cluster were statistically different
from normal at the $p < .001$ level. Due to non-normal distribution, the same procedures as the
ones above were used employing non-parametric tests. Non-parametric probes, combined with
Bonferroni correction, were used. The $p$ value in post-hoc analyses reported below was adjusted
from 0.05 to 0.016 (0.05 divided by the number of post-hoc tests [n=3] for each attribute). There
was a statistically significant difference between Web site clusters on Alexa traffic rank score,
Kruskal-Wallis $H(2) = 39.80, p < .001$. Post-hoc probing indicated that traffic of cluster 1 Web
sites ($Mdn = 12,709$) did not differ from traffic of cluster 2 Web sites ($Mdn = 19,393.50$), Mann-
Whitney $U = 1316.00, p = .49$. Cluster 1 Web sites ($Mdn = 12,709$) had lower traffic than cluster 3 Web sites ($Mdn = 483$), Mann-Whitney $U = 511.00, p < .001$. Finally, cluster 2 Web sites ($Mdn = 19,393.50$) had lower traffic than cluster 3 Web sites ($Mdn = 483$), Mann-Whitney $U = 233.00, p < .001$.

In other words, cluster 3 Web sites were significantly more trafficked than cluster 2 or 1 Web sites. Cluster 1 Web sites were somewhat more trafficked than cluster 2 Web sites, but such traffic difference was not statistically significant. The above analyses illustrated that the obtained three-cluster solution does offer some preliminary theoretical benefits.

Cutting the dendrogram resulted in the identification of a new classification of the political UGC Web sites based on the presence of the five technological attributes. Descriptive characteristics of each cluster are shown in Table 6A and 6B. Cluster 1 appears to be characterized by relatively high information retrieval and manipulability scores. Cluster 2 appears to be characterized by relatively low scores on most attributes. Finally, cluster 3 appears to be characterized by relatively high scores on most attributes.
<table>
<thead>
<tr>
<th>Information Retrieval</th>
<th>Customizability</th>
<th>Manipulability</th>
<th>Participation Facilitation</th>
<th>Community Orientation</th>
</tr>
</thead>
</table>

Rescaled mean, median, mode

<table>
<thead>
<tr>
<th>Cl. 1</th>
<th>Cl. 2</th>
<th>Cl. 3</th>
<th>Cl. 1</th>
<th>Cl. 2</th>
<th>Cl. 3</th>
<th>Cl. 1</th>
<th>Cl. 2</th>
<th>Cl. 3</th>
<th>Cl. 1</th>
<th>Cl. 2</th>
<th>Cl. 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>50.64</td>
<td>34.18</td>
<td>55.65</td>
<td>13.49</td>
<td>10.32</td>
<td>40.24</td>
<td>41.72</td>
<td>16.77</td>
<td>37.97</td>
<td>0.00</td>
<td>7.06</td>
<td>4.26</td>
</tr>
<tr>
<td>45.45</td>
<td>36.36</td>
<td>54.54</td>
<td>12.50</td>
<td>12.50</td>
<td>37.50</td>
<td>42.86</td>
<td>14.28</td>
<td>42.85</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>36.36</td>
<td>36.36</td>
<td>63.64</td>
<td>12.50</td>
<td>12.50</td>
<td>37.50</td>
<td>42.86</td>
<td>14.29</td>
<td>28.57</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Rescaled SD

<table>
<thead>
<tr>
<th>Cl. 1</th>
<th>Cl. 2</th>
<th>Cl. 3</th>
<th>Cl. 1</th>
<th>Cl. 2</th>
<th>Cl. 3</th>
<th>Cl. 1</th>
<th>Cl. 2</th>
<th>Cl. 3</th>
<th>Cl. 1</th>
<th>Cl. 2</th>
<th>Cl. 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>20.95</td>
<td>11.99</td>
<td>15.50</td>
<td>10.83</td>
<td>8.45</td>
<td>19.06</td>
<td>14.12</td>
<td>6.93</td>
<td>14.84</td>
<td>0.00</td>
<td>13.60</td>
<td>17.59</td>
</tr>
<tr>
<td>11.43</td>
<td>16.66</td>
<td>24.99</td>
<td>11.49</td>
<td>7.14</td>
<td>0.00</td>
<td>11.56</td>
<td>0.00</td>
<td>14.28</td>
<td>0.00</td>
<td>11.49</td>
<td>54.35</td>
</tr>
</tbody>
</table>

Notes: Comparison is made between three clusters of Web sites obtained through cluster analysis. The scores were rescaled to simplify interpretation of results. The following rescaling procedure was used: Each score within a given attribute was divided by the largest score on that attribute (among all Web sites) and then multiplied by 100. The resulting rescaled scores range from zero (if the lowest score on that specific attribute was zero) to 100 (the highest actual score on that specific attribute).

Table 6. Descriptive Characteristics for Attribute Scores.
### Table 6 Continued

<table>
<thead>
<tr>
<th>Information Retrieval</th>
<th>Customizability</th>
<th>Manipulability</th>
<th>Participation Facilitation</th>
<th>Community Orientation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>Cl. 1</td>
<td>Cl. 2</td>
<td>Cl. 3</td>
<td>Cl. 1</td>
</tr>
<tr>
<td>median</td>
<td>5.57</td>
<td>3.37</td>
<td>6.12</td>
<td>1.08</td>
</tr>
<tr>
<td>mode</td>
<td>5.00</td>
<td>4.00</td>
<td>6.00</td>
<td>1.00</td>
</tr>
<tr>
<td>SD</td>
<td>Cl. 1</td>
<td>Cl. 2</td>
<td>Cl. 3</td>
<td>Cl. 1</td>
</tr>
<tr>
<td>2.30</td>
<td>1.32</td>
<td>1.70</td>
<td>0.86</td>
<td>0.67</td>
</tr>
</tbody>
</table>

*Notes:* Comparisons are done by using raw (unmodified) scores.

B

The distributions of attribute scores in each cluster are shown in Figure 6. Visually, the histograms appear to indicate that there is a consistent deviation from normalcy in all clusters for all attributes (perhaps with the exception of information retrieval).
<table>
<thead>
<tr>
<th>Cluster 1</th>
<th>Cluster 2</th>
<th>Cluster 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Information Retrieval</strong></td>
<td><strong>Information Retrieval</strong></td>
<td><strong>Information Retrieval</strong></td>
</tr>
<tr>
<td><img src="image1" alt="Histogram" /></td>
<td><img src="image2" alt="Histogram" /></td>
<td><img src="image3" alt="Histogram" /></td>
</tr>
<tr>
<td><strong>Customizability</strong></td>
<td><strong>Customizability</strong></td>
<td><strong>Customizability</strong></td>
</tr>
<tr>
<td><img src="image4" alt="Histogram" /></td>
<td><img src="image5" alt="Histogram" /></td>
<td><img src="image6" alt="Histogram" /></td>
</tr>
<tr>
<td><strong>Manipulability</strong></td>
<td><strong>Manipulability</strong></td>
<td><strong>Manipulability</strong></td>
</tr>
<tr>
<td><img src="image7" alt="Histogram" /></td>
<td><img src="image8" alt="Histogram" /></td>
<td><img src="image9" alt="Histogram" /></td>
</tr>
<tr>
<td><strong>Participation Facilitation</strong></td>
<td><strong>Participation Facilitation</strong></td>
<td><strong>Participation Facilitation</strong></td>
</tr>
<tr>
<td><img src="image10" alt="Histogram" /></td>
<td><img src="image11" alt="Histogram" /></td>
<td><img src="image12" alt="Histogram" /></td>
</tr>
<tr>
<td><strong>Community Orientation</strong></td>
<td><strong>Community Orientation</strong></td>
<td><strong>Community Orientation</strong></td>
</tr>
<tr>
<td><img src="image13" alt="Histogram" /></td>
<td><img src="image14" alt="Histogram" /></td>
<td><img src="image15" alt="Histogram" /></td>
</tr>
</tbody>
</table>

**All scores = zero**

**Notes:** The Y-axis represents percent of Web sites within the cell. The X-axis represents the rescaled score of Web sites. Each score within a given attribute was divided by the largest score on that attribute (among all Web sites) and then multiplied by 100. The resulting rescaled scores range from zero (if the lowest score on that specific attribute was zero) to 100 (the highest actual score on that specific attribute).

**Figure 6:** *Rescaled Frequency Distributions of Attribute Scores in Clusters.*
Despite the differences among clusters suggested by the descriptive statistics and histograms, it is important to see if such differences are statistically significant. However, before performing analyses, it was important to determine if data met the normal-distribution assumption. Formal tests of deviations from normal distribution (Kolmogorov-Smirnov and Shapiro-Wilk tests) showed that distribution of scores for each cluster on each attribute was statistically different from normal. Due to consistently similar results and due to K-S test being more commonly utilized, only K-S test results are reported below. For cluster 1, distributions of all attributes were significantly non-normal: for information retrieval, $D(63) = 0.18$, $p < .001$; for customizability, $D(63) = 0.20$, $p < .001$; for manipulability, $D(63) = 0.29$, $p < .001$; for participation facilitation the score did not vary and analysis could not be performed; and for community orientation, $D(63) = 0.26$, $p < .001$.

As for cluster 1, for cluster 2, the distributions of all attributes were significantly non-normal: for information retrieval, $D(46) = 0.16$, $p < .01$; for customizability, $D(46) = 0.27$, $p < .001$; for manipulability, $D(46) = 0.46$, $p < .001$; for participation facilitation, $D(46) = 0.45$, $p < .001$; and for community orientation, $D(46) = 0.28$, $p < .001$.

Finally, for cluster 3, the distributions of all attributes also were significantly non-normal: for information retrieval, $D(41) = 0.18$, $p < .01$; for customizability, $D(41) = 0.19$, $p < .01$; for manipulability, $D(41) = 0.20$, $p < .001$; for participation facilitation, $D(41) = 0.52$, $p < .001$; and for community orientation, $D(41) = 0.18$, $p < .01$.

Due to non-normal distribution, the same procedures as the ones above were used employing non-parametric tests. Non-parametric probes, combined with Bonferroni correction, were used. The $p$ value in post-hoc analyses reported below was adjusted from 0.05 to 0.016 (0.05 divided by the number of post-hoc tests [$n=3$] for each attribute). In the current analyses Bonferroni correction resulted in no pairs of clusters being regarded as statistically
indistinguishable, while the difference between the clusters had statistical significance with $p < 0.05$.

Table 7 shows how each cluster of UGC Web sites differed from one another in terms of their scores on the five attributes. Figure 7 visually illustrates such differences. More detailed description of these data is offered below.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cluster 1</strong></td>
<td>5.00***</td>
<td>1.00***</td>
<td>3.00***</td>
<td>0.00***</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cluster 2</strong></td>
<td>4.00***</td>
<td>1.00***</td>
<td>1.00***</td>
<td>0.00***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>0.50</strong>*</td>
</tr>
<tr>
<td><strong>Cluster 3</strong></td>
<td>6.00***</td>
<td>3.00***</td>
<td>3.00***</td>
<td>0.00n.s.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>3.00</strong>*</td>
</tr>
</tbody>
</table>

**Notes:** Entries represent medium scores. Significance tests are based on post-hoc Mann-Whitney non-parametric test, which compared score of each UGC cluster with one another for each attribute separately. Bonferroni correction method was utilized to avoid inflation of Type I error (two-tailed $p$ value in post-hoc analyses was set to 0.016).
1. Letter “a” signifies difference from cluster 1, “b” from cluster 2, and “c” from cluster 3.
2. Cluster 1 Web sites had a lower score on participation facilitation attribute than cluster 2 Web sites. **< .01; *** < .001.

Table 7. Median Scores of Web Site Clusters on Attributes
Information Retrieval

Participation Facilitation

Customizability

Community Orientation

Manipulability

Notes:
The bars represent 95% confidence interval for the mean. The Y-axis represents the rescaled scores of Web sites. Each score within a given attribute was divided by the largest score on that attribute (among all Web sites) and then multiplied by 100. The resulting rescaled scores range from zero (if the lowest score on that specific attribute was zero) to 100 (the highest actual score on that specific attribute).

