

Viacheslav Lanin, Lyudmila Lyadova, Elena Zamyatina and Nikita Vostroknutov
AN ONTOLOGY-BASED APPROACH TO SOCIAL NETWORKS MINING

RESEARCH RELEVANCE AND PURPOSE

Social networks help to solve many tasks, and on the other hand, they become a source of problems. This led to interest in the study of social networks for various categories of researchers.

The greatest interest is caused by following tasks:

- Investigating dissemination of information on the network.
- Studying communities' formation in networks.
- Modeling behavioral patterns of different categories of users.
- etc.

Static and dynamic methods may be used for these tasks solving.

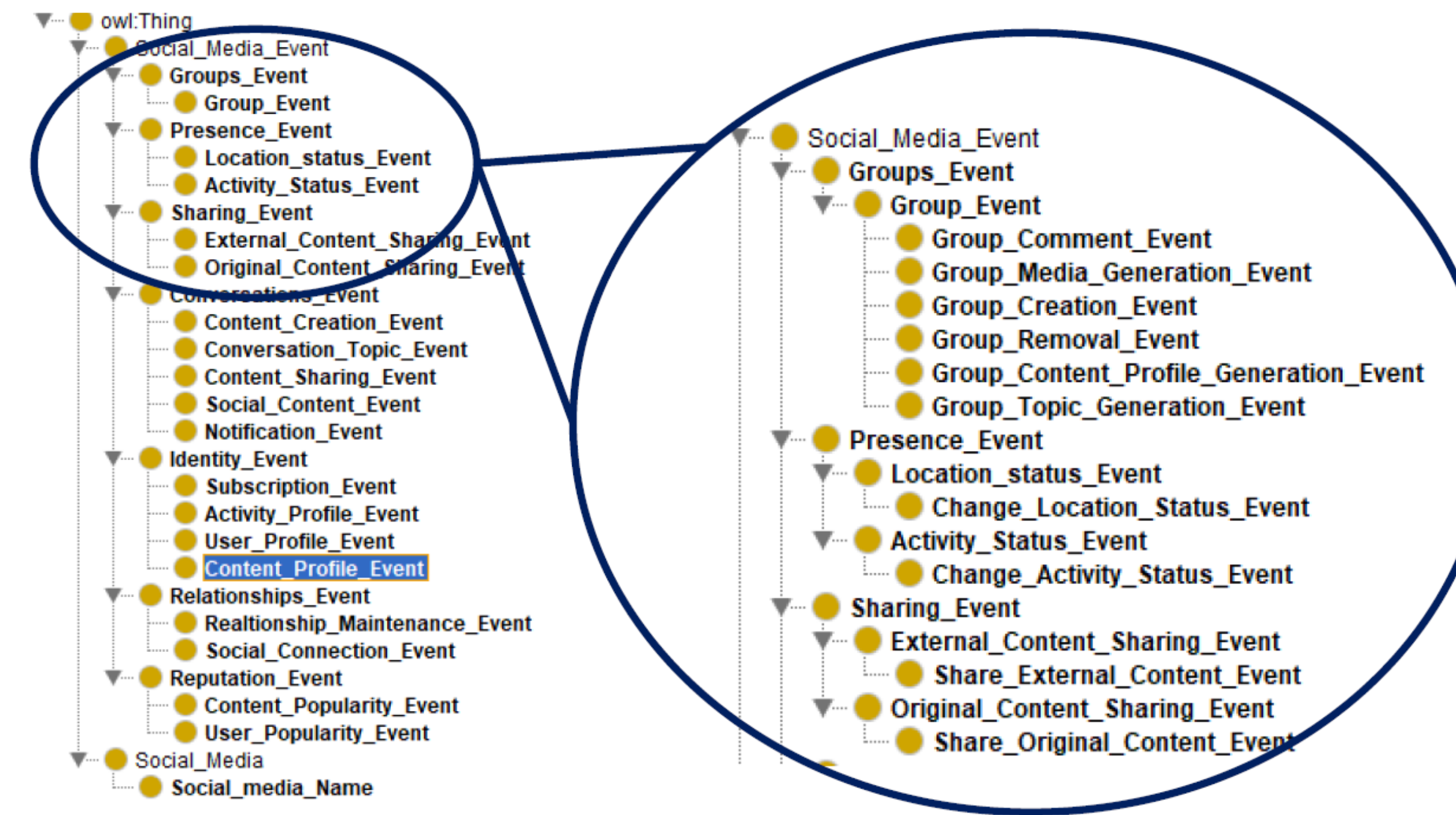
The methods of *static analysis* of networks using

- graph models,
 - statistical methods and
 - machine learning methods
- are most deeply developed.

