

Visually Exploring Social Participation in Encyclopedia of Life

Jae-wook Ahn*, Jennifer Hammock[‡], Cynthia Parr[‡], Jennifer Preece[†], Ben Shneiderman*, Katja Schulz[‡],
Derek Hansen[§], Dana Rotman[†], and Yurong He[†]

*Department of Computer Science & Human-Computer Interaction Lab, University of Maryland, Email: {jahn,ben}@cs.umd.edu

[†]College of Information Studies, University of Maryland, Email: {preece,drotman,yrhe}@umd.edu

[‡]Smithsonian Institution, Email: {hammockj,cparr,schulzk}@si.edu

[§]School of Technology, Brigham Young University, Email: dlhansen@byu.edu

Abstract—Social network visualization is useful for understanding the complex structure of collaborative efforts such as citizen science projects. It has been widely accepted by social network analysts for exploring and analyzing networks by visually showing their members, the relationships among them, and their behaviors and attributes. The strength of social network visualization can be increased even further, by incorporating the time dimension of evolving networks. We analyzed the conversation network of a collaborative citizen science web platform called the Encyclopedia of Life using dynamic network visualization methods. This paper shows how the temporal visualization was applied to the social collaboration analysis of EOL and presents the findings. We found that some EOL web site features increased the interactive as well as individual member activities. We also found evidence that EOL curator activities encouraged the activities of other members.

I. INTRODUCTION

Citizen science is a collaborative effort involving members of the public in scientific research [1]. Unlike traditional scientific inquiries conducted by professional scientists, collaborative citizen science projects include citizen scientists who contribute voluntarily and lack formal credentials and professional positions [2]. It is a form of collective intelligence or mass collaboration, as the project goals are achieved by a large number of people and it frequently involves large scale data. Several models have been suggested to facilitate collective intelligence research [3]–[6] which stressed the potential of social network analysis including visualization [6].

Network visualization is a useful for understanding the complex structure and characteristics of networks. It has been widely used for analyzing various forms of networks including social networks, citation networks, affiliation networks, and disease transmission networks. Most recently, with the popularity of social media, many social communications are studied using network visualization methods. Network visualization has clear advantages compared to other traditional social network analysis methods. It can help users quickly understand large and complicated network structures, investigate complex relationships, identify important members or sub-communities, and discover anomalies and interesting patterns.

Because of these advantages, many professional social network analysis software packages support network visualization as their core feature. Most recently, an open source Microsoft

Excel extension for network visualization called NodeXL (<http://nodexl.codeplex.com>) [7] was introduced, which made network visualization tasks much easier for general users. However, despite the popularity and ubiquity of network visualization, few approaches support analyzing dynamic networks. Most network visualizations focus on static networks that simply show a network snapshot in a single time point or an aggregated overview of multiple time points. Those static networks can hardly investigate dynamic social interaction. This limitation can be critical for analyzing social participation networks. Therefore, we investigated an interactive method to explore temporal network evolution over time.

Using temporal network visualization, we analyzed the social network of a web-based citizen science site called Encyclopedia of Life (EOL, <http://eol.org>). EOL is a “content curation community” with the goal of constructing an online database and website with authoritative data about every biological species on the planet [8]. It houses diverse forms of information including species descriptions and photos that are collected and aggregated from various sources, including contributions by volunteer citizen scientists. It is important to collect species data and encourage member contribution for protecting biodiversity, so we focused on the online communications of the EOL citizen scientists and tried to find out what could encourage their participation using network analysis methods. We were able to show that new software features added to the site encouraged interactive participation and that the key members who managed and helped others were important.

Section II describes related work on network analysis and network visualization. Section III introduces EOL and its building blocks. Section IV and V show the research questions and data collection method. The EOL conversation network analysis and the network visualizations are presented in Section VI. The final section concludes this paper.

II. RELATED WORK

Network analysis has been widely applied to various biology or health sub-fields, such as protein structures [9], genetic networks [10], neural networks of the brain [11], epidemics or

disease transmission networks [12]. Saraiya et al. [13] evaluated five popular visual microarray visualization tools in order to identify the characteristics of insights and provide guidelines to biologists. They found that larger software packages worked consistently across different datasets whereas more focused tools provided better results with one kind of data. Powell and White [14] analyzed the collaboration network in the life sciences. Christakis and Fowler [15] analyzed a large social network in order to understand the spread of obesity.

In particular, social network analysis and visualization is useful for managing online participants in large open structures such as Wikipedia [4]. Therefore, it has been actively applied to studying and supporting highly user participation-oriented social media sites [16]–[18] and citizen science projects. Alabri and Hunter [19] exploited social network analysis tools for enhancing citizen science data quality. Sullivan et al. [20] studied the bird observation network of citizen scientists even though they didn't directly utilize conventional social network analysis techniques. Howison et al. [21] studied static and dynamic communication networks of an open source software community using social network analysis methods.

Visualization has long been used for showing the topological structure of networks and the relationships among the network nodes. Various networks have been analyzed using visualization techniques. Moreno's sociogram [22] has been used from the very early days of social network analysis. There are multiple visualization techniques for social network analysis, including the popular force-directed placement, adjacency matrix, or heat map [23]–[25].

Despite its popularity, many network visualization techniques remain in static network analysis. However, in order to better understand the dynamic nature of networks, it is becoming more important to visualize temporal network changes. For social network analysis, Moody [26] pioneered a way to use dynamic network visualization techniques. In the SoNIA system, he used dynamic network animation in order to analyze sociology problems such as social balance, fraternity networks, and social communication in classrooms. Powell visualized the affiliation network of life science institutions [14]. TimeMatrix was used for temporal analysis of inter-organizational collaboration networks [25]. C-Group visualized the change of affiliation networks of co-authors [27]. The life sciences and biology have also paid attention to temporal network visualization for analyzing the life science organizations [14] or spread of disease [12], [15].