Figure 7: Error Bars Plotting Rescaled Attribute Scores by Cluster and Web Site Group.
There was a statistically significant difference between different Web site clusters on information retrieval, Kruskal-Wallis $H(2) = 36.23$, $p < .001$. Post-hoc probing indicated that cluster 1 Web sites ($Mdn = 5.00$) scored higher than cluster 2 Web sites ($Mdn = 4.00$), Mann-Whitney $U = 725.50$, $p < .001$. However, cluster 1 Web sites ($Mdn = 5.00$) did not statistically differ from cluster 3 Web sites ($Mdn = 6.00$), Mann-Whitney $U = 1025.00$, $p = .07$. Finally, cluster 2 Web sites ($Mdn = 4.00$) were lower than cluster 3 Web sites ($Mdn = 6.00$), Mann-Whitney $U = 275.00$, $p < .001$.

There was a statistically significant difference between different Web site clusters on customizability, Kruskal-Wallis $H(2) = 67.36$, $p < .001$. Post-hoc tests indicated that cluster 1 Web sites ($Mdn = 1.00$) did not differ from cluster 2 Web sites ($Mdn = 1.00$), Mann-Whitney $U = 1219.50$, $p = .14$. However, cluster 1 Web sites ($Mdn = 1.00$) were lower than cluster 3 Web sites ($Mdn = 3.00$), Mann-Whitney $U = 264.00$, $p < .001$. Finally, cluster 2 Web sites ($Mdn = 1.00$) were lower than cluster 3 Web sites ($Mdn = 3.00$), Mann-Whitney $U = 111.00$, $p < .001$.

There was a statistically significant difference between different Web site clusters on manipulability, Kruskal-Wallis $H(2) = 78.18$, $p < .001$. Post-hoc tests indicated that cluster 1 Web sites ($Mdn = 3.00$) were higher than cluster 2 Web sites ($Mdn = 1.00$), Mann-Whitney $U = 118.00$, $p < .001$. Also, cluster 1 Web sites ($Mdn = 3.00$) did not differ from cluster 3 Web sites ($Mdn = 3.00$), Mann-Whitney $U = 1136.00$, $p = .27$. Finally, cluster 2 Web sites ($Mdn = 1.00$) were lower than cluster 3 Web sites ($Mdn = 3.00$), Mann-Whitney $U = 202.50$, $p < .001$.

There was a statistically significant difference between different Web site clusters on participation facilitation, Kruskal-Wallis $H(2) = 17.62$, $p < .001$. Post-hoc tests indicated that cluster 1 Web sites ($Mdn = 0.00$) were lower than cluster 2 Web sites ($Mdn = 0.00$), Mann-Whitney $U = 1102.50$, $p < .001$. Also, cluster 1 Web sites ($Mdn = 0.00$) did not differ from cluster
3 Web sites ($Mdn = 0.00$), Mann-Whitney $U = 1197.00$, $p = .05$. Finally, cluster 2 Web sites ($Mdn = 0.00$) did not differ from cluster 3 Web sites ($Mdn = 0.00$), Mann-Whitney $U = 795.50$, $p = .06$.

There was a statistically significant difference between different Web site clusters on community orientation, Kruskal-Wallis $H(2) = 72.14$, $p < .001$. Post-hoc tests indicated that cluster 1 Web sites ($Mdn = 1.00$) did not statistically differ from cluster 2 Web sites ($Mdn = 0.50$), Mann-Whitney $U = 1325.50$, $p = .41$. Also, cluster 1 Web sites ($Mdn = 1.00$) were lower than cluster 3 Web sites ($Mdn = 3.00$), Mann-Whitney $U = 152.50$, $p < .001$. Finally, cluster 2 Web sites ($Mdn = 0.50$) were lower than cluster 3 Web sites ($Mdn = 3.00$), Mann-Whitney $U = 138.00$, $p < .001$.

Based on the above analysis, the new UGC classification system appears to differ from the conventional UGC classification that relied on grouping Web sites into blogs, wikis, content-sharing sites, discussion forums, and participatory journalism sites. The full dendrogram describing each Web site (located at: http://goo.gl/V7jOV) shows that each cluster and most sub-clusters (i.e., lower-level clusters) generally contained Web sites of all types. Figure 8 provides numerical information, and confirms that each of the three main clusters consisted of all five conventional UGC types. In fact, there was a relatively even distribution of Web sites in cluster 1 (no Web site group exceeded 32% or fell below 11%) and cluster 2 (no Web site group exceeded 33% or fell below 9%). In cluster 3, however, discussion forums were relatively dominant (46%), and blogs relatively scarce (2%).
Cluster 1  
14 blogs  
11 wikis  
11 content-sharing sites  
20 participatory journalism sites  
7 discussion forums

Cluster 2  
15 blogs  
15 wikis  
5 content-sharing sites  
7 participatory journalism sites  
4 discussion forums

Cluster 3  
1 blog  
4 wikis  
14 content-sharing  
3 participatory journalism  
19 discussion forums

Notes: Percentage indicates a proportion of specific group out of all Web sites in the cluster.

Figure 8: Distribution of Web Site Groups across Three Clusters.

Also, based on the above analysis, UGC groups should be described as follows: (1) Cluster 1: high information retrieval and manipulability with participatory journalism sites and blogs being the largest Web site groups. (2) Cluster 2: consistently low scores on all attributes - except participation facilitation, with blogs and wikis being the largest Web site groups. (3) Cluster 3: consistently high scores on all attributes - except participation facilitation, and
primarily composed of content-sharing sites and discussion forums (two group accounting for 
80% of all Web sites in cluster 3).

*Hypothesis 2.* The second hypothesis proposed that “The presence of task-specific 
information retrieval performance and information environment customizability will be positively 
correlated across all UGC Web sites.” Due to the data skewness, a non-parametric Spearman 
correlation analysis was performed (Field, 2005). Spearman \( r \) is interpreted similarly to Pearson’s 
\( r \) (Sheskin, 2004). The results of a one-tailed bivariate correlation analysis suggested that the 
hypothesis was supported due to presence of a positive correlation between information retrieval 
and customizability attributes on UGC Web sites, Spearman \( r = 0.31 \), \( p \) (one-tailed) < 0.001.

*Hypothesis 3.* The third hypothesis proposed that “The presence of direct participation 
facilitation and of community orientation will be positively correlated across all UGC Web sites.” 
Results of a one-tailed bivariate correlation analysis indicated that this hypothesis was rejected 
due to the absence of a statistically significant correlation between participation facilitation and 
community orientation, Spearman \( r = 0.07 \), \( p \) (one-tailed) = 0.21.
Chapter 9: Discussion

This section begins by reiterating the study’s goals. It then proceeds to describe several unexpected results. Elaboration on these surprising findings and on general attribute distribution across UGC Web sites is offered next. Finally, study’s limitations, implications, and directions for future research are discussed.

The present study is a part of a broader project aiming to demonstrate the benefits of the MOA approach for systematizing knowledge about media and their effects. The goals of this broader project are to: (a) explicate the technological attributes of media important to understanding media’s various effects; (b) use UGC to illustrate which attributes might be regarded as important and how the attribute explication process should proceed. The present study specifically attempted to:

1. Contribute clarity to the conceptual understanding of political UGC.
2. Identify a set of technological attributes that could serve as a theory-driven operationalization of political UGC, which could be a useful foundation for understanding the nature of political UGC and effects it exerts on a specific variable of interest.
3. Document the presence of attributes on the most popular UGC and traditional news media Web sites.
4. Develop a theoretically- and empirically-grounded classification of political UGC for political participation research.
5. Demonstrate how the MOA approach can be used for explicating media through their attributes.

The study appears to have successfully achieved these five specific goals. However, before addressing these goals more thoroughly, it is important to elaborate on several surprising findings.

*Unexpected Findings*

It might be useful to begin with the unexpected findings that represent relatively easy cases, and then proceed to more difficult ones. Therefore, hypothesis 3 findings are discussed first, and followed by hypotheses 1 and other findings.

*Hypothesis 3*. An expected correlation between participation facilitation and community orientation was absent. The correlation was expected due to a rather intuitive positive impact both attributes could have on promoting political participation, which should encourage Web site administrators willing to promote political participation to extensively utilize both attributes. Examination of the data suggests that a possible explanation for this null finding has to do with the presence of community orientation on blogs. Table 3 and Figure 2C and Figure 3 show that blogs had very low scores on community orientation. The scores of blogs did not differ from the reference group. Blogs, a group of explicitly political Web sites, unsurprisingly had the highest scores on participation facilitation. All remaining UGC Web sites had relatively high scores on community orientation and relatively low scores on participation facilitation.

This pattern in the data suggests that perhaps the impact of community orientation on political participation is not as intuitive as was suggested, and that administrators of political blogs did not recognize community orientation as an effective tool for engaging users in political action. In other words, blog administrators might have been unwilling to implement community orientation features because they underappreciated the potential positive impact of this attribute.
Another possible explanation of low community orientation on blogs has to do with the specific attribute features and how they fit in the overall design and function of blogs. Community orientation was theorized to help Web site users better understand who the other users on the Web site were and what groups existed on the Web site. This attribute should be particularly valuable when there is a substantial amount of horizontal (i.e., user-to-user) communication. Although not coded as part of this study, subjective impressions are that blogs have relatively little of such communication, especially when compared to discussion forums, which generally contain nothing but such communication.

The operational model of many popular blogs appears to be anchored or centered on the blogger, who posts entries on which blog readers are sometimes given an opportunity to comment. The only possible horizontal communication can occur among readers in this comments section. It appears that the dominant practice on most blogs is for the blogger to talk to blog readers and occasionally engage in interactive exchanges with the readers, as opposed to provide a variety of communication options for blog readers to interact with one another. This can be contrasted with the practice of Wikipedia to provide a discussion space for each entry, and with the practice of its members organizing into groups to tackle specific Wikipedia projects. It can also be contrasted with potential horizontal communication that occurs on content-sharing Web sites, centered on the content uploaded by users (e.g., commenting, rating, private messaging). Therefore, community orientation does not appear to fit the blog model particularly well, especially when compared to other Web site types, and webmasters of blogs that follow the above-described model might be not interested in utilizing what community orientation attribute offers.

A blogger survey conducted by Pew (2006), although somewhat dated, provides a thorough description of bloggers’ motivations for blogging, which are relevant to the present
discussion. The in-depth survey of a random sample of 233 bloggers asked them to indicate major and minor reasons for choosing to blog. These reasons were:

1. “To express yourself creatively” (Pew, 2006, p. 8) was mentioned by 77% of bloggers (with 52% citing it as a major reason and 25% as a minor reason).

2. “To document your personal experiences or share them with others” - 76%.

3. “To share practical knowledge or skills with others” - 64%.

4. “To motivate other people to action” - 61%.

5. “To entertain people” - 61%.

6. “To influence the way other people think” - 61%.

7. “To stay in touch with friends and family” - 57%.

8. “To network or to meet new people” - 50%.

9. “To store resources or information that is important to you” - 49%.

10. “To make money” - 15%.

Examination of this list is informative and shows that the top six reasons appear to indirectly support the model of blog operation proposed above: The blogger is the center of attention, the blogger produces a vast amount of content, and blog readers might be given only limited opportunities to interact with one another and to contribute content. Also, more than 60% of bloggers indicated their desire “To motivate other people to action” (p. 8). This partially helps explain the high score blogs received on participation facilitation. It should be acknowledged that the survey did not focus on the most popular political blogs, and that motivation for blogging might have changed in the last five years. However, the results of the study appear to indirectly corroborate the above reasoning about why blogs had low community orientation and high participation facilitation. An even earlier and more exploratory study largely confirmed Pew’s findings by showing that four of five major motivations for blogging were: “documenting one’s
life; providing commentary and opinions; expressing deeply felt emotions; articulating ideas through writing” (Nardi, Schiano, Gumbrecht & Swartz, 2004, p. 43). However, this study did find that the fifth motivation for blogging was a creation of a community forum.

