On the Genesis of an Assistive Technology Crowdsourcing Community

Christopher Michael Homan
Rochester Institute
of Technology
Rochester, NY, USA
cmh@cs.rit.edu

Akshai Prabhu
Rochester Institute
of Technology
Rochester, NY, USA
ap4170@rit.edu

Jon I. Schull
Rochester Institute
of Technology
Rochester, NY, USA
jonchull@me.com

CHI'17 Extended Abstracts, May 06 - 11, 2017, Denver, CO, USA

ACM 978-1-4503-4656-6/17/05..$15.00.
http://dx.doi.org/10.1145/3027063.3053356

Abstract
The e-NABLE movement is a global confederation that designs, builds, and distributes free, 3D-printed, upper limb assistive devices to children born without fingers and hands. It has been called one of the most inspiring philanthropic efforts of the 21st century. We use social network analysis and natural language processing on the original e-NABLE Google+ community to understand the challenges and opportunities in organizing a rapidly growing real-world social entrepreneurship venture via social media. Our results provide important lessons and benchmarks for similar communities.

Author Keywords
Social media; crowdsourcing; micromanufacturing; time-series analysis; social network analysis; computational linguistics.

ACM Classification Keywords
H.5.3 [Information interfaces and presentation]: Group and Organization Interfaces—computer-supported cooperative work

Introduction
After losing his fingers to a shop accident in 2011, a South African carpenter saw a YouTube video of a giant cable-driven “monster hand” created by an artist from

Washington State. The carpenter contacted the artist and the two then collaborated online on an inexpensive, 3D-printable, prosthetic-like, partial hand. Inspired by a YouTube video of their story, Jon Schull (a coauthor of this paper) posted a comment under the video, inviting makers and those needing prosthetics alike to identify themselves on a Google map mashup he created to gauge interest in scaling up production and delivery of these devices. Six weeks later, there were 70 pins on the map and the e-NABLE movement was born.

It soon moved to a Google+ private community, which subsequentally multipled into approximately 50 local (on- and off- line) communities, several non-profit groups such as the Enable Community Foundation, and 200 classrooms in the United States that have begun developing project-based learning activities based upon the fabrication and assembly of e-NABLE devices (in some cases designed for children or adults in their own local community) [2, 8].

In this case study, we tell the story of the original Google+ community through data. Since this is the earliest comprehensive study of the data from this site, and to provide some basis of comparison to similar research, we chose methods from among those most commonly found in studies of other social media, including LIWC (Linquistic Inquiry and Word Count) and analyses of the structure of social interactions in the community.

Specific Motivations

The reasons we in particular were interested in looking at the data at this time are three-fold:

Vanilla Google+ limitations. The Google+ communities platform is well designed and supported, but it lacks support for the many needs specific to a community-governed, globally distributed micro -design and -manufacturing

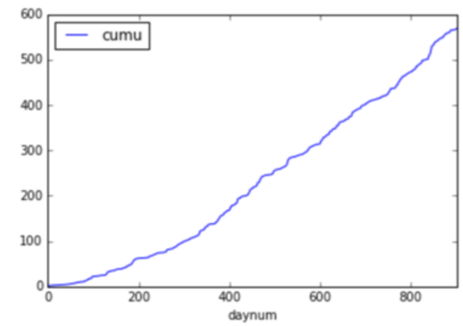


Figure 1: Cumulative mentions of e-NABLE in various mass media shows that public interest in the community remains constant. Data were obtained from news.google.com using keywords such as “e-NABLE” “3D-printed prosthetics”, “enablingthefuture” and “Enable Commmunity Foundation”

community. While we are aware of many such special needs (including those around governance [7, 3] and information recall) we suspect that quantitative analytics may help uncover less obvious needs.

Assessment. To date, the Enable Community Foundation has counted over 2000 artificial hands delivered to needy participants. But what impact do these hands have on the quality of life of their recipients? Can activity on the Google+ community provide any insights?

