BPM Academic Initiative
Fostering Empirical Research

Matthias Kunze, Philipp Berger, and Mathias Weske

Hasso Plattner Institute at the University of Potsdam
Prof.-Dr.-Helmert-Strasse 2-3, 14482 Potsdam,
{matthias.kunze,mathias.weske}@hpi.uni-potsdam.de
philipp.berger@student.hpi.uni-potsdam.de

Abstract. The BPM Academic Initiative strives to support education and research in business process management. This paper announces a platform to be used by researchers to download process models, providing data to be used in empirical research. This paper presents a web portal, where process models can be filtered by various criteria and downloaded, and a research platform that facilitates analysis of the downloaded models by means of a small show case.

1 Introduction

We started the BPM Academic Initiative (BPM AI) together with colleagues from the BPM community in late 2009. The goal of this endeavor has been twofold. First, to support education in business process modeling and analysis by providing a professional software tool free of charge together with assignments to be used by lecturers. Secondly, to strengthen research in our area by showcasing recent research results. Today, the system is used by more than ten thousand students, lecturers, and researchers world wide. At the same time, several research prototypes have been integrated with the platform, including soundness checking, structural analysis of process models, and business process simulation.

In this paper we add to these goals by opening a new service to the BPM community: to provide process models for empirical research in business process management. We hope this service will be as successful as the other parts of the BPM Academic Initiative are today. We are quite confident, since in recent years empirical BPM research has become more and more prominent. However, in many cases researchers find it hard to get access to process models. There are a handful of process model collections that have been used in empirical research on process models. However, these model collections reveal different internal formats, so that researchers need to develop software to access them.

This paper introduces and announces the availability of a platform to filter and to download process models, with accompanying software to process them, such as parsers. In the current version, models of the BPM Academic Initiative collection are provided, but future versions shall also provide additional process model collections. Thereby, researchers are provided with process models and with

software that can help processing them. As a result, researchers can concentrate on their particular research questions, which we hope will strengthen empirical scientific work related to process models and process model collections.

The paper is organized as follows. After a brief discussion of the background of the BPM AI and the offered process models in Section 2, we present a web portal to filter and to download the models in Section 3. Finally, we sketch a show case that is also provided as a short screen cast at `http://vimeo.com/43098307`, in Section 4.

2 Background

The BPM Academic Initiative offers a number of services to support education and research, depicted in Fig. 1: a collection of teaching material, a professional, web-based process model editor and collaboration platform, and the process model provisioning introduced in this paper. We presented the former two components in an earlier demonstration [4], whereas the latter component provides a recent addition to the services of our initiative.

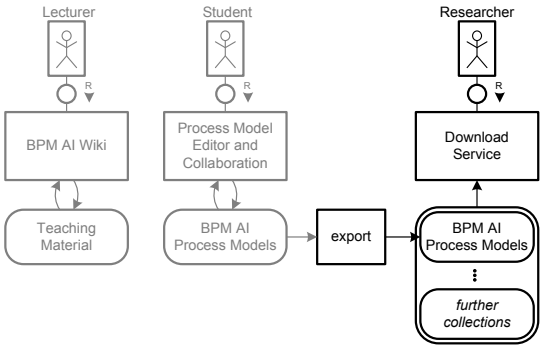

Fig. 1: Overview of services provided by the BPM Academic Initiative

Teaching material made available through the *BPM AI Wiki*[1] comprehensively captures topics of the field of business process management. All materials are offered publicly under the terms of a Creative Commons license and lecturers are invited to contribute to the content. The wiki provides further information about the BPM AI and references related publications.

To facilitate process modeling as part of teaching activities, assignments, or research, a professional, web-based *process modeling platform* from our industry partner Signavio is provided to academics free of charge. The platform sports a holistic set of process modeling languages, e.g., BPMN (including conversation and choreography diagrams), EPC, and Petri Nets. Work from many contributors has made its way into the platform and is offered to the users, as will future

[1] BPM Academic Initiative wiki, cf. `http://bpmai.org`

developments, especially with regard to new scientific features and modeling languages. For instance, LoLa soundness checker [2] and bpstruct [5], a tool to structure spaghetti-like process models automatically, are integrated with the platform; future enhancements are welcome. Additionally, this tool comes with collaboration features that allow multiple process designers to jointly create a process model and discuss revisions by comments.

