Behind the myths of citizen participation: Identifying sustainability factors of hyper-local information systems

Claudia López, Universidad Técnica Federico Santa María, Chile
Rosta Farzan, University of Pittsburgh, US
Yu-Ru Lin, University of Pittsburgh, US

Various information systems have emerged to facilitate citizen participation in the life of their communities. However, there is a lack of robust understanding of what enables the sustainability of such systems. This work introduces a framework to identify and analyze various factors that influence the sustainability of so-called "hyper-local information systems." Using longitudinal observations of participation from 35 online neighborhood discussion forums over six years, we analyze the relationship between sustainability and online-offline community characteristics. Our results not only show patterns consistent with previous observations but reveal the dubious influences of member heterogeneity and network structure. Design insights are discussed.

ACM Reference Format:
Claudia López, Rosta Farzan and Yu-Ru Lin, 2016. Sustainability of information systems for citizen participation: Identifying sustainability factors of hyper-local information systems *ACM Trans. Internet Technol.* V, N, Article A (January YYYY), 28 pages.
DOI: 0000001.0000001

1. INTRODUCTION

While there is growing interest in creating participatory information systems for local communities, little is known about what makes these kinds of systems thrive and become sustainable over time. Prior research has argued that certain characteristics of target communities, such as the community's social capital, can influence the survival of participatory information systems [Kavanaugh and Patterson 2001; Williams and Durrance 2008]. Still, evidence about information systems capable of maintaining sustainable participation is scarce, and robust understanding of what and how community characteristics contribute to sustainability of such systems remains an open issue for researchers and developers of these systems.

Information systems that rely on their users to contribute content face sustainability challenges [Resnick et al. 2012; Kraut and Fiore 2014]. Their existence relies on their ability to engage enough users to generate resources, which will in turn attract other users [Kraut et al. 2010]. Heavy reliance on users' participation renders these systems as complex social structures, with their users serving as both resources and source of resource-generation. Within these systems, the concept of sustainability is defined as the ability of these social structures to "continue providing benefits for members over the long term" [Butler 2001].

In this work, we examine so-called *hyper-local information systems* – systems that focus on providing information services to a specific locale. Hyper-local information systems tend to rely heavily on residents to create locally-oriented content [López and Farzan 2015]. Their dependence on small subsets of user population who are connected within physical boundaries brings further challenges in understanding the design and technical needs for supporting sustainability.

Grounded in prior research in social computing, work teams, and volunteer associations, we introduce a new analysis framework to investigate whether sustainability of participatory information systems for local communities is associated with (1) offline aspects of the target communities, and/or (2) collective characteristics of the online interactions that take place in the participatory information systems. The framework serves as a basis to guide a six-year longitudinal observational study of the sustainability of 35 online discussion forums for neighborhoods and districts that are hosted by the E-Democracy platform. These forums serve local communities in the cities of Minneapolis and St. Paul in Minnesota, USA.

Our study involves a mixed research design. First, we used public data about the forums' target urban communities in order to represent their main characteristics. Second, we quantify collective measures of the sustainability, online activity and social structures of the forums from archival data about their members and posts.

The results provide evidence on the importance of considering the demographics of the target communities in design and in studies of participatory information systems. Among demographic factors, specifically, residential instability of the neighborhoods populations is most related to all measures of sustainability. The results also indicate that online forums that have a mix of newcomers and old-timers among their contributors are more sustainable. Drawing from these results, we provide design lessons for information experts and technology developers seeking to study or develop information systems for citizen participation.

The key contributions of this work include:

(1) We introduce a new analysis framework to identify and analyze both online and offline aspects of local communities and their residents that can affect the sustainability of their participatory information systems.
(2) We present a comprehensive, longitudinal empirical analysis that uses the proposed framework to study the collective aspects of the sustainability of long-term online forums across multiple urban communities in the US. Our results reveal interesting factors, including neighborhood demographics and online network structure, that are not systematically examined in prior work.
(3) We further provide a set of evidence-based design guidelines that capitalize on lessons learned in order to better tackle the challenge of sustainability of local information systems. The discussion offers suggestions to new designs and techniques needed for supporting sustainability in similar systems. This work brings new understanding of sustainability factors, and how future technological solutions may potentially affect or support sustainability, which is fundamental to the emerging research on technologies for *connected communities*.

2. MOTIVATION: BRINGING NEIGHBORS TOGETHER THROUGH INFORMATION SYSTEMS

The popularity of social media, big data and open government has motivated a number of technology practitioners to become involved in technological endeavors that target city residents. Many of these technological efforts have been categorized under the term "hyper-local", which aims to emphasize their focus on bounded geographical communities as opposed to a more global worldwide scope [Radcliffe 2012].

Early versions of hyper-local information systems can be found in community networks, which often included a major software component that enabled residents to become both producers and consumers of local information (e.g. [Schuler 1994]). Over time, different technologies have been used to implement participatory media for local communities, such as electronic bulletin boards [Colstad and Lipkin 1975], discussion forums [Rogers et al. 1994; Dahlberg 2001], local email lists or listservs [Carroll and Rosson 1996; Hampton and Wellman 2003] and question-and-answers sites [Beamish

1995]. With the development of the Web, social computing technologies [Button and Partridge 2007; Carroll 2012] were embraced to implement the participatory mechanisms for content creation. This trend can be observed in current systems such as social networking sites for neighborhoods,[1] [2] online spaces for community conversation[3] or local activism,[4] marketplaces to foster offline interactions among neighbors,[5] and applications that integrate locally-relevant content from other social media sites [Hu et al. 2013; Xia et al. 2014].

Prior research has shown that hyper-local information systems can help local communities achieve goals such as enhanced democracy, social capital, and a strong sense of community [O'Neil 2002]. Among them, social capital [Coleman 1988] has been widely investigated. Social capital refers to the value associated with the social relationships among people. It has been defined as the "*resources embedded in a social structure which are accessed and/or mobilized in purposive actions*" [Lin 1999]. To study the impact of hyper-local online communities on social capital, individuals' number of social ties have been used to operationalize the concept of social capital [Simpson 2005; Williams and Durrance 2008]. Studies have shown that participation in these systems can be associated with increases in social capital [Hampton and Wellman 2003; Hampton 2007; Kavanaugh and Patterson 2001]. Furthermore, these systems can create opportunities to develop local social interactions [Hampton 2002] and reduce barriers to collective action and community mobilization [Hampton and Wellman 2003].

Although there is potential for these systems to have a positive social impact on local communities, maintaining sustainable hyper-local information systems has remained difficult in practice. Even though the urban setting provides a rich context for information-oriented technology endeavors [Carroll and Rosson 2013], many hyper-local information systems that rely on user-generated content, such as discussion forums, social networking sites and digital public displays, struggle to survive over time [Carroll 2012]. Indeed, lack of enough participation and content has led several for-profit local community sites to close[6] [7] or become part of bigger multi-purpose systems[8]. Research-funded projects to build systems for local communities have rarely reported what has happened after the research was over. It remains unknown if and how these systems have become sustainable "in the wild" [de Moor 2009].

The challenge of attracting and maintaining a reliable stream of content and contributors is an issue not specific to information systems for local communities. Online systems with global reach struggle to become self-sustainable and many attempts have failed to do so [Resnick et al. 2012; Kraut and Fiore 2014]. However, the challenge is even more acute for hyper-local systems because their potential audience is limited to people who live in or are visiting a specific place. Furthermore, creating enough relevant local content to attract this audience is not trivial. The frequency of new local information is generally low (few new items per day) and the community interaction that discusses this information online is often insufficient [Carroll 2012].

There is scarce evidence about how hyper-local information systems achieve sustainability. Literature reviews have shown that not all projects have found evidence

[1] https://nextdoor.com/
[2] http://www.neighbortree.com/
[3] http://www.everyblock.com/
[4] https://neighborland.com/
[5] http://neighborgoods.net/
[6] http://www.nearbie.com/
[7] http://www.sharesomesugar.com/
[8] http://nabewise.com/

that community networks met their goals [O'Neil 2002; Simpson 2005; Williams and Durrance 2008]. Much of that research has been heavily dominated by case studies [de Moor 2009], and each case study is very dependent on the particular characteristics of the served community [de Moor 2009]; as a result, it is difficult to compare findings across studies and discover patterns that can explain success or failure in more general terms [Williams and Durrance 2008]. There are a few studies investigating the sustainability of local information systems from the lens of factors related to their content. They concluded that discussion of politics and local problems generated more participation in local discussion forums [Millen and Patterson 2002]; and community poverty level does not relate to the length of discussions and the variety of discussion topics [Gad et al. 2012].

A potential influence of community characteristics on the sustainability of hyper-local information systems has been discussed in prior research [Kavanaugh and Patterson 2001; Williams and Durrance 2008], but has not yet been systematically explored. Moreover, it is commonly assumed that once a technology is made available to a local community, all residents will be equally inclined and able to engage with it. While the concept of the digital divide is widely studied, most studies on participatory information systems refrain from reporting their contribution to addressing digital inequalities within local communities. Both early and recent research on the digital divide has revealed that there have been systematic inequalities in access and use of the Internet across different populations in the US and worldwide [Joiner et al. 2015; Novak 1998; Schradie 2012]. While the access gap has narrowed in the US [Perrin and Duggan 2015], inequalities in Internet use and content production across different platforms remain [Hargittai and Shaw 2015; Joiner et al. 2015; Schradie 2012]. In particular, Internet access both at home and at work, as opposed to at one location or none at all, is significantly related to the production of online content at the individual level [Schradie 2012]. This effect could have an aggregated effect on the sustainability of participatory information systems in a local setting. Nevertheless, current evidence cannot yet confirm this hypothesis.

In this work, we seek to understand how these various factors relate to sustainability of hyper-local participatory information systems.

3. RELATED WORK

Our work is situated at the intersection of research on sustainability of participatory information systems with global reach, small teams and volunteer associations.