Tracking the evolution of the community. Figure 1 shows that global media coverage on eNABLE has remained relatively constant over time, yet Figure 2 shows a steady decline in the volume of activity on the original Google+ community. Moreover, the overall volume of activity over all e-NABLE-related online sites known to us has fluctuated but appears to be stable. Is the decline in activity on the original site a sign of weakening in the community or evidence of a

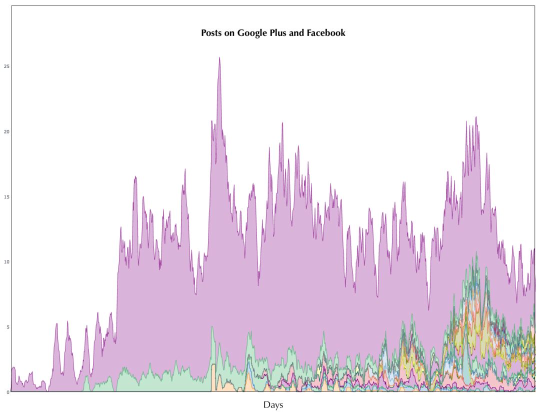


Figure 2: A stacked graph of the volume of posts on various e-NABLE sites per unit time (a moving 9-day window). The Google+ community is the largest, mauve region. The remaining sites are Facebook-based sites that now account for roughly the same amount volume of posts as google plus. There are approximately 30 other known e-NABLE related websites not pictured represented here. While e-NABLE has thus expanded well beyond the Google+ community, there is no other collaborative site that brings so many e-NABLE stakeholders together.

falling off of activity that naturally occurs in many online communities [10]. Conversely, is the proliferation of local communities a sign of health and specialization in the movement or an early sign of disintegration?

Data

Because Google does not offer an API for private Google+ communities, we used the splinter python API[1] to scrape all posts and comments from this website made between July of 2013 and August 2016 1. As a typical Google+ site, it

[1] https://splinter.readthedocs.io/en/latest/

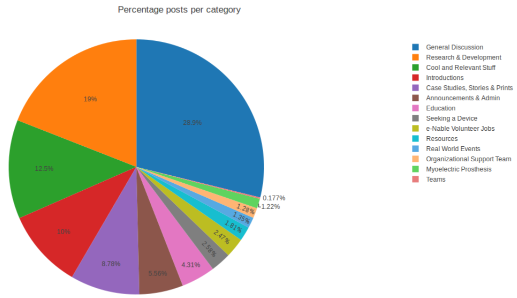


Figure 3: Pie chart showing different categories of posts, the second most popular of which (beyond "general discussion") was "research and development." Note that, over the last summer three new categories were added. We report on only those that existed from near the founding of the site. All posts labeled with one of these newer categories counts here as "general discussion."

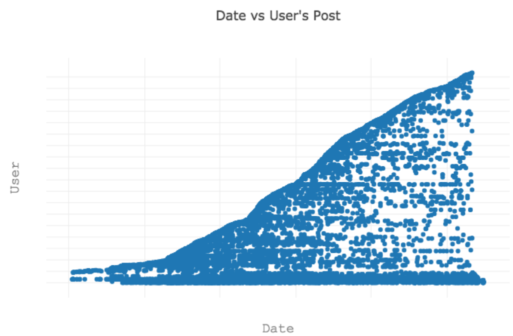


Figure 4: Posts by all users over time. This is a prototype for a clickable interface that will eventually color-code posts by topic and tag, and allow participants to recall posts and patterns of posts that currently get buried under the ongoing stream of new posts.

8,500	Participants
2,815	Active Participants
9,708	Posts
51,530	Comments

Table 1: Summary of data from e-NABLE's Google+ page. "Participants" consists of all registered members of the site. "Active participants" includes only those who have ever posted or commented on the site.

presents each user with a vertical stream of members' *posts*. Each post contains a brief text and/or other media, which the author can assign to one of 20 different topic categories (Figure 3). Figure 4 shows all posts made over time by all community members. Participants may then *comment* on any post or *"Plus-One"* posts and comments.

