Idea Management Communities in the Wild

An exploratory study of 166 online communities

Jorge Saldivar*, Marcos Baez†, Carlos Rodriguez‡, Gregorio Convertino§ and Grzegorz Kowalik¶
*Catholic University ”Nuestra Señora de la Asunción”, Asunción, Paraguay, jorge.saldivar@uc.edu.py
†University of Trento, Trento, Italy, baez@disi.unitn.it
‡University of New South Wales, Sydney, Australia, carlos.rodriguez@unsw.edu.au
§Informatica Corporation, Redwood City, CA, USA, gconvertino@informatica.com
¶Polish-Japanese Academy of Information Technology, Warsaw, Poland, grzegorz.kowalik@pjwstk.edu.pl

***Abstract*—Idea Management (IM) communities have the potential to transform business and communities through innovation. However, building successful communities is a difficult endeavor that requires a significant amount of both community management and technological support. Doing this requires a good understanding of how IM systems are used and how users behave, as these are fundamental aspects for the design of effective technological support as well as devising community management strategies.**

In this paper, we study 166 IM communities in the "wild" — communities openly available on Ideascale, one of today's leading IM software platforms— to better understand *how* they are used in practice, and by *whom*. We do this via i) a qualitative analysis of community properties to identify community archetypes; ii) a quantitative analysis of user activity logs to identify patterns of collective and individual user behavior.

***Keywords*—Collaborative Open Innovation, Collective Intelligence and Crowdsourcing**

I. Introduction

The increasing competitiveness of the markets forces organizations to sustain a continuous process of innovation fueled with ideas originated from managers, employees and, for some time now, from people outside the organization. Idea Management (IM) is the process of requesting, collecting, selecting and evaluating ideas to develop new, innovative products, services or regulations, or to improve existing ones [1]. The goal of IM is to capture ideas that can deliver benefits to the organization by generating innovations or by solving specific problems [2].

The emergence of social and collaborative web-based technologies has transformed the physical suggestion boxes — the former preferred method to listen to customers— into dedicated IM systems, which lets people propose ideas, as well as rate and place comments on other users' suggestions [3]. Examples of popular IM systems are IdeaScale (http://ideascale.com), Crowdicity (http://crowdicity.com), Spigit (http://www.spigit.com).

The adoption of IM practices and systems has empowered various innovation initiatives around the world. Almost 200,000 people have participated in My Starbuck Idea, the world-wide IM initiative conducted by Starbucks to collect ideas from its customers about future products and services [4]. Similar participation rates can be found when analyzing Idea Storm, the IM initiative sponsored by the computer company Dell [5]. But, its application has not been limited solely to commercial domains. In the political and civic domain, the Icelandic participatory constitution-writing process represents an emblematic case. Here, the population at large has been invited to contribute to the constitution draft with suggestions, proposals, and ideas [6].

IM has the potential to benefit organizations and businesses by allowing them to discover valuable ideas that can lead to innovations. In this context, contributions of participants to provide valuable ideas are seen as strategic assets in the success of IM initiatives [5]. The larger the community of participants, the more diverse views are likely to appear. More diversity increases the chances of producing valuable ideas [7].

However, building successful online communities is a challenge. It requires an understanding of the people and their needs, as well as setting up the proper technology and policies to match the characteristics of users and purpose of the community [8]. Success then depends as much on proper management as it does on proper support. By gaining a better understanding on how organizations and users make use of IM communities, platforms and systems can better accommodate their designs to serve these needs and facilitate the management of the overall community.

In this paper we explore *how* and by *whom* IM systems are used in practice. We do so first by qualitatively analyzing and classifying IM communities and then by quantitatively analyzing collective and individual behaviors of users. We explore these questions on a dataset of 166 openly available communities in IdeaScale. This research work contributes to the state of the art on IM as follows:

- *Characterization of IM communities on the same platform*. We perform a qualitative analysis of a large set of IM communities that share the same technology platform and derive a set of community archetypes. These archetypes tell us how and by whom IM systems are used.
- *Identification of collective and individual behavior patterns from user actions*. We study four types of user actions (i.e., registering as member, posting ideas, commenting, voting) and identify a set of individual and collective patterns of behaviors.

In the next section, we give an overview of Idea Management in general, and Idea Scale in particular, the platform we

Figure 1. (a) IdeaScale's community website; (b) Idea submission features; (c) Detailed view of an idea, commenting and voting functions

use in this study. We then provide an overview of the related works and then switch to the presentation of the methods we use in this paper and the actual study. We close this paper with a discussion of the main findings of our study.

II. Background

A. Idea Management: Process and System

Idea Management (IM) is a process that organizations use to promote innovation. Thorough IM, an organization can leverage its communities of clients, employees, suppliers, or interested stakeholders to (1) request ideas, (2) collect and (3) evaluate them, and (4) select the most promising ones to source their innovation needs or to address a defined organization's problem [9].

The execution of IM processes can be supported by dedicated software tools known as IM systems. These systems allow an organization to describe an innovation problem it wants to solve (e.g., innovate the public transportation system) and setup campaigns through which they collect the proposed solutions. At the same time, IM systems let users suggest ideas as well as evaluate and place opinions on other users' ideas.

We focus on IdeaScale[1] as the IM system of interest for this study. IdeaScale is one of today's leading technologies for supporting the execution of IM processes and used by big companies like Microsoft and Xerox and government institutions such as NASA and the White House. Apart from being a popular commercial platform in the market of IM systems, IdeaScale offers publicly accessible data that can be collected for research purposes through dedicated Web APIs[2] — an important facilitator for conducting research on these IM communities.