Dynamic network visualization is a good tool for analyzing online social networks as well as traditional offline social networks. Mutton's PieSpy visualized IRC communication [28]. Gloor built a system called iQuest and visualized dynamic online communication networks including email, phone records, blogs, etc [29]. Recently, microblogs or Twitter became a good target for dynamic network visualization too in order to show temporal growth [30], political message propagation [31], change of important people and events in Twitter networks [32], [33], and temporal change of popular music [34]. Ahn et al. created a taxonomy of temporal network analysis tasks

using three dynamic network dimensions: entities, properties, and temporal features [35].

These dynamic network visualization techniques can be classified by various criteria. One interesting criterion is how to deal with the dynamics of the network layout: static or dynamic layout. Some approaches dynamically change the position of the network nodes within the visualization according to the evolution of networks [12], [26], whereas others make the positions fixed [36], [37]. These approaches have both benefits and shortcomings, so Peterson proposed an approach to keep the balance of the benefits of keeping node positions relatively static while allowing some layout adjustment [38]. A comprehensive review of dynamic network visualization layout techniques can be found in [38].

III. ENCYCLOPEDIA OF LIFE

The Encyclopedia of Life (EOL) is a collaborative effort that aims to build an online database and provide global access to knowledge about life on Earth. It gathers, generates, and shares knowledge in an open, freely accessible and trusted digital resource. As of August 2012, it has collected information for over one million taxa, and has more than 60,000 registered members. It includes the descriptions of the taxa, their images (over 1.7 million), biological classification hierarchies, and various comments from its participants. The members can contribute by adding comments on the data objects (e.g. species text, images, or taxa), rating objects, adding text, creating collections, joining communities, or becoming a curator.

A. EOL Objects

EOL involves a variety of activities and objects. Below is the list of the most representative information types of EOL (<http://eol.org/info>). Table I shows the statistics of them.

(1) **Taxon Pages** – EOL gathers species information from various content partners and displays them on appropriate taxon pages. Taxon means a group of organisms related by lineage.

(2) **Data Objects** – The information objects about species such as images, sound, text, and movies. The data objects have their own pages and EOL members can post comments on them.

(3) **Collections** – EOL members can create their own virtual collections of EOL objects that they can name, annotate, and share. EOL Collections can include just about anything in EOL: Taxon Pages, images, articles, even other EOL Collections, EOL Communities and individuals. For example, a collection named “Birds of D.C.” includes birds found in the D.C. area by aggregating 146 taxon pages, 46 articles, 18 images, 2 sounds, and one video. Four EOL members act as managers in the collection.

(4) **Communities** – People who share a common interest in a particular aspect of living nature come together in EOL communities to ask questions, share their expertise, and identify opportunities to improve EOL.

(5) **Members** – Members of EOL. Each member has her own profile page. Special members called curators are explained in more detail in the next section.

TABLE I
EOL OBJECT STATISTICS

Taxon Pages	Data Objects	Collections	Communities	Members
3,366,427	20,023,376	28,425	149	67,151

TABLE II
EOL CURATOR STATISTICS

Curator Level	Assistant	Full	Master
Count	136	947	2

B. Curators

Curators are special members who manage EOL content quality. Curators are differentiated from regular members, in that they help review and evaluate resources that EOL members have contributed to the service, answer their questions, and examine whether the resources have any incorrect information. Any member can participate as a curator but three different curator levels are distinguished in order to match curator responsibilities with each curator’s expertise and experience: (1) Assistant curators, (2) Full curators and (3) Master curators. Assistant curators do not require any credentials but have limited curator privileges. Full curators are professional scientists or EOL community members promoted from assistant curators. Master curators are experienced EOL curators who have demonstrated an exceptional level of engagement with the project. As of July 2012, the numbers of curators by their levels are as in Table II.

IV. RESEARCH QUESTIONS

Through discussion with EOL staff members, we defined the following research questions.

(RQ1) Do new EOL web site features such as forwarding Flickr comments to the EOL site, newsfeeds, and notifications increase member activities?

(RQ2) Do special EOL members such as curators encourage other member activities by commenting on their posts or administering contents?

Certain people or facilities can help EOL members and promote their activities. Among other factors, the existence of curators could be unusually important compared to other collaborative activities. The curators are experts and can influence non-curator member activities. With their intervention, other EOL members could participate more actively.

In order to answer the first question, we identified three events that each added a new feature to EOL (Section VI-B). If an activity increase appears after these events, we could justify the positive effect of the new features. For the second question, we examined the activities of two types of EOL members regarding their connection to the curators in the conversation network (Section V).

(1) Members who participated in the conversations with the curators in the conversation network.

(2) Members who participated in the conversations only with the non-curator members.

TABLE III
NUMBER OF COMMENTS IN EOL

Collection	Community	Data Object	Taxon Page	Member	Total
539 (3.4%)	927 (5.9%)	8,366 (53.0%)	4,651 (29.5%)	1,306 (8.3%)	15,789

V. CREATING THE EOL SOCIAL NETWORK

Various types of social activities happen on EOL. Members share species information, exchange opinions about them, and those with more expertise can review the other members’ collaborative efforts. Among them, we chose the conversation activities on taxa and data objects in order to construct the social network in EOL. It is the most active type of contribution with sufficient information and can reflect better what really happens among the members than other social activities. EOL generously gave us their dataset containing the conversation data.¹ It includes data from January 2009 to July 2012.