3.1. Sustainability of participatory information systems

A rich body of research has been studying sustainability of participatory information systems with global reach such as Usenet, Wikipedia, Facebook and Twitter. This work suggests that sustainability is related to measures at both the collective and individual levels.

Regarding collective aspects, prior research provides evidence that aggregated measures of users' characteristics and the dynamics of their online interactions are correlated to the survival of their online groups. The factors include heterogeneity of users [Chen et al. 2010; Raban et al. 2010], membership overlap across online groups [Wang et al. 2013; Solomon and Wash 2014; Zhu et al. 2014a; Zhu et al. 2014b], number of people who participate [Butler 2001; Butler et al. 2014; Solomon and Wash 2014], and heterogeneity of content that people share online [Butler 2001; Wang and Kraut 2012].

Among the factors that influence individual decisions, the probability of joining an online group was positively related to the number of social ties among current members of the online group and the level of connectivity among them (i.e., embeddedness in the social structure) [Backstrom et al. 2006]. Several other factors are associated

with user retention. These factors can be classified as the level of online response or feedback received by a user [Arguello et al. 2006; Brzozowski et al. 2009; Burke et al. 2009; Joyce and Kraut 2006; Lampe and Johnston 2005; Sarkar et al. 2012], the kinds of content shared in the system [Wang et al. 2012; Wen and Rose 2012], and characteristics of the user's online behavior [Yang et al. 2010].

Additionally, a number of research projects have focused on assessing whether system design has an effect on user behavior, and hence, on the sustainability of the systems. An evidence-based approach for the design of participatory information systems has been undertaken [Kraut et al. 2010]. For example, studies have implemented alternative design decisions to socialize newcomers [Choi et al. 2010; Farzan et al. 2012] and leverage contribution rates [Beenen et al. 2004; Kraut and Resnick 2012] in order to compare their effectiveness regarding specific measures of sustainability.

3.2. Small groups and teams

Research on small groups [Levine and Moreland 2012] has been frequently used as a source of theories and evidence to guide the investigation of the dynamics of group interactions in participatory information systems. The study on work groups has been especially relevant with regard to the way online groups function as a collective and what makes them more effective and productive. The mechanisms that allow teams to work effectively can also help online social structures to work better, and in turn, ensure their sustainability [López et al. 2015]. Team effectiveness has been measured in terms of team performance, satisfying the members' needs, and the members' desire to remain in the team (i.e., viability) [Kozlowski and Ilgen 2006]. These measures are affected both by the team composition (i.e., members, their characteristics and resources) as well as the activities that the team members undertake (i.e., processes). Among the various factors that have been studied as antecedents of team effectiveness, we focus on the social ties among team members. Social ties are a common operationalization of social capital, which is a perspective that has been often used to study local information systems. At a collective level, teams with strong bonds among their members (i.e., group cohesiveness [Hogg 1993]) are associated with high performance [Beal et al. 2003]. Density is also positively and strongly associated with team viability [Balkundi and Harrison 2006]. Additionally, a match between network centrality of formal ties (i.e., leadership) and network centrality of informal ties is strongly and positively related to performance [Balkundi and Harrison 2006].

Within the networks of teams, an individual's network position can also be associated with team outcomes [Burt 1999]. Network brokers facilitate knowledge transfer between sub-groups in organizations [Tortoriello et al. 2012] and more central individuals contribute more to the group than less-central ones [Wasko and Faraj 2005], thus positively affecting team performance.

3.3. Volunteer associations

Research on volunteer associations is also closely related to the study of participatory information systems [Butler 1999]. A volunteer association is a group of people "*relatively freely organized to pursue mutual and personal interests or to achieve common goals, usually non-profit in nature*" [Scott 1957]. Beyond resemblances between this definition and that of online social structures that emerge from participatory information systems, they also present similarities on dimensions such as size of membership, rates of growth and loss of membership, and communication activity [Butler 1999]. Moreover, hyper-local participatory information systems are also comparable in that they target a population within a geographically-defined area.

Volunteer associations compete for members within a locale and vary their composition as a result of the dynamics of this competition [McPherson and Rotolo 1996].

Atypical members and those who are also the target of other associations are more likely to leave [Popielarz and McPherson 1995]. In aggregate levels, town heterogeneity (such as diversity in race and education) has been found to be related to the number of memberships held by residents. Prior research has also found that volunteer organizations are larger in larger cities [Mc Pherson 1983]. Given the analogies between volunteer associations and hyper-local information systems, these factors might also be associated with the composition of members of these online groups.

As will be described in the following sections, our work is positioned at the intersection of previously mentioned research areas. We have borrowed from these existing research methods and research findings to propose an analysis framework to examine the sustainability of participatory information systems for local communities.

4. ANALYSIS FRAMEWORK: SUSTAINABILITY OF HYPER-LOCAL INFORMATION SYSTEMS

Grounded in related research on online communities, work teams and volunteer associations, we propose an analysis framework (see Figure 1) to investigate the sustainability of participatory information systems for local communities. Research on online communities with global reach suggests that their sustainability depends on (1) the availability of resources, and (2) the ability to convert resources into tangible and intangible benefits for the members [Butler 2001]. In the context of small groups, viability is highly relevant to team effectiveness and their performance [Balkundi and Harrison 2006]. In our earlier work [López et al. 2015], we have proposed that these two perspectives can be connected.

In participatory information systems such as those targeting local communities, the main resources are the users that can contribute content to the system. Group sustainability can be operationalized as the ability to attract new members and to retain current members, which in turn ensures resource availability in the online communities. Performance can be measured as the team's ability to take advantage of the available resources to achieve the team's goals of gathering content and influencing users. Hence, by incorporating these two different theoretical perspectives, we characterize sustainability through the following aspects:

(1) *Attraction*: ability to attract new members.
(2) *Retention*: ability to retain existing members.
(3) *Performance*: ability to gather content and generate an impact.

Hyper-local information systems exist in an underlying social context where the target users have already developed dynamics of coexistence and communication that might affect the adoption of another communication media. At the same time, the system design can also affect the way users behave on the system. Thus, both the offline context and the system design are expected to play a role on the sustainability of hyper-local information systems. These factors can be categorized at the collective and individual levels. At the collective level, we include these factors:

— Community aspects: population size and heterogeneity of population.
— Online activity: tenure heterogeneity and membership overlap.
— Social network: size, connectedness, and centralization of the network.
— Content: content distribution and online responsiveness.

With regard to analogous categories, we consider the following factors at the individual level:

— Individual aspects: roles and demographics.
— Network position: centrality and brokering position.
— Online activity: tenure and membership overlap.

— Content: type of content and level of online responsiveness.

Prior research provides evidence about the influence of these factors; however, they have been often assessed in isolation and not all evidence is conclusive. On the collective level, for example, heterogeneity of users was positively related to survival of Internet Relay Channels (IRC) over time [Raban et al. 2010]. However, it had a curvilinear relationship with membership withdrawal in WikiProjects [Chen et al. 2010]. Evidence about the effect of the number of active users is also mixed. Some projects have provided evidence of its positive effect on attraction of new users in subsequent periods of time [Butler 2001; Solomon and Wash 2014], but others have found that the number of active users is not a significant predictor of sustainability of online groups [Raban et al. 2010].

4.1. Proposed framework and research questions

Taking into account prior research, we propose the following analysis framework to guide systematic studies of participatory information systems for local communities. Figure 1 provides an overview of the proposed framework and the specific measures to operationalize each concept in the framework.

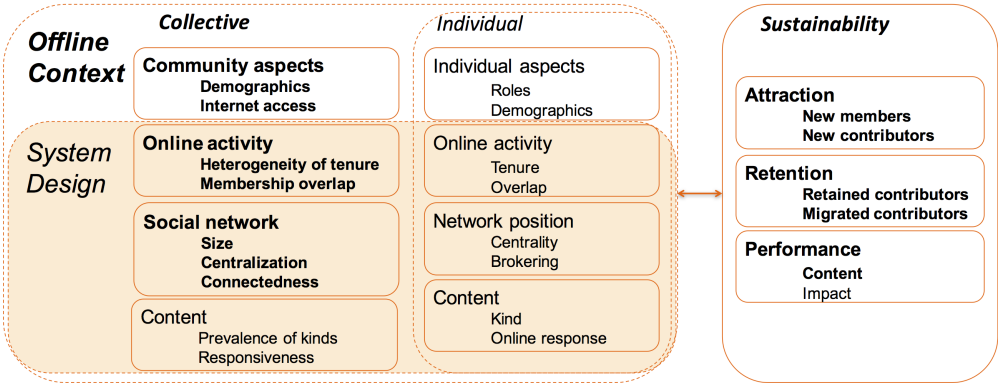

Fig. 1. Measures to analyze sustainability of information systems for citizen participation

While the framework includes similar factors at the individual and collective level, in this article, we focus on assessing how collective factors relate to the sustainability of hyper-local participatory information systems, with exception of those content factors which have been already reported in [López and Farzan 2015]. Accordingly, we have highlighted the new measures to be assessed in this study in bold font in Figure 1 and formulated the following research questions:

(1) What is the relationship between collective characteristics of urban communities and sustainability of hyper-local information systems?
(2) What is the relationship between the collective users' online activity, online network structure, and sustainability of hyper-local information systems?

5. RESEARCH DESIGN, OPERATIONALIZATION, AND MEASUREMENT

5.1. Study platform: E-Democracy.org

We conducted our research on data from the E-Democracy platform. E-Democracy is a non-profit organization that has provided online discussion forums for local communities since 1994. Its mission is "to harness the power of online tools to support par-

ticipation in public life, strengthen communities, and build democracy."[9] The platform was built and maintained by volunteers. Over time, this virtual space initially designed for political discussion during elections became a stable discussion forum that had attracted more than 1,300 residents. With that user base, E-Democracy launched their web version and created an announcement-only email list to allow residents to promote their civic events without the need to follow political discussions.

