Designing During Use: Modeling Communities of Practice

Zinayida Petrushyna∗, Ralf Klamma†, Milos Kravcik‡
Information Systems and Databases
RWTH Aachen University, Aachen, Germany
petrushyna@dbis.rwth-aachen.de∗
klamma@dbis.rwth-aachen.de†
kravcik@dbis.rwth-aachen.de‡

Abstract—We examine digital social media ecosystems such as blogs and mailing lists from the perspective of communities of practice. We observe behaviors of agents and specify prevalent patterns of collaborations in digital media ecosystem. We describe several techniques important in patterns specification such as social network analysis and content analysis. The conceptualization of our observations was done by the goal-oriented modeling framework i∗. Models of digital social media ecosystems include agents causing changes over a time and detect disturbances coming from outside. As a proof statement we use the models of our case study that define three different communities of practice based on discovered patterns. In the end we present what is expected to be done in the future in the area of digital media ecosystems modelling.

I. INTRODUCTION

Forms of organizations such as networks or communities of practice do not automatically evolve smoothly and well in a self-organized fashion. While there are some communities one can label as ‘successful’ in a mostly ad-hoc defined manner, many communities suffer from pitfalls and shortcomings. Some of these problems like a low number or absence of workers in social capitals of executives [20], [21], [28] are well known and sometimes also well understood. But others such as engagement or motivation of community members in collaborations are unclear either due to lack of digital literacy or due to lack of management and analyzing tools alerting about good and bad practices within communities. For example, in communities of practice we observe a culture of distinctiveness between experts defining the core of the communities and the newcomers defining the periphery of the community.

One of the biggest problems for CoPs on the Web is how to deal with a shared passion, emergence of sub-communities, mutual engagement or not sharing practices. Managers and other members of communities often don’t have answers on how to motivate users in collaborations, how to filter good and bad contributors and to make community members agree and help each other. To find the answers, one has to operate on the abstract level of a community, particularly with a model that depicts main agents of communities, agent properties and dependencies.

As information changes rapidly on the Web, static models are no more appropriate while these should be continuously remodeled to suit the current situation. This is a time-consuming and complex task. Even sometimes it is preferable to start establish new models than to rewrite old ones. If the information changes rapidly, why can’t its models? Why can’t the models continuously reflect the current situation. Our intention to create evolvable models based on continually changing information is not initial. The idea of designing-during-use was explained in [16] and applied to the software engineering discipline. This design approach gave birth to adaptive systems [18], that adjust themselves if a modification of the system is triggered. In end user development [13] the system defines automatically the needs of users or users point out preferred extensions. Software evolves not only under the influence of users but also triggered by changes in the environment. If no Acrobat Reader, Quicktime Player and Windows Media Player had appeared on the market, the browsers would not have changed by presenting plug-ins for those programs allowing to open documents, play music and watch clips directly from browsers.

In this paper we describe what disturbances in the environment influence the evolution of digital social media ecosystems. We differentiate between models of ecosystems based on discovered patterns within ecosystems. Our contributions in the research of digital ecosystems include an approach for modeling of ecosystems as CoPs. As the solution considers changes over time, if anything unexpected happens digital ecosystems must re-model their models. The unexpected events are those that are not described in a standard model of an ecosystem. Moreover we present three standard models that differ from each other by a number of discovered patterns.

The rest of the paper is structured in the following way: in the following section we discuss related work. Then we explain the concept of community of practice and apply it to ecosystems of digital social media. Further we present the i∗ modeling framework and its components and depict CoP using the framework. In section 5, we discuss the role of patterns in ecosystem models and explain three case models that differ from each other by patterns and roles of ecosystem agents. At the end, we summarize our observations and discuss the future work in the last section.
II. RELATED WORK

Relationships in communities of practice have often a master-apprenticeship pattern. Learners observe experts in legitimate peripheral participation and become experts over time. In the Web 2.0 world long-lasting learning processes which include social identity building and relying on experiences of others are not well supported yet. Communities which are seeking support from tools like wikis and blogs are often not aware of all the shortcomings coming with these fantastic collaboration tools. A recent example is Twitter.