The EOL member conversations are found in comments. Members can post comments on five object types (Section III-A): Collection, Community, Data Object, Taxon Page, and Member. They can post an opinion on a species photo (Data Object), an organism itself (Taxon Page), or on another members profile page (Member). Figure 1 shows examples of data object comments and taxon page comments. Table III shows the number of comments posted on each object. EOL members most frequently posted comments on Data Objects and Taxon Pages, which comprise more than 92% of the entire comments. Therefore, we used the comments posted on these two most common objects for creating the conversation network.

In addition to comments, there is another source of conversation, particularly from the curators. Curators can *trust* or *untrust* EOL content objects and these actions are visible to the other members. Even though it is not a typical verbal communication, we decided to include it as a part of the conversation because it constitutes a statement that either (1) “I positively approve this object’s quality” or (2) “I question this object’s quality or disagree with the species identification.”

We defined the conversation network as the links between members who posted comments and/or trusting actions on the same object. For example, if member A posted a comment on a picture of a bird and member B posted another comment on the same picture, we assumed that they were in a conversation and connected them in the network (Figure 2). The choice of time window size in which conversations took place dramatically influences the analysis. Two people are connected only when their comments fall within the time window [39]. After consulting with EOL staff members, we set the size of the *time window* as one week. That is, if multiple members commented on the same object in the same week, we accepted that they were in the same conversation and socially connected. One should be careful about the size of the time window because it can strongly influence the analysis.

¹The member identities were anonymized in order to protect their privacy.

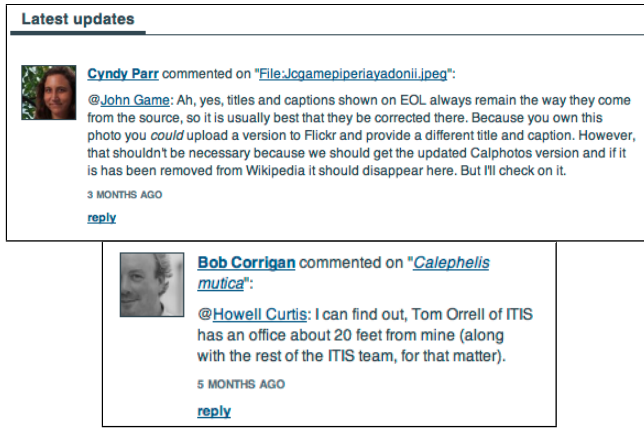


Fig. 1. Comment examples – on a data object and a taxon page

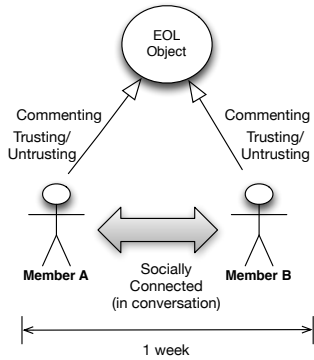


Fig. 2. Creating the conversation network of EOL. Multiple members should comment on the same object in the same week to create a link in the network.

We created the network and visualized it using a standard Force-directed layout method in Section VI-A using NodeXL. In order to better understand the dynamics of the network, we also created weekly snapshots of the network evolution and visualized it using a tool called TempoVis (Section VI-C) [36]. The visualizations and analysis results are presented below.

VI. ANALYSIS

A. Overview of the EOL Comment Network

Figure 3 shows the conversation network of EOL created using NodeXL. In EOL, members and curators can freely add comments on data objects or taxa and make conversations like in online forums. If multiple members add comments on the same target in the same time window (1 week), they are considered connected in the network. This network includes curator activities in addition to pure commenting behaviors. Curators can approve the target object quality by *trusting* or *untrusting* them. The four small dyads on the lower left are not connected to the main component because they just had conversations between themselves. The nodes represent curators and the non-curators. The node shape discriminates the level of membership: (1) disks (●) are full curators, (2) squares (■) are assistant curators, and (3) triangles (▲) are non-curator members. The curators (assistant or full) are painted in blue too. Node size and link thickness are proportional to node betweenness-centrality and comment frequency respectively.

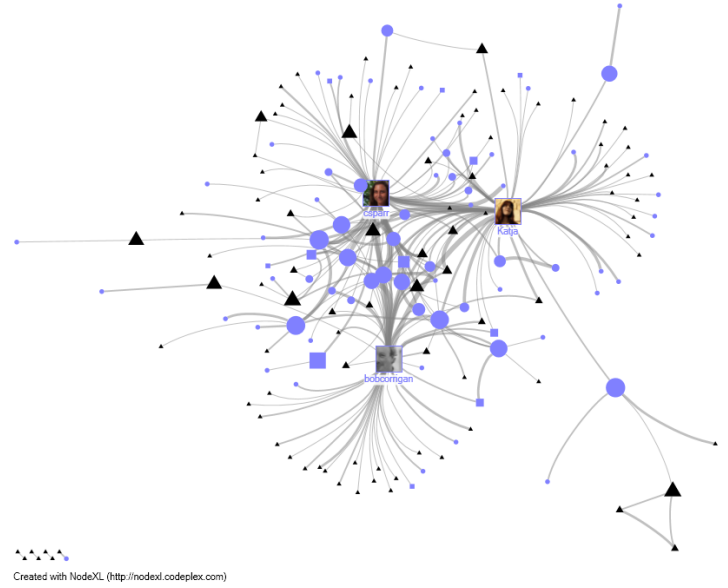


Fig. 3. EOL comment network visualization: Disks and squares are full and assistant curators. Triangles are non-curators. Node size and link thickness are proportional to node betweenness-centrality and comment frequency.