In IdeaScale, ideation initiatives are created by setting up a community website in which organizers describe the goals of the initiatives and define campaigns through which ideas are collected. Figure 1 (a) illustrates the main interface of an IdeaScale's community website.

Figure 1 (b) shows the empty template used to submit ideas on this website. When submitting an idea, a user, who previously registered as member of the community, provides a title and a description of the idea and associates the idea to a campaign. Optionally, the user can categorize the idea using tags and attach an image or file to enrich the description.

Users can also comment and assign positive or negative votes to others' ideas and comments. They can also reply to existing comments. Such functionalities enable users to contribute arguments in favor or against an idea or a previous comment. This helps the authors with refining the content and the organizers with selecting and growing the best ideas. Figure 1 (c) introduces an example of an idea together with the features to vote and comment.

III. Related Works

IM systems are playing a key role in enabling grassroots innovation initiatives [11]. In this context, IM platforms have proven able to properly instrument campaigns for soliciting ideas from large-scale crowds, in business and public sectors [12].

The discussions on how to extend and improve online IM platforms have taken different directions among industrial and academic researchers. These include ways to improve features of IM systems (e.g., techniques to display streams of ideas, assess ideas, and find promising ideas) and empirical studies about different phases of the IM process (e.g., idea submission, evaluation, selection, and implementation) [9].

A. Applied Research

Deliberation maps have been presented in [13] to structure participants' contribution as problem trees containing the problem to solve, potential solutions, and arguments for and against proposed solutions. The use of semantic technologies has been proposed by Westerski et al. to organize, link and classify the proposed ideas using meta data annotations [14]. Improving scoring methods used to rate the ideas has been the goal of Xu et al. who have proposed a reference-based scoring model as an alternative to the traditional thumbs up/down voting systems [15]. Faridani et al. have introduced a two-dimensional visualization plane as an approach to address the filter-bubble effect —narrowing the exposure to recent, popular, or controversial information— of linear listings used to display opinions in online sites [16].

Convertino et al. have targeted information overload in the evaluation phase by employing natural language processing methods to automatically identify the core of the proposals

[1] https://ideascale.com

[2] APIs: set of functions through which a system can be programmatically accessed [10]

[17], and by providing novel organization tools to facilitators in IMS [18]. Along this line, Bothos et al. have introduced the application of information aggregation markets to facilitate the evaluation of the ideas [19]. From an analytical perspective, [20] has employed social network analysis techniques to study the relationship between the quality of the ideas and the connectivity (degree centrality) of the contributors.

B. Empirical Studies

By conducting an empirical study on Dell's and Starbucks' IM initiative, Hossain and Islam have analyzed the factors that influence the selection and implementation of ideas [21], [22]. Studying Dell's case, Di Gangi and Wasko have investigated the attributes that characterized the ideas that end up being pushed further in the innovation pipeline [5]. For his part, Bayus has taken also data from Dell IdeaStorm and discovered that the commenting activity of participants has positive effects on the possibility of participants to submit valuable ideas [23].

Saldivar et al. conducted empirical study on IM communities for civic engagement where they analyzed the effectiveness of social sharing features, e.g., share and tweet button — the preferred approach to integrate IM platforms and social networking sites— and the role of social networking sites, such as Facebook and Twitter, in the IM process [24].

Although previous empirical works have provided useful insights about IM, the state of art lacks studies that investigate properties and characteristics of large numbers of IM communities and discover regularities in the behavior of the members and the community as a whole.

IV. Methods

A. Research questions

In this research work we address the following research questions:

RQ1. What type of communities emerge in Idea Management Systems? The goal is to understand what types of communities live in IM systems by identifying relevant properties that characterize such communities.

RQ2. What individual and collective behaviors emerge in Idea Management Systems? The goal is to identify common patterns of behavior by looking at how users and communities as a whole participate in the ideation process.

Understanding how communities work in practice can help i) researchers identify potential gaps between current theory and practice, and ii) practitioners design solutions that fit better the needs of users and communities.

B. Data

The data set used in this research consists of public-access IdeaScale communities, available as of October 2015. It contains data from 166 communities generated through the main actions supported by the platform (registering as member, submitting ideas, posting comments and voting), which collectively account for 50,187 registered members, 24,403 ideas, 32,592 comments, and 217,933 votes[3]. The number of members, ideas, comments, and votes are distributed across the 166 communities following right skewed distributions, as outlined in Figure 2.

[3]Datasets and R scripts of this study are available at https://github.com/joausaga/collective-behavior-im-communities

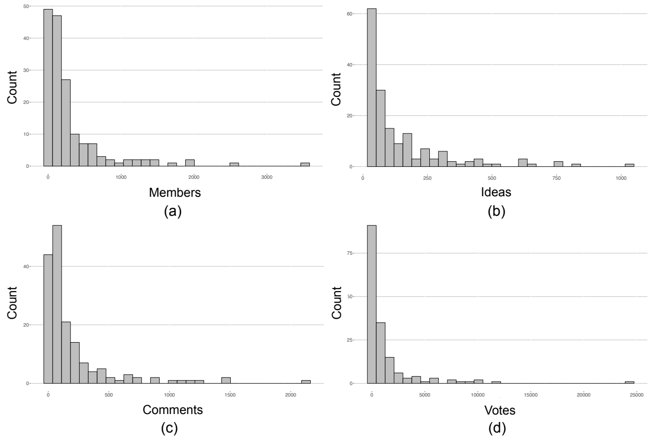

Figure 2. Distribution of members (a), ideas (b), comments (c), and votes (d) across the 166 communities

C. Qualitative analysis of community archetypes

To address **RQ1**, we conducted a qualitative analysis of the 166 communities in our data set. For each community, the content analyzed was the main IM community page and a few of the most prominent (e.g., most voted) ideas. The analysis consisted of the following steps:

Step 1. Two independent coders analyzed a random sample of 20 communities using an open coding method [25], [26]. Then, the coders shared the results and agreed on a common coding scheme of six descriptive dimensions, where each dimension takes one of a bounded set of possible values. For example, when coding a community, the first dimension "Type of organization" could take one of these values: "Business", "Governmental", "NGO" or "Community".