By 2000, E-Democracy launched online discussion forums for a few cities in the state of Minnesota. The platform adopted a metaphor of an online town hall to describe the dynamics and goals of the forums. These city-wide forums still had a strong focus on political discussion and deliberation about local issues. By 2007, E-Democracy raised funding from multiple sources to launch a new kind of forum that would target neighborhoods in addition to cities. These smaller-scope forums covered daily topics of life in the neighborhoods that ranged from community events to garage sales to crime reports.

Several characteristics of the E-Democracy forums make this platform appropriate for our research purposes. All forums have used the same interface, have followed a similar process of creation, and have enacted similar moderation rules. Creating and maintaining a forum is free. A new forum is created when at least 100 residents sign up to create their local forum.[10] To post to a forum, users need to register in the platform and subscribe to the particular forum. Although registering and explicitly joining a forum are required to post, the content shared in the forums is public and available to unregistered users as well. Users are requested to provide their real full name when registering as a way to encourage trust among users (i.e., neighbors). Every forum has a volunteer forum manager who acts as a moderator when needed, but also encourages participation in several other ways. At the same time, all forums constrain the number of daily posts that a user can add in order to avoid a few members dominating the forum activity [Dahlberg 2001]. The daily maximum number of posts per user varies from two to six across different forums.

E-Democracy currently hosts more than 40 online forums across three countries. Both city-wide and neighborhood forums co-exist in the platform. With the goal to share their knowledge and lessons learned about connecting communities through technology, E-Democracy has provided us their archival data of its local forums.

The rest of this section explains how the factors in the proposed analysis framework were operationalized in variables and the methods used for data collection. Table I and II sum up the main variables of this study.

5.2. Measuring sustainability

Our dataset includes all the posts that were exchanged in 35 E-Democracy online forums for neighborhoods or districts in the cities of Minneapolis and St. Paul in Minnesota. We segmented the data into calendar quarters as the observation period for our longitudinal analysis. Our panel dataset of posts included data from the first quarter of each forum until the second quarter of 2014. Given that not all forums have been active since the same year, our panel dataset was unbalanced. The oldest forums in our dataset were created in the first quarter of 2008 and the newest forums were initiated in the second quarter of 2013. Thus, the tenure of the sampled forums ranged from one to six years. Additionally, we had the membership data of the forums. However, this data was only available since last quarter of 2010, not for the whole lifecycle of all forums. Overall, we had 481 observations in our longitudinal data of posts and 402

[9]http://forums.e-democracy.org/about/
[10]http://blog.e-democracy.org/posts/280

observations in the membership data. For each quarter in a forum, we computed the following measures to represent attraction, retention and performance:

Table I. Independent variables

(1)	Community aspects
	(a) Size
	(b) Diversity
	(c) Instability
	(d) Full access to Internet
(2)	Online activity
	(a) Membership overlap
	(b) Heterogeneity of tenure
(3)	Social networks
	(a) Size
	(b) Connectedness
	(c) Centralization

Table II. Dependent variables

(1)	Attraction
	(a) Number of new members
	(b) Number of new contributors
	(c) Proportion of new contributors
(2)	Retention
	(a) Number of retained members
	(b) Proportion of retained contributors
	(c) Proportion of migrated contributors
(3)	Performance
	(a) Posts
	(b) Productivity
	(c) Productivity change

— *Attraction* is measured using three variables: (a) number of new users who joined the forum in the quarter; (b) number of users who posted for the first time to the forum (i.e., new contributors); and (c) proportion of new contributors in the quarter to the total number of contributors in the previous quarter.
— *Retention* comprises three measures: (a) number of contributors from the previous quarter who continued to post in the current quarter; (b) ratio of retained contributors (those who kept contributing from the last quarter) to the number of contributors in the prior quarter; and (c) proportion of users who migrated (i.e., stopped contributing to the original forum and contributed to another one) to the total number of contributors in the previous quarter.
— *Performance* is represented by three metrics: (a) number of posts in the quarter; (b) volume of posts divided by the number of contributors in the quarter (i.e., productivity); and (c) productivity change between two consecutive periods of time. The change in productivity is the productivity of the current quarter minus the productivity of the prior quarter. This measure could have positive (increase in productivity) or negative values (decrease in productivity).

While the absolute measures (numbers of new members, new contributors, retained contributors and posts) had right-skewed distributions, most of the other dependent variables that represent relative measures (proportion of new contributors, proportion of retained contributors, productivity and productivity change) showed roughly normal distributions. The exception was the proportion of migrated contributors, which had a right skewed distribution and was log-transformed to more closely match a normal distribution. We used Poisson regressions to model the absolute measures (count variables) and linear regressions to estimate the relative measures. As we have repeated measures of the dependent variables over time, we used *xt commands* in Stata 14 that analyze panel data while controlling for the correlation of measures within subjects.

5.3. Community aspects: public data

The sampled E-Democracy forums target neighborhoods or districts with diverse demographics, according to the data made available by Minnesota Compass.[11] [12] This dataset comprises data collected from the 2010 US Census, the 2009 Local Employment Dynamics data, and the 2005-2009 American Community Survey. The dataset includes a set of 166 variables, each of which can be considered a demographic feature.

To identify the major demographic features, we conducted principal component analysis using Stata 14. The results indicate that 81.15% of the data variance is explained by the first three components. After removing redundant demographic features, the main three components revealed groups of variables that we conceptualized as follows:

— *Size* of the neighborhood includes total population, housing units, number of employed residents, and number of jobs in the area.
— *Diversity* of the neighborhood includes the percentage of the population that is identified as people of color, the percentage of the population that is seventeen or younger, and the percentage of the population that is 25 and over whose education level is less than high school graduate.
— *Instability* represents how transient a neighborhood's population is and contains the percentage of rented occupied households, the percentage of occupied households that moved into the neighborhood in 2005 or later, and the percentage of households with an annual income of less than $35,000.

For each component, we chose the most representative factor to characterize the component in our statistical analyses. Size is represented by the total population of the neighborhood. Diversity is measured as the proportion of the population that is identified as people of color. Instability is characterized by the percentage of residents who had moved into the neighborhood in 2005 or later.

We complemented these demographics features with data from the Minneapolis Community Technology Survey.[13] The Information Technology Department of the city of Minneapolis conducted this annual survey in 2012, 2013 and 2014 to obtain data about the access to and use of computers, mobile technology and Internet among the city residents. The survey was distributed by mail to a random sample of the residents across the eleven communities in the city. We used the raw data of this three-year survey to obtain a measure of Internet access in the E-Democracy forums' target urban communities. We defined the following variable:

— *% with full Internet access* measures the proportion of survey respondents from each urban community who had access to the Internet both at home and at work.

5.4. Online activity: archival data analysis

Using the archival dataset of posts, we measured the characteristics of users' online activity in the forums in terms of the following two measures:

— *Diversity of tenure,* defined as the coefficient of variation in tenure of the users who posted in a forum during a given quarter. For each forum in a given quarter, we computed the tenure of all contributors as the number of minutes since the user's first contribution until the last day of the quarter. Following [Chen et al. 2010], diversity

[11] http://www.mncompass.org/profiles/neighborhoods/minneapolis-saint-paul\#!areas
[12] The fact that the neighborhoods were located in the Twin Cities provides analysis advantage as we could control for demographic variation among neighborhoods and assume that all of our subjects (neighborhoods) shared a similar context (e.g., located in a developed country, with a majority of English speakers).
[13] http://www.minneapolismn.gov/it/inclusion/WCMS1P-118865

of tenure was computed as the ratio of the standard deviation of contributor tenure to the mean of the same variable.

— *% membership overlap,* defined as the ratio of contributors who had posted in two or more forums on the platform at a given quarter. For each contributor to a focal forum, we assessed if the contributor had posted content to another E-Democracy forum in the same quarter. We labeled the contributors who had done so as users with membership overlap. To consider this measure at a collective level, we computed the proportion of contributors with membership overlap to the total number of contributors to the forum in the quarter.

5.5. Social structures among contributors: social network analysis

We modeled a social network of users according to a procedure we had discussed in [López et al. 2015]. We created a bipartite network of contributors and threads to describe the interactions in the online forums. When a user added a message to a thread, a connection between a user node and a thread node was created in the network. When two users participated in the same thread, two user nodes were linked to a single thread node. A projection of this bipartite network over the user nodes illustrates the social network of users within a forum. Based on this projected user-to-user network (undirected and unweighted), the following measures were computed.

— *Network size:* The number of user nodes in a forum's network in the current quarter.

— *Network connectedness:* In order to assess the connectedness of a forum's social network, we first computed individual users' clustering coefficients in the network, and the connectedness was computed as the mean of all users' clustering coefficients in the network. Unlike the network density measure, which is sensitive to the entire size of the network, this measure reflects an average of local connectivity in the social structure; i.e., how embedded each node is in the network. This measure varies from zero to one. Larger scores denote more connected social structures.

— *Network centralization:* To capture the centralization or inequality of a forum's social network, we first measured individual users' centrality in the network by the degree centrality - the number of connections the users have with other users in the network. Then, following [Lin et al. 2014], we measured the level of structural concentration using the Gini coefficient of the degree centrality of the forum users in a quarter. As suggested in [Lin et al. 2014], the Gini coefficient is a measure for identifying preferential patterns in general networks, as opposed to measures such as power-law exponents that can only apply to networks following power-law distribution. This measure captures the inequality of users' centrality in the social structure. This coefficient can have values from zero to one. Larger scores signify more unequal degree distributions, thus representing more centralized social structures.

6. ANALYSIS RESULTS

6.1. Descriptive statistics

The archival dataset of the 35 sampled forums includes 75,374 posts, organized into 32,903 discussion threads and posted by 5,207 unique users. These posts were collected during the entire lifecycles of the forums through the second quarter of 2014.