The graph shows a typical scale-free network, where a small number of members play dominant roles and act as hubs. As can be seen in Figure 4, which shows a power law distribution, a very small number of members are responsible for the most of the conversations and the remaining members made just one comment. 40 members made 35,996 comments (90.5%) out of all 39,757 comments. The mean number of comments made by curators and non-curators were 32.38 and 6.74 respectively, which suggests that the curators are the hubs. We can observe a similar pattern from the network topology in Figure 3. The master curators (disks) tend to locate in the center of the graph while the assistant curators or non-curators (squares and triangles) are in the periphery of the network. The curators show higher betweenness-centrality (size) and degree (number of connections) than the non-curators. Even among the curators, there appear three curators (with EOL member images) who are responsible for a significant number of conversations in the comment network. We can identify their EOL usernames (under the images) and can verify that they were playing significant roles as staff on the EOL project.

B. Comparing Events and Conversations

In order to test the first research question, we compared the EOL member activities with three important events. The EOL staff members suggested examining the following time points.

(1) **Feb 25, 2011:** EOL comments on image objects provided from flickr begin to be forwarded from EOL to Flickr.

(2) **Sept 5, 2011:** EOL version 2 launches; this entails newsfeeds of activity for each member, on each taxon page and data object, and on the EOL homepage. The homepage feed in particular is quite visible (lots of very invested members, curators, etc., arrive on the homepage and might see something they want to react to.)

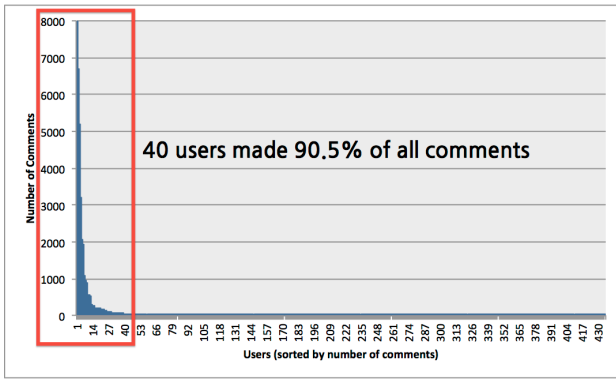


Fig. 4. Comment activity distribution. Each vertical bar is a member and the height represents her comment count. It shows a power law distribution where a few people make most of the comments (left-most peaks) and many other people make just one or two comments (flat line close to 0).

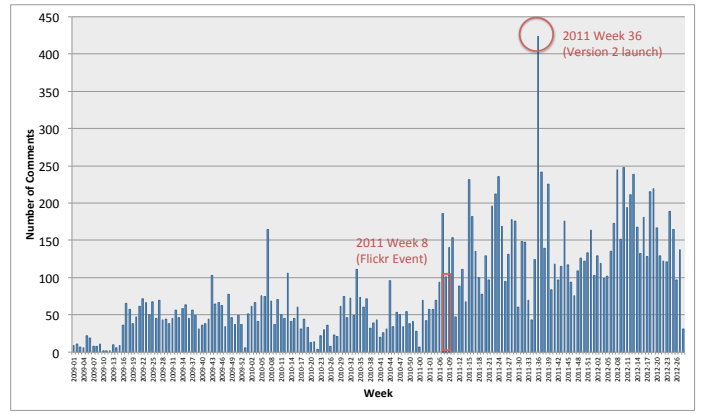
(3) **May 2012:** Email notifications begin. EOL activity in an individual member's newsfeed is sent to the member in immediate or (daily/weekly) digest emails.

We tried to answer whether these events would stimulate the member activities and we would be able to observe sustained increases in the comments activity timelines. Figure 5 compares the three events. Each bar represents the number of comments or conversations per week. They are sorted chronologically from 2009 week 1 (labeled “2009-1” on the x-axis, January 2009) to 2012 week 26 (“2012-26”, July 2012).

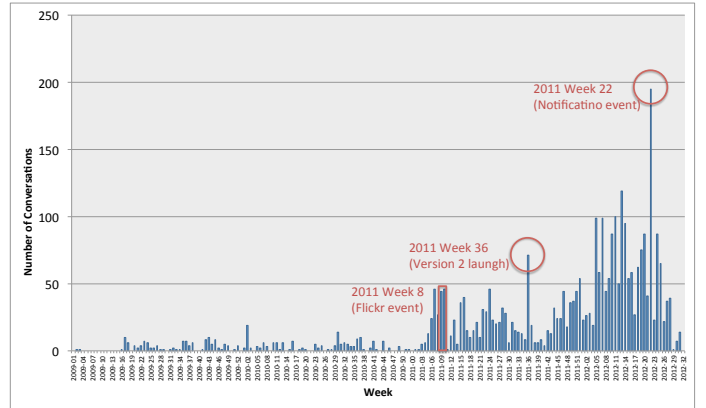
In the first graph (Figure 5(a)), we counted all comments made by the EOL members each week. Here, a sustained increase is observed beginning in 2011 week 8, corresponding with the first event, flickr comment forwarding. A peak is observed in 2011 week 36 (marked in red), which exactly matches with the launch of EOL version 2. However, there is no corresponding sustained increase. No definitive activity change was observed regarding the email forwarding events.

This graph (a) counted the absolute number of comments regardless of whether they are comments of a single member or conversations between multiple members. However, in the second graph (Figure 5(b)), we only counted the conversations of multiple members. They are comments posted on the same object in the same week, as explained in Section V. Here, we can observe similar patterns. A sustained increase in activity is still visible beginning in 2011 week 8. Peaks in activity are also visible. The third event (email notifications) is more evident while the second event (version 2 launch) is still significantly visible.