Hypothesis 1. The most counterintuitive results were in regards to the extensive implementation of information retrieval and customizability on the Web sites belonging to traditional news organizations. The explanation appears to be that news organizations and their Web sites have a capacity for change. Historically, news organizations have evolved to reflect cultural or technological changes. For example, the development of the telegraph allowed newspapers to deliver more timely news across greater distances (Fang, 1997). Also, when television emerged as a dominant medium, radio had to take several drastic steps to remain competitive: (1) improving sound quality with the invention of the FM frequency; (2) becoming portable (Fornatale & Mills, 2011); (3) beginning to sound more personal than before (Fang, 1997; Rodman, 2006); and (4) refocusing from prime time to “drive time” (Fang, 1997, p. 150). Finally, the development of the Internet facilitated media convergence, in which various platforms of information delivery were utilized by news organizations (Briggs & Burke, 2002).

Media evolve for many reasons, one of which is user preferences. Web sites of traditional news organizations must respond to preferences and requests of their users. When users of a traditional news organization’s Web site express a strong preference for some features (i.e., a greater ability to customize their experience on that site or to be able to search content more efficiently), the Web site can either respond to its customers or risk losing them. A recent Pew (2010) survey found that 42% of Internet users consider the ability to customize what information they receive from news sites to be important. News organizations are likely to respond to such preferences of users even if these preferences do not agree with the news organization’s philosophy (i.e., customizability is in conflict with the traditional news organizations’ view of
themselves as gatekeepers of information, see Dylko et al., 2009). Therefore, it is possible that news organizations have incorporated information retrieval and customizability because their users value the specific features representing these two attributes. Numerous other factors can encourage news organizations’ management to adopt information retrieval and customizability attributes, including: reduced cost and improved functionality of communication technology that creates new opportunities for information delivery (Bimber, 2003); changing norms and culture among professional journalists and audiences (Hermida & Thurman, 2008); novel opportunities for generating profits by news organizations (Deuze, Bruns & Neuberger, 2007); and competition from the novel Internet-only media organizations (Dimmick, Che & Li, 2004; Lowrey & Mackay, 2008). Without knowledge of the proprietary deliberations taking place among the management of news organizations, it is impossible to know for sure what specific considerations guided decisions in regards to various specific changes. Some limited evidence does indicate that today’s news executives pay substantial attention to Web site usage preferences of their audiences (Hermida & Thurman, 2008).

However, despite some disappointment associated with obtaining results opposite to what was expected in regards to the presence of information retrieval and customizability on UGC versus traditional Web sites, such results might paradoxically support the main thesis of this work and of the MOA approach. These findings reemphasize that instead of focusing their attention on specific media forms (which have a robust capacity for transformation), researchers should focus their attention on the specific technological attributes. These attributes are argued to be substantially more stable are resilient to technological change. Information technology will inevitably change over time, making it likely that specific attribute features might change as well. However, the conceptual nature of the attributes and theorized effects of attributes should not change. In other words, while attribute operationalization might be susceptible to influence of
future technological transformation, the conceptualization is not. As was proposed in the beginning of this work, different media are likely to continue evolving rapidly and drastically in the near future. Treating such media as discrete and stable units of study can be unproductive, at best, and misleading, at worst. How might it be misleading? The rapid pace of media transformation might produce a conceptual confusion, as the discussion of hypothesis 1 results demonstrated (assuming that transformation of traditional Web sites was indeed responsible for greater presence of the two attributes on traditional news Web sites). Media forms conceptualized to possess certain characteristics at time 1 might be found to have different characteristics after some time passes, at time 2. In other words, despite theorizing at time 1 being accurate and empirical results at time 2 being valid, they might appear to contradict one another due to rapid and substantial media transformation.

The general pattern of attribute distribution also showed that the five attributes exist across various media forms (i.e., UGC and traditional news Web sites). Participation facilitation was the only attribute entirely absent from traditional news Web sites. Such a finding confirms a major proposition of the MOA approach: Media forms differ from one another in a quantitative rather than qualitative sense. In other words, most media are generally composed of the same sets of attributes. What differentiates one media forms from another are the scores on specific attributes and not the presence of an entirely different set of attributes. The absence of one attribute out of five (i.e., participation facilitation) from traditional news Web sites is an exception that confirms the general MOA rule: Qualitative differences among media forms are rather rare, while quantitative are much more common.

It is also possible that there are some traditional news Web sites with participation facilitation present, and that these Web sites simply did not make it into the sample analyzed in the present study. In this case, participation facilitation would not be absent from traditional news
Web sites, and instead traditional news Web sites would simply have a very low score on participation facilitation. If this proved to be the case, it would be an even stronger support for the MOA approach proposition about media differing quantitatively rather than qualitatively.

*Consistently low scores on attributes.* Another surprising finding was that the presence of all five attributes on UGC Web sites was relatively low. As shown in Figure 2B, 2C, and 3, the vast majority of UGC Web sites had low scores. Such a consistent trend in the data might suggest that the five attributes do not accurately represent the essence of UGC phenomenon. It is argued that this is not the case, and another factor explains why the scores were relatively low.

Another possibility of low scores is that all or most features within any given attribute are functional alternatives of one another. For example, adding Web site users to one’s friends list might be viewed as tantamount to having the Web site recommend some users to be one’s friends. If both features are viewed as accomplishing exactly the same goal, they can be viewed as functional alternatives of each other. If that, in fact, is the case, this would provide a logical explanation of low attribute scores. However, the features do not appear to be functional alternatives. For example, sorting YouTube content on view count should return different results than sorting content by rating, despite both features representing information retrieval. The view count sorting can return content that is either extremely “bad” or “good” in quality, or funny, or ridiculous, or offensive. The rating sorting is likely to return content that is mostly “good” in quality, or “funny,” or has some other “positive” characteristics. Therefore, sorting content by view count and rating do not appear to be functional alternatives of one another. The presence of diverse information retrieval features on the most popular search engine, Google, might provide indirect evidence of this reasoning. If various sorting and search options were functional alternatives of one another, Google would likely reduce their number to simplify and streamline the user experience. Simplifying and streamlining its Web site is Google’s major priority (Jarvis,
A similar analysis applies to other attributes: All features appear necessary for any given attribute to be considered fully implemented, indicating that these features are not functional alternatives of one another.

The most likely explanation for this positive data skewness is methodological, and has to do with how the specific technological features, comprising each attribute, were coded. As the codebook (see Appendix) shows, many different features were combined into subgroups within each attribute to enable as reliable identification and coding of these features as possible. For example, the second feature of customizability allows users to place different persons into one’s online social network and to subscribe to different information streams on a particular Web site. Any Web site could receive a maximum score of only 2 for this feature (if the feature was located on the homepage), even if the Web site enables users to do all of the following: (a) “watch” a specific wiki page; (b) subscribe to a specific discussion forum thread; (c) “favorite” a given piece of content; (d) subscribe to specific email updates; (e) add Web site users to one’s “friends list;” and (f) subscribe to updates from a specific group on the Web site. Despite resulting in loss of valuable detail, combining several features into subgroups emerged as a necessary step during the coder training process. It became clear that some specific features, despite being different from one another in theory, could not be reliably differentiated from one another in practice. For example, in some cases it was impossible to differentiate features (d) from (f), (f) from (e), and (b) from (d).

Such a method of coding makes it likely that attribute indexes rather conservatively represent how extensively attributes are actually present on the Web sites. Therefore, the positive skew in the distribution of attribute scores does not indicate that the attributes were an incorrect representation of UGC, but is rather an artifact of the coding procedure.
Having found a reasonable explanation for low attribute scores, it was also interesting to find any other factors that could uniformly affect how extensively attributes were present. An examination of outliers in Table 4 suggests that many of these Web sites, which by definition had very high scores on attributes, were also relatively more popular than other Web sites. Can this imply a relationship between the degree of attribute implementation and Web site popularity? If yes, what can explain such a relationship?

It appears that all attributes except participation facilitation require substantial technical expertise (and thus, substantial financial resources) in order to be implemented extensively. There is no hard data available about the type of financial and technical resources possessed by the Web sites in the present sample. However, it appears that there should be at least some variation given the presence of vast commercial enterprises (e.g., YouTube, Flickr), the presence of non-profit Web sites (e.g., Wikipedia), and the presence of many relatively obscure participatory journalism Web sites (i.e., there are 11 participatory journalism Web sites with traffic rank greater than 1,000,000). These attributes, in essence, represent specific sophisticated database management tools. Because of the technological sophistication required to implement the attributes, it can be suggested that Web site traffic score should be correlated with the presence of all attributes except participation facilitation. The mechanism behind this prediction is as follows: To implement features, Web sites need to hire professional software developers. Hiring such developers has high financial costs. The costs might be manageable if the organization owning a Web site has substantial financial resources. These financial resources might often be an outcome or a correlate of the traffic the Web site receives. Therefore, Web site traffic should have a positive correlation with the presence of attributes. Although the Web sites included in the present study represent the most trafficked Web sites in each of their respective groups, there is still a substantial variance in their traffic score (as described above). Therefore, at least some
correlation between traffic score and presence of technologically sophisticated attributes should be expected. However, the strength of the correlation might be smaller in the present sample compared to a true random sample that would include Web sites of all popularity levels.

To empirically examine this possibility a non-parametric post hoc analysis was conducted to determine if Alexa traffic rank, described above (low values represent high traffic), was correlated with attribute scores among UGC Web sites. Evidence of a negative correlation would confirm this expectation. Results of the analysis were as follows: Alexa rank was negatively correlated with information retrieval, Spearman $r = -.16, p$ (one-tailed) $< 0.05$, with customizability Spearman $r = -.59, p$ (one-tailed) $< 0.001$, with manipulability, Spearman $r = -.14, p$ (one-tailed) $< 0.001$, and with community orientation, Spearman $r = -.44, p$ (one-tailed) $< 0.001$. Alexa rank was not correlated with participation facilitation, Spearman $r = .01, p$ (one-tailed) $= 0.40$. Such results fit the above proposition exceptionally well. However, the reverse causality is also plausible: More technologically sophisticated Web sites (i.e., those with high attribute scores) might attract more traffic because such Web sites possibly offer more content, more ways of using such content, and other functionality appealing to users.

In regards to participation facilitation, which is technologically substantially easier to implement compared to other attributes, one explanation of low scores appears likely. The ease of posting links that can facilitate participation of Web site users in politics does not guarantee that Web site administrators and fellow users will post such links. For example, excessively posting links that attempt to recruit users into political action can make such users (a) annoyed with the Web site as a whole or with the fellow users who posted such links and (b) reluctant to click on these links. Also, because the American public generally has relatively low levels of participation in politics (Dylko, 2010; Valenzuela, Park, & Kee, 2009; Vecchione & Caprara, 2009; Verba et al., 1995), it might be unreasonable to expect Web site users to click on many available
participation facilitating links. Instead, Web site users and administrators might want to maximize the likelihood of the users clicking on the links by making very *infrequent* requests for involvement accompanied with the links representing participation facilitation. All of the above factors should make Web site administrators and fellow users likely to abstain from posting many links, leading to a relatively low presence of participation facilitation.

The lack of motivation of news editors appeared to be the main reason responsible for the low presence of mobilizing information (a phenomenon similar to participation facilitation, which was discussed above) on pages of alternative newspapers (Stanfield & Lemert, 1987). The desire to explicitly promote participation among the audience, as opposed to providing readers issue information, was associated with the presence of mobilizing information.

Several other studies showed that one of the reasons why mobilizing information was absent was the fact that it was deemed expendable and dull detail (Lemert, 1984; Lemert, Mitzman, Seither, Cook & Hackett, 1977, p. 725). Similarly, three decades later, posters of political UGC might be cautious about making their readers bored or irritated with frequent suggestions for political participation. Extensive implementation of participation facilitation might therefore create a risk that readers migrate to more exciting and less pushy and annoying online political news destinations.