Sustainability measures. The measures of sustainability vary considerably across the neighborhood forums (see Table III). The average forum attracts 34 new members (range: 1 - 345) and 9 new contributors in a quarter (range: 0-60). On average, about a third of the forum contributors in a quarter are new contributors who had never posted before. From one quarter to the next, a neighborhood forum retains slightly less than half of its contributors. Another 12% of them stop contributing to the forum

and instead contribute to other E-Democracy forums. In a quarter, the forums garner a mean number of 152 posts, with a productivity of 2.8 posts by contributor, and a productivity increase of 0.024.

Table III. Dependent variables that represent sustainability of the neighborhood forums

Dependent variable	Mean	Std. Dev.	Min	Max
# joined (new members)	34.308	39.381	1	345
# new contributors	9.593	11.638	0	60
% new contributors	34.641	23.187	0	100
# retained contributors	23.519	34.932	0	171
% retained contributors	48.932	23.446	0	100
% migrated contributors	12.4805	13.168	0	100
# posts	152.645	215.039	1	1180
Productivity	2.845	1.037	1	9
Productivity change	0.024	1.313	-6.250	6.667

Community aspects. The forums serve local communities with distinct characteristics (see Table IV). Communities' population sizes range from 2,833 to 36,255 inhabitants. The ratio of people of color goes from 10.2% to 86.3%. The proportion of households with new residents (who moved in 2005 or later) varies from 16% to 60.2%. The communities are also heterogeneous in terms of Internet access rates (range: 24.32% - 69.27%).

Table IV. Demographics of the forums' target urban communities

Independent variable	Mean	Std. Dev.	Min	Max
Size: Population	12,412	8.527	2,833	36,255
Diversity: % of people of color	42.89	22.31	10.2	86.3
Instability: % moved in 2005 or later	35.53	10.82	16	60.2
% full Internet access	52.09	12.18	24.32	69.27

Online activity. Among the online measures, there is also a great deal of variation (see Table V). The average forum has considerable tenure diversity among their contributors, with the standard deviation of tenure being larger than the average (coefficient of variation = 1.12). On average, slightly less than half of the contributors to a forum also contribute to other E-Democracy forums.

Table V. Online measures of neighborhood forums by quarter

Independent variable	Mean	Std. Dev.	Min	Max
Div. tenure	1.12	0.34	0.13	2.28
% membership overlap	47.60	24.76	0.00	100.00

Social networks. The size of the social networks of contributors varies from a single user to 284 users in a quarter (see Table VI). Given that the values for centralization and connectedness can go from zero to one by definition, we can say that the average forum has mid-levels of centralization and slightly higher levels of connectedness.

Table VI. Social network measures of neighborhood forums by quarter

Independent variable	Mean	Std. Dev.	Min	Max
Network size	53.40	64.06	1.00	284.00
Network connectedness	0.63	0.21	0.00	1.00
Network centralization	0.49	0.26	0.00	0.97

Data considerations. We had to consider the following data adjustments for an appropriate analysis model. First, we dropped the network size variable due to high levels of multi-collinearity among the network size and the measure of centralization (larger networks were more centralized) and the population of the neighborhoods (the more inhabitants, the larger the network).

The Internet access data was only available for the neighborhoods in Minneapolis. Therefore, we ran a separate analysis to assess the impact of Internet access. We did not use this measure as an independent variable in all analyses in order to be able to use all data from St. Paul in the remaining analyses.

Furthermore, we log-transformed the population variable to adjust its range with the rest of the variables. For relative-measure dependent variables (e.g., proportion of new contributors or proportion of retained users), we constrained our analysis to observation periods with at least four contributors in the prior quarter to avoid high relative measures for too small absolute values.

6.2. Collective factors and sustainability

To examine the impact of the neighborhood forums' offline and online characteristics on their future sustainability, we used longitudinal models to estimate the association between the online factors in a given quarter (time *t*) and the measures of attraction, retention and performance in the following quarter (time *t+1*) while controlling for the effect of the offline characteristics of the target neighborhoods, the forums' time of creation and the forums' tenure at each observation period.

Given that our analyses involve a large number of variables (see Table I), we have organized the presentation of the results into subsections that focus on each aspect of sustainability: attraction, retention, and performance. In each subsection, we describe the most consistent patterns of influence of the independent variables on the specific dependent variable that measures each aspect of sustainability.

Each section includes three tables that are organized according to the three dependent variables. Each table describes the results of the regressions based on the following independent variables: (1) demographic community aspects, (2) all community aspects,[14] (3) demographics and online activity, and (4) demographics and both online activity and social network factors. In each of these regression analyses, we controlled for the time of creation of the forums and their tenure at the period of observation.

6.2.1. Attraction.

Community characteristics. Among the community demographics, size and instability of the neighborhoods are related to the absolute values of attraction (new members and new contributors), but not the relative metric (ratio of new contributors) (see Tables VII, VIII, and IX). Population size is positively related to the number of new members and new contributors; i.e., forums belonging to larger neighborhoods attract more people. The instability of the neighborhoods' populations is negatively associated with the same variables; i.e. forums belonging to neighborhoods with larger proportions of

[14]These regressions had fewer observations because the Minneapolis Community Technology Survey data was only available for neighborhoods in the city of Minneapolis.

new residents attract fewer people. The relationship between population size and absolute measures of attraction is often, but not always, statistically significant. Generally, the relationship becomes statistically insignificant when accounting for variables that represent or might encourage interaction in the forums (e.g., network connectedness, and Internet access). Population diversity was not a significant factor for any of the measures of forum attraction. Adding the measure of Internet access, we found that wider access, at home and at work, in the neighborhoods is significantly related to higher attraction of new members.

Online participation. The next analyses of attraction assess the role of online aspects. As we had longitudinal data, we evaluated how the attraction measures in a quarter relate to online measures in the prior quarter (denoted as L. in the Tables). We controlled for the effect of the same variable as measured in the prior quarter. We made this decision to control for the effect of hidden variables that could affect the decision to join a forum in every quarter, but that we were not able to measure. An analogous approach was taken in the remaining regression analyses.

Regarding online participation, the analyses indicate that diversity of tenure among the forums contributors is positively related to all variables of attraction, but is only consistently significant for the absolute measures of sustainability. Neighborhood forums made up of a combination of old-timers and new contributors are better at attracting new members and encouraging active participation in the subsequent period.

The ratio of members who had participated in other E-Democracy forums has a mixed influence on attraction. While larger membership overlap helps to bring new members to a forum in the next quarter, it has a negative relationship with the relative measure of attraction of new contributors. Membership overlap helps to broaden the audience of the forums, but can harm contribution levels in the long-term.

Network characteristics. The results are less conclusive for the role of connectedness and centralization of a forum's social networks. Connectedness is negatively related to attraction of new members and new contributors; however, it is only statistically significant with regard to attraction of new members. Forums that exhibit more connected social networks engage fewer new members in a subsequent quarter. In turn, even though only the positive relationships between centralization and attraction are statistically significant, the sign (positive or negative) of the association varies across different variables of attraction. As this association is not supported by prior research either, we conclude that further research is needed to confirm it.

Controlling factors. Across all of the reported regressions, we controlled for the forums' time of creation and tenure at the observation period. Consistently, we found that these two variables are significantly associated with attraction. Forums that were created in 2011 or later are less likely to engage new users and new contributors. Newer forums also had smaller proportions of new contributors. Overall, newer forums struggle more than older forums to attract residents. This can be due to the larger number of alternative participatory information systems that have become available for local communities in recent years. Older forums might have had the chance to create an initial user base before there were a large number of competing sites, which in turn helped to attract more people over time. Newer forums might have more trouble attracting such an initial, thriving core of users.

Tenure has a more nuanced relationship with attraction. There is evidence that forums attract larger numbers of new contributors over time, but they engage fewer new members and smaller proportions of new contributors. This behavior might be attributable to the fact that more new members are engaged in early phases of the forums' lifecycle and, over time, more of these members contribute content to the sys-

tems. However, the rate to which members become contributors (transitioning from readers to producers of content) decreases over time.

Table VII. Attraction: Number of new members

	(1) Joined	(2) Joined	(3) Joined	(4) Joined
Log population	2.762***	1.931	2.796*	1.354
% of color	1.001	1.003	0.999	0.997
% moved 2005+	0.977**	0.981*	0.963*	0.963**
% full int. access		1.024*		
L.Joined			0.998***	0.998***
L.Div. user tenure			2.517***	2.666***
L.% overlap			1.021***	1.020***
L.Connectedness				0.645***
L.Centralization				1.288**
Created 2011+	0.514***	0.306***	0.240***	0.319***
F. tenure	0.976***	0.988***	0.970***	0.982***
Observations	435	253	335	305
AIC	11447.8	2874.7	8109.0	6562.2
BIC	11476.3	2902.9	8147.1	6606.9

Exponentiated coefficients
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table VIII. Attraction: Number of new contributors

	(1) # new	(2) # new	(3) # new	(4) # new
Log population	3.782*	4.893	2.585*	1.193
% of color	1.004	1.016	1.001	1.001
% moved 2005+	0.934***	0.927***	0.955**	0.959**
% full int. access		1.037		
L.# new			1.008***	1.008***
L.Div. user tenure			1.596***	1.627***
L.% overlap			1.001	0.998
L.Connectedness				0.886
L.Centralization				0.956
Created 2011+	0.308**	0.139***	0.314***	0.425***
F. tenure	0.999	1.001	1.012***	1.015***
Observations	481	291	375	344
AIC	3548.0	1972.2	2856.2	2600.3
BIC	3577.2	2001.6	2895.4	2646.4

Exponentiated coefficients
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

6.2.2. Retention.

Community characteristics. A neighborhood's population size has a weaker association with retention as compared to attraction. The relationships are often positive but only significant in few cases (See Tables X, XI and XII). Diversity and instability of the neighborhoods' populations are consistently related to absolute and relative measures of contributor retention. Forums belonging to more diverse neighborhoods retain significantly more of their contributors and at larger proportions. On the contrary, forums of neighborhoods with larger proportions of new residents (who moved into the neighborhood in 2005 or later) keep fewer contributors and at smaller rates.