In Twitter ecosystem there is an ongoing discussion on how to define popularity, mastership or more neutral rankings of tweeters. While some argue for a simple count of followers or followings or a ratio of both, others argue for counting the number of tweets. The basic idea of PageRank [3] to calculate the authority of a tweeter from the authority of those re-tweeting her is another line of argument. If the authority of the tweeter is high, the information she is distributing is highly trustfull. Community structures of the Twitter users can facilitate the identification of roles of community members. Wherever the discussion will lead to, the need of community structures is high. Recently Twitter began to support groups of users called lists.

In another popular Web 2.0 tool, FriendFeed, the creation of community structures was already included in the design. In FriendFeed you can find so-called rooms. A room is topic-oriented and can be created by any user within seconds. Also everybody can join an existing room. The barriers are extremely low. However, until now there are no existing tools to support community awareness in the rooms well, e.g., a member cannot see which other members are online.

The existing work on examination of social media [25] considers the diversity between participants of communities but does not examine what is the reason or motivation for their behavior. The results of the work are recommendations about further social media community development. Another work about social media describes social information foraging (SIF) model [6] that allows to model the behaviors of agents within digital ecosystems. According to SIF, agents try to benefit as much as possible but in a collaborations with others. The application of the SIF model on the computational ecology focuses mostly on human agents or groups of human agents [23].

III. ECOSYSTEMS AS COMMUNITIES OF PRACTICE

Digital ecosystems are numbers of living or non-living agents, which are situated in associated environment in a specific place [24]. Agents of ecosystems find a consensus in interacting with each other so that the system they organize is consistent and stable. Behaviors of agents and their interaction strategies change over a time and under the influence of the environment. The consideration of ecosystems through the prism of CoP was described in [14]. CoPs incorporate agents interacting with each other and changing their statuses within a time. Dermott wrote that a CoP is like a living organism growing and evolving over time with organic growth [15]. CoPs is a concept of a system organization [29] similar to ecosystems. However in CoPs all agents are human beings. Models of ecosystems can not always reflect the current situation because reality is dynamic. If we can respond to the changes in reality rapidly, this can give us a proper understanding of the ecosystem’s current state.

Here we propose our solution how to understand digital ecosystems better. The simplified model in Figure 1 depicts a gap between real life situation (Self-analysis) and ideal situation (Self-reflection). The gap is represented as a loop between two states: Self-analysis state examines the current situation of an ecosystem using Social Network analysis (SNA), Knowledge metadata and Activities analysis; Self-reflection state adopts the ecosystem according its purposes with the help of Pattern analysis, Community awareness and Recommendations. The loop allows us to make ecosystems adaptive to mutable environments. Changes (disturbances) may be positive (an expert in a main topic of a community has joined the community), negative (the expert has left the community) and with unclear influence (the community changed the focus of work). We state that considering communities changes using the loop allows to keep community models updated. Unless anything unexpected (not covered in a model) happened we have enough knowledge to understand a community. As soon as unforeseen events occur (disturbances), the state of the community is unclear and should be reflected. For example, Facebook was a closed social network with a limited class of users for a long time. After opening access to everybody, this social network site had to deal with different types of users - non-students and non-educators. Facebook extended its communities and types of users. In doing so Facebook enhanced the roles of users. Numerous applications emerged after the extension to support the changes in Facebook and to help find suitable niches for users. The same idea is depicted in the loop: to be clear about the situation in digital ecosystems, their models should reflect accordingly changes of the environment.
IV. Modeling Communities of Practice with I*

A. Introduction to the i* modelling framework

Yu introduces the modeling technique named later i* [30]. The i* language is an agent-oriented high-level modeling approach. It was developed for the design of business processes on a strategic level. The core of i* models are motives, interests, options and constraints of agents in a business process. Cooperation among agents is needed for fulfillment of common goals. Due to dependencies between goals, tasks, resources and the agents themselves agents become also vulnerable. In i* we find the following premises: agents act intentionally, because they are following their goals, have beliefs, competences, commitments, needs and desires. Agents are strategic actors, because they have to proceed according to their plans. Intentional dependencies can be made explicit with i* to disclose the reasons behind observable processes. Questions to an i* model are: Who is deciding what? Which options does an agent have? How can I downplay critical dependencies?