Over the nearly three years of data we collected, the e-NABLE community progressed through three distinct phases. In the first year, the community discovered its identity, and remained at a size where everyone knew everyone else. At the beginning of the second year, Johns Hopkins University hosted a conference for members of the community, and the resulting coverage led to an explosion in media coverage and a period of intense growth in the community. Finally, the last year was characterized by strife and discord as the community split into factions with competing philosophies.

Social Network Analysis

Figure 4 shows the distribution of posts by (anonymized) user over time. We performed time series analysis on the structure of these interactions. For each sliding 20-week window we construct a graph over all the community members where there is an edge between each pair of users where one of them comments on a post made the the other.

On each of these graphs we measure:

1. **Diameter.**
2. **Average shortest path length**. A more robust alternative to diameter. Lower values in either are associated with "small world" graphs, which are a hallmark of most natural social networks.
3. **Largest connected component (lcc)**.
4. **Number of connected component (ncc)**.
5. **Average degree**.
6. **Average number of triangles**.
7. **Average clustering coefficient**: for a given node x, the clustering coefficient is the proportion of pairs of x's friends who themselves are friends.
8. **Average core number**: The k*-core* of graph is a maximal subgraph in which each vertex has at least degree k. The *core number* of a vertex is k if it belongs to the k-core but not to the $(k+1)$-core. It is as a computationally tractable alternative to the *clique number*.
9. **Closeness centrality**: for a node u this is defined the reciprocal of the sum of the shortest path distances from u to all other nodes in the graph, times the length of the longest possible path on all graph of the same size [4].
10. **Core number centralization**: The sum of differences in coreness between each node and the node of maximum coreness in the graph, divided by the sum of differences in coreness between each node and

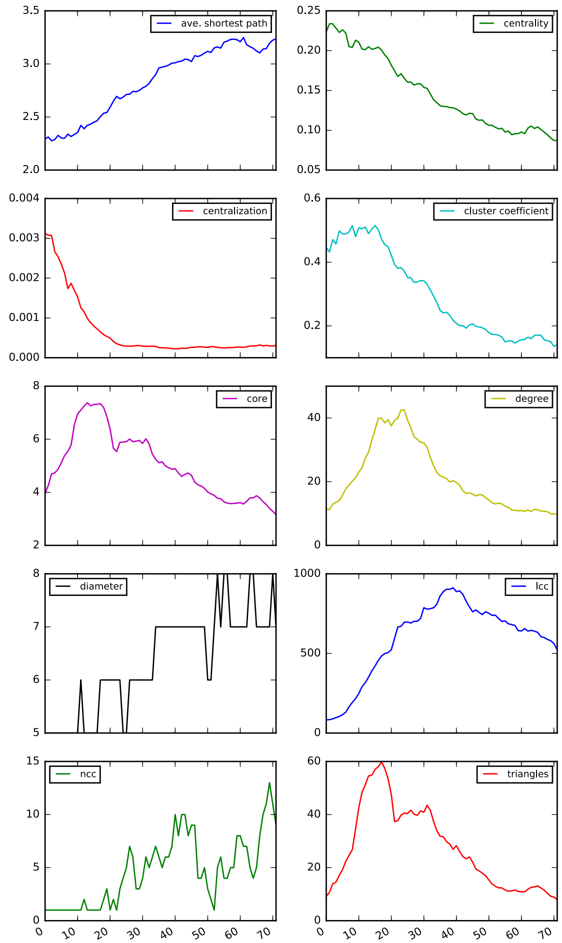
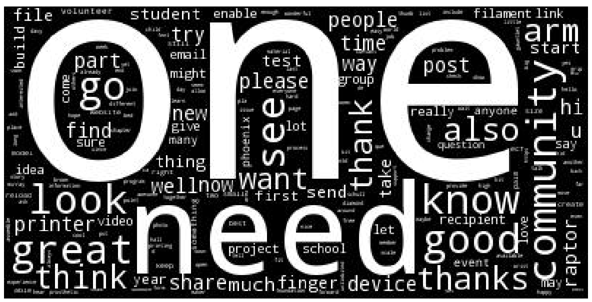


Figure 5: Statistics for the time series on the social graph. Each unit on the x-axis represents two weeks.

the maximum coreness possible over all graphs of the same size [4].