The goal of the research project presented here is to develop an ontology-based approach to social networks analysis with Process Mining tools.

THE SOCIAL MEDIA EVENT ONTOLOGY

The first part of ontology deals with events. The identification of functional components for structural components, listed above, leads to information for analyzing events and their description.

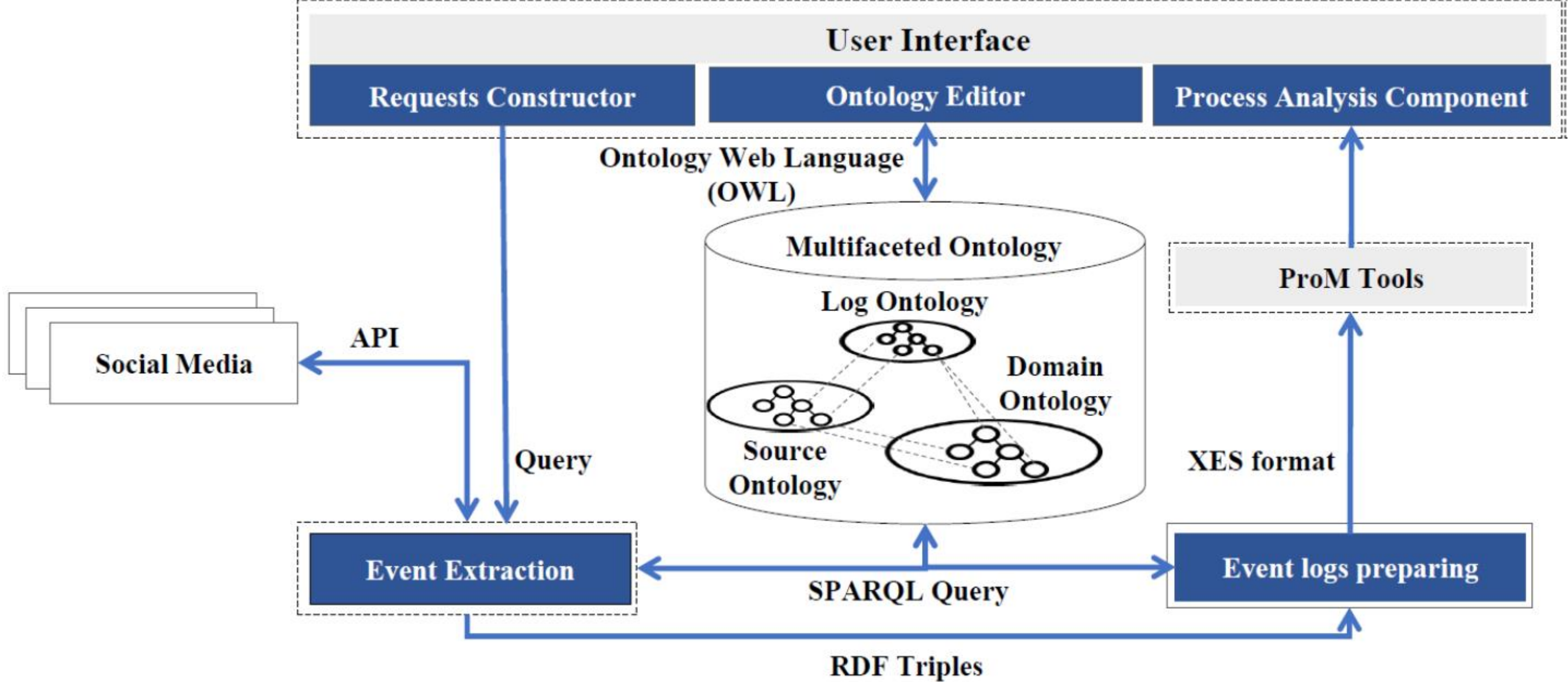


Typically, in social media, several elementary events are closely related to each other. For example, the event of sending a message by one user is associated with the event of creating a notification for other users to whom the message is addressed. For modeling, such events class *Complex_Event* is added. For example, when a user clicks the *Subscribe* button on Youtube, two different events are triggered within the system. The first one is adding the channel to the user subscription list, and the other one is a notification of the channel owner about a new subscriber. The following processes can be described using two classes stored inside the ontology: *Update_User_Popularity_Event* and *Add_Subscription_Event*.