Step 2. Three independent coders (the previous two coders plus a third coder) categorized the 166 communities using the coding scheme described in Table I. The inter-coder agreement was 83%. For each case where there was a disagreement the three coders met and reached consensus on the final categorization.

Step 3. The results of the categorization were then used to cluster the communities based on emerging archetypes, i.e., groups of communities where tuples of values tended to co-occur frequently among the dimensions. Due to insufficient information two of the six dimensions, "Contributor" and "Can act?", were excluded from the analysis (see results below).

D. Quantitative analysis of collective behaviors

In answering **RQ2**, we investigated common patterns around the following four types of actions: *idea submission*, *community member registration*, *comment posting*, and *vote casting*. We assumed that communities behave differently at different stages of their lifecycle. Particularly critical for the success is for example the behavior of the community after it

TABLE I. COMMUNITIES CODING SCHEME

Type of organization.
Type of organization running the ideation process. **Business.** Profit organizations (e.g., a company). **Governmental.** Organizations such as government agencies. **NGO.** Non-profit, non-governmental organizations. **Community.** Individuals running a community, without conforming a formal organization (e.g., gamers community).
Domain of the organization
Domain in which the community is operating **Technology.** Related to software and hardware. **Civic.** Organizations seeking civic participation. **Education.** Organizations such as universities and schools. **Bureau.** Related to the financial, legal, political and military sector. **Leisure.** Related to entertainment and hobbies (e.g., tv, games) **Retail**, including food & drinks (e.g., shops, restaurants, wineries). **Other.** related to other sectors not described above.
Contributor
Participants of the ideation process in relation to the organization **External.** People external to the organization (e.g., clients). **Internal.** Members of the organization (e.g., employees).
Scope
The location of the target contributors. **Local.** A country or local community (e.g., Serrenti county, Italy). **Global.** Any country or region of the world.
Purpose
Reason for running the community **Feedback.** The purpose is to gather requirements and feedback over a product or service (e.g., suggestions to improve a service). **Innovation.** The purpose is to gather ideas for new products and services (e.g., school reforms in a local community). **Coordination.** The purpose is to coordinate actions (e.g., for events). **Discussion.** The purpose is to discuss (e.g., priest replying to questions about faith).
Can act?
The inicitative owner is able to implement the idea and take actions **Yes.** The ideation and deliberation are actionable. **No.** The ideation and deliberation are *not* actionable.

is launched. To mitigate the effect of time and maturity of the community, we limited our analysis to its first year of life. For each type of action, we performed the following: *i)* the actions performed in the first year were partitioned into quarters; and *ii)* the proportion of actions performed in each quarter in relation to the yearly total was computed. In addition, we computed the relative number of ideas, votes and comments per member to cancel the effect of community size. As a result, for each community we obtained a four feature vector, with one feature per action type. The first feature contained the proportion of member registrations in each quarter, and the remaining three contained the proportion of ideas, votes, and comments by members in each quarter. We used a K-means clustering algorithm [27] to group communities according to the similarity of their feature vectors. We iteratively tested the algorithm with different number of clusters until we were satisfied with the grouping. The satisfaction criteria we used were simplicity and clearness. Next, for each cluster, we drew the evolution of user actions (e.g., member registrations, idea generation) within communities over the first year of life, thus, obtaining a set of patterns that describes the collective behavior of communities within that period. These patterns help us address questions such as when we should expect the majority of member registrations and how user action evolve over time.

E. Quantitative analysis of individual behaviors

We also analyzed the individual behavior of members to address **RQ2**. To do so, we selected all the actions recorded in the 166 communities that have authors with known registration dates. We found 173,433 action records meeting this criteria.

In this analysis, our aim was to find what the typical “lifetime” of a member in a community is — what their first actions are and how long they remain active. To this end, we computed the percentage of actions performed during the day of registration, day after, two day after, etc. In addition, we investigated what type of action seemed to motivate people to join a community. By “joining” we refer here to the registration date of a member and we used the first action of that member after the registration as the “first reason” for joining the community. Finally, we analyzed the individual user behaviors against the archetypes described previously.

V. RESULTS: COMMUNITY ARCHETYPES

In this section we first present the results of our characterization of the online communities according to the coding scheme, and then the emerging community archetypes. These analyses are summarized in Figure 3.

A. Communities according to the coding scheme

Exploring each dimension of the coding scheme we have observed the following general trends:

Type of organization. The majority of communities are run by companies (*Business* 48%) followed by self-driven communities (*Community* 21%), i.e., communities without the backing of a formal organization. Closely behind we have communities run by non-for-profit / non-governmental organizations (*NGO* 17%), and by governmental organizations (*Governmental* 14%) (see Table I and Figure 3).

Organization domain. Most organizations running the communities are related to the *Technology* domain (54%), followed by *Civic* (15%) and *Education* (10%), with fewer communities from the other domains (see Table I and Figure 3).

Contributor. As we were limited to publicly available communities, most of them involved *External* actors. Since we were not able to reliably determine the type of contributors, this third dimension was excluded from the analysis.

Scope. Both local and global communities were frequent. Communities appear to be somehow equally distributed between local and global audiences. *Local* (57%) communities are the most common, mostly consisting of civic communities, while the *Global* (43%) ones are more technology-oriented focusing on product and services available worldwide (see Table I and Figure 3).