(a) First 40 weeks



(b) Last 40 weeks

Figure 6: Wordclouds of e-NABLE data for the 40-week windows at the beginning and ending of the study period. The following were the most frequent across all time scales and were removed: 'use', 'http', 'like', 'just', 'get', 'make', 'work', 'hand', 'will', 'design', 'can', 'print', 'help.'

Figure 5 shows clear and dramatic trends. For instance, the largest connected component—an index of the portion of the community that functions as an interdependent information economy—increases monotonically until around

$t = 40$ (each time unit is two weeks long; since the windows from which edges are drawn are each 20 weeks wide, there is significant overlap between the edges at each step step). During that same period of growth, two measures of network density–triangles and core number–increase at first, but then drop precipitously at $t = 20$, even as the size of the graph increases and average degree holds steady. This shows that as the network increases we might expect community members to feel less like they are part of an integrated whole.

One explanation for these changes involves Dunbar's number [6], and has been observed in models of social network growth [5]. Before $t = 20$, the network, with only 500 active participants, did not substantially exceed the maximum number of social relations that any individual can maintain. We speculate that once the size of the network became substantially greater than this cognitive limit, members became more selective in establishing links to other members; as a result, the overall graph became more sparse, friend clusters did not overlap much as as they did before, and there were fewer triangles.

Linguistic Analysis

We begin with a naive ranking of the most frequently used words over time. For each forty-week time windows, we normalize and aggregate all posts and comments into a single text document. Looking at only the first and last time window there is considerable variation (Figure 6). Part of this is no doubt due to the long-tailed nature of the distribution of words in texts, but some of it appears to reflect trends over time. For instance, "arm" becomes more frequent, which could be because e-NABLE originally focused on hand designs only and later branched out to arms. "Community" also becomes more frequent over time, a trend that may reflect an increasing amount of discussion we observed on the site about the future of the community.

LIWC (Linguistic inquiry and Word Count) [9] is a collection of approximately 80 lexicons for linguistic categories such as pronouns, anger, positive emotion, etc. For each time window we construct two documents, one from all posts and the other from all comments (Figure 7—we report only on those categories for which LIWC scores were above 1% and ranged over 10% of the average score and omit the comment scores due to lack of space).

We next used Latent Dirichlet allocation (LDA) *topic modeling* to further characterize the community's collective discourses. [1]. We concatenate each post and the comments under it into a single document. Starting with 19 (the number of categories shown in Figure 3), we ran LDA on the corpus of documents formed in this manner and used gradient descent to estimate the most likely (in terms of perplexity) number of topics.

Figure 8 shows the results for 19 topics. They do seem to be suggestive of real categories of concern in the community. Topic 1 is related to the conventional protocol a new participant uses when he or she has produced her first hand. Topic 2 is about underserved international populations. Topic 3 is Portugese. Topic 6 is about the manufacturing process. Topic 7 seems to be about materials and components. 9 is about printing. 10 is about design features 12 is about schools and classrooms (many schools participate in eNABLE). 15 is expressions of enthusiasm (perhaps about the unlimbited arm model, another e-NABLE design). 16 is a plea for volunteers for specific service to the community. 17 is about the pieces used to make the arm more comfortable to participants. 18 was planning about an e-NABLE conference at Johns Hopkins. 19 is about design software.