In i* we use the strategic-dependency-model (SDM) with a focus on the intentions of agents. It serves an abstract view on the external relations of an agent. It delivers very concrete focus on the intentions of agents. It serves an abstract view strategic-dependency-model In i* we use the dependencies? Options does an agent have? How can I downplay critical dependencies?

In i* we use the strategic-dependency-model (SDM) with a focus on the intentions of agents. It serves an abstract view on the external relations of an agent. It delivers very concrete focus on the intentions of agents. It serves an abstract view strategic-dependency-model In i* we use the dependencies?

Here we present a short description of i* conceptual modeling components and use it later in Figure 2, 3, 4 and 5. A so-called dependee is an agent that influences other agents performance, e.g., Members in Figure 2. The other agents are dependers, e.g., Sense of belonging. The object through which the dependency is defined is called dependum, e.g., Interact. The dependencies can be one of the following four types - resource dependency, task dependency, goal dependency and softgoal dependency. Resource dependency like Policies, rules exists in a case when the target provides a resource to the source. There is a task dependency such as Interact when the source expects the target to execute a task and goal dependency indicates the case when the source expects that the target to achieve a goal, e.g., Contribution in Figure 3. In general a softgoal dependency like change is used when the goal is unable to be described clearly, though a dependency exists.

B. Community of Practice model

An ecosystem is always influenced by environment and supporting technologies. Therefore, these factors must be taken into account during community modeling. Wenger defined a community with shared domain (Shared repertoire - SR), collaborations (Mutual engagement - ME) and practices (Joint enterprises - JE) [29]. We consider digital social media ecosystems as communities of practice, where practices are performed using technologies (social media such as mailing lists and blogs). With the model in Figure 2 we conceptualize digital ecosystems based on the CoP methodology. The model [19] includes several actors. The main actors in an ecosystem are its members, shown in the middle of figure. Usually they collaborate with each other and if not we should deeply think if they are contributors of a community of the ecosystem. Are their knowledge and intentions acceptable for all others in the community? It is reasonable to assume that this is not the case. Any community member must have a sense of belonging to the community.

A sense of belonging to the community must be possessed by a community member. The sense of belonging is considered as an agent in our model as it plays a key role of being together with others in the community. However, members of the ecosystem can only have a sense of belonging if they interact with each other. This is represented through the task dependency between Members and Sense of belonging.

The dependency from Members to Technology&Enterprise is based on rules and policies the ecosystem sets. Consistency of the ecosystem can only be reached if the members follow its policies. A stable ecosystem has a goal and every agent of the ecosystem more or less helps to achieve the goal, e.g., in a forum users have to follow the rules of politeness and mutualism (the interaction of two organisms and each of them benefit from the interaction). The dependency from Members to Knowledge repository identifies the knowledge domain of the ecosystem. Members of the ecosystem either understand of what it is talking about in the community or they are out of the community.

The iterative nature of the ecosystem is represented by dependencies between a process (in our case the Learning process) and other actors. As soon as learning happens within the ecosystem, other actors are influenced. For example, new members enter and others get to know them, pieces of information operated by ecosystem members change, and somebody shares his experiences that may add new ideas and rules.

V. Patterns in Modeling

When collaborations of ecosystem members are visualized as a graph, we can notice who is who in the ecosystem. The graph consists of nodes, which represent ecosystem community members and edges, that are relationships between the members supported by different Web technologies, such as e-mails, entries and comments in wikis, blogs, and forums.
We can easily see who is more talkative in a community. Members that collaborate with each other more often than others are more central, otherwise they are far away from the graph center. Node properties such as closeness centrality or betweenness centrality may be calculated [26]. These properties originate from graph theory and were applied to SNA. Dustdar and Hoffman evaluate SNA efficiency by comparing discovered roles of workers in a company with SNA properties of networks of workers built on their e-mail communications [8]. The authors conclude that it is efficient to use SNA techniques as it simplifies the analysis as there is no need for complicated experiments. Moreover SNA provides additional information such as the person who communicate the most per e-mail.