Purpose. The dominant purpose of the communities is collecting *Feedback* (65%) followed by *Innovation* (25%) and to a lesser extent *Discussion* (6%) and *Coordination* (4%) (see Table I and Figure 3). For example, a common case is that of communities focusing on software products where members report bugs and request features (feedback).

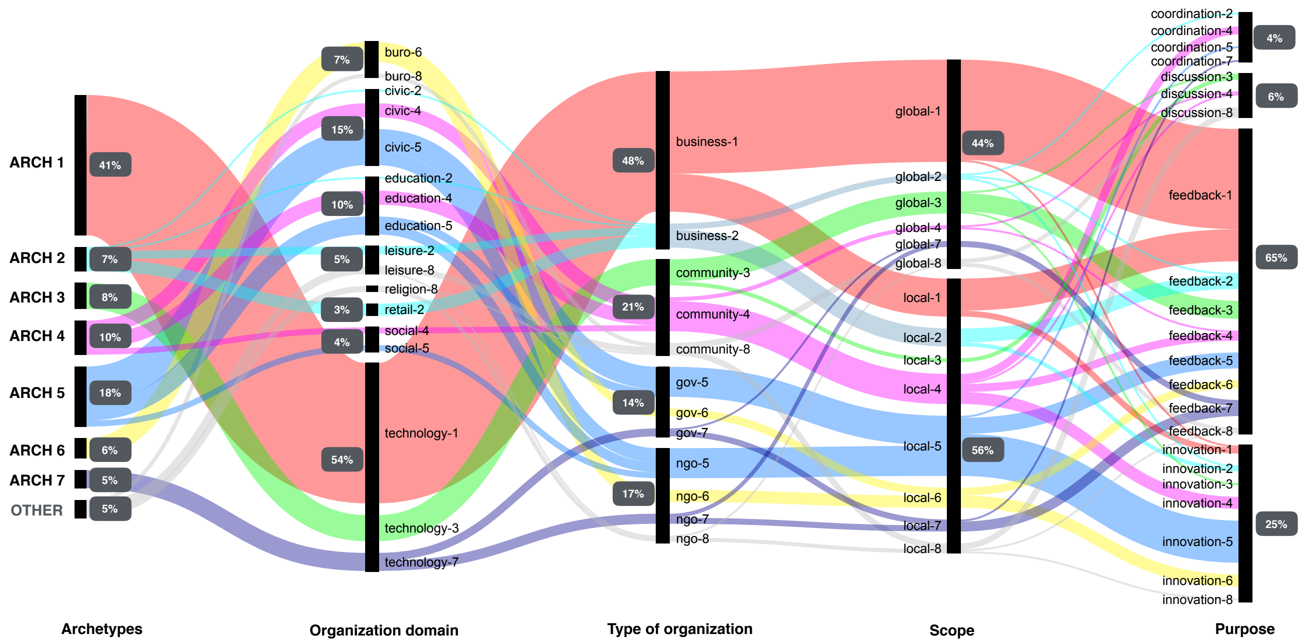

Figure 3. Alluvial chart illustrating the emerging community archetypes. The percentage represents the distribution of communities for each dimension of the coding scheme

Can act?. The capacity of communities to act on the results of the deliberation was difficult to assess. This is partly due to the lack of information on the communities and the misuse of the different phases in the ideation process. For this reason, this dimension was excluded from the analysis.

B. Communities archetypes

Based on the categorization done using our coding scheme (shown in the previous section), in this section we focus on identifying community archetypes (see Figure 3). We use the desriptive construct of community archetypes to categorize types of IM communities. An archetype is defined as a frequently observed tuple of values along the four coding scheme dimensions.

ARCH 1. Communities run by companies in the technology domain. This archetype was the most frequent in the data set (70). Communities belonging to this archetype were mostly seeking feedback from users and customers on their technology-related products and services. A representative example is `QuestionPro Feedback`[4], a community where users report on bugs and request features their product.

ARCH 2. Communities run by companies in other domains. This archetype clusters the remaining communities run by companies (11). The domains of these companies include leisure, retail, food & drinks, civic and education. For example, the `The Beerenberg Family Farm`[5] is a community run by a food processing company on its products.

ARCH 3. Self-driven communities on the technology domain. This archetype represents communities without the backing of a formal organization, run by its own members, on topics related to technology (13). These communities are similar to communities of practice, a type of communities frequently investigated in previous research [28]. This cluster combines the community-driven nature with the dynamics of software products and services. As in *ARCH 1*, the dominant purpose is feedback, although we also observed a much higher number of cases with a focus on discussion. An example of this cluster is `Vivo Open Source`[6], a community on an open source software managed by the community itself.

ARCH 4. Self-driven communities in civic, education and social domains. This archetype represents communities without the backing of a formal organization, run by its own members and focusing on topics related to their civic life, education and other social themes (16). This archetype combines the self-driven nature of the communities, focus on social impact, and local scope. Here, we see innovation as the prominent purpose, followed closely by feedback. An example of this cluster is `Rescatar a Lois`[7], a community run by concerned citizens on how to save a local factory from a crisis.

ARCH 5. Communities driven by a formal organization focusing on civic, education and social domain. This archetype groups communities run by either governmental or non-profit organizations (Governmental, NGO) on topics that relate to the civic life, education and other social causes (30). This is the second most frequent archetype and it combines the local scope with the presence of governmental or non-profit organization as drivers of the communities. Compared to *ARCH 4*, innovation is by far the most dominant purpose here. An example of

[4]https://questionpro.ideascale.com
[5]http://beerenberg.ideascale.com
[6]http://vivo.ideascale.com/
[7]http://rescataralois.ideascale.com/

this cluster is `HoCoInnovations`[8], a community run by a county on ideas to improve the school system.