Table IX. Relative attraction: Proportion of new contributors

	(1) % new	(2) % new	(3) % new	(4) % new
Log population	0.156	−0.067	0.069	0.043
% of color	0.053	−0.245	0.068	0.094
% moved 2005+	−0.046	0.293	−0.026	−0.061
% full int. access		−0.021		
L.Div. user tenure			0.238*	0.213
L.% overlap			−0.201**	−0.089
L.Connectedness				0.008
L.Centralization				0.128*
Created 2011+	−0.188	−0.306	−0.250**	−0.211*
F. tenure	−0.498***	−0.565***	−0.358***	−0.443***
Observations	392	235	382	351
R^2	0.137	0.142	0.275	0.277

Standardized beta coefficients
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Similar to the results about attraction, wider full Internet access (both at home and work) in a community is significantly associated with higher absolute and relative measures of retention in its online forum. This confirms the key role of full Internet access in sustaining participation of residents as content providers.

Online participation. There is weak evidence that tenure diversity plays a role in retention. Membership overlap has a more consistent association with retention. Neighborhood forums with large membership overlap retain fewer of their contributors from one quarter to the next. From those who leave these forums, larger proportions do not leave the platform but migrate to other forums. Together with the results of attraction, the analysis reveals that membership overlap helps to attract new members, but also brings the risk of lowering the number of contributors in the long-term.

Network characteristics. The forums' social network has a mild effect on retention. Centralization is not a significant factor in the measures of this aspect of sustainability. Connectedness is only significantly associated with the proportion of people that migrate to other forums in a subsequent quarter. Forums that have more connected social networks have larger proportions of members moving to another forum within the E-Democracy platform. This means that while connectedness might not help to retain users in a focal forum, it positively influences retention within the platform.

Controlling factors. Time variables are related to absolute, but not relative, measures of retention. Newer forums are expected to retain significantly fewer contributors than older forums and, over time, the forums are expected to maintain more contributors from one quarter to the next. However, none of these effects are significant with regard to the ratio of retained contributors.

Overall, the count and relative measures of the retention of contributors are strongly related to several community and time aspects, but have very weak relationships with the measures based on online interaction among users in the forums. Conversely, the proportion of contributors who migrate from one E-Democracy forum to another is not influenced by any of the community aspects. The behavioral measures that describe the online interactions among users are the only factors that show significant associations with this dependent variable.

6.2.3. Performance.

Table X. Retention: Number of retained users

	(1) # retained	(2) # retained	(3) # retained	(4) # retained
Log population	1.909	4.557*	1.416	1.063
% of color	1.016*	1.036***	1.011*	1.009
% moved 2005+	0.932***	0.922***	0.964**	0.962**
% full int. access		1.046**		
L.# retained			1.007***	1.007***
L.Div. user tenure			1.042	1.038
L.% overlap			0.991***	0.991***
L.Connectedness				0.867
L.Centralization				1.217
Created 2011+	0.340***	0.204***	0.433***	0.478***
F. tenure	1.046***	1.043***	1.022***	1.021***
Observations	457	278	342	314
AIC	3757.8	2557.8	2604.0	2491.1
BIC	3786.7	2586.8	2642.3	2536.1

Exponentiated coefficients
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table XI. Relative retention: Proportion of retained contributors

	(1) % retained	(2) % retained	(3) % retained	(4) % retained
Log population	0.001	0.195*	−0.030	−0.086
% of color	0.183*	0.564***	0.172*	0.134
% moved 2005+	−0.267**	−0.542***	−0.187	−0.110
% full int. access		0.326***		
L.Div. user tenure			−0.132	−0.066
L.% overlap			−0.101	−0.116
L.Connectedness				−0.034
L.Centralization				−0.100
Created 2011+	−0.319***	−0.276***	−0.273***	−0.299**
F. tenure	−0.115	−0.088	−0.179**	−0.160*
Observations	373	226	365	336
R^2	0.140	0.265	0.167	0.159

Standardized beta coefficients
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table XII. Relative retention: Proportion of migrated contributors

	(1) Log % migrated	(2) Log % migrated	(3) Log % migrated	(4) Log % migrated
Log population	−0.034	−0.113	0.080	0.108
% of color	0.046	0.061	0.005	0.079
% moved 2005+	0.083	0.070	0.017	−0.101
% full int. access		−0.194		
L.Div. user tenure			−0.227*	−0.189*
L.% overlap			0.392***	0.429***
L.Connectedness				0.101*
L.Centralization				−0.079
Created 2011+	0.084	0.051	0.071	−0.000
F. tenure	0.199***	0.174**	0.051	0.070
Observations	373	226	365	336
R^2	0.017	0.075	0.314	0.371

Standardized beta coefficients
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Community characteristics. Similar to the case of retention, diversity and instability of the neighborhoods' populations have consistently significant relationships with performance. Compared to homogeneous neighborhoods, more diverse neighborhoods are expected to have forums with a higher volume of posts and more positive productivity changes from one quarter to the next. On the contrary, forums belonging to neighborhoods with high population instability struggle with poor performance in generating content. The more newly-arrived residents in a neighborhood, the less content their forums garner in a quarter. More unstable neighborhoods are also associated with less productive forums and less positive changes in productivity between two consecutive quarters. There is weak evidence that a neighborhood's population size has a significant relationship with performance. We also found that the forums serving neighborhoods with wider Internet access achieve higher volume of posts.

Online participation. Higher diversity in tenure and smaller percentages of membership overlap are associated with higher volume of posts in a subsequent calendar quarter. The first finding is aligned to the positive effect of tenure diversity on attracting members and contributors. The last result confirms the negative relationship between membership overlap and sustainability in terms of contribution to the forum.

Network characteristics. More centralized and more connected social networks are expected to generate a smaller number of posts in the next period of observation. The last finding does not comply with prior research that indicates that a more connected social structure performs better than a less connected one in the context of small groups. It is possible that the different scale of group size (teams vs. local online communities) explains this unexpected result. The evidence regarding centralization provides initial data to argue for a negative impact of network centralization on content generation.

Controlling factors. Time, again, plays a significant role. Compared to older forums, newer forums are associated with a lower volume of posts. Altogether, newer forums have more difficulty in regard to attraction and performance. Tenure has a mixed impact. Over time, the forums are likely to garner more posts; however, they are also expected to be less productive and have less positive changes in productivity.

In general, the results indicate that the number of posts is influenced by time as well as offline and online aspects. However, we found no evidence that productivity and changes in productivity are influenced by any online measures that we considered. We conclude that there might be other hidden variables, which we are not measuring, that more reliably determine both productivity and productivity change.

7. SUMMARY OF RESULTS AND DISCUSSION

A summary of the results is presented in Table XVI. The positive and negative signs indicate the kinds of consistent relationships between the dependent (rows) and independent (columns) variables. The colored cells in the table denote significant results.

7.1. Impact of community aspects

Together, the results indicate that community aspects are significant factors on all three dimensions of forum sustainability. Population size is relevant for attracting new people but is not significantly related to the production of content or retention over the long-term. The first result is aligned to prior research that has found that larger communities tend to have larger volunteer associations [Mc Pherson 1983].

Diversity of a neighborhood's population does not significantly relate to attraction, but it is significantly and positively associated with retention and performance. Forums of more diverse neighborhoods sustain more of their contributors from one quar-

Table XIII. Performance: Number of posts

	(1) Posts	(2) Posts	(3) Posts	(4) Posts
Log population	2.347	5.598*	1.812	1.180
% of color	1.018*	1.036***	1.013*	1.012*
% moved 2005+	0.928***	0.915***	0.955***	0.951***
% full int. access		1.049**		
L.Posts			1.001***	1.001***
L.Div. user tenure			1.080***	1.152***
L.% overlap			0.996***	0.996***
L.Connectedness				0.823***
L.Centralization				0.898*
Created 2011+	0.354***	0.198***	0.428***	0.494**
F. tenure	1.039***	1.036***	1.022***	1.023***
Observations	481	291	383	351
AIC	17068.4	11119.6	12681.6	11959.6
BIC	17097.6	11149.0	12721.1	12005.9

Exponentiated coefficients
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table XIV. Relative performance: Productivity

	(1) Productivity	(2) Productivity	(3) Productivity	(4) Productivity
Log population	−0.014	0.200	−0.031	−0.013
% of color	0.446***	0.766***	0.463***	0.489***
% moved 2005+	−0.349**	−0.700***	−0.288*	−0.242
% full int. access		0.259*		
L.Div. user tenure			−0.061	−0.038
L.% overlap			−0.083	−0.049
L.Connectedness				−0.086
L.Centralization				0.106
Created 2011+	0.011	0.066	0.050	0.019
F. tenure	−0.089	−0.113	−0.149	−0.179
Observations	393	236	383	351
R^2	0.138	0.238	0.161	0.201

Standardized beta coefficients
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table XV. Relative performance: Change in Productivity

	(1) Prod. change	(2) Prod. change	(3) Prod. change	(4) Prod. change
Log population	−0.076	−0.054	−0.102	−0.073
% of color	−0.056	−0.069	−0.069	0.010
% moved 2005+	0.007	−0.116	0.001	0.032
% full int. access		−0.080		
L.Div. user tenure			0.180*	0.212**
L.% overlap			0.034	0.066
L.Connectedness				0.211**
L.Centralization				0.039
Created 2011+	0.015	0.083	−0.042	−0.072
F. tenure	−0.138***	−0.136**	−0.065	−0.032
Observations	383	231	383	351
R^2	0.019	0.020	0.027	0.082

Standardized beta coefficients
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

ter to the next (they also do so at higher rates) and they have more active streams of content. Prior research suggests that diversity at the collective and individual levels tends to harm retention in volunteer associations [Rotolo 2000; Popielarz and McPherson 1995; McPherson and Rotolo 1996]. A potential explanation for our unexpected result is that the users of the neighborhoods' forums might not be as diverse as the neighborhood population and hence, the consequences of heterogeneity within the forum do not develop. Our future work aims to explore this issue.