*Presence of Attributes across UGC Web Sites*

One of the goals of this project was to describe the pattern of attribute presence across UGC Web sites. Because no specific hypotheses could be derived, an examination of the pattern based on Figure 3 and Table 8 is offered in this section. Whereas Figure 3 visually illustrates differences between various conventional groups of Web sites on five attributes, Table 8 supplements this information with formal statistical tests. Table 8 reports post-hoc non-parametric tests used to determine if the conventional UGC groups differed on the attributes. The *p* value in
post-hoc analyses reported below was adjusted from 0.05 to 0.005 (0.05 divided by the number of
post-hoc tests [n=10] for each attribute). A Kruskal-Wallis test was used as an omnibus test of
difference, and a Mann-Whitney test was used to probe differences between each pair of UGC
groups. Due to the number of all tests (n=45), detailed results are not reported, and instead Table
8 summarizes the results.
<table>
<thead>
<tr>
<th>Information Retrieval</th>
<th>Customizability</th>
<th>Manipulability</th>
<th>Community Orientation</th>
</tr>
</thead>
<tbody>
<tr>
<td>blogs = wikis</td>
<td>blogs = wikis</td>
<td>blogs = wikis</td>
<td>blogs = wikis</td>
</tr>
<tr>
<td>blogs = cont.-shar. sites</td>
<td>blogs &lt; cont.-shar. sites***</td>
<td>blogs &lt; cont.-shar. sites**</td>
<td>blogs &lt; cont.-shar. sites***</td>
</tr>
<tr>
<td>blogs &lt; part.-journ. sites**</td>
<td>blogs = part.-journ. sites</td>
<td>blogs &lt; part.-journ. sites</td>
<td>blogs = part.-journ. sites</td>
</tr>
<tr>
<td>blogs &lt; disc. forums**</td>
<td>blogs &lt; disc. forums***</td>
<td>blogs = disc. forums</td>
<td>blogs &lt; disc. forums***</td>
</tr>
<tr>
<td>wikis &lt; cont.-shar. sites***</td>
<td>wikis &lt; cont.-shar. sites**</td>
<td>wikis = cont.-shar. sites</td>
<td>wikis = cont.-shar. sites</td>
</tr>
<tr>
<td>wikis &lt; part.-journ. sites***</td>
<td>wikis = part.-journ. sites</td>
<td>wikis &lt; part.-journ. sites</td>
<td>wikis = part.-journ. sites</td>
</tr>
<tr>
<td>wikis &lt; disc. forums***</td>
<td>wikis &lt; disc. forums**</td>
<td>wikis = disc. forums</td>
<td>wikis &lt; disc. forums***</td>
</tr>
<tr>
<td>cont.-shar. sites = part.-journ. sites***</td>
<td>cont.-shar. sites &gt; part.-journ. sites***</td>
<td>cont.-shar. sites = part.-journ. sites</td>
<td>cont.-shar. sites = part.-journ. sites</td>
</tr>
<tr>
<td>cont.-shar. sites = disc. forums</td>
<td>cont.-shar. sites = disc. forums</td>
<td>cont.-shar. sites = disc. forums</td>
<td>cont.-shar. sites = disc. forums</td>
</tr>
<tr>
<td>part.-journ. sites = disc. forums</td>
<td>part.-journ. sites &lt; disc. forums</td>
<td>part.-journ. sites = disc. forums</td>
<td>part.-journ. sites &lt; disc. forums</td>
</tr>
<tr>
<td>Notes: Significance tests are based on post-hoc Mann-Whitney non-parametric test, which compared score of each UGC cluster with one another for each attribute separately. Bonferroni correction method was utilized to avoid inflation of Type I error (two-tailed $p$ value in post-hoc analyses was set to 0.005). Results for participation facilitation are not reported, due to absence of statistically significant omnibus difference between the groups, as indicated by the Kruskal-Wallis test. ** $&lt; .005$; *** $&lt; .001$.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Before addressing the pattern of attribute presence it should be acknowledged that numerous factors affect what attributes and to what degree are implemented across various groups of Web sites. The following factors repeatedly emerge as likely explanations for attribute presence: (a) the type of users, (b) resources available to web administrators operating the Web sites; and (c) the motivations of web administrators. Each attribute is likely affected by a unique set of these reciprocally causal factors.

Figure 3 shows that for information retrieval, the highest-scoring UGC groups appeared to be content-sharing sites, participatory journalism sites, and discussion forums. Blogs and wikis had somewhat lower scores. This pattern is largely confirmed by formal tests reported in Table 8. Such a pattern of attribute presence is not surprising because there appears to be a positive correlation between the volume of information contained in the Web site’s database and the presence of information retrieval functionality. The greater the amount and diversity of information in a database (such as the voluminous databases of many content-sharing sites, participatory journalism sites, and discussion forums), the more useful different information retrieval options can be expected to be. However, many wikis (most prominently, Wikipedia) have vast information stores, and yet, appear to have the lowest information retrieval score.

Figure 3 shows that for customizability, the highest-scoring UGC groups were content-sharing sites and discussion forums. The lowest-scoring UGC groups were blogs and participatory journalism sites. Again, tests reported in Table 8 largely validate this observation. One possible explanation of the customizability presence pattern is a plausible correlation between diversity of sources or content on a Web site and the number of customizability features. Herring, Scheidt, Kouper, and Wright (2007) found that blogs are most often written by one individual (e.g., Instapundit), or by several individuals who often share the same ideology or interest (e.g., PowerLine). Because such a structure produces relatively homogenous content (e.g.,
there is lack of oppositional content), it appears that customizability options might be relatively less useful on blogs. However, on content-sharing sites or discussion forums, a great diversity of existing sources and content makes customizability options more useful, enabling users to screen out unwanted sources and content in a systematic and efficient fashion. However, such reasoning does not transfer to participatory journalism Web sites, which are presumably characterized by diversity of sources and content, and yet, have low customizability scores.

Figure 3 shows that for manipulability, the highest-scoring UGC group was participatory journalism sites. Blogs and wikis had somewhat lower scores. Tests reported in Table 8 largely confirm this observation. On the one hand, it is understandable why blogs have low scores: As was discussed above, many blogs might offer commenting as the only way for users to contribute content (e.g., *HotAir*, *CrooksAndLiars*). Such a mode of operation can be viable if the blogger is famous or widely read, who blogs to express his or her own views, and who will be followed regardless of how involved in the blog-writing process the readers are. Also, although wikis are almost synonymous with a widespread bottom-up content production (Bruns, 2008), it appears that wiki Web site administrators might believe that offering only a few ways for users to manipulate content (upload and edit various forms of content on a wiki entry) is sufficient. On the other hand participatory journalism sites have relatively high scores. Although participatory journalism has many forms (Domingo et al., 2009), it generally is based on the model of operation that relies heavily on content contribution of active users (Deuze et al., 2007). Therefore, the decision by participatory journalism Web site administrators to offer their users many ways to contribute content appears quite logical.

Figure 3 shows that for participation facilitation, the highest-scoring UGC group was blogs, and the lowest-scoring group was discussion forums. Results of formal tests for
participation facilitation are not reported due to the absence of a statistically significant omnibus difference between the groups, as indicated by the Kruskal-Wallis test.

It is quite clear why blogs had a high score on participation facilitation: The sampling procedure utilized in this study resulted in blogs being the most explicitly politically-oriented groups of UGC Web sites. The key factor determining the presence of this attribute appears to be the motivation of the content creator or poster. If the motivation is to mobilize other Web site users, then this attribute will be more extensively present. Many political blogs (e.g., CrooksAndLiars, Instapundit) often circulate petitions, or post requests to donate money or contact their representatives and senators (for similar evidence in non-political blogs, see Pew, 2006).

Additionally, the presence of this attribute might be correlated with who the Web site users are in terms of their political interest and partisanship strength. If the users are only somewhat interested in politics and do not have a strong commitment to any political ideology, it might be hard to expect any mobilization attempts directed at them to be successful. This reasoning was clearly established in the scholarly debate about whether requests to participate increase the likelihood of one’s participation, or whether it is people who are more likely to participate in the first place are being more frequently asked to participate (Fisher & Boekkooi, 2010; Lim, 2010; Verba et al., 1995). Therefore, the nature of the Web site audience might be important to whether Web site administrators implement participation facilitation more or less extensively. There is some evidence that readers of political blogs are more ideologically extreme than nonreaders (Davis, 2008; Eveland & Dylko, 2007). Most of blogs included in the sample were explicitly political and likely targeted politically interested and active audiences. This idiosyncratic characteristic of blogs is discussed in the Limitations section. Had more explicitly
political discussion forums (or other groups of Web sites) been included into the study, their scores on participation facilitation would likely be higher.

Figure 3 shows that for community orientation, the highest-scoring UGC groups were content-sharing sites and discussion forums. Blogs were the lowest-scoring UGC group on this attribute. This pattern of community orientation distribution across different Web site groups is largely confirmed in Table 8. One theoretically plausible explanation is that whenever the formation of communities enhances the user experience on the Web site, features tapping community orientation should be more extensively implemented. For example, some discussion forums might focus on very intimate topics, such as terminal illnesses, STDs, emotional problems, or other sensitive topics (Mo & Coulson, 2008). The creation of a community in which members grow to know one another, and where there are strong norms of civility, respect, and support, become essential for the successful operation of this type of a discussion forum. Content-sharing Web sites also incorporate many community-orientation features, as they are trying to connect users to one another. However, many blogs are centered on their prominent, and often opinionated, blogger (e.g., Michelle Malkin, Markos Moulitsas Zúñiga). The development of a sophisticated community structure might be unnecessary for such blogs to operate successfully, as was suggested above.

Finally, several other valuable conclusions can be drawn from Table 8. Table 8 shows that two pairs of Web site groups – blogs and wikis on the one hand, and discussion forums and content-sharing Web sites on the other – are virtually indistinguishable on all five attributes (i.e., blogs and wikis do not differ, and discussion forums and content-sharing Web sites do not differ). Table 8 also shows that wikis and discussion forums, blogs and discussion forums, and blogs and content-sharing sites differ from one another on three out of five attributes. All remaining pairs differ on a maximum of two out of five attributes.
This suggests that for blogs and wikis, and for discussion forums and content-sharing Web sites, the conventional groupings of UGC Web sites clearly do not work, because such groupings do not coincide with their attribute scores. This data generally suggests a very limited usefulness of using conventional UGC classification as a proxy for attribute scores. This conclusion is corroborated by Figure 8, which shows that each of the three clusters consists of all conventional Web site groups, indicating absence of a robust association between Web site type and attribute scores. If the theorizing about attribute effects detailed in the description of each attribute is correct, using the conventional UGC classification to examine such effects might be a suboptimal research strategy. Instead, as suggested by the MOA approach, focusing on the attributes themselves appears to be more appropriate.

**Limitations**

As with any studies, the present one has several limitations. The first limitation is a relatively low reliability obtained for one of the attributes – content manipulability (Krippendorff's alpha = 0.69). An in-depth examination of the content manipulability attribute indicated that feature number four, described in Appendix, brought the reliability level down. Removal of this feature from the attribute index increased the Krippendorff's alpha level to 0.86. A discussion with the coder responsible for coding content manipulability confirmed that both this researcher and the coder were interpreting codebook instructions for this feature identically. The discussion did not provide any specific clues as to why this feature was not coded particularly reliably. Also, specific cases where two coders disagreed on this specific feature were examined. Both coders appeared to follow coding instructions equally well. Therefore, it remains unclear why the reliability for this particular attribute was relatively low.

Some have suggested that Krippendorff's alpha, like many chance-correcting reliability coefficients, is a rather conservative measure of intercoder reliability (Lombard, Snyder-Duch &
Bracken, 2002). Obtaining a high level of Krippendorff's alpha is challenging particularly for variables with low variability of scores (Krippendorff, 2004). The manipulability attribute data used in reliability assessment was examined to determine whether it lacked variability. However, data from both coders appeared to be sufficiently variable (for coder 1, the minimum value of the attribute was “1,” maximum was “5,” $M = 3.28$, $SD = 1.17$, variance = 1.38; for coder 2, the minimum value was “2,” maximum “5,” $M = 3.33$, $SD = 0.97$, variance = 1.41). Therefore, it appears that the above methodological consideration cannot explain the relatively low level of the reliability coefficient for manipulability.