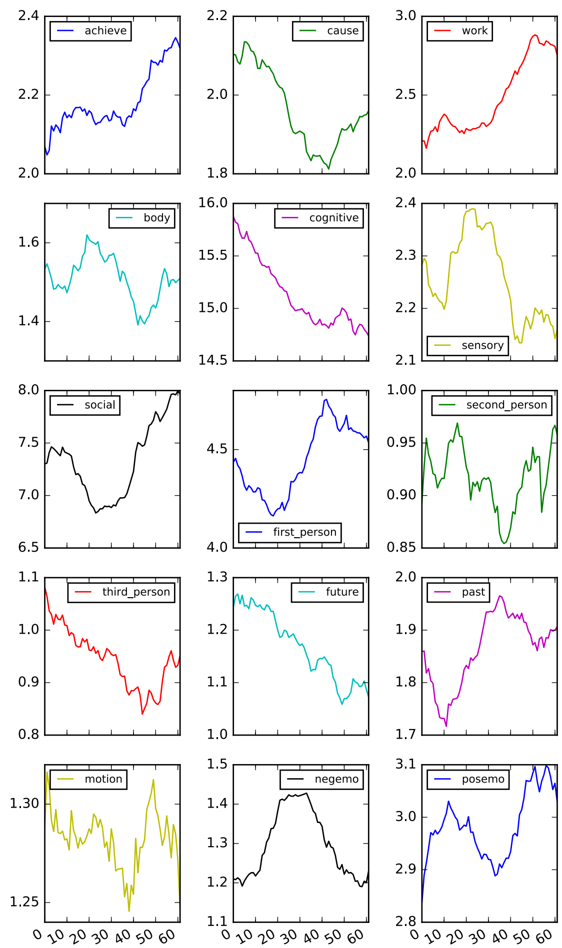


Figure 7: LIWC scores of all posts occurring during a sliding 40-week time window. Each unit on the x-axis represents two weeks.

1. hand can get help email know just need work group print welcome match send people one form will sure make
2. translate veteran downtown viz torn disability delivery spanish china nonprofit translation nigeria asset dhl structural truck savvy shed maxence quadriplegic
3. de que la y en para e o el reportage se wiltgen con una un por brazilian brazil los si
4. device hand use like one cost prosthetic can need think say get make child sensor time http control really lot
5. change measure use matcher hello thank learn stl slicker size per can raptor ma file e hand sono hug weed
6. work can use finger hand get way tuner like see make cast will mould material just video look mold grip
7. elastic hand nylon print cable work finger one joint try screw just wire hinge like design use filaflex wood also
8. lego acetone paint vapor polish nail ab dye dark smooth epoxy biscuit victor strand merry cylinder dim christmas http hydraulic
9. print printer use filament try will part can get just good pla support look think like material see work one
10. hand design finger think can use make wrist work idea thumb like also will might grip just want motion much
11. hand file use can will print make raptor need one scale just get http work size fit look arm forearm
12. design work student school enable base also can thanks http see one r project will teacher d good interested make
13. hand help can work make arm get design like will thank look video hi know see http great please need
14. logo hand graphic font great first france paris thierry la hour thanks french le can please print community un one
15. work great red arm today beautiful job thanx share fantastic wonderful color first unlimbited love wiltgen davy team story blog
16. can get work people community will make http need help like go post u want just one see please think
17. leather van shirt de osprey sock bracer argentina wear vote stud ik shin het en swim fin erik een padding
18. design printer hand get just think will know great go send ross kane print time people conference one make love
19. fusion design free bastian parametric resident software solid autodesk solid-works model will use body sandbox just thanks please watch andreas

Figure 8: The topics discovered over the postings corpus, and the 20 most probable words in each topic

We next ask whether LDA can tease apart the various *roles* that users play. We created one document per Google+ participant by concatenating together all posts and

1. use hand can like work design make think get one just will finger good also look need see print well
2. de hi o thank regard que e eu help para je em brown um raptor brazil com fantastic good Ãľ
3. de que la y en el para con e una un se por los si te lo como hola mi
4. http recipient form intake match email matching welcome system team enable enablematcher design add free circle complete volunteer r contact
5. print use hand can get will work just make good one printer http file look need part try see finger
6. http hand can get one like post people design community help work see just will enable u print well need
7. student school enable hi work educator great hand thanks http teacher see project will please get exchange love also can
8. ashok bilan dash ot therapy sri niĀśa peace sl pune disability hope south downloads lanka j teresa tks activity server
9. di e per il che con la sono ho mi un una se grazie ciao le della del protesi mano
10. can hand will get help thank thanks work make know need just great one like see go look send email

Figure 9: The topics discovered over the users corpus, and the 20 most probable words in each topic.

comments made by each and ran LDA on this corpus, fixing the number of topics to the number of roles we have identified. Topics 2 and 3 are Portuguese and Spanish, respectively. Topic 4 is about policy and procedures. Topic 5 is about the printing process. Topic 6 is about the community itself. Topic 7 is education related. Topic 8 is about southeast Asia and topic 9 is Italian. Topic 10 seems to be and "newbies."