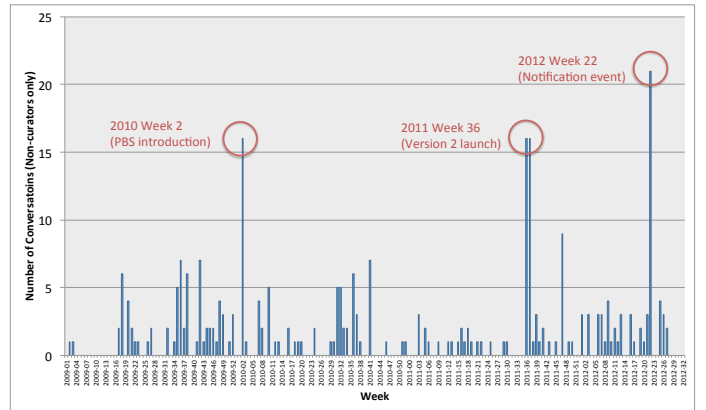
The last graph (Figure 5(c)) shows the conversation of non-curator members. Again, peaks in activity during the second and the third events are clearly visible from the timeline. One interesting incident here is a peak in 2010 week 2, which was not visible at all in the previous two graphs. We suspect that it matches with a re-airing of the EOL introduction on PBS, which drew the attention of non-curator members. No change in non-curator conversations is visible following the first event, flickr comment forwarding, and no important sustained changes in activity are visible anywhere in this timeline. From



(a) Number of weekly comments – 2011 week 2 (Flickr event) and 2011 week 36 (EOL version 2 launch)



(b) Number of weekly conversations – 2011 week 2 (Flickr event), 2011 week 36 (EOL version 2 launch), and 2012 week 22 (Notification added)



(c) Number of weekly conversations (non-curator members) – 2010 week 2 (PBS introduction), 2011 week 36 (EOL version 2 launch), and 2012 week 22 (Notification added)

Fig. 5. EOL comment/conversation frequencies

these three graphs, we find the following observations.

(1) The EOL events are more closely related to the interactive conversations of EOL members, rather than isolated individual commenting activities (Figure 5(b)).

(2) Among the three events, the EOL version 2 launch and notification events did not have a sustained effect on activity rates. The Flickr event does precede a sustained increase in

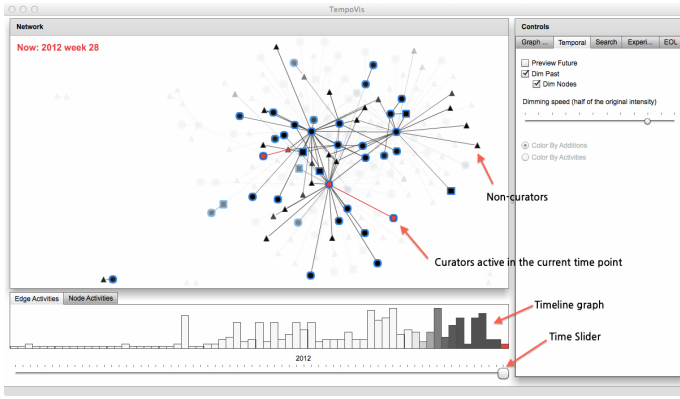


Fig. 6. Dynamic network visualization using TempoVis (2012 week 28). Using the time slider on the bottom, one can explore through time and see a network snapshot of the corresponding week. The nodes mean the EOL members, the links mean the conversation between them, and the blue bordered nodes mean the curators. Red nodes and links mean that their activities are in the current time period (2012 week 28).

total comments and in conversations, but it had no effect on conversations that did not involve curators.(Figure 5(b), 5(c)).

(3) The activity peaks were not solely caused by the EOL curators. Other regular members contributed significantly to the peaks too. However, sustained activity increases depend upon the involvement of curators (Figure 5(c)).

C. Time-based network analysis

Dynamic Network Visualization. Even though the visualization in Section VI-A provides an interesting observation about the curators' importance in the conversation network, it does not provide insights about the temporal change of the network. We need to learn about how the network structure changes and how the EOL members' behaviors change over time according to their different roles. Therefore, we created weekly conversation network snapshots and loaded them into a tool called TempoVis [36] that supports the exploration through time using a time-slider and shows the dynamic addition and aging of network nodes/links (Figure 6).

In the main window, the EOL conversation network is visualized in a specific week (2012 week 28). As in Figure 3, the nodes represent the EOL members and the links show the connected nodes had conversations in the same week. The red nodes and links show that the activities were in the current week (2012 week 28) and the black to gray nodes and links show that the activities happened in the past weeks. If the gray intensity is lower, it means that the corresponding activities happened further in the past.