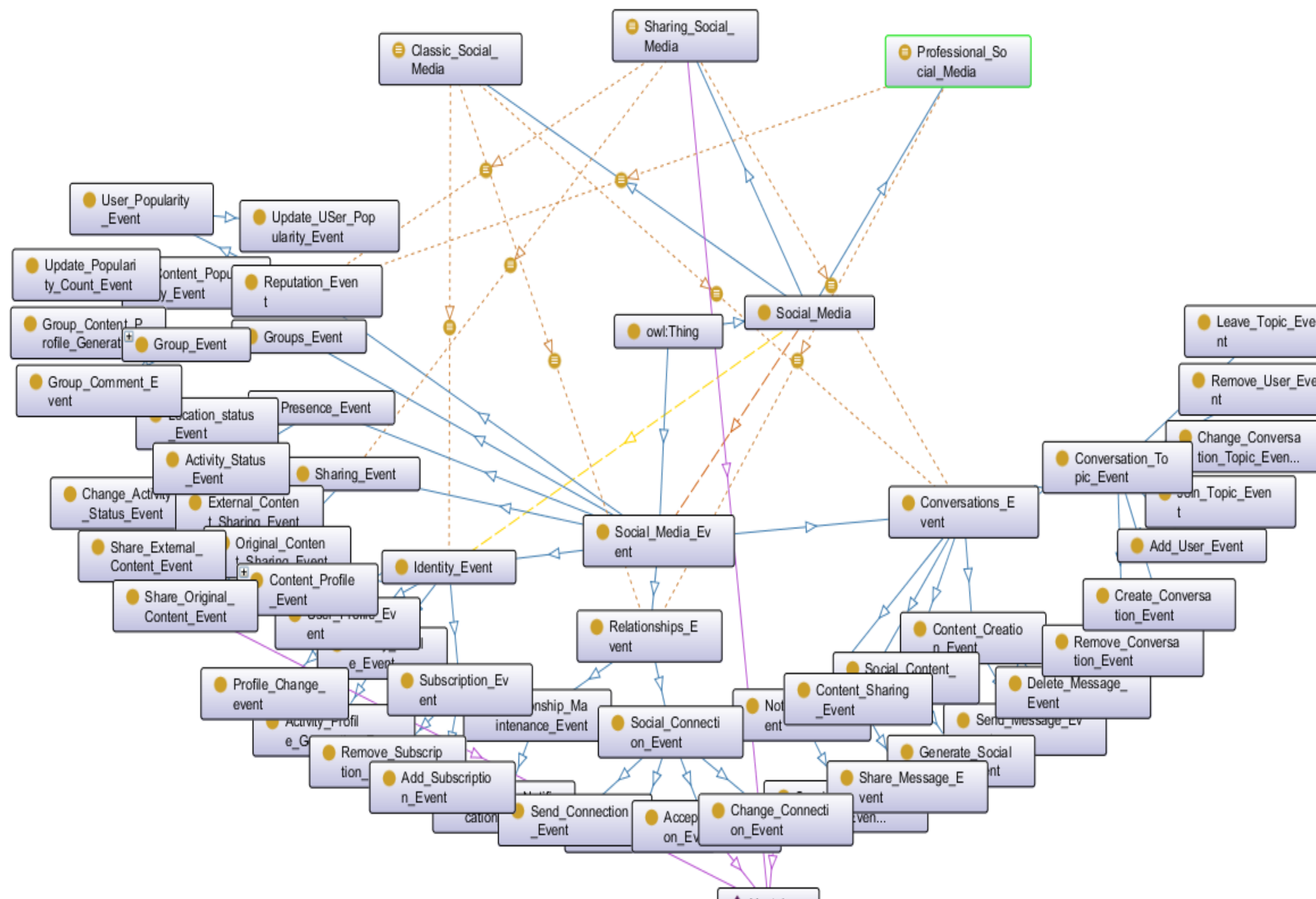
API OF SEVERAL POPULAR SOCIAL NETWORKS

| | Facebook API | VK API | Twitter API |
|--------------------------|--|---|--|
| Data availability | Allows to receive the data that the user has access to on his account. You need to get special permissions to work with group events | Allows to receive data to which the user has access to his account | Allows to receive publicly available data after sending and confirming the application |
| Identity | You can get information about your content profile | You can get information about your content profile; you can get recommendations for the user (used to generate an activity profile) | You can get information about your content profile |
| Discussion | You can receive user-accessible text messages, audio, photos | You can receive user-accessible text messages, audio, photos | You can receive user-accessible text messages, audio, photos |
| Community | You must obtain special permissions from Facebook to work with these groups | You can get publicly available group data | The system lacks community functionality |
| Content sharing | You can get information about likes; you can get information about reposts | You can get information about likes; you can get information about reposts | You can get information about likes; you can get information about reposts |
| Location status | You can get the status, location of the user | You can get user status | You can get user status |
| Relationship | You can receive friend requests | You can receive friend requests | You can receive information about requests |
| Reputation | You can get data about likes and subscriptions | You can get data about likes and subscriptions | You can get data about likes and subscriptions |

