COLLABORATIVE BUSINESS PROCESS ELICITATION THROUGH GROUP STORYTELLING

João Carlos de A. R. Gonçalves, Flávia Santoro and Fernanda Baião
NP2TEC – Research and Practice Group in Information Technology, Department of Applied Informatics
Federal University of the State of Rio de Janeiro (UNIRIO), Rio de Janeiro, Brazil

Keywords: Business Process Modelling, Knowledge Elicitation, Computer-Supported Collaborative Work.

Abstract: Business Process Modelling remains a costly and complex task for most organizations. One of the main difficulties lies on the process elicitation phase, where the process analyst attempts to extract information from the process' participants and other resources involved. This paper describes a case study in which a previously proposed Story Mining method was applied. The Story Mining method and its supporting tool, ProcessTeller, makes use of collaborative storytelling and natural language processing techniques for a semi-automatic extraction of BPMN-compliant business process elements from text.

1 INTRODUCTION

Business Process Modeling is a very time-consuming and expensive task. Nevertheless, it is vital for most organizations, since it enables people to explicit and to share their knowledge about how business tasks are actually performed in day-to-day activities. The main factor that brings so much difficulty to this area is the knowledge issue.

Knowledge can be defined as "the combination of data and information to which expert opinion, skills and experience are added" (Davenport and Prusak, 1998). There are two main types of knowledge: explicit knowledge, externalized and able to be used, and tacit knowledge, very difficult to register or externalize (Nonaka and Takeuchi, 1995).

One of the main problems with business modeling resides on the fact that the source of knowledge is usually the activity performers of the process to be modeled. Even if documentation or other sources of information are available, generally they are incomplete or outdated. Although many techniques were developed and used for extracting the information from the performers and participants of the organization's tasks, process modeling (and specially process elicitation) remains complex.

One of the most widely-used techniques for process and knowledge elicitation is the interview. Based on a series of questions chosen by the analyst, who focuses on the relevant knowledge for the modeling and elicitation of a given process, the analyst interviews people involved with the tasks that are executed during the process. Later, the analyst transcripts the answers and analyzes them, in order to extract the process knowledge.

Nevertheless, the interview technique presents some drawbacks. First, both the selection of the questions and the interviewees can be biased. Second, further interpretation of the interview transcript by the analyst will be inevitably restricted to his viewpoint about the entire process. Alvarez (2002), during a research involving the analysis of requirements elicitation based on interviews, evidences the possibility of incomplete knowledge from the user due to the analyst's bias.

Based on previous research (Leal et al., 2004; Freitas et al., 2003), we argue that collaborative technique based on storytelling is more effective in collecting knowledge about work process. Thus, we propose a Group Storytelling approach, supported by the Story Mining method (Gonçalves et al., 2009) and a groupware tool, named ProcessTeller. A case study applying this tool in a real environment was conducted and its results are presented in this paper.

This paper is structured as follows: Section 2 briefly describes the Story Mining method, Section 3 presents ProcessTeller tool, Section 4 describes the case study and Section 5 concludes the paper.

Carlos de A. R. Gonçalves J., Santoro F. and Baião F. (2010).
COLLABORATIVE BUSINESS PROCESS ELICITATION THROUGH GROUP STORYTELLING .
In *Proceedings of the 12th International Conference on Enterprise Information Systems - Information Systems Analysis and Specification*, pages 295-300
DOI: 10.5220/0002910002950300

2 A METHOD FOR COLLABORATIVE PROCESS ELICITATION

Process Discovery techniques and methods can be classified in two distinct groups: (i) the first group emphasize the "human factor" of business process, focusing on the people that participate on the activities regarding a given process; (ii) the second group focus on the automated and non-human factors of a process, including mining process from information system logs and automatic techniques for knowledge acquisition. Both groups have different advantages and disadvantages.

We proposed the Story Mining method (Gonçalves et al., 2009) which intends to bridge the gap between these two groups, making it able to access the richness of knowledge present at the process' participant, while tapping on the swiftness of automatic proposals. Leal et al. (2004) states that a story is a natural way to make the knowledge explicit. They proposed the Group Storytelling technique, which is based on collective narrative, and the construction of stories by groups of people for knowledge elicitation in organizations. Based on free-form narrative, knowledge about different subjects can be found on a story. We argue that its application for process elicitation enhances its results. However, the large amount of collected information that is not related to the process may become a problem.