Table XVI. Summary of results: Collective online and offline aspects

	Attraction			Retention			Performance		
	#join	#new	%new	#ret	%ret	%migr	#posts	prod	△prod
Log population	+	+		+			+		-
% of color		+		+	+		+	+	
% moved 2005+	-	-		-	-		-	-	
Created 2011+	-	-	-	-	-	+	-	+	
F. tenure	-		-	+	-	+	+	-	-
% full int. access	+	+		+	+		+	+	
L.Div. user tenure	+	+	+	+	-	-	+	+	+
L.% overlap	+		-	-	-	+	-	-	-
L.Centralization	+	-	+	+	-	-	-	+	+
L.Connectedness	-	-	-	-	-	+	-	+	+

The most influential and consistent demographic factor for sustainability of the E-Democracy forums is residential instability in the neighborhood. Neighborhoods that have larger proportions of residents who had arrived in 2005 or later are expected to have less sustainable forums. They engage fewer new members, fewer new contributors and smaller proportions of new contributors. Forums of more unstable neighborhoods also retain fewer contributors both in absolute and relative numbers. Furthermore, these forums also tend to receive less content, be less productive and have less positive changes in productivity from one quarter to the next. The concept of social capital can offer a feasible explanation for the impact of residential instability on sustainability. Prior research provides evidence that residential stability is highly and positively connected to social capital in geographical communities [Kang and Kwak 2003]. Unstable populations have less social capital, which in turn negatively affects the neighborhood's ability to make their participatory information systems function effectively. This reasoning is aligned to prior research that speculates about the need for social capital as a key antecedent of the success of community networks [Kavanaugh and Patterson 2001; Williams and Durrance 2008]. Although our results cannot support such a hypothesis, they provide additional evidence in favor of such a direction.

7.2. Impact of Internet access

We were also able to assess the impact of the level of Internet access in the neighborhoods on the sustainability of their local information systems. Prior research has documented that Internet access both at home and at work, as opposed to at one location or none at all, is a significant predictor of the production of online content at the individual level [Schradie 2012]. Our results confirm this relationship at the collective level in online forums for local communities. Neighborhoods with higher ratios of residents with Internet access, both at home and at work, maintain forums that not only generate more online content, but also attract more new members and have better retention. It seems obvious that wider Internet access leads to more online participation. However, this relationship might be forgotten in the age of the Internet, mobile technologies and big data, especially in developed countries. While Internet access seems

pervasive, there are still differences in the availability and quality of Internet access. Content production requires digital skills that can be better developed when people have full Internet access. At the neighborhood level, this becomes a critical factor for the sustainability of hyper-local participatory information systems.

7.3. Impact of online activities

Diversity in the tenure of the contributors has a positive effect on attraction and content generation. However, its impact on retention is not statistically significant. The results indicate that a mix of old-timers and newcomers is a healthy combination for local online forums. They make the forums look more attractive for residents to join and they also bring larger numbers of new contributors every quarter. Moreover, tenure diversity helps in achieving larger volumes of content. The positive relationship between diversity of contributors and sustainability is aligned to a prior finding [Wen et al. 2016], which revealed that a lower concentration of users (i.e., a more diverse group) encouraging offline events in an online group was significantly associated to attraction and retention os users.

Membership overlap is more of a double-edged sword for the sustainability of local information systems. While it creates mobility of users within the E-Democracy platform by bringing more new members to the forums and having larger proportions of people migrating to other forums instead of abandoning the platform, larger membership overlap is also related to lower levels of contribution in several ways. Forums with larger proportions of contributors who add content to other E-Democracy forums are likely to retain fewer of their contributors and garner fewer posts in a subsequent period of time. This means that membership overlap increases exposure to the forums, but does not contribute to building sustainable streams of content over the long-term.

7.4. Impact of network structure

After controlling for demographics and online participation, the impact of the network structure of contributors is somewhat dubious. The average level of embeddedness of contributors in a social network has a negative relationship with the attraction of new members and performance in absolute numbers, but a positive association with migration. These results can be explained as more connected social networks can be perceived as more closed groups that seem less welcoming to new members and new, diverse content. Nevertheless, when contributors decide to leave these groups, a larger proportion of them do not leave the platform completely but decide to contribute elsewhere. This can be seen as a positive impact of connectedness on the sustainability of the platform, even though it does not help the sustainability of the focal forum.

Network centralization has a positive impact on attraction of new members and proportion of new contributors, but it is negatively associated with the number of posts that a forum garners. Centralization is a measure of inequality in degree centrality among the contributors; i.e. a representation of central vs. peripheral members. Unlike what has been found in Wiki Projects, where centralization is associated with less growth [Solomon and Wash 2014], our results indicate that forums with apparent central leaders, who are connected to larger number of the members than the others, become more successful in attracting newcomers. A less desirable consequence of centralization is that it harms performance in absolute terms. Centralized networks seem to attract new people as readers, but generate less content as a group. Additional mechanisms need to be developed to deal with this negative impact of centralization.

7.5. Sustainability measures

Our analysis also reveals that the volume of contributors and posts in the E-Democracy forums for local communities is rather low. On average, a neighborhood forum engages

53 contributors and gathers 152 posts in a calendar quarter. The most active neighborhood forum in E-Democracy has attracted 284 contributors and received 1,180 posts in a quarter. This confirms prior research that argues that virtual spaces for local communities hardly generate new content on a daily basis and have difficulty encouraging online discussions [Carroll 2012]. Nevertheless, some of the sampled forums have been active for six years. This suggests that users perceive value in the forums, even though the stream of content is not as active as other social media. This raises questions about the suitability of using volume as a measure of success in local information systems. Although these concerns have been discussed in prior literature [Carroll and Rosson 2001; Resnick 2001], further research is needed to explore alternative measures of performance that might better represent the effectiveness of local information systems.

7.6. Limitations

This work has limitations consistent with any research endeavor. We have used archival data from a sample of forums in a specific state in the US, which might not be representative of forums for local communities in different cultural and social contexts. While we strove to control for neighborhood demographics, we could not control for all neighborhood variables that could have affected the results, such as the average level of civic engagement in the neighborhood. Additionally, we were not able to control for the number of contributors due to multicollinearity among the independent variables. Finally, our forums' social network measures were based on data about who responded to whom and not in an underlying social connections, as in the studies on Facebook and Twitter. This might explain why we were not able to replicate results from prior work regarding social network measures.

8. TOWARDS DESIGN GUIDELINES FOR HYPER-LOCAL INFORMATION SYSTEMS

Based on the results of our empirical study, we offer four design insights that may contribute to the design and study of hyper-local systems.

Big audience myth vs. hyper-local communities. Hyper-local information systems need to engage people as content contributors. However, engaging enough contributors is challenging. Targeting very restricted user audiences, such as residents of a specific neighborhood, can seem formidable. Nevertheless, our data analyses suggest that while population size of a neighborhood is important for attracting new people to neighborhood forums, it is not critical for performance and retention. Our results suggest that online forums for small geographically-bounded communities can maintain streams of content that make them significant for the communities. An active E-Democracy neighborhood forum targets an area of about 11,000 people on average. Thus, we provide evidence here that focusing on hyper-local communities such as neighborhoods can create active-enough streams of content to become sustainable over time. This finding has implications for technology designers. While focusing on neighborhoods might not seem to be a good decision because it can be hard to achieve critical mass, we found that other aspects of the target community might be more critical to sustainability than population size. Therefore, the success of citizen participatory systems, from social networking sites [Masden et al. 2014] to local marketplaces for human sensing data [Cao et al. 2016], lie in new designs and techniques that can pinpoint the optimal boundary of geographically-confined communities.

The challenge of residential instability. Residential instability is a solid threat to sustainability. It negatively affects attraction, retention, and performance. Sharing the same geographical context for longer period of time leads to developing stronger social ties among neighbors. Our results suggest that such social ties in the neighborhood are correlated to sustainability of hyper-local information systems. This finding

is aligned with related work that argues that a level of initial social capital in the community [Kavanaugh and Patterson 2001; Williams and Durrance 2008] is necessary for technology to have a positive impact that ensures sustainability on the long term.

We speculate that another reason behind the importance of this variable can be the forums' strong focus on community involvement. Possibly, community involvement is more appealing to those who have lived longer in, and are more attached to, the neighborhoods. However, not everyone seems to be motivated by the same goal. Satisfying goals of, and engaging, distinctive residents has been found to be a key aspect of larger volunteer associations in urban communities [Babchuk and Booth 1969]. New designs and techniques of urban computing to adapt to the needs of a diverse population of local residents is necessary to design sustainable hyper-local information systems that support such populations.

Mixed effects of membership overlap. Local information systems can focus on different geographical areas. They can target neighborhoods, cities or other small geographically-bounded areas. Residents might want to participate in more than a single forum, such as forums for adjacent neighborhoods or its home city's forum. We found evidence that about a third of the users in E-Democracy participate in more than one local forum. The existence of this membership overlap has divergent influences on sustainability. Collectively, membership overlap increases attraction: more people feel that they can be part of the forums. However, membership overlap negatively influences retention and performance. Thus, membership overlap increases the audience of a focal forum and encourages mobility across forums in the platform, but it harms other aspects of the focal forum's sustainability. New designs and techniques can capitalize on attraction powers of membership overlap and may seek to identify methods to better support the activities of overlapping members to keep them engaged.