Different motives make people collaborate or prevent collaborations [22]: lack of time, disinterest in the community, personal negative attitude to or impression about other members. Moreover there are different reasons for a community content creation, spam or troll-based discussions. To find these and other features of people interactions, we consider patterns of collaborations between people with the help of SNA. In [11] the authors defined the spammer, the troll, the answering person, the questioneer, the communicator and some other patterns. While constructing the ecosystem models we reuse the concept of patterns to refine and define new patterns.

A pattern consists of a description of a problem, a context where it is applied, forces which influence appearance of the problem and possible solutions to the problem [2]. Why are patterns produced: because of technologies, community colleagues, content, or anything else? What are disturbances that produce patterns? What techniques can be used to discover the disturbances?

SNA is not a perfect and not the only technique for defining disturbances. The flaws of SNA are engendering trust, dispelling the illusion of accuracy, and taming the expert mindset. The most critical problem from our point of view is data accuracy. It is possible that a studied group of 200 people may organize a network consisting of 400 or more people. From those the half is out of case studies. However, we can operate user-generated content (UGC), so that we extend the methodology for community analysis by examination of UGC. Experiments utilizing UGC in [17] show us how different distributions of content are. But it is hard to explain the reasons for this phenomenon. UGC has been intensively analysed by data mining techniques to find a number of patterns. Such techniques are successful in predicting customer preferences. It is advised that the more entities in a data set, the more precise are the patterns established.

We are sure that both, data mining and SNA, need to be considered for patterns definition. One of the reasons is obvious: the experts in a community topic, who are not collaborating often, can be defined by precise analysis of the content they create. As an example of efficient text mining, [27] Webb et al define an effective algorithm for spam detection looking only into mails headers. Outcomes of that algorithm and other SNA properties collect heterogeneous information about the situation and may provide more precise results in spam detection. Nevertheless, we should not forget that spam definition may differ from community to community.

VI. CASE MODELS OF COMMUNITIES

In this section we consider different models of digital social media ecosystems that we design during observation of the ecosystems. Every model is based on digital social media in real life, its members with their goals and characteristics that transmit information about roles of the members, their behaviours, alone and in a group, static and in the course of time.

The social media we examine is stored in a so called Mediabase [10]. We extract the data of different social media by crawlers that are adopted for each type of media [9]. For creation of ecosystem models we are operating on analysis data of mailing lists but plan examination of other media such as blogs, wikis, newsletters and forums. In the following we describe four types of ecosystem communities that we discovered during observations.

The question-answer community is shown in Figure 3. The case of its application are FAQs - a question is asked and an answer is given and in most cases no discussion is developing further. Such types of collaborations appear between experts and novices, that might be teachers and students, developers of a software product and its customers, consultants and users of a system.

Ecosystem members of the community have an Attitude to an ecosystem domain: it is a hobby or devoted to a
work or study. The *Attitude* engages members to collaborate (*Reasoning for collaboration* task). The community members differentiate: inactive and active members. The inactive members are those who are lurking (*Lurk* task); the active members are those who are interacting (*Interact* task). The active members can be separated into *Experts* and *Novices* (the separation can be done with the inactive members too, but there is no reason to examine those members).

In the whitepaper [1] the forces and motivations of social media users were discussed. How far do people go to achieve their goals? The main reason that engages people to collaborate is doing something good for the world (world contribution). For others it is a community recognition, in other words identification of themselves within an ecosystem. *Experts* have global goals (*Recognition* and *Contribution* goal dependencies from *Social media*). The *Experts* are the community nodes that have deep knowledge regarding to the community topic (*Knowledge goal dependency on Domain of ecosystem*). The knowledge can be measured by contribution/participation in other communities with similar topics (*Contribution/participation* task dependency with *Outside ecosystem*). However, if the members are *Lurking or interacting continuously* within the community they often become community experts [12]. The novices are those who have no idea about the community topic (*No knowledge goal dependency on Domain of ecosystem*).