The colors match with the bar color of the graph below. Using the time-slider beneath the graph, one can see other snapshots of different time periods (Figure 7 shows 2011 week 36). Node shape represents the EOL member level just like in Figure 3 and blue nodes represent that they are curators. Figure 6 shows that in 2012 week 28, three curators (blue disk-shaped nodes with red center) and one non-curator member (small red triangle) are active. The two curators in the center (red nodes with blue borders) are interacting with each other,

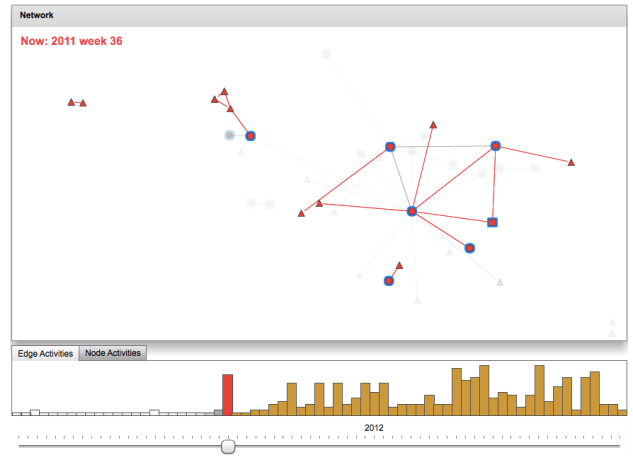


Fig. 7. Dynamic network visualization using TempoVis (2011 week 36). The color codings of the nodes and the edges are identical to Figure 6.

whereas the curator to the left is having conversations with a non-curator (small red triangle).

Visual Exploration of Temporal Network Changes. Using the time-slider and the time-based color-coding in the EOL network in TempoVis, we can explore how the EOL conversation activities evolved. The EOL dataset includes more than 100 weeks of network activities, so it is not easy to show every example in this paper. Therefore, we chose to focus on an important time point and to make in-depth observations. The most prominent peak from Section VI-B was picked up: EOL version 2 launch (September 2011). Figure 8 shows the corresponding 4 week snapshots (from 2011 week 36 to 48). From them, we can observe the following trends.

(1) Most noticeable are the three most active curators, which were visible in the static visualization too (Figure 3). They formed a triangle in the center of the network (marked with dashed lines in Figure 8) and connected the other curators and non-curator members. They sometimes completely connected to each other.

(2) The number of active links did not show steady growth or decrease in this short term period. Rather, it was more like fluctuating between active and less active states. In the active state, the non-curator members were also active.

(3) Even though the three main curators showed strong influences, the activities did not entirely concentrate on them. The group of active nodes and links moved twice from the center to the periphery area (week 36 to 40; week 44 to 48). This suggests the flow of activities from the three main curators to their neighboring curators or non-curator members.

(4) The curators closer to the center were more consistent across time and the ones in the periphery were less consistent.

(5) Non-curator members were mostly connected to curators, rather than having conversations by themselves.

These observations imply the importance of the three most active curators. They were involved in the conversations with non-curator members as well as with other curators. However, they did not dominate the conversations. The remaining curators also showed active conversations with curators and non-curator members, even though they are less connected than

the former three curators. It is also interesting to notice conversation flows from the core curators to the other members.

Another important implication is that the non-member conversations happened mostly with the curators. It confirms again the importance of curators to encourage EOL member participations (research question 2). In order to confirm whether this is the general trend not limited to this specific time point, we compared the average number of members who participated in conversations with the curators and the members who conversed without the curators (Table IV). On average, there were 5.56 members per week who participated in conversations with curators, whereas only 1.52 members had conversations without curators. This difference was statistically significant (One-way Anova, $F(1, 92) = 29.9$, $p < 0.001$).

TABLE IV
COMPARISON OF AVERAGE NUMBER OF CURATOR CONNECTED MEMBERS
VERSUS NOT CONNECTED MEMBERS

Curator connected members	Not connected members
5.56	1.52

VII. CONCLUSIONS

This paper introduces the social network analysis of a collaborative citizen science project, the Encyclopedia of Life. We raised questions whether new EOL web site software features could increase member activities (RQ1) and whether special EOL members could encourage other member activities (RQ2). The analysis was conducted by analyzing the log data from EOL and by visualizing the social network extracted from the log data. We also created a dynamic network visualization in order to trace the temporal changes of the network. We found activity patterns which show that new EOL web features did increase member activities. In particular, the social participation that involved multiple members experienced greater impact than isolated individual participation. By visually exploring the dynamic network changes and statistically analyzing the log data, we found evidence that activity by EOL curators encouraged other members to be more active. We also found a flow of activity from a small number of high-impact curators to other curators or non-curator members (Section VI-C, item (3)). These are encouraging in that they provide hints about how to encourage citizen scientist participation and enrich the participation-based projects. However, we make it clear that the results should not be over-generalized. We cannot say that every feature change will produce increased activity. At the same time, more studies are needed to understand which curator activities have the most impact and which members are most inspired by the activities.

Future research plans include the analysis of more diverse social networks, such as rating activities and member contributions to EOL contents. It is possible to create a co-rating network of EOL members and a network of people who contributed resources from partner services such as Flickr. These networks will reveal different aspects of EOL member participation than the simple comment networks. We