Temporal and information access conditions. Our last implication is that context matters considerably when assessing the use of hyper-local information systems for citizen participation. We found that the time a local forum was created is a significant factor for almost all sustainability measures. Generally, more recently created forums were associated with lower sustainability. We speculate that this effect is related to competition with a larger number of social media sites that have become available more recently. Tenure is often significant as well. As time goes by, some aspects of sustainability become better and others worsen. Furthermore, not only are demographics of the target neighborhoods significant, but the level of full Internet access in the local community is also an influential factor for the contribution levels of a local forum. Our results emphasize the importance of considering the local context in the study and design of local information systems that are often overlooked in studies of online systems due to challenges of acquiring them. This calls for new designs and techniques that can distinguish different collective context and access conditions (e.g., from low-bandwidth Internet connection to new Internet-of-Things access) to facilitate seamless communications within and between communities of different conditions, and to more accurately interpret portrayals of city dynamics based on the existence (or lack) of local social media content [Lathia et al. 2012; Silva et al. 2014].

9. CONCLUSIONS

This article proposes an analysis framework to investigate the factors associated with sustainability of participatory information systems for local communities. Based on the proposed framework, we report on a longitudinal study of a sample of 35 long-tenure online forums for neighborhoods and districts in the US. The study explores a set of online and offline collective factors that can influence sustainability in terms of attraction of new users, retention of existing users, and content production perfor-

mance. The results provide evidence that sustainability is indeed correlated to a mix of these factors; therefore, providing support for the elements of our proposed framework as basis of analysis of sustainability of hyper-local information systems.

In particular, our findings provide evidence that the sustainability of hyper-local information forums is significantly associated with the demographics of their target local communities. Among these factors, residential instability of a neighborhood plays the most influential role in sustainability. Discussion forums in the neighborhoods with more transient population are less successful in attracting new users, retaining the current ones, and generating new content. Other demographics measures have less prominent associations with the success of the online forums. Population size is positively related only to attraction of new users, and racial diversity is associated with better retention and performance. Furthermore, our analysis also reveals that online measures have reliable relationships with sustainability of hyper-local forums. For example, membership overlap across different hyper-local forums has a positive relationship with attracting new users and a negative relationship with user retention and content generation during a subsequent period of time. We hope our design guidelines for each of these factors contribute to a more effective design of hyper-local participatory technologies.

Future work can explore other aspects of the analysis framework, such as the relationship between online and offline individual aspects of users and sustainability. Further research can also design experimental studies to test the effectiveness of our design lessons in a controlled setting. Overall, we believe that our contributions can help to systematically develop evidence-based guidelines to design more sustainable hyper-local information systems to encourage citizen participation.

ACKNOWLEDGMENTS

The authors would like to thank E-Democracy.org for providing its data and domain knowledge. The authors are grateful to Dr. Brian Butler and Dr. Peter Brusilovsky for their feedback on conducting this work. The work of the first author was supported by CONICYT Chile, under grant Conicyt-Fondecyt Iniciación #11161026. The second author thanks the support provided by a Google Research Faculty Award. The third author would like to acknowledge the support by NSF Grants #1634944 and #1637067.

REFERENCES

Jaime Arguello, Brian S. Butler, Elisabeth Joyce, Robert Kraut, Kimberly S. Ling, Carolyn Rosé, and Xiaoqing Wang. 2006. Talk to Me: Foundations for Successful Individual-group Interactions in Online Communities. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '06)*. 959–968.

Nicholas Babchuk and Alan Booth. 1969. Voluntary association membership: A longitudinal analysis. *American Sociological Review* (1969), 31–45.

Lars Backstrom, Dan Huttenlocher, Jon Kleinberg, and Xiangyang Lan. 2006. Group formation in large social networks: membership, growth, and evolution. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD '06)*. 44–54.

Prasad Balkundi and David A Harrison. 2006. Ties, leaders, and time in teams: Strong inference about network structure's effects on team viability and performance. *Academy of Management Journal* 49, 1 (2006), 49–68.

Daniel J Beal, Robin R Cohen, Michael J Burke, and Christy L McLendon. 2003. Cohesion and performance in groups: a meta-analytic clarification of construct relations. *Journal of Applied Psychology* 88, 6 (2003), 989.

Anne Beamish. 1995. *Communities on-line: Community-based computer networks*. Master's thesis. Department of Urban Studies and Planning. Massachusetts Institute of Technology.

Gerard Beenen, Kimberly Ling, Xiaoqing Wang, Klarissa Chang, Dan Frankowski, Paul Resnick, and Robert E. Kraut. 2004. Using social psychology to motivate contributions to online communities. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work (CSCW '04)*. 212–221.

Michael J. Brzozowski, Thomas Sandholm, and Tad Hogg. 2009. Effects of Feedback and Peer Pressure on Contributions to Enterprise Social Media. In *Proceedings of the ACM 2009 International Conference on Supporting Group Work (GROUP '09)*. 61–70.

Moira Burke, Cameron Marlow, and Thomas Lento. 2009. Feed Me: Motivating Newcomer Contribution in Social Network Sites. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*. 945–954.

Ronald S Burt. 1999. The social capital of opinion leaders. *The Annals of the American Academy of Political and Social Science* 566, 1 (1999), 37–54.

Brian S. Butler. 1999. *The dynamics of cyberspace: Examining and modeling online social structure*. Ph.D. Dissertation. Carnegie Mellon University. Chapter: When is a group not a group: An empirical examination of metaphors for online social structure.

Brian S. Butler. 2001. Membership Size, Communication Activity, and Sustainability: A Resource-Based Model of Online Social Structures. *Info. Sys. Research* 12, 4 (2001), 346–362.

Brian S Butler, Patrick J Bateman, Peter H Gray, and E Ilana Diamant. 2014. An Attraction-Selection-Attrition Theory of Online Community Size and Resilience. *Mis Quarterly* 38, 3 (2014), 699–728.

Angela Button and Helen Partridge. 2007. Community networks today: Analysing new media for local social networking and community engagement. In *Prato CIRN Conference: Communities and Action*. 5–7.

Tien-Dung Cao, Tran-Vu Pham, Quang-Hieu Vu, Hong-Linh Truong, Duc-Hung Le, and Schahram Dustdar. 2016. MARSA: A Marketplace for Realtime Human Sensing Data. *ACM Trans. Internet Technol.* 16, 3, Article 16 (May 2016), 21 pages. DOI:http://dx.doi.org/10.1145/2883611

John M. Carroll. 2012. *The Neighborhood in the Internet*. Routledge, Chapter What are community networks?, 23–47.

John M. Carroll and Mary Beth Rosson. 1996. Developing the Blacksburg electronic village. *Commun. ACM* 39, 12 (1996), 69–74.

John M Carroll and Mary Beth Rosson. 2001. Better home shopping or new democracy?: evaluating community network outcomes. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '01)*. 372–379.

John M. Carroll and Mary Beth Rosson. 2013. Wild at Home: The Neighborhood as a Living Laboratory for HCI. *ACM Trans. Comput.-Hum. Interact.* 20, 3, Article 16 (2013), 16:1–16:28 pages.

Jilin Chen, Yuqing Ren, and John Riedl. 2010. The effects of diversity on group productivity and member withdrawal in online volunteer groups. In *Proceedings of the 28th International Conference on Human Factors in Computing Systems (CHI '10)*. 821–830.

Boreum Choi, Kira Alexander, Robert E. Kraut, and John M. Levine. 2010. Socialization tactics in wikipedia and their effects. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work (CSCW '10)*. 107–116.

James S Coleman. 1988. Social Capital in the Creation of Human Capital. *Amer. J. Sociology* 94 (1988), 95–120.

Ken Colstad and Efrem Lipkin. 1975. Community memory: a public information network. *SIGCAS Comput. Soc.* 6, 4 (1975), 6–7.

Lincoln Dahlberg. 2001. Extending the public sphere through cyberspace: The case of Minnesota E-Democracy. *First Monday* 6, 3 (2001).

Aldo de Moor. 2009. Moving community informatics research forward. *The Journal of Community Informatics* 5, 1 (2009).

Rosta Farzan, Robert Kraut, Aditya Pal, and Joseph Konstan. 2012. Socializing volunteers in an online community: a field experiment. In *Proceedings of the ACM conference on Computer Supported Cooperative Work*. 325–334.

Samah Gad, Naren Ramakrishnan, Keith N. Hampton, and Andrea Kavanaugh. 2012. Bridging the Divide in Democratic Engagement: Studying Conversation Patterns in Advantaged and Disadvantaged Communities. In *Proceedings of the 2012 International Conference on Social Informatics*. 165–176.

Keith N. Hampton. 2002. Place-Based and IT Mediated "Community". *Planning Theory & Practice* 3, 2 (2002), 228–231.

Keith N. Hampton. 2007. Neighborhoods in the Network Society the e-Neighbors study. *Information, Communication & Society* 10, 5 (2007), 714–748.

Keith N. Hampton and Barry Wellman. 2003. Neighboring in Netville: How the Internet Supports Community and Social Capital in a Wired Suburb. *City & Community* 2, 4 (2003), 277–311.

Eszter Hargittai and Aaron Shaw. 2015. Mind the skills gap: the role of Internet know-how and gender in differentiated contributions to Wikipedia. *Information, Communication & Society* 18, 4 (2015), 424–442.

Michael A Hogg. 1993. Group cohesiveness: A critical review and some new directions. *European review of social psychology* 4, 1 (1993), 85–111.

Yuheng Hu, Shelly D. Farnham, and Andrés Monroy-Hernández. 2013. Whoo.Ly: Facilitating Information Seeking for Hyperlocal Communities Using Social Media. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. 3481–3490.

Richard Joiner, Caroline Stewart, and Chelsey Beaney. 2015. Gender Digital Divide. *The Wiley Blackwell Handbook of Psychology, Technology and Society* (2015), 74.