As a solution to this new problem, the use of Text Mining and Natural Language Processing (NLP) techniques was proposed, in order to provide support for the analyst at the interpretation of the story's content and the structuring of process elements in a process model. The proposed Story Mining method is divided into three main phases. The first phase is the storytelling part, where the tellers are chosen and the collaborative process happens. After the story is collectively created, the second phase applies NLP and Text Mining techniques on the story text elements. The techniques involved at this stage are out of scope of this paper and will not be described. The third and final phase automatically builds a formal representation of the process from the story. Using the elements extracted by the method's second phase as a guide, the analyst can model the process using the desired notation and tools, and discusses it with the participants for further improvements and, hopefully, reach a consensus about how the process should be depicted.

In order to support the Story Mining method, the ProcessTeller tool was created, and is described in the next section.

3 THE PROCESS TELLER TOOL

Group Storytelling is a technique based on collaborative free-form narrative, aiming at collective constructing and sharing knowledge among the storytellers and other participants.

This section describes ProcessTeller, a supporting tool for the Story Mining method previously proposed in (Gonçalves et al., 2009).

ProcessTeller is a web application, implemented as an extension of the TellStory tool (Leal et al., 2004), for process-oriented Group Storytelling that allows users to collaboratively create stories. Stories are composed of a chain of events (textual excerpts of the story). It allows alternative flows of events in the same story, additional story elements (such as documents linked to events) and group management features (such as polls).

The new functionalities of ProcessTeller are described below.

3.1 Story Groups

The original tool allowed the user to list all stories, regardless of theme or other criteria. As process elicitation is focused on a "theme" (the process itself), the tool was modified in order to show a directory structure of stories classified by Story Groups.

3.2 Document Linking to Story Events

Tellstory allows documents related to the story being told to be stored and linked to it. However, in order to benefit from Text Mining and NLP techniques as much as possible, ProcessTeller enables its users to provide specific information about every excerpt of a process description, by linking documents to each of the story events. Users are now able to upload relevant documents and determine which events are related to each document.

3.3 Event-level Character Detection

Characters are important elements of a story, having active roles during the description of a narrative. They come closer to the performer of a task in a business process (which may be represented as lanes in BPMN (BPMN, 2010) notation).

In order to improve the quality of the story being told, while avoiding direct interference with the storytelling collaborative process, at the time of a new event's insertion, if its text does not contain known characters, ProcessTeller detects it and suggests the user to register a new character or modify the event's content.

3.4 Starting Event

A story that describes a business process is not just a narrative, it usually has a starting event that enables it to happen. Nevertheless, free-form stories told by users may not present events in chronological order, or may not be initiated by its first event. Therefore, ProcessTeller enables its users to arbitrarily establish the starting event of a story. A Start Event is defined as "where a particular process will start" (BPMN, 2010). Thus, a new attribute "Starting Event" was added to each story, so as to enable the ProcessTeller to know.

4 CASE STUDY AT DIA/UNIRIO

4.1 Description

A Case Study using ProcessTeller was conducted at the Department of Applied Informatics (DIA) of the Federal University of the State of Rio de Janeiro (UNIRIO). The case study aimed to evaluate the viability of the Story Mining method (Gonçalves et al., 2009) and its supporting tool, ProcessTeller, as well as to extract evidences and new insights about the knowledge issues involving the collaborative storytelling process.

The chosen process to be modelled was Course Enrollment, since it is an extremely common process in educational institutions and known by a broad range of "storyteller candidates", including students, professors and university staff.

Although there were different processes at the institution, due to the presence of bachelor and master's degree courses at the same department, a decision was made to not divide the two processes in two different narratives but, instead, to allow participants of both contexts to tell their viewpoints and experiences in a single story, in order to assess the richness of knowledge and the differences present in the narrative.

Invitations to participate on the case study were sent through e-mail to 18 people, including undergraduate students, university staff, graduated students (at the Master's Degree level) and professors. Additionally, an open invitation was sent to the mailing lists of all students and professors.

During a full month, the users told their experiences in Course Enrollment using ProcessTeller. They were also able to read and comment each other´s contributions, include additional elements such as characters, upload new documents and relate them to the story events.

At the first week of the case study, the first user, a bachelor course student, created 8 new events, describing his entire view of the process. The analysis of the case study results pointed out that those 8 initial events functioned as a "skeleton view" of the basic process components, as other participants commented his events and added new ones, complementing the story as a whole.

As the case study progressed, additional 18 story events were added to the main story flow, and 51 new comments were made by the users. As the narrative grew larger and richer, some users have chosen to contribute only with their comments on previously created events, leaving their opinions and insights on other participants' contributions. After the end of the second week of the study, the number of event creations slowed down and the rate of adding new comments on existing events increased.

At the end of the method's first phase, the main flow of the story told by the participants and the auxiliary information registered (characters, documents, among others) were used for the second phase, the application of text mining and natural language processing algorithms.