ARCH 6. Communities driven by a formal organization in the "Bureau" domain. This archetype groups communities run by either Governmental or NGO organizations on topics that relate to financial, legal, political and military matters (10). These are local communities that tend to have very structured contributions around campaigns. In some cases they have more complex organizational structures: the median number of campaigns per community in this archetype was higher (median = 6) than in the other archetypes (median = 4). An example of such communities is `Martellago Cinque Stelle`[9], a community run by a political party in an Italian town on local programs and actions.

ARCH 7. Communities driven by a formal organization in the technology domain. This archetype groups communities run by both governmental and non-profit organizations (gov, ngo) on technology-related areas (9), in contrast to *ARCH 1* and *ARCH 3*, which are run by companies or the communities themselves. However, similar to *ARCH 1*, these communities are predominantly focused on feedback. This cluster combines the nature of technology-related products and services, with the dynamics of NGOs and governmental agencies. An example of such communities is `API Developers Forum`[10], a community run by the US Census Bureau on the API for accessing their data. The above archetypes give us some interesting insights about how and by whom IM systems are used: (i) Communities related to technology largely focus on incremental or corrective feedback; (ii) communities on social themes tend to seek for more innovative ideas; (iii) communities run by its own members tend to incorporate more discussion; (iv) communities run by organizations on "bureau" tend to have more structured campaigns.

VI. Results: Collective Behavior

This section of the paper focuses on describing how communities act collectively. We found five patterns that shape the development of *member registration*, *idea submission*, *commenting*, and *voting* in communities. Also, we observed that these behavioral patterns are apparently influenced by the intervention of moderators. Finally, we did not observe a clear correlation between behavioral patterns and archetypes, except for voting behaviors.

A. Behavioral Patterns

After applying the k-means algorithm with different number of clusters, we found five behavioral patterns, i.e., trends over 1 year for one of four types of actions (see Figure 4).

For most communities (142 out of 166, 85%), the evolution of registrations over the first year of their life follows patterns 1, 3, or 5 (see Table II for the list of patterns). In **behavioral pattern 1**, which we call *Q1 peak and gradual decent*, 55 (33%) of the communities show to have a burst of registrations during the first three months of the year and then the number of new members gradually decreased or remained somehow

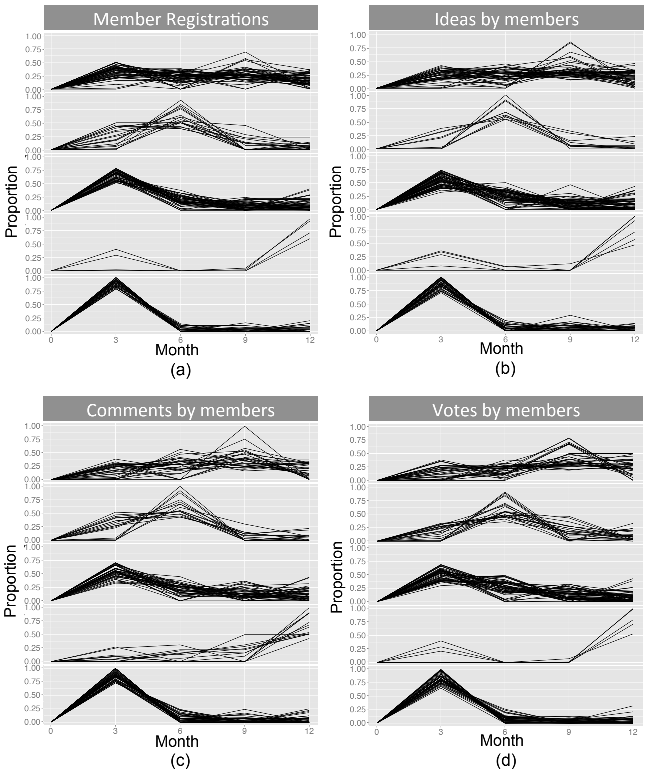

Figure 4. Patterns in the evolution of member registrations (a), idea submissions (b), comment posting (c), and vote casting (d) over first year of life, respectively. X-axis indicates the month of the year while Y-axis shows the proportion of the actions done in the different months

constant until the end of the period. Communities that follow **behavioral pattern 3**, which we call *Q1 peak and rapid decent*, (53 out of 166, 32%) show, however, a more prominent peak of registrations during the first quarter. In fact, between 50 and 75% of registrations occurred in that period of time. Then, from the second quarter on, the proportion of registrations falls remaining stable around 25%. **Behavioral pattern 5**, which we call *Q1 peak and super rapid decent*, represents a more extreme case of pattern 3. Here, between 75 and 100% of member registrations happened in the fist quarter. Then the number of member registrations decays drastically and remains very low until the end of the period.

A quite different pattern is followed by 13% of the communities, which corresponds to **behavioral pattern 2**, which we call *Q2 peak and very rapid decent*. Instead of having large proportions of registrations at the beginning, they concentrate their registration activities during the second quarter (from month three to half-year). After that period, the registration of members falls down to quite low levels. Finally, very few communities (4 out of 166, 2%) show peaks of registrations towards the end of the year (**behavioral pattern 4**, which we can call *Q4 latter peak*). This type of behavior could be considered more an outlier than a pattern.