Elisabeth Joyce and Robert E. Kraut. 2006. Predicting Continued Participation in Newsgroups. *Journal of Computer-Mediated Communication* 11, 3 (2006), 723–747.

Naewon Kang and Nojin Kwak. 2003. A Multilevel Approach to Civic Participation Individual Length of Residence, Neighborhood Residential Stability, and Their Interactive Effects With Media Use. *Communication Research* 30, 1 (2003), 80–106.

Andrea L Kavanaugh and Scott J Patterson. 2001. The Impact of Community Computer Networks on Social Capital and Community Involvement. *American Behavioral Scientist* 45, 3 (2001), 496–509.

Steve WJ Kozlowski and Daniel R Ilgen. 2006. Enhancing the effectiveness of work groups and teams. *Psychological science in the public interest* 7, 3 (2006), 77–124.

Robert E Kraut and Andrew T Fiore. 2014. The role of founders in building online groups. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*. 722–732.

Robert E. Kraut, Mary Lou Maher, Judith Olson, Thomas W Malone, Peter Pirolli, and John C Thomas. 2010. Scientific Foundations: A Case for Technology- Mediated Social- Participation Theory. *Computer* 43, 11 (2010), 22–28.

Robert E Kraut and Paul Resnick. 2012. Encouraging contribution to online communities. In *Building successful online communities: Evidence-based social design*. MIT Press Cambridge, MA, 21–76.

Cliff Lampe and Erik Johnston. 2005. Follow the (Slash) Dot: Effects of Feedback on New Members in an Online Community. In *Proceedings of the 2005 International ACM SIGGROUP Conference on Supporting Group Work (GROUP '05)*. 11–20.

Neal Lathia, Daniele Quercia, and Jon Crowcroft. 2012. The hidden image of the city: sensing community well-being from urban mobility. In *Pervasive computing*. Springer, 91–98.

John M Levine and Richard L Moreland. 2012. A history of small group research. *Handbook of the history of social psychology* (2012), 383.

Nan Lin. 1999. Building a Network Theory of Social Capital. *Connections* 22, 1 (1999), 28–51.

Yu-Ru Lin, Brian Keegan, Drew Margolin, and David Lazer. 2014. Rising tides or rising stars?: Dynamics of shared attention on Twitter during media events. *PloS one* 9, 5 (2014), e94093.

Claudia López and Rosta Farzan. 2015. From Community Networks to Hyper-local Social Media. In *Proceedings of the 18th ACM Conference Companion on Computer Supported Cooperative Work & Social Computing (CSCW'15 Companion)*. 239–242.

Claudia López and Rosta Farzan. 2015. Lend Me Sugar, I Am Your Neighbor!: A Content Analysis of Online Forums for Local Communities. In *Proceedings of the 7th International Conference on Communities and Technologies (C&T '15)*. 59–67.

Claudia López, Yu-Ru Lin, and Rosta Farzan. 2015. What Makes Hyper-local Online Discussion Forums Sustainable?. In *Proceedings of the 48th Hawaii International Conference on System Sciences (HICSS '15)*. 2445–2454.

Christina A Masden, Catherine Grevet, Rebecca E Grinter, Eric Gilbert, and W Keith Edwards. 2014. Tensions in scaling-up community social media: a multi-neighborhood study of nextdoor. In *Proceedings of the 32nd annual ACM Conference on Human Factors in Computing Systems*. ACM, 3239–3248.

Miller Mc Pherson. 1983. The size of voluntary organizations. *Social Forces* 61, 4 (1983), 1044–1064.

J Miller McPherson and Thomas Rotolo. 1996. Testing a dynamic model of social composition: Diversity and change in voluntary groups. *American Sociological Review* (1996), 179–202.

David R. Millen and John F. Patterson. 2002. Stimulating social engagement in a community network. In *Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work (CSCW '02)*. 306–313.

Thomas P Novak. 1998. Bridging the racial divide on the Internet. *Science-AAAS-Weekly Paper Edition* 280, 5362 (1998), 390–391.

Dara O'Neil. 2002. Assessing community informatics: a review of methodological approaches for evaluating community networks and community technology centers. *Internet Research* 12, 1 (2002), 76–102.

Andrew Perrin and Maeve Duggan. 2015. *Men catch up with women on overall social media use*. Technical Report. Pew Research Center.

Pamela A Popielarz and Miller McPherson. 1995. On the edge or in between: Niche position, niche overlap, and the duration of voluntary association memberships. *Amer. J. Sociology* (1995), 698–720.

Daphne R. Raban, Mihai Moldovan, and Quentin Jones. 2010. An empirical study of critical mass and online community survival. In *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work (CSCW '10)*. 71–80.

Damian Radcliffe. 2012. *Here and Now UK Hyperlocal media today*. Technical Report. Nesta.

Paul Resnick. 2001. Beyond bowling together: Sociotechnical capital. *HCI in the New Millennium* 77 (2001), 247–272.

Paul Resnick, Joseph Konstan, Yan Chen, and Robert E Kraut. 2012. Starting New Online Communities. In *Building successful online communities: Evidence-based social design*. MIT Press Cambridge, MA, 231–280.

Everett M. Rogers, Lori Collins-Jarvis, and Joseph Schmitz. 1994. The PEN project in Santa Monica: Interactive communication, equality, and political action. *Journal of the American Society for Information Science* 45, 6 (1994), 401–410.

Thomas Rotolo. 2000. Town heterogeneity and affiliation: A multilevel analysis of voluntary association membership. *Sociological Perspectives* 43, 2 (2000), 271–289.

Chandan Sarkar, Donghee Yvette Wohn, and Cliff Lampe. 2012. Predicting Length of Membership in Online Community "Everything2" Using Feedback. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work Companion (CSCW Companion '12)*. 207–210.

Jen Schradie. 2012. The trend of class, race, and ethnicity in social media inequality: Who still cannot afford to blog? *Information, Communication & Society* 15, 4 (2012), 555–571.

Doug Schuler. 1994. Community networks: building a new participatory medium. *Commun. ACM* 37, 1 (1994), 38–51.

John C Scott. 1957. Membership and participation in voluntary associations. *American Sociological Review* (1957), 315–326.

Thiago H. Silva, Pedro O. S. Vaz de Melo, Jussara M. Almeida, Juliana Salles, and Antonio A. F. Loureiro. 2014. Revealing the City That We Cannot See. *ACM Trans. Internet Technol.* 14, 4, Article 26 (Dec. 2014), 23 pages. DOI:http://dx.doi.org/10.1145/2677208

Lyn Simpson. 2005. Community Informatics and Sustainability: Why Social Capital Matters. *The Journal of Community Informatics* 1, 2 (2005), 102–119.

Jacob Solomon and Rick Wash. 2014. Critical Mass of What? Exploring Community Growth in WikiProjects. In *Proceedings of the International Conference on Weblogs and Social Media (ICWSM '14)*.

Marco Tortoriello, Ray Reagans, and Bill McEvily. 2012. Bridging the knowledge gap: The influence of strong ties, network cohesion, and network range on the transfer of knowledge between organizational units. *Organization Science* 23, 4 (2012), 1024–1039.

Xiaoqing Wang, Brian S Butler, and Yuqing Ren. 2013. The Impact of Membership Overlap on Growth: An Ecological Competition View of Online Groups. *Organization Science* 24, 2 (2013), 414–431.

Yi-Chia Wang and Robert Kraut. 2012. Twitter and the development of an audience: those who stay on topic thrive!. In *Proceedings of the 2012 ACM Annual Conference on Human Factors in Computing Systems (CHI '12)*. 1515–1518.

Yi-Chia Wang, Robert Kraut, and John M. Levine. 2012. To stay or leave?: the relationship of emotional and informational support to commitment in online health support groups. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work (CSCW '12)*. 833–842.

Molly McLure Wasko and Samer Faraj. 2005. Why should I share? Examining social capital and knowledge contribution in electronic networks of practice. *MIS Quarterly* (2005), 35–57.

Miaomiao Wen and Carolyn Penstein Rose. 2012. Understanding Participant Behavior Trajectories in Online Health Support Groups Using Automatic Extraction Methods. In *Proceedings of the 17th ACM International Conference on Supporting Group Work (GROUP '12)*. 179–188.

Xidao Wen, Di Lu, Yu-Ru Lin, and Claudia López. 2016. The Roles of Information Seeking Dynamics in Sustaining the Community Participation. In *Collaboration and Internet Computing (CIC), 2016 IEEE 2nd International Conference on*. IEEE, 357–362.

Kate Williams and Joan C Durrance. 2008. Social Networks and Social Capital : Rethinking Theory in Community Informatics. *The Journal of Community Informatics* 4, 3 (2008), 1–20.

Chaolun Xia, Raz Schwartz, Ke Xie, Adam Krebs, Andrew Langdon, Jeremy Ting, and Mor Naaman. 2014. CityBeat: Real-time Social Media Visualization of Hyper-local City Data. In *Demo in International World Wide Web Conference (WWW Companion '14)*.

Jiang Yang, Xiao Wei, Mark S Ackerman, and Lada A Adamic. 2010. Activity Lifespan: An Analysis of User Survival Patterns in Online Knowledge Sharing Communities.. In *Proceedings of the International Conference on Weblogs and Social Media (ICWSM '10)*.

Haiyi Zhu, Jilin Chen, Tara Matthews, Aditya Pal, Hernan Badenes, and Robert E Kraut. 2014a. Selecting an effective niche: an ecological view of the success of online communities. In *Proceedings of the 32nd annual ACM Conference on Human Factors in Computing Systems*. ACM, 301–310.

Haiyi Zhu, Robert E Kraut, and Aniket Kittur. 2014b. The impact of membership overlap on the survival of online communities. In *Proceedings of the 32nd annual ACM Conference on Human Factors in Computing Systems (CHI '14)*. 281–290.

Received April 2016;