BPMNCommunity.org: A Forum for Process Modeling Practitioners – A Data Repository for Empirical BPM Research

Alexander Grosskopf, Jan Brunnert, Stefan Wehrmeyer, and Mathias Weske

Hasso-Plattner-Institute, Potsdam 14482, Germany,
`{alexander.grosskopf,mathias.weske}@hpi.uni-potsdam.de,{jan.brunnert,`
`stefan.wehrmeyer}@student.hpi.uni-potsdam.de`,
WWW home page: `http://bpt.hpi.uni-potsdam.de`

Abstract. This short paper reports on an online platform where modelers can gather, share and discuss knowledge around BPMN. The models, ratings, comments and discussions are no longer based on pictures but related to actual process models that can be edited in the web browser. After two months 97 registered users developed and shared 166 process models and 372 revisions. In this paper we introduce the platform, show its architecture and provide samples for data extraction and analysis. We invite researchers to use the available data to conduct further empirical studies.

Key words: Online Communities, Oryx, Process Modeling, BPMN

1 Introduction

Business process modeling is at the heart of modern organizations, since process models capture how work is performed in the organization and how business goals are reached. Recently, the Business Process Modeling Notation (BPMN) became the de facto standard in process modeling. We observed that, dealing with this new language, people desire to discuss the best or most adequate model to represent a common real-world situations. We want to support these discussions.

In this paper we introduce BPMNCommunity.org, a web collaboration platform that supports discussions about process models. Other than traditional wikis, the process modeling environment Oryx [1] is embedded, so that process models can be created and modified with standard web browsers. All models in the community are public knowledge. Together with tags, ratings, descriptions, comments, and the revision history, it creates a diverse data pool that can be leveraged for empirical research in business process management.

This paper is organized as follows. Section 2 introduces the major use cases supported by the platform, before Section 3 sketches its software architecture. Section 4 looks at the available data for empirical research, before the paper concludes with project information.

## 2 Use Cases and Functionality for the Community

We started with the goal to create a user-driven community. Following the wiki principle all content can be read and changed by all users. Models can be described, tagged, rated, and comments can attached to modeling elements. This allows discussions to take place directly at the models. We tailored the platform to meet three main use cases: learn BPMN, jointly create a processes, and discuss good modeling style.

BPMN novices can learn the language using the tutorial section. Tutorials and exercises are created by users for users. Solutions to exercises can be submitted and rated. To create models together users can use the groups section of the platform. One group can have multiple members and processes that relate to the topic. At present, groups have formed, e.g. to show the Workflow Patterns [2] modeled in BPMN, discuss reference processes or the relation of EPC and BPMN. Discussions about modeling alternatives and good style are captured in the best practice modeling section. Users can start a topic similar to a group. Alternative solutions can be modeled and marked as a good or bad sample to model a given situation in BPMN. By now, best practices were created on topics such as multiple start events, reactive processes and batch processing of events.

Additionally, the platform provides a dash board with widgets that aggregate information from the platform and the internet. Widgets show new or most active users, recent news, blog posts or twitter messages related to BPMN, and more information that can be customized by the user to have a cockpit into the BPMN world.

3 Platform Architecture

The web platform was built with Django, a Python Web framework. As mentioned before, we integrate the Oryx Editor [1], a web modeling tool. The process models are stored in the Oryx data base. They are accessible through HTPP interfaces in various formats, such as ERDF, JSON, PNML (Petri Nets Markup Language), and picture formats. For BPMNCommunity.org we created an API to easily access, navigate and comment models in websites.
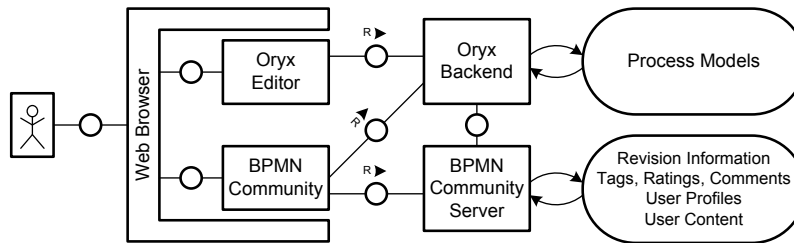

Fig. 1. System architecture of the BPMN Community

To edit a model, users are redirected to the Oryx modeling environment. Upon saving a new model is created and this information is passed back to the

community platform. There we manage the revision links, the user profiles and all other content, except for the process models themselves. Figure 1 depicts available information in the community database.