This stage generated the proto-model in two notations, BPMN (BPMN, 2010) and XPDL (XPDL, 2010), therefore two output files were generated. BPMN was chosen since it is an OMG standard, while XPDL was used due to the fact that it is the visualization format adopted by the most popular Business Process Modelling tools. Also, an additional XML file was generated by ProcessTeller, in which each process element (Actor, Activity and Parameter) is described in more detail.

The files generated by ProcessTeller should not be taken as the final version of the process model but, instead, as an intermediary version to be validated and improved. The graphical visualization of the generated model enables the analysis to easily assess the knowledge present at the story and to modify it to achieve the final representation.

Meanwhile, these files may also be used by story participants to visualize the exposed process. This variety of ways for the analyst to assess the captured knowledge and achieve the final model composed

the third and final phase of the method, where a final formal representation of the process is built.

At the end of the case study, a questionnaire was applied to the participants, where they reported their opinions and perceptions about the experience of using the ProcessTeller tool and how close they found the final story was to the Course Enrollment process.

4.2 Analysis of Results

The results achieved by the method can be divided in two groups: The extracted activities by the method (Table 1 and Table 2), and the analysis of the Questionnaire answers from the case study participants (Table 3).

The final process model resulted from the case study (which we will call PM1) was compared to another version of the process model, which was manually created using interviews (PM2 model), in order to evaluate the precision of the automated knowledge extraction. The results of this comparison are in Table 1, depicting the coincident and non-coincident activities between the two models.

Table 2 groups the extracted activities and classifies them in three groups, based on thieir content: (i) General activities (common activities of a Course Enrollment process); (ii) Master's Degree activities (specific activities belonging to Master's degree course enrollment); and (iii) Bachelor's Degree activities (specific activities belonging to Bachelor's degree course enrollment) and "Special" Students (activities belonging to the course enrollment of visiting students at UNIRIO).

Table 1: Process models comparison.

Statistics	PM1	PM2
Total # of activities	21	51
Total # of coincident activities	8	21
Total # of non-coincident activities	13	30

Table 2: Extracted activities by group.

Group	# of Activities
General	35
Bachelor's degree	4
Master's degree	6
"Special" Student	6
Total of # Activities	**51**

The main difference between the Story Mining activities and the pre-existent model activities are the perceptions of the people involved. For instance, activities that are related to the usage of information systems or that are mainly administrative were depicted on the traditional model, while activities that were clearly visualized by the participants were present at the Story Mining's activities.

A number of suggestions can be raised from this observation about the activities. Although characters are already presented on a story, special elements of the process, such as systems and business rules, can be described in different ways, stimulating the tellers to describe activities involving them. Table 2 also shows that there is a tendency for an increasing capture of activities common to a "generic" Course Enrollment process, composed by activities like "Student selects a course" and "Professor offers course". However, the activities related to alternative flows (for example, activities specific to each course type) are also captured as well, even if a smaller number of participants may have been in contact with them. The answers from the questionnaire are summarized in Table 3. A selected number of questions were evaluated, regarding the participants' opinions about several characteristics of the study and the tool itself.

Table 3: Questionnaire results.

Question	Agree	Indifferent	Disagree
The tool allowed the expression of your viewpoint regarding course enrollment?	10	1	1
Telling a story in a collaborative way allowed the easy expressions of your viewpoint about course enrollment?	9	2	1
The event document attachment functionality was useful to complete your viewpoint of the event?	7	5	0
Automatic detection of character-less events stimulated the register of new characters?	6	4	2
Polling was useful to solve incoherence and conflicts?	2	10	0
Event operation "Switch Places" was useful for a better flow of story events?	9	2	1
Event operation "Join Events" was useful for a better flow of story events?	9	3	0
Event operation "Break Event" was useful for a better flow of story events?	9	3	0

Table 3 shows a concise view of the case study. The first questions were very important, since they are related to the validity of the method for process elicitation and the ProcessTeller tool. The majority of users confirmed its usefulness, while a small group disagreed. The main causes of disagreement were: the blend between master's degree and bachelor's degree course enrollment events on the same story and the inability, at a case study level, for each user to tell a separate story.

Three questions had less than 75% approval (i.e. 9 replies on "Agree") from the users. The first two regards character detection and the attachment of documents to events. They may be explained due to the reduced number of new events as long as the story developed and the process depicted achieved its completion, as some users weren't able to experience the character detection feature of ProcessTeller, explaining the high number of "Indifferent replies".

The last question is about the poll feature and has the highest number of "Indifferent" replies as well as zero "Disagree". This fact points out that polls were not used during the case study and the users arguably preferred to use comments to discuss issues related to the events.