Interestingly, for the rest of the actions, i.e., idea submissions, comment posting, and vote castings, communities follow the same patterns. However, the distribution of communities

[8]http://hocoinnovations.ideascale.com/
[9]http://martellago-m5s.ideascale.com
[10]http://apiforum.ideascale.com/

per pattern is different as shown in Table II. Although the distribution of communities in each pattern show to be different from action to action, a general trend can be seen: patterns 1, 3, and 5 are followed by the majority of the communities. Pattern 2 depicts the behavior of about 6 to 15% of the communities for each action while pattern 4 is rather negligible.

TABLE II. NUMBER AND PERCENTAGE OF COMMUNITIES AFFECTED BY THE PATTERNS FOR EVERY ACTION

Behavioral Pattern	Action: Member Reg.	Action: Idea Submission	Action: Comment Posting	Action: Vote Casting
1	55 (33%)	48 (29%)	32 (19%)	34 (20%)
2	20 (13%)	11 (6%)	18 (11%)	24 (15%)
3	53 (32%)	61 (37%)	48 (29%)	56 (34%)
4	4 (2%)	6 (4%)	13 (8%)	5 (3%)
5	34 (20%)	40 (24%)	55 (33%)	47 (28%)

A general finding is that a main peak is present in each of the patterns. The peak indicates a localized period of predominant activity, which could be explained by external events, such as dissemination events that trigger it. Except for pattern 4, the level of activity decreases after the peak.

One third of communities (55 out of 166) follow the same collective behavior for all of the action types. Such commonality suggests overall attention peaks, where contributions —in all forms might— follow member registration. We will go in depth on these results in the next section.

B. *Influence of moderation in collective behavior*

Different factors may influence the collective behavior of communities. We have no information about the external ones, such as promotional events, incentives, or other public events because they are not registered in our data set. Other factors are internal and in particular previous research has shown the benefits of having organizers and moderator interventions on the quality of IM processes [29].

In this analysis, we investigated if there was a relationship between moderator interventions and behavioral patterns, understanding moderator intervention as all submissions (ideas, comments and votes) performed by moderators and organizers of communities. The analysis was limited to actions related to content creation because we assume that the actions by moderators within the communities have little influence on attracting new members.

Interestingly, communities that follow patterns 1 and 3 are at the same time those that show the strongest presence of moderators. On average, moderators intervened 2.5 times (69.71 vs. 27.92 interventions in average) more in communities in which their ideation actions are shaped by patterns 1 and 3 than in communities that follow patterns 2, 4, and 5. Similar numbers were found when studying the participation of moderators in communities where commenting and voting are governed by these patterns.

By splitting interventions into quarters, we observed that periods with high level of activity correspond to quarters of high activity by moderators. For every pattern, significant

TABLE III. DISTRIBUTION OF COMMUNITIES ARCHETYPES PER VOTING PATTERNS

Behavioral Pattern	ARCH1	ARCH 2	ARCH 3	ARCH 4	ARCH 5	ARCH 6	ARCH 7	ARCH 8
1	19	3	0	4	3	2	3	0
2	9	2	6	0	4	1	1	1
3	25	3	3	3	12	5	2	3
4	0	0	0	1	1	0	2	1
5	14	3	4	9	10	2	1	4

correlations ($\alpha = 0.05$) were found between interventions and productivity of ideas, comments, and votes (idea submission: Person r=0.89, $p < 0.001$, commenting: Person r=0.55, $p < 0.05$, and voting: Person r=0.73, $p < 0.001$). In light of previous research [30], these results confirm that in our communities a higher number of interventions by moderators is associated with higher activity levels by the community.

C. *Patterns and archetypes*

We did not observe associations between behavioral patterns and archetypes, except for the patterns for voting. By conducting Pearson's Chi-squared tests, we found that archetypes are associated with the patterns of behavior for casting votes ($X^2 = 48.52, df = 28, p < 0.01$). That is, some archetypes exhibit distinctive behavioral patterns for voting.

Voting in 66% (44 out of 67) of the communities in ARCH 1 is shaped by patterns 1 and 3. More than half of the communities in ARCH 4 (9 out of 17) follow pattern 5 when casting votes. Voting follows patterns 3 and 5 in about 75% of communities belonging to ARCH 5. For the rest of the archetypes (2, 3, 6, 7 and 8) the voting action is homogeneously distributed among patterns.

The nature of voting action —which requires much less effort compared to ideation, commenting, or registering— may explain why groups of archetypes are associated to patterns. It might be that low-effort actions are more easily shaped by common patterns than more time-consuming actions, which may be more influenced by external factors. Further research is needed to better understand the reasons behind this association.

VII. RESULTS: INDIVIDUAL BEHAVIOR

This section contains analyses of community members actions on individual level. We found that most of members perform only one action and that action happens normally during the first day after registration.

A. *Number of actions per member*

To study the number of actions per member, we included only active community members (13,619 members, 27%), defining "active" members as those who performed at least one action, i.e., submit idea, cast vote, or post comment.

The majority of community members did just one action of each type (idea, comment, vote). The median of action per member is 1 idea, 1 comment and 2 votes. There is a very small group (10%) of more active members with more than 3 ideas, 4 comments, and 23 votes.

B. Time of actions

In our analysis of community member actions, we computed the day in which they were performed since author registration. Results are summarized in Table IV. A large part of actions was performed some time between the day of registration or the day after (0 means the registration day, 1 means day after, etc.). About 50% of ideas, 20% of comments, and 40% of votes were submitted in this time window. Probably, patterns of registration, ideas, comments and voting show similar shapes because these actions are performed within a short time window (usually within the first few days). See Figure 4.