4 Community Data Analysis - An Example

We assume that the data created in the community is interesting to researchers for analysis. As an example, we used the Kettle ETL tool [1] and custom scripts to load and transform the data from the community database into an analysis database. We extracted the models as ERDF and PNML format using existing Oryx web interfaces. This model data can be used to calculate process metrics.
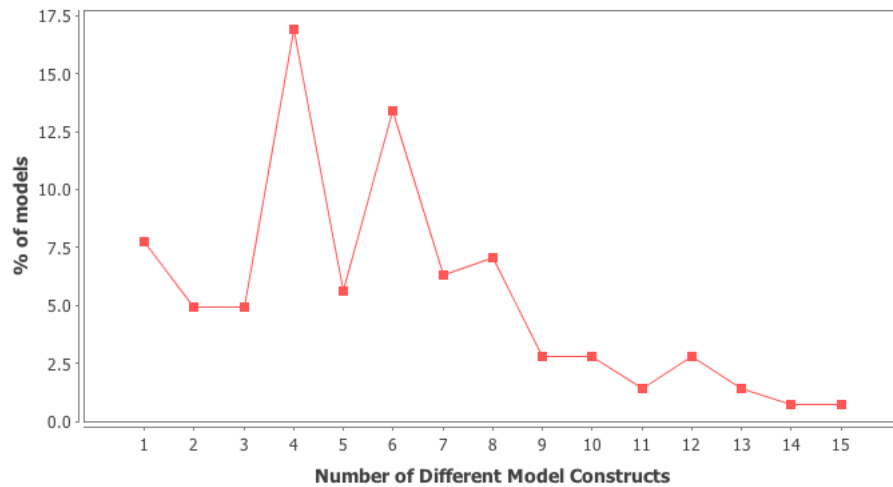
Fig. 2. Syntax Complexity Graph for 166 head revision models (26th of June 2009)

To demonstrate this we created a sample script that counts the number of model elements.Then we replicated Figure 5 from the 2008 paper 'How much language is enough' [3] using the community data, see Figure 2. The other graphics from the same paper can be replicated alike. For sophisticated metrics calculation, we used the PNML exported models and a modified version of ProM [4] that can bulk-process PNML models. This allows us to reuse the existing Petri-Net complexity calculation implementations in the ProM framework yielding metrics such as Density [5] or CFC [6].

These formal metrics could be related to the ratings, comments, revision history and other information given in the community. This is yet to be investigated. As an example, one could relate structural complexity metrics to the number of distinct editors, the editor experience or some kind of editor trust metric. Furthermore, as a process model is a result of a social process, meta-data gathered

---

[1] http://kettle.pentaho.org

during the evolutionary development in a community context may prove valuable in a holistic approach to researching the characteristics of process models.

5 Conclusion and Future Directions

This paper introduced BPMNCommunity.org a web platform to learn BPMN, co-create models and discuss them. We outlined the functionality and architecture. We believe that the gathered information is valuable not only for BPMN practitioners but also for empirical research. By example, we showed how data can be accessed and used for calculations.

Within the first two month the platform attracted more than ninety registered users. We aim to expand its reach and get more people involved in the discussions around BPMN modeling. With the new standard revision, BPMN 2.0[7], to be finalized soon we expect even more need for a place to meet, discuss, and develop knowledge for BPMN modelers.

The steadily growing raw data base offers an interesting playground for empirical studies. We propose a common data model and tool chain to extract, transform, load, analyze and compare process models and meta-data, which is applicable not only to our community. We invite other researchers to join hands on this and use the available data from BPMNCommunity.org to assess existing metrics or test new ones.

References

1. Decker, G., Overdick, H., Weske, M.: Oryx–An Open Modeling Platform for the BPM Community. In: Proceedings of the 6th International Conference on Business Process Management, Springer (2008) 382–385
2. Russell, N., Ter Hofstede, A., van der Aalst, W., Mulyar, N.: Workflow control-flow patterns: A revised view. BPM Center Report BPM-06-22, BPMcenter. org (2006) 06–22
3. Zur Muehlen, M., Recker, J.: How much language is enough? theoretical and practical use of the business process modeling notation. 465–479
4. van Dongen, B., de Medeiros, A., Verbeek, H., Weijters, A., van der Aalst, W.: The ProM framework: A new era in process mining tool support. Application and Theory of Petri Nets **3536** (2005) 444–454
5. Mendling, J.: Testing density as a complexity metric for EPCs. In: German EPC workshop on density of process models. (2006)
6. Cardoso, J.: Complexity analysis of bpel web processes. Software Process: Improvement and Practice **12**(1) (2007)
7. OMG: Business Process Model and Notation (BPMN) Specification 2.0 V.0.9.12. (May 2009)