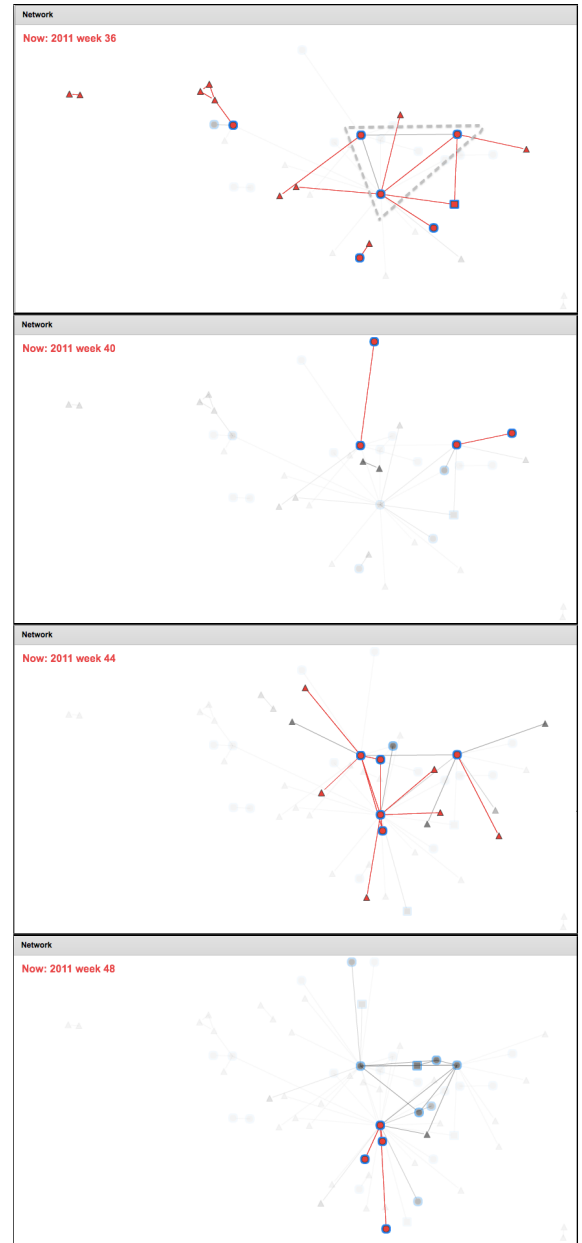


Fig. 8. Dynamic network visualization (from 2012 week 36 to week 48).

defined the time window as one week when we created the conversation network in this study but tracking the sequence of individual activities within the time window is expected to provide more insights for understanding EOL social activities. We are also interested in extending the current visualization tools, by incorporating the individual activity sequence analysis and by embedding statistical analysis into visualization.

ACKNOWLEDGMENT

We appreciate Encyclopedia of Life for providing the dataset and Microsoft External Research for the development of the free, open source NodeXL (<http://www.codeplex.com/nodexl>). We also appreciate for National Science Foundation Award SES 0968546 – “Biotracker - Melding human and machine intelligence to create large-scale collaborative systems,”

plus National Science Foundation-Computing Research Association Computing Innovation Fellow, Postdoctoral Research Grant for Jae-wook Ahn.