Whenever all agents are important to consider, we pay the main attention on *Members*. How interactive they are? Simple graph visualizations help to define active and inactive nodes. For discovering experts and novices SNA and text mining come on the stage. Structural holes denote members - experts - who are participants in other communities[4], [7]; dynamic analysis of a community graph - DNA[5] - help to define members who were novices but after a period of time and due to their activity/awareness become experts; experts are in most of cases answering persons that post in an existing thread, while novices are questioneers that create new threads, where answering persons and questioneers are patterns of social media[11]. Similarity of community topics can be solved by named entity recognition with such a system as Open Calais 1 that extracts main topics of ecosystem domain such as people, events, organizations, products and so on.

The next model we consider represents the ecosystem of innovation community on Figure 4. The innovation community is one of the most important components of a successful organization. It may exist together with heterogeneous nodes. For discovering experts and novices SNA and text mining come on the stage. Structural holes denote members - experts - who are participants in other communities[4], [7]; dynamic analysis of a community graph - DNA[5] - help to define members who were novices but after a period of time and due to their activity/awareness become experts; experts are in most of cases answering persons that post in an existing thread, while novices are questioneers that create new threads, where answering persons and questioneers are patterns of social media[11]. Similarity of community topics can be solved by named entity recognition with such a system as Open Calais 1 that extracts main topics of ecosystem domain such as people, events, organizations, products and so on.

Whenever all agents are important to consider, we pay the main attention on *Members*. How interactive they are? Simple graph visualizations help to define active and inactive nodes. For discovering experts and novices SNA and text mining come on the stage. Structural holes denote members - experts - who are participants in other communities[4], [7]; dynamic analysis of a community graph - DNA[5] - help to define members who were novices but after a period of time and due to their activity/awareness become experts; experts are in most of cases answering persons that post in an existing thread, while novices are questioneers that create new threads, where answering persons and questioneers are patterns of social media[11]. Similarity of community topics can be solved by named entity recognition with such a system as Open Calais 1 that extracts main topics of ecosystem domain such as people, events, organizations, products and so on.

The heterogeneousity of people, their brokerage and connections with other communities is easier to find analyzing the global graph that include all the players of a community and theirs connections with other communities. Text mining should be used to secure that topics of the other communities are not only the same as a topic of the community. Precise NLP-based end emotional analysis of community content can be done to say if communications are self-regulated and non-stressful, easy and semi-official.

An expert community is organized as within an organization as well as because of hobbies of community members (*Hobby, Work and Study* attitudes). There is no place for novices, because the community members possess special community-dependent thesaurus, dependent on a *Domain of ecosystem*. Even if new members come into the community, it is considered that they are on a particular level of the domain understanding (*Knowledge goal dependency on Domain of ecosystem*) and acquired enough knowledge to be participants.

1http://opencalais.com
in the expert community (Continuous lurking or interacting task dependency on The ecosystem). The ecosystem mostly includes the members that being a long time within the community (Experts) and span structural holes within other networks with similar topics (Broker dependent on Other ecosystems through Participate task dependency).

VII. CONCLUSIONS AND FUTURE WORK

We explain the methodology of i* modeling framework and how it can be used for the iterative modeling of digital ecosystems. Afterwards we apply the methodology to our test cases and identify several types of ecosystem models that focus on different roles of users. In future we are planning to integrate SNA, data mining, NLP, emotions and other characteristics that we mentioned in models description to have more precise models and to follow the minor changes within them. The models can be as well useful for future predictions of user behaviours, community evolution, ecosystem problems and so on.

VIII. ACKNOWLEDGEMENTS

The research work described in this paper is funded through the ROLE Integrated Project; part of the Seventh Framework Programme for Research and Technological Development (FP7) of the European Union in Information and Communication Technologies.

REFERENCES

[21] T. C. Turner, M. A. Smith, D. Fisher, and H. T. Welser. Picturing Usenet: Community (Continuous lurking or interacting task dependency on The ecosystem). The ecosystem mostly includes the members that being a long time within the community (Experts) and span structural holes within other networks with similar topics (Broker dependent on Other ecosystems through Participate task dependency).

Fig. 5. Expert community