What are the implications of the relatively low reliability coefficient for manipulability? Some suggest that conclusions based on data with coefficients between 0.667 and 0.80 should be treated as tentative (Krippendorff, 2004). However, due to the fact that only one out of five features was responsible for the relatively low reliability (i.e., there was no uniformly low reliability across all features of this attribute), one can have at least moderate levels of confidence in the conclusions drawn about this attribute. Two other considerations make it appropriate to use a relatively low reliability coefficient in the present case. First, a lower alpha level (such as .69 for manipulability) is “appropriate for research that is breaking new ground with concepts that are rich in analytical value” (Riffe, et al., 1998, p. 151). Second, lower reliability levels are also appropriate when using conservative reliability coefficients, such as Krippendorff's alpha (see Lombard, Snyder-Duch & Bracken, 2002, p. 600).

The second limitation also concerns the issue of intercoder reliability. As was mentioned above, various similar features were combined into subgroups within each attribute. Without this step reliable coding of content would be impossible. However, as was acknowledged, such method of coding might have produced an across-the-board reduction in attribute scores. It is unclear how future researchers can overcome this challenge. Of course, it is theoretically possible
to hire coders who represent outliers in terms of frequency of using Web sites the study focuses
on (i.e., they use UGC Web sites much more frequently than the average), who are experts in the
technology being investigated (e.g., have training in web development, or web usability), and
who can intensively train for an extended period of time. These costly measures should improve
reliability. Such measures, however, might be out of reach for most researchers operating under
standard financial and time constraints. Also, even such measures cannot guarantee that (1) no
novel technological executions of an old feature would arise during the training process, or that
(2) the training will cover all the numerous ways specific features are technologically
implemented in one’s sample. Both of these factors (particularly the second one) can render
precise discrimination among features very challenging.

The third limitation is the nature of the UGC sample, in which explicitly political Web
sites (i.e., blogs) were analyzed alongside not explicitly political Web sites (i.e., all other groups).
It is important to reiterate why the sample was constructed in this manner. This study focused on
popular (in terms of audience size) UGC groups that could be content analyzed and that had
implications for individual-level political participation. Political blogs were important to include
because they met both criteria and because they are often considered to be some of the earliest
and most important political UGC Web sites (Tremayne, 2007). It was important to include all
other types of political UGC because they also met both criteria. However, all other popular UGC
groups were generally not explicitly political, although they contained substantial volumes of
popular explicitly political content. For example, despite being not explicitly political, YouTube
contains millions of popular political videos (see Dylko et al., 2009).

What are the implications of this particular aspect of sampling for the results and
conclusions of this study? It appears likely that if explicitly political discussion boards, wikis,
content-sharing sites, and participatory journalism Web sites were included, the presence of
participation facilitation would be similar to those of blogs. However, any attempt to sample only explicitly political Web sites of all those four groups might either (a) be impossible or (b) result in a sample of Web sites which are not particularly popular. This researcher is not aware of any lists containing only explicitly political content-sharing sites, or wikis, or participatory journalism sites. However, even if such lists exist, it appears that those Web sites might be relatively obscure – unlike the list of political blogs.

One potential solution would be to include the most popular blogs of all topical orientations. This would make all Web site groups similar in the degree of explicit focus on politics. However, politics is one of the most popular topics in the blogosphere (Pew, 2006), which means this step likely would only marginally reduce the participation facilitation score on blogs. However, future research can examine value of this sampling strategy change.

Due to differences in the degree of political orientation explicitness, analyses were performed after having removed blogs from the entire sample. Removing blogs had negligible impact on the study results. For example, the cluster solution remained largely unchanged: 100% of the Web sites in the newly obtained cluster 3 came from the original cluster 3. Similarly, 100% of the Web sites in the newly obtained cluster 1 came from the original cluster 1. However, only 66% of Web sites in the newly obtained cluster 2 came from the original cluster 2 (three Web sites from the original cluster 3 migrated into the new cluster 2; and 14 Web sites from the original cluster 1 migrated into the new cluster 2). The “migration” of Web sites from the original cluster 1 into the new cluster 2 was due to a relatively high degree of similarity between Web sites in cluster 1 and cluster 2 (especially compared to cluster 3), and was not particularly surprising.

A comparison of descriptive statistics (i.e., means, medians, standard deviations) between the original and new cluster 1, the original and new cluster 2, and the original and new cluster 3,
indicated that the new cluster solution was substantively identical to the original. Similarly, the results of the Kruskal-Wallis test (representing the omnibus test of differences among three clusters on five attributes) were the same for original and new cluster solutions. Finally, the results of the Mann-Whitney test (probing the differences between each pair of clusters) were largely the same: 12 of 15 tests produced identical results, whereas the remaining three tests differed marginally in terms of the size of the $p$-value. Therefore, it is safe to conclude that the removal of blogs from the sample did not substantively matter; the results and conclusions described above remained overwhelmingly unchanged. Also, the fact that removing a substantial portion of the sample (30 blogs) did not substantively alter the cluster solution can be viewed as the third validation of the obtained cluster solution (in addition to the split-sample validation and the external-criterion validation described above).

The fourth limitation is a narrow focus on political participation as an ultimate outcome of interest. Although the attributes explicated in this work do have implications for other variables (e.g., political knowledge, social capital) and actually operate through mechanisms involving various other variables, researchers should continue expanding the set of politically-relevant technological attributes. The present study relied very heavily on the literature aimed at identification of independent variables or mediators that facilitate political participation (Bimber, 2003; Eveland & Hively, 2009; Eveland & Scheufele, 2000; Garrett, 2006; McLeod, et al., 1999; Mutz, 2002, 2006; Prior, 2007; Scheufele, 2000; Scheufele, Nisbet, & Brossard, 2003; Scheufele, Nisbet, Brossard, & Nisbet, 2004; Stein, 2009). This literature was instrumental to the conceptualization of the five attributes, and to the formulation of the proposed causal mechanism by which attributes can influence political participation.

Similar work can be done by researchers whose primary interests deal with understanding causal mechanisms that increase or decrease political knowledge, social capital, political efficacy,
political interest, political discussion, and other variables. Although this study used some of these variables, these variables have numerous correlates, antecedents, mediators and moderators that could not be examined here.

The present work illustrates how technological attributes specifically relevant to a certain dependent variable of interest can be explicated. The inductive and deductive methods utilized in this work are transferrable to explication of other attributes. The more technological attributes are explicated, the more systematic our understanding of media and their effects will eventually become, and the more comprehensive, predictive, and therefore, useful theoretical models dealing with media effects will become.

The fifth limitation stems from the nature of today’s Web sites, which are constantly and significantly evolving. Specifically, how valuable will the results of this study be several years from now, when the content and technological configuration of UGC Web sites will be drastically different? How likely is it that the findings of this study will persist over time? Is it important for all of the findings of this study to persist? It is necessary to sort out which conclusions should be expected to remain relevant several years from now, and which ones should not. One set of findings that might not be particularly relevant in the future is the particular distribution of the attributes within each conventionally-used group of UGC Web sites (i.e., blogs, wikis, etc.). If some new groups of UGC Web sites are developed (which is highly likely), or if the existing ones evolve and become characterized by different mixes of attributes, it is likely that the group of Web sites of interest to researchers will have to be content analyzed again. This does not appear to be particularly troublesome if researchers move away from relying on conventional groupings of UGC Web sites and instead focus on examining attributes, or at least focus on groups of Web sites suggested by the cluster analysis performed in this study. Refocusing attention appears to be a sound research strategy because conceptual explication of the attributes and (to lesser extent)
the theoretically- and empirically-derived Web site classification that was developed in this study should remain relatively resilient to technological change, as the MOA approach suggests and as this study argues.

Another finding that appears to be rather resilient to future technological transformation is that the conventional classification of UGC sites and a more theoretically and empirically grounded one, are not interchangeable. Unless each group of Web sites is developed with a specific (mix of) clearly explicated attributes in mind, the “lay” classification and the theoretically and empirically derived classifications will remain different.

Yet another insight that should remain relevant in the future is the conceptualization of the five attributes. Assuming that the traditional news Web sites will continue adopting more and more UGC elements, the value of the technological attributes explicated in this work should not diminish, and might in fact increase as such attributes are adopted by a greater number of popular Web sites. Although some additional content analyses might be needed, the attributes will continue serving as useful and familiar lenses through which researchers can understand media and their effects in the future. The specific technological features making up each attribute might change in the future (i.e., novel ones might appear and existing ones might disappear). However, the conceptualization of the attributes themselves should remain the same, or may undergo minor modifications, if theoretically justified.

The sixth limitation stems from the particular type of classification this study attempted to develop. There are two types of classifications: (1) General-purpose (or natural) classifications that can be used in numerous contexts and that can serve a great variety of scientific purposes; (2) Specific-purpose (or artificial) classifications that are the opposites of the general-purpose classifications, and which are developed with a specific use in mind (Romesburg, 2004). The periodic system of chemical elements and the tree of life classification of organisms are the
classifications of the first type. These classifications organize knowledge about their respective phenomena in a way useful to many research goals. One can argue that both systems represent or capitalize upon respected scientific theories or perhaps even laws (i.e., law of periodicity and theory of evolution). Each of these laws or theories synthesizes a substantial amount of research and philosophy relevant their respective phenomena. Therefore, each provides valuable insights into the nature of the phenomena (e.g., pressure of external factors on an organism to adapt, random variation in genes), making classification systems built upon these laws equally valuable. However, these systems might not be useful for every conceivable purpose. For example, if one wants to classify animals based on whether their heads resemble a triangle or a rectangle, tree of life classification might not be particularly useful. The same can be said of the periodic system of chemical elements: “… there are many forms of the periodic table… Whereas a chemist might favor a form that highlights the reactivity of the elements, an electrical engineer might wish to focus on similarities and patterns in electrical conductivities” (Scerri, 2007, p. 20).

Classification of books by topic in a library is an example of the second type. Unlike the previous classification systems, which were based on respected scientific laws or theories, the topical book classification communicates nothing profound or universally useful about the nature of the books, other than what the topic of the book is.

In the present study, an attempt was made to develop a specific-purpose classification that is useful for research on political participation. What this means is that the developed classification might have little-to-no relevance to researchers studying concepts other than political participation. Due to lack of MOA studies in communication literature, there is a lack of attributes that can be used to formulate laws or theories of communication that would have a wide applicability. Therefore, currently it appears impossible to develop a general-purpose classification of media.
However, as will be described below, the attributes explicated in this work might have implications for a variety of political and non-political communication research areas. Indeed, when the attributes have been conceptualized, an effort was made to focus on a communication phenomenon (i.e., UGC) and propose theoretically useful dimensions (i.e., attributes) of this phenomenon. Only after identification of such attributes these attributes were examined to determine whether they might be related to political participation. Such dual purpose – identifying attributes that describe the nature of UGC phenomenon and that have implications for political participation – creates a potential that these attributes might be relevant beyond the narrow purpose of predicting effects on political participation.

Implications

This section provides a concise summary of the key takeaways of this study. Some of the points discussed below are based mostly on the results of this study, whereas others are largely based on informed speculation and need further empirical validation.

Low attribute presence. This study showed that all five technological attributes are present (although they have relatively low presence) both on traditional news Web sites and on UGC Web sites. Participation facilitation has the lowest presence. Such low presence is likely due to either (a) the novelty of the attributes that have only recently began to diffuse throughout the universe of the Web sites, or (b) the method of coding presence of attribute features, or (c) both. The five attributes appear to be able to serve as useful “lenses” for understanding the nature of media and their effects, and should continue to be examined in the future research.