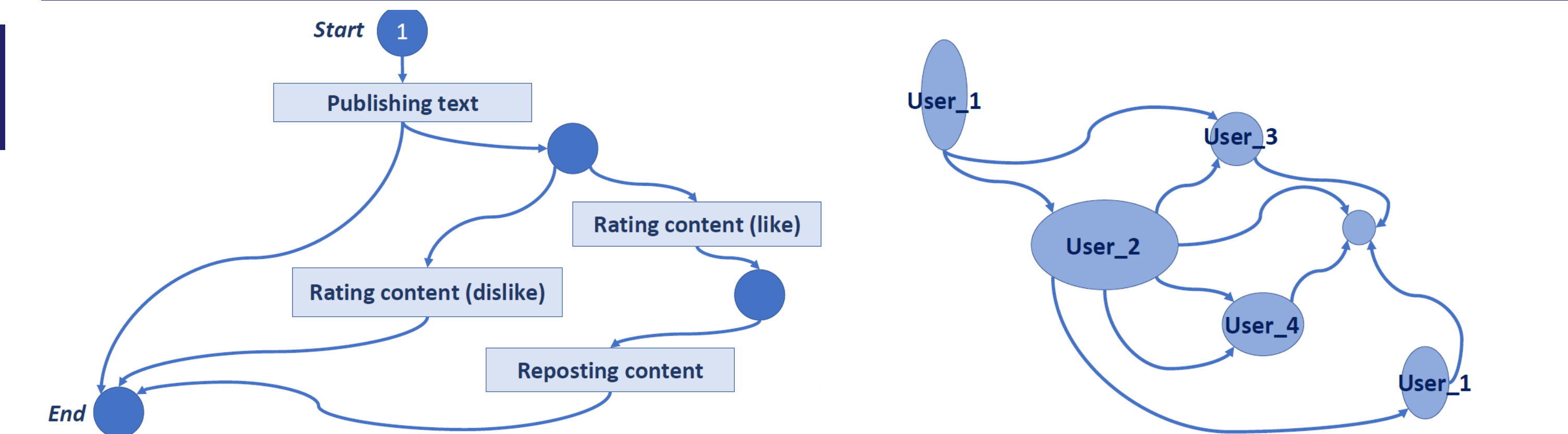
SOCIAL NETWORKS MINING TOOLS: GENERALIZED SYSTEM ARCHITECTURE



THE SOCIAL NETWORK ONTOLOGY: FULL ONTOGRAPH



SOCIAL NETWORK MODELING BASED ON EVENT LOGS



The simplified example of the Petri Net model of the processes in the social network (pattern of the users' behavior). It can be built with ProM tools on the base of event log (plugin "Mine Petri net with inductive Miner").

The simplified example of the model showing network users, relations between them, as well as the degree of activity of each of the them. It can be built with ProM plugin "Mine for Handover-of-Work Social Network".

SOCIAL NETWORKS STUDYING

All social networks have a property called user-oriented design. The basic principles of user-oriented design are

- interaction between users;
- recognition (individuality) of community members.

These properties allow to solve the tasks of generating event logs, where we need to identify not only events, but also objects and processes and their cases. The main *structural elements*, characterizing social networks, are based on the principles of user-oriented design:

1. Identity (possibility of self-expression of the social network user).
2. Conversations (ways of communication between social network users, both personal and group).
3. Sharing (methods of obtaining, distributing and modifying public content).
4. Presence (the ability of users to have the location of other users).
5. Relationships (different connections between social network users).
6. Reputation (the ability of users to confirm their popularity or competence).
7. Communities (user interaction with communities (groups) and sub-communities).

CONCLUSION

The main result of this stage of research is the ontology of events, which can be used as a basis for the development of process models and their research. The presence of the described information in the event logs allows to analyze the connections between various events in social networks, to identify patterns of user behavior, to evaluate the "intensity of work" and "workload" of users. Research prototype of the tools to create event logs is developed. Event logs are prepared in a standard format for export to analyze with using external applications (Process Mining tools, for example, ProM).