Also, three open questions were present, regarding their viewpoint about the process depicted and the users' contribution to the story, the final story generated, the tool itself and other general comments about the elicitation task as a whole. Specific trends were observed from the users, based on their answers to these questions:

Many users could not assert that the story reflected the complete or correct Course Enrollment process, due to the fact that they had a limited viewpoint of it. But, on the other hand, they were confident that their contributions were correct and that their experiences with the process were reflected on it.

Another group of users expressed their surprise with details about the process that they were not aware of, because they had contact with just a small part of its activities.

An abundance of knowledge was also noticed by the participants, as some of them stated that "there is a mix of master degree and bachelor degree events" at the story. The majority of them agreed that they expressed their limited knowledge, but were satisfied by the final story, as being the reunion of these smaller contributions in a collaborative way.

The fact, highlighted at the case study's description, of a single user registering many events, triggered an increase of contributions (new events, comments, as well as new characters and documents). It reinforces the value of the collaborative element of Group Storytelling, as people probably would not recall so many process elements, without the aid of the first user's narrative elements.

Finally, the increasing usage of event comments was unexpected, specially in cases where the comment feature was used as an "ad-hoc forum" for discussions on a specific event's contents.

5 RELATED WORK

Indulska et al. (2009) carried out a study in order to identify a research agenda for process modeling. They include the following items: the value of process modeling and expectations of stakeholder groups involved in process modeling. One of their conclusions was that group design is an advantageous approach.

Ryan and Heavey (2006) argue four requirements for a collaborative approach in process modeling: have a low modeling cognitive load and therefore be capable of being used by non-specialists; present modeling information at a high semantic level so that personnel can rationalize with it; have good visualization capabilities; and, support project teamwork. Our approach is aligned with all of them, although that graphic visualization will only occur at the end of the interaction when model is generated.

Freitas et al. (2003) propose a cooperative graphic editor (CEPE - Cooperative Editor for Processes Elicitation) that supports the building of the knowledge about the current process and intends to support the reporting of associated problems.

6 CONCLUSIONS

The case study conducted and described in this work suggests that the Group Storytelling technique may be successfully applied for process elicitation. The participants' responses about the extracted process support this observation, since most of them agreed that the story reflected the desired process.

The main advantages of the proposed technique, compared to the traditional approach, is the lessening of analyst's bias as well as the free expression of its own knowledge by the participants, regardless of their level of involvement with the process.

Additional features could improve the tool and the elicitation method, as deeper detailing on characters, the use of more complex relationships between events and the linking of extracted activities and story events, making it able for the analyst to cross-reference the method output with the original part of the story.

For the Text Mining and NLP techniques phase of the method, the increased usage of text comments on event, surpassing the number of events of the story brings up the need for a future consideration on the comments' text use as input for these algorithms, broadening the range of its application from the story events to all textual elements available.

ACKNOWLEDGEMENTS

The authors would like to thank Petrobras for supporting this project. Flávia Maria Santoro is partially supported by CNPq (Brazil) under the grant 305404/2008-3 and João Carlos de A. R. Gonçalves is supported by CAPES.

REFERENCES

Alvarez R., 2002. Discourse Analysis of Requirements and Knowledge Elicitation Interviews. In *Proc. 35th Hawaii International Conference on System Sciences*.

BPMN – Business Process Modeling Notation, http://www.bpmn.org/Documents. Last accessed on January 2010.

Davenport, T. and Prusak L., 1998. *Working Knowledge,* arvard Business School Press.

Gonçalves, J. C. A. R., Santoro, F. M., Baião, F. A, 2009 Business Process Mining from Group Stories. In: *Proceedings of 13th International Conference on Computer Supported Cooperative Work in Design*, 161-166.

Leal, R. P., Borges, M. R. B., Santoro, F. M., 2004. Applying Group Storytelling in Knowledge Management. In *IX International Workshop on Groupware (CRIWG),* Lecture Notes in Computer Science 3198 , 34-41

Freitas, R., Borges, M. R. B., Santoro, F. M., Pino, J. A., 2003. Groupware Support for Cooperative Process Elicitation. *Lecture Notes in Computer Science*, Paris, v. 2806, 232-246.

Indulska, M., Recker, J., Rosemann, M., Green, P, 2009. Business Process Modeling: Current Issues and Future Challenges. *Advanced Information Systems Engineering*, 501-514.

Nonaka, I. and Takeuchi H., 1995. *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation*, Oxford University Press.

Ryan, J., Heavey, C., 2006. Process modeling for simulation. *Computers in Industry*, 57 (5), 437-450.

XPDL – XML Process Definition Language. http://www.wfmc.org/xpdl.html, last accessed on January 2010.