Web site traffic correlates with attribute presence. There is an association between Web site popularity and the degree to which technologically sophisticated attributes are present on Web sites. Reciprocal causality might underlie this correlation. On the one hand, popular Web sites have financial resources to implement technologically sophisticated functionality. On the
other hand, the presence of sophisticated functionality might make a given Web site more popular. Therefore, incorporating the attributes on one’s Web site might not just facilitate political participation, but might contribute to the Web site’s bottom line by increasing its traffic. This suggests that the five attributes explicated in this study have implications beyond just political participation.

Attribute presence on UGC versus traditional news Web sites. Community orientation, participation facilitation, and manipulability are more extensively present on UGC Web sites, whereas information retrieval and customizability are more extensively present on traditional news Web sites. High information retrieval and customizability scores among traditional news Web sites can be explained by the capacity of such Web sites to change and adopt successful features of UGC Web sites (assuming these attributes were originally more extensively present on UGC Web sites and later became more extensively present on traditional news Web sites). Due to the rapid transformation of Web sites, researchers should focus on technological attributes rather than on loosely defined categorizations of Web sites because attributes might be the only theoretical entities that can remain stable and relevant over time. It appears that even separation of Web sites into UGC versus traditional news Web sites might be no longer adequate because it is likely that the three attributes currently more extensively present on UGC Web sites (i.e., community orientation, participation facilitation, and manipulability) might be soon adopted by the traditional news Web sites.

Quantitative rather than qualitative difference. Traditional news Web sites differ from UGC Web sites quantitatively, rather than qualitatively. Both groups are characterized by the presence of all technological attributes, with the single exception of participation facilitation, which is only present on UGC Web sites. The major difference between various groups of Web
sites is in \textit{how extensively} different attributes are present. This provides empirical support for a major proposition of the MOA framework.

\textit{Conventional and new UGC classification.} Blogs do not differ from wikis in terms of scores on attributes. The same similarity is present between discussion forums and content-sharing sites. In other words, in studies investigating how media affect political participation, political blogs and wikis with political content can be treated as a single group, and discussion forums and content-sharing sites containing political content can be treated as a single group. This finding is useful for researchers studying effects of UGC on political participation because it suggests a novel perspective on blogs, wikis, discussion forums, and content-sharing sites. Conventionally, each of these four groups has been conceptualized as very different and distinct media form. The present research suggests that for the purposes of understanding implications for political participation, it is more appropriate to differentiate between two groups: blogs plus wikis, and content-sharing sites plus discussion forums.

Putting political participation aside, this finding is a vivid illustration of an important point raised in the beginning of this work: today’s media are becoming increasingly similar and many UGC forms share numerous features making it almost impossible to meaningfully differentiate between those forms. Judging against the criteria of the five attributes, blogs and wikis appear indistinguishable, and discussion forums and content-sharing sites appear indistinguishable. If we refer back to the discussion of the goals this study attempted to achieve (i.e., development of a purpose-specific classification), it should be acknowledged that blogs and wikis, and discussion forums and content-sharing sites should not \textit{always} be viewed as indistinguishable. Such conclusion is only warranted in research on political participation that considers five attributes examined here as the most important independent variables. This, of course, presumes that the effects on political participation are driven entirely (or to a significant
degree) by the attributes examined here, and not other factors (e.g., content, uses of technology, perceptions of technology). However, if attributes influence those other factors (e.g., community orientation affects the type of content available on the website), then the above statement is even more valid. How important attributes are, compared to other factors, and whether attributes influence those other factors are important empirical questions. Cluster analysis confirms similarity between blogs and wikis, and between discussion forums and content-sharing sites. Blogs and wikis combined make up more than 64% of Web sites in cluster 2, whereas discussion forums and content-sharing Web sites make up more than 80% of Web sites in cluster 3.

In other contexts, these four Web site groups can be potentially very different. For example, if one suggests that long articles are less effective than short articles in facilitating learning, then blogs and wikis should probably be viewed as very different – presuming that content analysis demonstrates that an average blog post is shorter than a wiki entry. Similarly, if one posits that the degree of social interaction matters for development of a cohesive community, then it might be necessary to view content-sharing sites and discussion forums as different – again presuming content analysis demonstrates that there is more social interaction on discussion forums versus content-sharing sites. This directly relates to the above discussion on the difference between general- versus specific-purpose classifications. Various specific-purpose classifications that use different variables for classifying Web sites will likely classify the same set of Web sites very differently. Only if one general-purpose classification is used across all contexts (which is highly unlikely) can specific clusters of Web sites be treated as universally useful.

One can wonder: Why are blogs similar to wikis, and why are discussion forums similar to content-sharing sites? In regards to the first pair, it appears that blogs and wikis are similar primarily because attributes generally have very low scores on both groups (the only exception is participation facilitation). In regards to the second pair, it appears that their similarity is largely
driven by uniformly high attribute scores (again, with the exception of participation facilitation). Factors that can potentially explain why the first pair has low attribute scores and second pair has high attribute scores were extensively addressed above. However, it does not appear that Web sites in each pair were intentionally designed to be similar on the five attributes explicated here. Instead, it appears that such similarity is a byproduct of various different groups of factors (e.g., volume and diversity of information available on a site, who the site users are, usability decisions by site administrators).

Cluster analysis suggests that the conventional classification of UGC Web sites (e.g., blogs, wikis, discussion forums) is a poor proxy for classification obtained by examining actual attribute scores. Instead of using imprecise conventional groupings, UGC Web sites should be thought of as belonging to three different groups or clusters: (1) High information retrieval and manipulability with participatory journalism sites being the largest group. (2) Consistently low scores on all attributes – except participation facilitation – with blogs and wikis being the largest groups. (3) Consistently high scores on all attributes – except participation facilitation – and primarily composed of content-sharing sites and discussion forums.

What does such a new classification mean for researchers? First, it was theorized that information retrieval and manipulability should each independently increase the depth of information processing. Because cluster 1 scored relatively high on both attributes, it is suggested that processing information on the sites belonging to this cluster should be as deep as for cluster 3 (which also scored high on both attributes), but deeper than cluster 2 Web sites (which scored low on both attributes). As was suggested in the attribute descriptions, deep information processing should increase political knowledge, which should increase political efficacy, and ultimately lead to greater political participation.
Second, cluster 2 Web sites scored higher on participation facilitation, but lower on all other attributes. This might suggest that Web sites belonging to this group might encourage political participation without affecting any other variables (e.g., political knowledge, attitude extremity, social capital, political efficacy). However, despite the fact that cluster 2 scored higher on participation facilitation compared to other clusters, its scores on participation facilitation were still very low (mean = 0.28, median = 0, and mode = 0). Therefore, predictions about the direct effects of cluster 2 on political participation should be considered tentative. Moreover, due to their low scores, it is quite reasonable to expect that use of Web sites from cluster 2 might produce none of the previously described effects.

Third, cluster 3 Web sites had relatively high scores on all attributes. This suggests that use of Web sites belonging to this cluster should lead to a full spectrum of effects proposed earlier: (1) increase in information processing depth; (2) increase in selective exposure; (3) increase in political efficacy; (4) increase in the size of one’s social network; (5) increase in social capital; (6) increase in frequency of interpersonal discussion; and most importantly, (7) direct and indirect increase in political participation.

It should be acknowledged that only future survey and experimental research can test these predictions. Therefore, only future research can indicate how useful the new classification of UGC Web sites is.

Theoretical value of explicated attributes. Finally, and perhaps most importantly, the theoretical portion of this work provides novel and useful theorizing about technological attributes of media (not just UGC). Such attributes can be understood as lenses for understanding media and their effects. In this work, the major focus was on understanding how attributes might facilitate political participation. Several causal mechanisms connecting each attribute with political participation were detailed – each of which constitutes novel theorizing about important
media phenomenon (i.e., UGC) and normatively important outcome (i.e., political participation). However, these attributes might have implications beyond the narrow domain of research focusing on political participation.

First, information retrieval has implications for marketing and public relations research. Having found that this attribute is most extensively present on popular websites, it is logical to ask: What types of content gets emailed or shared the most, gets rated highest, gets more comments or clicks? Investigating these questions has substantial practical and theoretical importance.

Second, the customizability attribute was theorized to encourage greater selective exposure and attitudinal extremity. If the selective exposure mechanisms outlined above are correct, can this attribute undermine deliberation (Mutz, 2006)? Can it reduce exposure to alternative views and make people intolerant or less civil?

Customizability also has implications for research on agenda setting and framing (McCombs, 1972). Greater control over what information to let in or filter out of one’s information environment greatly diminishes media’s ability to influence “what people should think about” (agenda setting) and influence “how to think about it” (framing). The ability to conveniently select preferred sources and preferred framing of issues should undermine the power of mainstream media to set the agenda and determine how important political issues are framed.

Customizability also has implications for gatekeeping theory. Today, traditionally powerful news organizations have substantially less gatekeeping power than they had in the past (Williams & Delli Carpini, 2004). They share their gatekeeping power with audiences who are able to customize information flow. While traditional news organizations might still provide the bulk of information for audience consumption (Dylko et al., 2009), the audience can create their own gatekeeping structure to let only very limited information from traditional and nontraditional
news sources into their information environment. Thus, customizability can be viewed as a unique mechanism through which traditional news organization lose their gatekeeping power.

Third, manipulability might have implications for spiral of silence research (Noelle-Neumann, 1974). People can express minority views without fear of socially isolating themselves, when they express themselves openly in homogenous-opinion (safe, others agree with them) communities, or when they express themselves anonymously in homogenous-opinion (dangerous, others disagree with them) communities.

Manipulability might also have implications for uses and gratifications research. An interesting question to examine is: What motivates users to contribute content of various forms? Existing research on motivation for blogging can be a useful starting place (Nardi et al., 2004; Pew, 2006). If UGC is growing in popularity and importance, content contribution motivations need to be understood as much as content consumption motivations.

Manipulability also has implications for research on news content, especially research on horse-race versus issue-oriented content. If we assume that people who contribute content are (a) interested in the topic and (b) knowledgeable about the topic (Hindman, 2009), can we expect more substantive issue-oriented (as opposed to horse-race-oriented) content on websites with extensive manipulability implementation compared to traditional news media?

Fourth, community orientation has implications for social information processing theory (Walther, 1992). An interesting question is: Can reputational information speed up the process of people getting to know one another in CMC context? Online reputational info about strangers does not appear to have a counterpart in the off-line world (i.e., there is no ready-to-use reputational information available about strangers in a real world, unless one knows somebody who knows that “stranger”).
Community orientation also has implications for research on opinion leadership (Weimann, 1994). Reputational information can be easily used to quickly identify opinion leaders within various communities. Such components of opinion leadership as knowledge of the subject, social connectedness, willingness to share information with one’s social network, psychological gregariousness are well represented by reputational information. Such information can be used very efficiently to identify and target opinion leaders in communication campaigns, which has strong implications for marketing communication, health communication, political communication, and other fields.

Community orientation also might have implications for the Elaboration Likelihood Model (Cacioppo, Petty, Kao & Rodriguez, 1986). It appears that the reputational information might serve as peripheral cues making messages more persuasive when processed via peripheral route. This information also appears to partially represent core components of a person’s credibility (i.e., expertise, trustworthiness, goodwill).

Finally, community orientation also might have implications for deliberation theory (Fishkin, 1997). Some relevant questions are: Does community orientation contribute to improving the quality of political discussion? Does it increase respect and reciprocity among discussants? Many of these questions can be investigated with experimental or survey data. These questions appear to indicate that the attributes explicited in this work have a substantial theoretical value.

Future Research

This section describes several studies of the greatest immediate theoretical and practical importance. The next most logical step is to experimentally test the theorized effects of the technological attributes. This program of research should begin by establishing the “main effects” of each individual attribute. In these experiments, each attribute should be operationalized as a
Web site focused on political content. In order to investigate possible curvilinear relationships, each experiment should have three groups: (1) attribute entirely absent, (2) attribute moderately present (the level of attribute presence would equal to median number of features for each attribute, as documented by the present study), and (3) attribute extensively present. The first group would serve as a control. The second group will serve as a proxy for many of the actual Web sites examined in the present study. The third group will include every theoretically necessary and technologically feasible feature representing a given attribute. The list of features for each attribute will likely expand, as more and more technological affordances become available.