TABLE IV. NUMBER OF DAYS THAT PASS FROM REGISTRATION TO FIRST ACTION

Percentile	First Idea	First Comment	First Vote
0.1	0	0	0
0.2	0	0	0
0.3	0	3	0
0.4	0	11	1
0.5	1	34	11
0.6	8	82	33
0.7	39	176	90
0.8	140	327	225
0.9	365	551	448
1	2192	2198	2111

Given the above results, next we try to understand in more details which action was the main driver for registration, i.e., which action was firstly performed after the person registered as member of the community. Results are shown in Table V. Almost half of community members posted their ideas as the first action after joining community. Interestingly, the a-priori "hardest" action was the main driver that attracted people to communities (the "easiest" and least time consuming action, voting, was the second one). Previous research has also found that people engage in this kind of initiatives mainly attracted by the possibility to disseminate their ideas [31].

TABLE V. FIRST ACTIONS OF USERS

Action	Number of users	Percentage of users
Idea Submission	6161	46.49%
Vote Casting	4853	36.62%
Comment Posting	2238	16.89%

C. Users action and archetypes

We also compared users action within each of the discovered archetypes. From Figure 5, we can see that there are communities with the majority of actions done within one day (ARCH 2,3,4,8) and those that have more active members during later days (ARCH 1,5,6,7). This is interesting, because ARCH 1,2,5,6,7 are more formal —they are supported by companies or formal organizations— while ARCH 3 and 4 are self-driven. In relation to the latter, we found that communities in ARCH 2 have more active members than 3 and 4 if results are analyzed in the 60-percentile level. It seems that company-driven or official communities have more success in keeping their members active for longer periods of time.

VIII. DISCUSSIONS AND CONCLUSIONS

The findings we report in this article reveal aspects of IM systems and communities to date scarcely studied. We expect

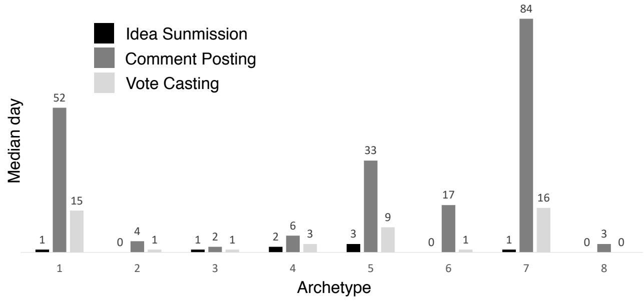

Figure 5. Median of days spent by communities in archetypes to perform their actions

that these results will help practitioners in the design and instrumentation of their IM initiatives.

Types of communities in IM systems. Most of the IM initiatives found in the platform are dominated by communities in the technology business and those that address civic, education and social issues. On the one hand, the civic communities are usually managed by for-profit organizations that use IdeaScale as a tool for collecting user feedback on their products and services. On the other hand, the education and social communities are either self-driven or driven by a formal organization, and they are characterized by its innovation nature and strong social impact. The rest of the communities have a lower prevalence and they typically relate to other domains such as leisure, food & drinks, military, politics, among other topics.

Collective behavior of communities. Overall, communities follow the same collective behavior pattern for all action types, i.e., for member registration, idea submission, comment posting, and vote casting. From the results that we reported earlier, patterns that show higher activity levels at the beginning of the life of communities prevail. This common behavior [32] might be the effect of the early enthusiasm occurring soon after the lauch of a community or the result of additional external factors such as the promotion of the initiative outside the IM platform or the incentive offered by the organizations to the participants. The implication of this behavior is that organizers or moderators who want increase the volume of interactions by members of the community should focus their efforts during this early period of high activity and high rate of member registrations, as opposed to leaving such efforts towards a later time.

Finally, our study on patterns and archetypes indicate that these two are not associated, except for the case of voting patterns. More concretely, the archetypes that include technology business and civic participation communities seem to be correlated with patterns that show high vote casting levels during the early stages of the initiatives.

Individual behavior of community members. Posting ideas seems to be the main reason that drive people to IM. In addition, we found that most postings occur during the same day of registration. In fact, we detected that members experience a quite active period right after registration and then become inactive. However, visible differences between archetypes were also discovered here: Members of communities supported by companies or official institutions remain active for longer periods than members in self-driven IM. We also found that the activity levels for the actions studied in this paper evolve following similar patterns (notice the similar pattern shapes in Figure 4). This may be explained by the short time that passes between user registration and the actions associated to content creation.

Limitations and future work. The findings we report in this paper are tightly connected to the platform we chose for our study (IdeaScale), and, of course, they should be interpreted within this context. We are also aware that the study is limited by its descriptive nature and we therefore could not investigate causal effects. The analyses we carried out in this work may also suffer from the lack of consideration for "lurking" variables, such as unattractive discussion topics, low promotion efforts, incentives, unclear participation rules, and timing of our observation. In spite of these limitations we hope that future research will draw on, refine, and further articulate the community archetypes identified in this paper.

An interesting question for future work that emerged from this research is the early identification of the point when the activity levels transition from an increasing phase to a decreasing one. In addition, researchers might investigate and understand what conditions may delay or speed up such phase transition, and how we can use such new knowledge to provide recommendations to organizers and moderators so that they can take corrective actions. We are also interested in exploring ways to leverage social networking sites, such as Facebook and Twitter, for communities organizers to increase participation in their IM communities.

Acknowledgment

This research was supported by the project Participa (14-INV-102) funded by CONACYT Paraguay.

References

[1] M. Klein and G. Convertino, "An embarrassment of riches," *Communications of the ACM*, vol. 57, no. 11, pp. 40–42, 2014.

[2] A. Westerski, C. A. Iglesias, and T. Nagle, "The road from community ideas to organisational innovation: a life cycle survey of idea management systems," *Web Based Communities*, vol. 7(4), pp. 493–506, 2011.