REFERENCES

- [1] A. Wiggins and K. Crowston, "From conservation to crowdsourcing: A typology of citizen science," in *System Sciences (HICSS), 2011 44th Hawaii International Conference on*, Jan 2011, pp. 1–10.
- [2] D. Rotman, J. Preece, J. Hammock, K. Procita, D. Hansen, C. Parr, D. Lewis, and D. Jacobs, "Dynamic changes in motivation in collaborative citizen-science projects," in *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, ser. CSCW '12. New York, NY, USA: ACM, 2012, pp. 217–226.
- [3] N. Bos, A. Zimmerman, J. Olson, J. Yew, J. Yerkie, E. Dahl, and G. Olson, "From shared databases to communities of practice: A taxonomy of collaboratories," *Journal of Computer-Mediated Communication*, vol. 12, no. 2, pp. 652–672, 2007.
- [4] J. Preece and B. Shneiderman, "The Reader-to-Leader Framework: Motivating technology-mediated social participation," *AIS Transactions on Human-Computer Interaction*, vol. 1, no. 1, pp. 13–32, 2009.
- [5] T. Malone, R. Laubacher, and C. Dellarocas, "The Collective Intelligence Genome," *Sloan Management Review*, vol. 51, no. 3, pp. 21–31, 2010.
- [6] B. Shneiderman, "Social discovery in an information abundant world: Designing to create capacity and seek solutions," *Information Services and Use*, vol. 31, no. 1, pp. 3–13, 01 2011.
- [7] D. L. Hansen, B. Shneiderman, and M. A. Smith, *Analyzing Social Media Networks with NodeXL: Insights from a Connected World*. Morgan Kaufmann, 2010.
- [8] D. Rotman, K. Procita, D. L. Hansen, C. S. Parr, and J. Preece, "Supporting content curation communities: The case of the encyclopedia of life," *Journal of the American Society for Information Science and Technology (JASIST)*, vol. 63, no. 6, pp. 1092–1107, 2012.
- [9] Y. Ho, A. Gruhler, A. Heilbut, G. Bader, L. Moore, S. Adams, A. Millar, P. Taylor, K. Bennett, K. Boutilier *et al.*, "Systematic identification of protein complexes in *saccharomyces cerevisiae* by mass spectrometry," *Nature*, vol. 415, no. 6868, pp. 180–183, 2002.
- [10] P. Ortiz, S. Ortiz, L. Rico, and A. Valencia. Gnom - Oracle Interface. [Online]. Available: <http://moebio.com/santiago/gnom/english.html>
- [11] E. Singer. The Brain Unmasked. [Online]. Available: <http://www.technologyreview.com/news/410560/the-brain-unmasked>
- [12] M. Morris, A. Kurth, D. Hamilton, J. Moody, and S. Wakefield, "Concurrent partnerships and HIV prevalence disparities by race: linking science and public health practice," *American Journal of Public Health*, vol. 99, no. 6, p. 1023, 2009.
- [13] P. Saraiya, C. North, and K. Duca, "An evaluation of microarray visualization tools for biological insight," in *Information Visualization, 2004. INFOVIS 2004. IEEE Symposium on*. IEEE, 2004, pp. 1–8.
- [14] W. Powell, D. White, K. Koput, and J. Owen-Smith, "Network Dynamics and Field Evolution: The Growth of Interorganizational Collaboration in the Life Sciences," *American Journal of Sociology*, vol. 110, no. 4, pp. 1132–1205, 2005.
- [15] N. Christakis and J. Fowler, "The spread of obesity in a large social network over 32 years," *New England Journal of Medicine*, vol. 357, no. 4, pp. 370–379, 2007.
- [16] A. Kittur, B. Suh, B. A. Pendleton, and E. H. Chi, "He says, she says: conflict and coordination in wikipedia," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, ser. CHI '07. New York, NY, USA: ACM, 2007, pp. 453–462.
- [17] F. Viegas and M. Smith, "Newsgroup crowds and authorlines: visualizing the activity of individuals in conversational cyberspaces," in *System Sciences, 2004. Proceedings of the 37th Annual Hawaii International Conference on*, jan. 2004, p. 10 pp.
- [18] B. Suh, E. H. Chi, A. Kittur, and B. A. Pendleton, "lifting the veil: improving accountability and social transparency in wikipedia with wikidashboard," in *Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, ser. CHI '08. New York, NY, USA: ACM, 2008, pp. 1037–1040.
- [19] A. Alabri and J. Hunter, "Enhancing the quality and trust of citizen science data," in *e-Science (e-Science), 2010 IEEE Sixth International Conference on*, Dec 2010, pp. 81–88.
- [20] B. L. Sullivan, C. L. Wood, M. J. Iliff, R. E. Bonney, D. Fink, and S. Kelling, "eBird: A citizen-based bird observation network in the biological sciences," *Biological Conservation*, vol. 142, no. 10, pp. 2282–2292, 2009.
- [21] J. Howison, K. Crowston, and A. Wiggins, "Validity issues in the use of social network analysis with digital trace data," *Journal of the Association for Information Systems*, vol. 12, 12/2011 2011.
- [22] J. Moreno, "Who shall survive? Foundations of sociometry, group psychotherapy and socio-drama." 1953.
- [23] T. M. J. Fruchterman and E. M. Reingold, "Graph drawing by force-directed placement," *Software: Practice and Experience*, vol. 21, no. 11, pp. 1129–1164, 1991.
- [24] N. Henry and J.-D. Fekete, "Matrixexplorer: a dual-representation system to explore social networks," *IEEE Transactions on Visualization and Computer Graphics*, vol. 12, pp. 677–684, 2006.
- [25] J. S. Yi, N. Elmqvist, and S. Lee, "TimeMatrix: Analyzing Temporal Social Networks Using Interactive Matrix-Based Visualizations," *International Journal of Human-Computer Interaction*, vol. 26, no. 11, pp. 1031–1051, 2010.
- [26] J. Moody, D. McFarland, and S. Bender-deMoll, "Dynamic network visualization," *American Journal of Sociology*, vol. 110, no. 4, pp. 1206–41, 2005.
- [27] H. Kang, L. Getoor, and L. Singh, "Visual analysis of dynamic group membership in temporal social networks," *ACM SIGKDD Explorations Newsletter – Special issue on visual analytics*, vol. 9, no. 2, pp. 13–21, December 2007.
- [28] P. Mutton, "Inferring and visualizing social networks on internet relay chat," in *Information Visualisation, 2004. IV 2004. Proceedings. Eighth International Conference on*. IEEE, 2004, pp. 35–43.
- [29] P. Gloor and Y. Zhao, "Analyzing actors and their discussion topics by semantic social network analysis," in *Information Visualization, 2006. IV 2006. Tenth International Conference on*. IEEE, 2006, pp. 130–135.
- [30] B. Arian. (2008) Growth of a twitter graph. [Online]. Available: <http://burak-arikan.com/growth-of-a-twitter-graph>
- [31] G. Miller, "Social scientists wade into the tweet stream," *Science*, vol. 333, no. 6051, pp. 1814–1815, 2011.
- [32] D. Shamma, L. Kennedy, and E. Churchill, "In the limelight over time: Temporalities of network centrality," in *Proceedings of the ACM 2011 conference on Computer Supported Cooperative Work (CSCW 2011)*. Hangzhou, China: ACM, March 19-23 2011.
- [33] —, "Peaks and Persistence: Modeling the Shape of Microblog Conversations," in *Proceedings of the ACM 2011 conference on Computer Supported Cooperative Work (CSCW 2011)*. Hangzhou, China: ACM, March 19-23 2011.
- [34] D. Smits, H. Smid, N. Hu, C. Meijer, M. Honig, A. Sardar, N. van Wijk, and J. Bajec. (2010) Twitter lyrics. [Online]. Available: <http://www.alimdardesign.com/twitterlyrics/index.htm>
- [35] J. Ahn, C. Plaisant, and B. Shneiderman, "A task taxonomy for network evolution analysis," University of Maryland HCIL Technical Report HCIL-2012-13, Tech. Rep., 2012.
- [36] J. Ahn, M. Taieb-Maimon, A. Sopan, C. Plaisant, and B. Shneiderman, "Temporal Visualization of Social Network Dynamics: Prototypes for Nation of Neighbors," *Social Computing, Behavioral-Cultural Modeling and Prediction*, pp. 309–316, 2011.
- [37] U. Khurana, V. Nguyen, H. Cheng, J. Ahn, X. Chen, and B. Shneiderman, "Visual analysis of temporal trends in social networks using edge color coding and metric timelines," in *IEEE International Conference on Social Computing (SocialCom)*. IEEE, 2011, pp. 549–554.
- [38] E. Peterson, "Time spring layout for visualization of dynamic social networks," in *Network Science Workshop (NSW), 2011 IEEE*, June 2011, pp. 98–104.
- [39] S. Bender-deMoll and D. McFarland, "The art and science of dynamic network visualization," *Journal of Social Structure*, vol. 7, no. 2, 2006.