The study participants, most likely college students, will be told that they were invited to evaluate a prototype of a yet-to-be launched political news Web site. Participants will be instructed to avoid reading, watching, or discussing political news for the entire duration of the study (two days). To replicate the real-world news Web site usage, participants will be asked to use the Web site several times. Such Web sites will likely not closely resemble any of the familiar formats (e.g., blogs, content-sharing sites), but will rather be representations of the three levels of specific attribute. Participants will come to the lab twice: During the first session, participants will be given a 10-minute tutorial describing functionality of the Web site. The tutorial will help make participants familiar with the Web site to replicate a real-world situation in which users get to know various aspects of a Web site across some extended period of time. Because a longitudinal design will not be feasible, the 10-minute tutorial appears to be the best possible approximation of the real-world Web site familiarization process. After the tutorial participants will be asked to use the Web site normally for the remainder of the first experimental session (35 minutes).
During the second session, which is to take place the next day, participants will browse Web sites for 35 minutes (the content will be the same as during the first session), then go through a 5-minute procedure that will help measure the various dependent variables of interest, and spend the remaining 5 minutes answering several survey questions. Whereas it is conceptually a rather straightforward procedure to design information systems with identical content and three levels of attributes implementation (given availability of financial resources obtained most likely via a research grant), it is far more difficult to properly operationalize dependent variables, unless, of course, one is using “intention-to-behave” self-report measures.

For each attribute, the first variable in the theorized mediation process is the dependent variable for the purposes of the experiment. In the first experiment, which investigates the effects of information retrieval, the dependent variable is the depth of political information processing. In the second experiment, which investigates the effects of customizability, the dependent variable is selective exposure to attitude-congruent content. In the third experiment, which investigates the effects of manipulability, the dependent variables are: (a) the depth of political information processing and (b) internal political efficacy. In the fourth experiment, which investigates the effects of participation facilitation, the dependent variable is online political participation. In the fifth experiment, which investigates the effects of community orientation, the dependent variables are: (a) the size of one’s social network, (b) social capital, and (c) the frequency of interpersonal discussion. The exact methods of operationalizing these seven dependent variables are yet-to-be determined. It should also be acknowledged that operationalizing several variables (i.e., size of one’s social network, social capital, and frequency of interpersonal discussion) might particularly hard in the context of a non-longitudinal study. An effort will be made to make sure these are observational measures, where participants’ actual online behavior can be tracked and measured. Appropriate corroborating self-report measures (e.g., frequency of discussion questions, standard
items measuring internal political efficacy) will be obtained from the questionnaire participants would fill out during the second experimental session.

If the proposed effects are found, researchers should move to examining effects of important attribute mixes. As it is impossible (and unnecessary) to find associations of every mathematically possible interaction combination of different variables contained in a survey dataset, so it is equally not recommended to “fish” for effects of every mathematically possible combination of the five attributes. Only \textit{a priori} interesting and important mixes of attributes should be examined. What might such mixes be?

First, it is important to investigate interactive effects of attributes that are currently most extensively present across various Web sites. Focusing on such attributes is justified because of the practical implications the findings might have for understanding the effects of today’s most popular UGC and traditional news Web sites. An examination of Table 2B suggests that information retrieval and manipulability are two such attributes. Second, it is theoretically important to examine the interaction among information retrieval, customizability, and manipulability. High information retrieval and manipulability were each theorized to independently stimulate deep information processing. Customizability was theorized to increase selective exposure (which in turn was theorized to increase attitude extremity). If the theorizing about the main effects is correct, an attribute mix where all three attributes are high might produce a situation in which one’s attitudinal extremity is amplified by confidence stemming from one’s having deeply processed political information. The result might be an even greater attitudinal extremity. Theoretically, such a possibility is important to examine because when citizens have extreme political attitudes and when such attitudes are strong (based on the confidence produced by having deeply processed information) they are likely to be very politically active and, at the same time, less tolerant of perspectives they disagree with. Therefore,
the emergence of such a citizenry might be a welcome development from the participatory
democracy perspective (Mutz, 2006), while being far less than ideal from the deliberative
democracy perspective (Delli Carpini & Keeter, 1996; Fishkin, 1997; Sunstein, 2002).

After concluding the examination of the “technological piece” of the larger media-effects
puzzle, researchers should add human (e.g., psychological and sociological) factors. It is the
position of this author that only after having examined how technology interacts with human
factors can researchers claim to have developed a comprehensive understanding of how a certain
medium affects an individual. The influence of social and human factors is undeniable, and future
research focusing on those factors will greatly complement the current work.

One important area to examine is how technology is actually used. Humans are known
for their creative, idiosyncratic and unpredictable use of technology (DeSanctis & Poole, 1994;
Fulk & Boyd, 1991), and such use can amplify or decrease the effects of the technological
attributes described in this work (Bachen, Raphael, Lynn, McKee, & Philippi, 2008). For
example, Facebook is an incredibly complex package of technological features that allows its
users to employ thousands of different configurations of those features. The same can be said
about YouTube, or any number of other popular Web sites. Therefore, use of community
orientation attribute features on these Web sites might encourage development of social capital,
whereas not using such features might have no impact on social capital (Valenzuela, Park, & Kee,
2009). Similar logic applies to all other attributes.

An interesting question, which touches upon the notions of technological versus social
determinism, is whether or not making certain attributes extensively present increases the
frequency with which features representing these attributes are actually used. If the degree of
attribute implementation causes more extensive use of the attribute, this finding would appear to
lend support to the technological deterministic perspective. If, however, there is no relationship
between extensiveness of attribute implementation and intensiveness or frequency of the attribute usage, this would support the social deterministic view.

In addition to addressing the technology-versus-social-determinism debate, investigating this question is of immediate interest to this study. For example, in the above theorizing on the potential effects of attributes, it was assumed that the act of making different technological features available to users should increase the likelihood of the features being used. However, as was discussed above, making participation facilitation features available does not guarantee that users will be motivated to use them. For users who are not very politically interested or whose attributes are not extreme, offering a small number of such features appears to be the most effective strategy of maximizing the number of clicks per political facilitation hyperlink.

Similarly, some Web site might allow users to import or add others users to their “friends” list, however, users ultimately choose whether or not they want to do it.

It appears that one important variable that has a direct impact on how technology is used, and thus has an indirect effect on many of the political variables examined in this work, is the profile of the users. If users have advanced technological skills, have income to acquire advanced technological hardware and software, have advanced understanding of the online culture, then they could use the full spectrum of features tapping all five technological attributes examined here, and could be susceptible to the full spectrum of suggested effects of those attributes. However, if some users lack those resources (see Bucy & Newhagen, 2004), then there will be a gap in terms of how some individuals versus other individuals use the attributes, which can create a gap in terms of the effects the technological attributes exert on these users (Chadwick, 2006; Cho & McLeod, 2007; Eveland & Scheufele, 2000).

Lastly, user motivations might moderate the effects produced by the use of different technological attributes. For example, some individuals might be more or less motivated to
selectively expose themselves to different information (Valentino, Banks, Hutchings, & Davis, 2009), which will moderate how the customizability attribute impacts on attitude extremity and political participation. Also, differences in pre-existing political interest and motivation to get politically involved might moderate the impact of the participation facilitation attribute on political participation.

The last future research direction is perhaps the most important and the most challenging – the explication of what a media “attribute” is and what a “medium” is. A conventional method of differentiating one medium from another is similar to conventional method of investigating media effects criticized above: It relies on a crude and imprecise grouping of increasingly similar information outlets – a method that was quite useful in the past, but is much less appropriate today. In much of this research, different media appear to be defined and differentiated by the mode (or modality, or channel) of information delivery (audio, textual, video, multimedia channel). For example Beck (1991) examined the role of television and newspapers in presidential election; Chaffee and Kanihan (1997) examined the potential of television and newspaper exposure to increase political knowledge; Tewksbury and Althaus (2000) studied the effects of online newspaper and print newspaper on political knowledge; Eveland and Scheufele (2000) examined impact of newspapers and television on political knowledge and participation gaps; Hollander (1997) examined effects of exposure to newspapers, television, and radio on political participation.

As suggested by the discussion of convergence and hybridization in the beginning of this work, many of today’s Internet-based information outlets combine all three forms of information delivery (Briggs & Burke, 2002). Nonetheless, the traditional approach still has some degree of usefulness as long as there are print-only, audio-only, or other uni-modality information outlets. However, the number of such outlets is likely declining rapidly, as the Internet is becoming more
affordable and widespread (OECD, 2011) and as more and more of these uni-modality
information outlets are moving online and begin using several modalities. Therefore, the
traditional method of separating one medium from another is applicable to a rather small fraction
of media outlets. If modality is not particularly useful for differentiating or defining numerous
popular media forms, a new set of criteria needs to be developed.

The MOA framework suggests that researchers should conceptualize media as mixes of
technological attributes. In one sense, this proposition is similar to the conventional approach of
how media are differentiated: Reliance on attribute(s). The conventional approach essentially uses
the modality attribute (mode of information delivery, such as video, audio, text, etc.) as the only
definitive attribute that can and should be used to differentiate media. What was called modality
above, was termed channel in the original articulation of the MOA approach (Eveland, 2003).
The MOA framework suggests a more sophisticated, but structurally similar approach: We should
use not just the channel (or modality) attribute, but many technological attributes to define or
describe a given medium. Such a prescription is useful for understanding the nature and effects of
a given medium, as has been argued throughout this work.

However, such a prescription has conceptual problems that need to be resolved. For
example, if we accept the proposition of the MOA approach that (a) what defines any given
medium is essentially a mix of technological attributes, then we should accept that (b) every
different attribute mix represents a different medium. If we accept “b” then almost every Web site
examined in this study should be considered a unique medium, since almost each one had a
unique mix of attributes (set of scores on the five attributes examined here). This is, obviously,
problematic. On a more abstract level, it can be argued that there are numerous technological
attributes characterizing each information outlet – some already explicated and most yet-to-be
explicated. This study explicated five attributes and Eveland (2003) proposed another six in the
original formulation of the MOA framework. If it is true that there are hundreds or even thousands of (technological) attributes, and that each attribute can vary from low to high, then one can argue that hypothetically every single information outlet or Web site can be regarded as its own medium because its combination of attributes is likely to be unique. Therefore, the MOA framework, although rich in the level of detail, analytically attractive, and significantly superior to the conventional method of defining and describing media, might not be parsimonious enough to be practically or theoretically useful for the specific task of defining what a medium is.

One way to make the MOA useful for this particular task is to make it more similar to a well-known multi-level biological classification system. In this system organisms (the lowest-level units of analysis) are organized in a hierarchical order, starting from small homogenous groups, called Species. Different Species are then combined into a larger group, called Genus; different Genus groups are then combined into an even larger group, called Family, and so forth. There are a total of eight levels of groups, with the largest set of three being called Domains (Campbell, Reese, & Simon, 2007; Mayr, 1968; Taxonomy, 2011).

For the purposes of classifying media, the lowest-level unit of analysis can be an information outlet - a source from which information can be obtained. Adoption of the multi-level classification systems inherently acknowledges that there are multiple legitimate levels of for examining information outlets: From micro, where information outlets are highly homogenous (e.g., each of the three clusters of websites described above is such a homogenous group), to macro, where there is a great degree of heterogeneity among information outlets (e.g., all UGC websites). For any particular task, an appropriate level of conceptualization can be used. The term “media,” with all of its existing imprecise connotations might be less useful than using different hierarchical levels of grouping information outlets.
Another question is: How can such information outlets be grouped into meaningful groups? The evolutionary classification system in biology uses several attributes to place organisms into similar groups: (1) physical appearance, (2) behavior, (3) ecology, (4) genetics, and (5) proximity in the evolutionary development tree (e.g., if organisms A and B are direct descendants of organism Z, the proximity between A and B is very high) (Avise & Johns, 1999; Mayr, 1968; Taxonomy, 2011). The fifth attribute appears to be the main organizing principle for classifying organisms, with the remaining attributes serving to validate or reject classification hypothesis based on the fifth attribute and refine the classification initially based on the fifth attribute (Campbell, Reese, & Simon, 2007). The fifth attribute is also considered a crucial theoretical principle making the evolutionary classification system into a coherent theoretical model.