[3] S. Hrastinski, N. Z. Kviselius, H. Ozan, and M. Edenius, "A review of technologies for open innovation: characteristics and future trends," in *HICSS 2010*, 2010, pp. 1–10.

[4] M. Schoultz, "My starbucks idea and the value of business crowdsourcing," http://www.digitalsparkmarketing.com/innovation/my-starbucks-idea, accessed: 16/06/2013.

[5] P. M. Di Gangi and M. Wasko, "Steal my idea! organizational adoption of user innovations from a user innovation community: A case study of dell ideastorm," *Decision Support Syst.*, vol. 48(1), pp. 303–312, 2009.

[6] H. Landemore, "Inclusive constitution-making: The icelandic experiment," *Political Philosophy*, vol. 23(2), pp. 166–191, 2015.

[7] J. Surowiecki, *The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business, economies, societies, and nations.* Doubleday & Co, 2004.

[8] A. J. Kim, *Community building on the web: Secret strategies for successful online communities.* Addison-Wesley, 2000.

[9] M. Klein and G. Convertino, "A roadmap for open innovation systems," *Journal of Social Media Organizations*, vol. 2, no. 1, 2015.

[10] J. Footen and J. Faust, *The service-oriented media enterprise: SOA, BPM, and web services in professional media systems.* CRC Press, 2012.

[11] H. J. Bakker, "Idea management: unravelling creative processes in three professional organizations," Ph.D. dissertation, Vrije Universiteit Amsterdam, 2010.

[12] B. P. Bailey and E. Horvitz, "What's your idea?: a case study of a grassroots innovation pipeline within a large software company," in *SIGCHI-Human Factors in Computing Systems 2010*, pp. 2065–2074.

[13] M. Klein and L. Iandoli, "Supporting collaborative deliberation using a large-scale argumentation system: the mit collaboratorium," 2008.

[14] A. Westerski, C. A. Iglesias, and F. T. Rico, "A model for integration and interlinking of idea management systems," in *Metadata and Semantic Research.* Springer, 2010, pp. 183–194.

[15] A. Xu and B. Bailey, "A reference-based scoring model for increasing the findability of promising ideas in innovation pipelines," in *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work.* ACM, 2012, pp. 1183–1186.

[16] S. Faridani, E. Bitton, K. Ryokai, and K. Goldberg, "Opinion space: a scalable tool for browsing online comments," in *SIGCHI-Human Factors in Computing Systems.* ACM, 2010, pp. 1175–1184.

[17] G. Convertino, Á. Sándor, and M. Baez, "Idea spotter and comment interpreter: Sensemaking tools for idea management systems," in *CCT*, 2013.

[18] M. Baez and G. Convertino, "Designing a facilitator's cockpit for an idea management system," in *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work Companion.* ACM, 2012, pp. 59–62.

[19] E. Bothos, D. Apostolou, and G. Mentzas, "A collaborative information aggregation system for idea management," in *Internet and Web Applications and Services, 2008. ICIW'08. Third International Conference on.* IEEE, 2008, pp. 289–296.

[20] J. Björk and M. Magnusson, "Where do good innovation ideas come from? exploring the influence of network connectivity on innovation idea quality," *Journal of Product Innovation Management*, vol. 26, no. 6, pp. 662–670, 2009.

[21] M. Hossain and K. Z. Islam, "Ideation through online open innovation platform: dell ideastorm," *Journal of the Knowledge Economy*, vol. 6, no. 3, pp. 611–624, 2015.

[22] ——, "Generating ideas on online platforms: A case study of my starbucks idea," *Arab Economic and Business Journal*, vol. 10, no. 2, pp. 102–111, 2015.

[23] B. L. Bayus, "Crowdsourcing new product ideas over time: An analysis of the dell ideastorm community," *Management science*, vol. 59, no. 1, pp. 226–244, 2013.

[24] J. Saldivar, C. Rodriguez, F. Daniel, F. Casati, and L. Cernuzzi, "On the (in) effectiveness of the share/tweet button: A study in the context of idea management for civic participation," *IEEE, Internet Computing*, 2016.

[25] S. Khandkar, "Open coding," http://pages.cpsc.ucalgary.ca/~saul/wiki/uploads/CPSC681/open-coding.pdf, accessed: 27/08/2016.

[26] C. S. De Souza and J. Preece, "A framework for analyzing and understanding online communities," *Interacting with Computers*, vol. 16, no. 3, pp. 579–610, 2004.

[27] J. MacQueen *et al.*, "Some methods for classification and analysis of multivariate observations," in *Berkeley Symp. on Math. Statist. and Prob.*, vol. 1(14). Oakland, CA, USA., 1967, pp. 281–297.

[28] E. Wenger, *Communities of practice: Learning, meaning, and identity.* Cambridge university press, 1998.

[29] T. Aitamurto, "Crowdsourcing for democracy: A new era in policy-making," *Publications of the Committee for the Future, Parliament of Finland*, vol. 1, 2012.

[30] T. Aitamurto, H. Landemore, D. Lee, and A. Goel, "Crowdsourced off-road traffic law experiment in finland. report about idea crowdsourcing and evaluation," *Publications of the Committee for the Future, the Parliament of Finland*, vol. 1, p. 2014, 2014.

[31] S.-K. Thiel, U. Lehner, T. Stürmer, and J. Gospodarek, "Insights from a m-participation prototype in the wild," in *Pervasive Computing and Communication Workshops (PerCom Workshops), 2015 IEEE International Conference on.* IEEE, 2015, pp. 166–171.

[32] C. Wang, M. Ye, and B. A. Huberman, "From user comments to on-line conversations," in *ACM SIGKDD 2012.* ACM, pp. 244–252.