Discussion and Conclusion

Recall that e-NABLE went through three phases of development, each lasting about one year; many of the changes in the network and LIWC scores seem tied to these one-year marks. For instance, the network shows a precipitous drop in clustering coefficient and triangles—signifying that groups of individuals in the graph are much less likely to all know each other than before—around the 1/3 mark, when the Johns Hopkins conference occurred. The largest connected component shows a drop just before the 2/3 mark, followed by a slower decline—a sign of outright social fragmentation. All of these signs are indicative of a community that grew beyond the size of a small organization, where self-governance and management are easy, to one that fragmented due in part to a lack of hierarchical organization.

Meanwhile, of the LIWC trends, the strongest seems to be that first-person language grew significantly in the middle year, but dropped in the outer years. This could be because in year one much of the discourse was on establishing a group identity and basic work practices. In the last year there was a focus on resolving tensions between factions in the community. In both of these outer years, members were thus perhaps focused less on themselves and more on the community than in the middle year.

Interestingly, negemo (negative language) peaked and plateaued early in the middle year. We had expected a peak in the final year, as community members aired grievances. Perhaps the early appearance of negative language revealed the beginning of tensions, and when these tensions were finally acknowledged and grievances aired in the last year community members were already choosing their words more carefully.

In sum, these social media measures seem to capture some of the basic contours of what we know to have happened in the first three years of the e-NABLE community. Certain measures, such as the shifts in clustering coefficient and triangles, and in the rise of negemo, might even constitute "early warning" signals for small communities as they growth to a scale where informal self-governance is less effective.

REFERENCES

1. David M Blei, Andrew Y Ng, and Michael I Jordan. 2003. Latent dirichlet allocation. *Journal of machine Learning research* 3, Jan (2003), 993–1022.
2. Erin Buehler, Shaun K Kane, and Amy Hurst. 2014. ABC and 3D: opportunities and obstacles to 3D printing in special education environments. In *Proceedings of the 16th international ACM SIGACCESS conference on Computers & accessibility*. ACM, 107–114.
3. Noshir S Contractor, Leslie A DeChurch, Jay Carson, Dorothy R Carter, and Brian Keegan. 2012. The topology of collective leadership. *The Leadership Quarterly* 23, 6 (2012), 994–1011.
4. Linton C Freeman. 1978. Centrality in social networks conceptual clarification. *Social networks* 1, 3 (1978), 215–239.
5. Emily M Jin, Michelle Girvan, and Mark EJ Newman. 2001. Structure of growing social networks. *Physical review E* 64, 4 (2001), 046132.
6. Jure Leskovec, Kevin J Lang, Anirban Dasgupta, and Michael W Mahoney. 2009. Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters. *Internet Mathematics* 6, 1 (2009), 29–123.
7. Kurt Luther, Casey Fiesler, and Amy Bruckman. 2013. Redistributing leadership in online creative collaboration. In *Proceedings of the 2013 conference on Computer supported cooperative work*. ACM, 1007–1022.
8. Kristen Oliver. 2014. Hall County teacher builds student a prosthetic hand. *Gainesville Times* (December 19 2014). `http://www.gainesvilletimes.com/archives/106701/`
9. James W Pennebaker, Ryan L Boyd, Kayla Jordan, and Kate Blackburn. 2015. The development and psychometric properties of LIWC2015. *UT Faculty/Researcher Works* (2015).
10. Clay Shirky. 2005. A group is its own worst enemy. In *The Best Software Writing I*. Springer, 183–209.