The attributes explicated and used in the present study appear to be hybrids of biological attributes “1” and “2” from the list above. The attributes examined here can often be observed by visually inspecting the Web site structure and content (hence the similarity to attribute “1” from above). For example, the feature allowing users to select topics to be displayed on the user’s homepage is visually apparent representation of the customizability attribute. Yet, often these same attributes characterize the Web site’s structure but are not visually apparent. However, observing “Web site behavior” can allow certain attributes to reveal themselves (hence the similarity to attribute “2” from above). For example, it is impossible to know whether or not a Web site customizes content for the user based on the user’s profile information. Only after performing a search (or reading the Privacy Policy) can this feature representing the customizability attribute become obvious.

The last set of questions is: What is an attribute, for the purposes of classifying and identifying media? What are the criteria for deciding which attributes should be used to describe
and classify media? In biological classification research, biologists use characteristics that exhibit some variability across organisms as attributes (Avise & Johns, 1999; Mayr, 1968, 1981; Taxonomy, 2011). Although the set of attributes can vary from one group of to-be-classified organisms to another (Dayrat, 2005), the most useful attributes for any particular situation appear to be those with the greatest degree of variability. This principle appears transferrable to the classification of information outlets. One should keep in mind, however, that the variability of any given attribute used to classify information outlets might change over time. This is especially likely in the situation in which a novel attribute is diffusing through the information outlets. In the present study several attributes have relatively low scores and rather small variability. This might suggest that these attributes might have just begun to diffuse through the universe of the Web sites and can be expected to become more extensively present and achieve greater variability in the near future.

As the above discussion shows, there are currently more questions than answers about how a “medium” and “attribute” should be conceptualized and operationalized. However, the work in this direction will greatly help systematize knowledge about media and their effects, and help make communication field a more “scientific” discipline.
References


Appendix A: Codebook

1. Coding sessions:
   a. Please code for about 1 hour per coding session, after which take a break. If you are
tired, sleepy, or are having trouble focusing before the hour is up, please take a
break. Some pages may be harder to code than others, so how long you code each
session may vary.
   b. Each day coding is done, these rules and portions of the codebook you will be using
should be reviewed in their entirety before beginning to code.
   c. When you start coding a given site, try to finish coding it on the same day. This helps
remember the site content.

2. Consulting one another: Please do not consult other coders during coding, and only refer to
the codebook and coding instructions when deciding how to code.

3. Your personal experience with sites: Please do not consult your personal memory/experience
with the sites. Sites frequently change, and features that were available when you used the
site may differ from the features that were available when the pages (that need to be coded)
were saved. Refer only to the saved pages.

4. Flow:
   a. Each site’s pages are located in a single folder. One document is titled “homepage”.
You should begin with the document “homepage” and then proceed to the other
documents, examining each very thoroughly.
   b. Examine if the first attribute’s features are present on the first page of the site, then if
the second attribute’s features are present on the first page of the site, and so on
before you finish examining if all attributes are present on the first page of the site.
Then move on to the second page of the site, and so forth.

5. Codesheet. Use one codesheet per site.

6. Ads. Disregard all explicitly identified ads (this includes description of ad policy on Privacy
Policy page).

7. Codes: For each site, each technological feature should be coded as follows:
   [ 1 ] Feature present on the homepage
      • If you find a feature on the homepage, you should code it “1” and move to
another feature.
      • If you don’t find a feature on the homepage, you should examine all other pages
of the site until you either (a) find the feature somewhere else and code it “2” or
(b) finish examining all the saved site pages and code the feature as “3”.
   [ 2 ] Feature present on the site, but not on the homepage
      • For example, a feature may be located on the first page of one of the site’s
sections, located on a page that could be accessed by clicking on the “advanced
search options” link, located in a box that must be expanded to see the feature.
   [ 3 ] Feature entirely absent from the site
Codes:
[ 1 ] – present on homepage
[ 2 ] – present, but not on homepage

Coder Name __________
Site ID __________

Task-specific information retrieval performance

Important notes:
- Exclude features allowing user to browse/sort through one’s own content.

Site user ability to sort/browse content by:

[ ] (1) creation date (recency, timeliness, upload date, year, etc.)
[ ] (2) modality (only video, audio, photo, and text). If a site has several distinct sections, where one section is dedicated to featuring photos, another to videos, another to textual articles, this feature should be coded as present.
- at least 2 modalities are needed for this feature to be coded as present (videos and photos, etc.)
[ ] (3) topic
- at least 2 topics are needed for this feature to be coded as present (technology and business, etc.)
- includes tag cloud or list of tags/labels
- excludes only “site index,” A-Z index, site map
[ ] (4) document type (PDF, jpeg, word, etc.)
[ ] (5) size or length (1mb, < 10mb, longer than 20 minutes, etc.)
[ ] (6) source/creator/user
- includes sections where you can see all posts/uploads/comments by a single user (usually, in user’s profile section)
- does not have to include a total number of posts/comments/uploads
[ ] (7) number of user-generated comments/replies a piece of content received
- includes such site sections as “most commented” articles/videos, etc.
[ ] (8) tone of user-generated comments/replies (positive, negative, etc.)
[ ] (9) user rating of the content (likes/dislikes, etc.)
[ ] (10) views/clicks/downloads count
- includes such site sections as “most read” or “most viewed” articles/videos/etc.
[ ] (11) number of times the content was shared (emailed/“favorited”/ “tweetted”, etc.).
- includes such site sections as “most emailed”, “most tweetted” content, etc.
Information environment customizability

Site user ability to:

[ ] (1) choose topical domains to be displayed OR ability to rearrange the order of information only on the site’s homepage or user’s profile page (moving “national news” section from the top to the bottom; and moving “tech news” to the top, etc.).

[ ] (2) pick/add “friends” or social network members OR subscribe to “channels,” “groups,” or other users within a site.
   - includes this nonexhaustive list:
     - “following” or “watching” an entry on a Wiki webpage
     - subscribing to discussion forum/message board threads
     - “favoritting” a video on YouTube
     - subscribing to receive emails from various groups/sections existing on a site (“click here to get content by email from I-Love-Obama group”, etc.)
   - excludes only the following:
     - situations where only a list of groups/users is present. Explicit indication that one can subscribe/join the group/users needs to be present.
     - RSS feeds, invitations to “Follow user on Twitter”
     - subscribing to receive emails from site, where the site’s general content is being emailed and where user cannot select which specific sections/groups on the site send content (“click here to subscribe to our email newsletter;” “click here to get our content by email,” etc.)
   - if this feature is coded “3”, features 3-5 should be automatically coded “3”.

[ ] (3) adjust (block/allow, reduce/increase) the amount of incoming communication from particular user(s) of the site, one’s “friends,” social network members, or organizations/groups one is a member of on the site.
   - includes ability to add site users to “ignore list” or ability to block any communication from site users.

Software code, utilized by the site, that:

[ ] (4) offers explicit recommendations of potential “friends” or network members (that user can subscribe/connect to) based on the preferences and information on the user’s profile page.
   - includes an option of adding one’s contacts from Gmail, Facebook, etc.

[ ] (5) offers explicit content recommendations based on the content preferences of user’s “friends” or social network members.

[ ] (6) offers explicit content recommendations (excluding recommendations of “friends”) based on the current or past searches OR based on one’s profile information.
   - includes “related” or “suggested” content that is displayed near the content that user is currently examining.
Codes:
[ 1 ] – present on homepage
[ 2 ] – present, but not on homepage

Content manipulability

Site user ability to:

[ ] (1) edit content, directly upload content or submit content for editorial review for possible future publication (make posts, start a discussion thread, post comments, upload video, audio, graphics, start/edit a wiki, etc.).
  - excludes only:
    - “send us a tip” or “tip us” functionality
    - ability to participate in chat rooms

[ ] (2) start a blog on the site.
  - when this feature is present, site generally (but not always) has a section dedicated to hosting user’s blogs.

Site user ability to:

[ ] (3) tag/classify/label or add “keywords” to describe content
  - this refers to the ability of the users of the site, not the creator of a particular piece of content, to tag/classify/label content

[ ] (4) vote on/rate content (this includes voting/rating other’s comments; liking/disliking content)

[ ] (5) apply to become a contributor/editor/poster on the site
  - excludes regular registration, where user automatically becomes a site contributor/editor/poster.
Codes:
[ 1 ] – present on homepage
[ 2 ] – present, but not on homepage

Direct participation facilitation

**Important notes:**
- Generally, you should be absolutely clear about the destination of the link when coding features below. In some cases, it will be necessary to read the entire sentence in which hyperlink is located, to determine the link’s exact destination.
- Examine only textual hyperlinks (they will be portions of text that are either in red color, or underlined, or both)
- Titles of articles/posts, “usernames” of users, and Privacy Policy page can be ignored.

A link (or hyperlink) to:

[ ] (1) an online form for contacting (sending email messages, etc.) to political person/group that users can click on to start writing an email message (see the definition of “political” below)
  - includes: a link that says “email FCC,” “contact your senator,” and other similar links.

[ ] (2) sites where users can make monetary donation in support/opposition of some political person/group
  - excludes donations to the site that you are currently examining

[ ] (3) sites where users can purchase products to show their support/opposition to some political person/group

[ ] (4) a political group (MeetUp or any other meeting-organizing site, etc.) that users can join OR a social network group in support/opposition of some political person/group

[ ] (5) sites where users can meet (in a “virtual” meeting sense) and interact with political persons
  - excludes all blogs

[ ] (6) sites where users can sign a political petition
  - excludes opinion polls where users can register their opinions

[ ] (7) sites where users can sign up to either volunteer or work for political person/group

**Definition of “political”**: politics and political have direct/explicit references only to the following:
- different branches of government (executive, legislative, judicial)
- laws and policies of the government (groups/persons trying to change/promote certain laws, etc.)
- employees of the government (bureaucrats, professionals, elected and unelected officials)
- candidates for public office
- political parties/philosophies/ideologies
Community orientation

Important notes:
- If a folder for some site does not contain pages whose titles begin with “Main User Profile” AND “User #” pages, code the next 6 features as “3” for that site.

Information available to the site user about:

[ ] (1) user’s “friends,” “fans,” “followers,” “subscribers” or other site users connected to him/her
   - excludes one’s own (Main User’s) friends, fans, followers, subscribers
   - includes information about the number and types of friends

[ ] (2) user’s seniority, length of being a site’s member, or join date

[ ] (3) user’s “rank” in the virtual hierarchy within a particular site.
   - includes visually represented hierarchy through "stars," "circles," or "bars".
   - excludes tags as "senior member"; indications whether a user is a site administrator or not; votes/ratings of individual posts/threads

Site user ability to:

[ ] (4) see/find different groups represented/existing on the site
   - includes sorting content by groups
   - if there are no groups represented/existing on the site, code features 29 and 30 as “3” for that site.

[ ] (5) sort/browse the existing groups by various criteria (topic, size, seniority, etc.)

Information available to the site user about:

[ ] (6) the number OR types of group members, followers, other organizations affiliated (formally or informally) with the groups existing on the site.
   - includes list of organizations recommended by the original group, etc.
### Tracking progress on a site

**Site ID**

<table>
<thead>
<tr>
<th>Pages Examined</th>
<th>For Attr. 1</th>
<th>For Attr. 2</th>
<th>For Attr. 3</th>
<th>For Attr. 4</th>
<th>For Attr. 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>about us</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>browsing content</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>content submission guidelines</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FAQs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>groups</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>homepage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>main user profile page</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>main user profile privacy page</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>main user profile settings page</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>main user profile subscriptions page</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>main user registration page</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>privacy policy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sample content</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>search options</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>search result</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>section</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sorting content</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>uploading content</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>user</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>