A multitude of models has been created using the process modeling platform, already. Now, we make these models available for research. Therefore, when signing up, every user of the platform agreed that their models may be reused for empirical research. A subset of these models has already been used towards understanding process modeling [3]; the authors also propose challenging research topics in the context of process model collections. Now, these models are offered to all interested researchers.

3 Service Description

The service to download process models addresses researchers that aim, for instance, at evaluating and validating their research results. As the BPM AI models have been created by students, lecturers, and researchers, they show a high heterogeneity in terms of the used natural language and modeling language, business domain, and quality. Hence, empirical insights that are derived using the BPM AI collection can be assumed to have a high external validity.

By June 2012, over 10000 users have adopted the academic initiative and created over 85000 models that comprise 290000 revisions.

Access and Download Process Models. Process models can be downloaded from a web portal, i.e., no specific software is required to access them.

At `http://bpmai.org/download/`, users are presented with a filter interface, shown in Fig. 2, that allows them to select a subset of models that match following criteria.

Modeling Language allows restricting downloaded models to one or several model types, e.g., BPMN models.

Language denotes the natural language of process models that we derived from the inscriptions of model elements. In particular, research that incorporates labels, e.g., process model similarity, is typically sensitive to the used language.

Connectedness is a quality measure that evaluates the size of the largest connected graph towards the size of the overall model. If a model consists of many disconnected fragments, this measure will be low.

Size counts the number of nodes in a process model and can be used as a simple means to address complexity.

Revisions. As the process model editor creates a new revision of a model each time it is saved, users can choose, whether they want to download all or only the latest revision of the diagrams.

In the filter interface, the user can obtain a random example chosen according to the current filter settings.

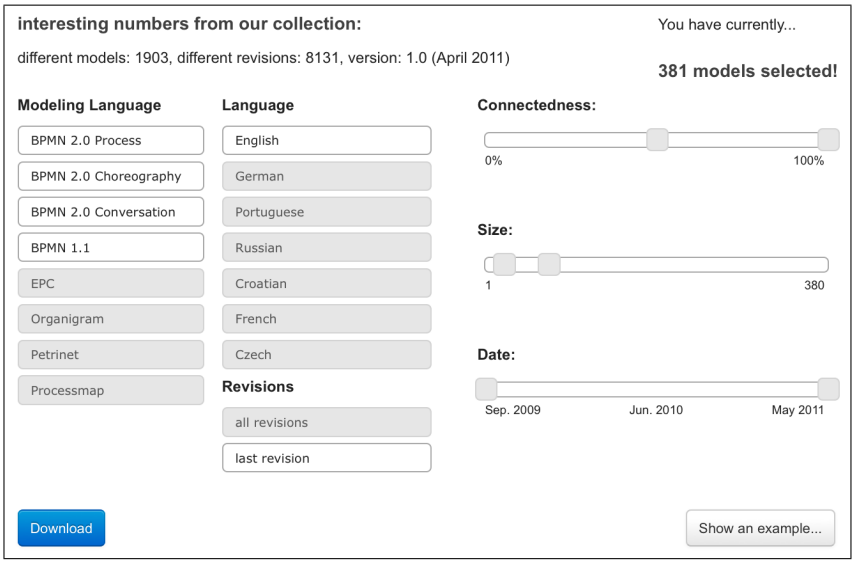

Fig. 2: Interface to filter process models by various criteria.

Once the user clicks the download button, they are presented a form that requests some information, i.e., the researcher's name, email address, affiliation, and a short research proposal to which the models shall contribute. The BPM AI core team will review the research proposal and grant access to earnest requests, thus avoiding abuse of the service. Also, before submitting the request, one has to accept a license agreement that restricts usage of the models to empirical research in non-commercial settings only.

When the request has been granted, the researcher will receive an email with a unique link to download their models. Upon choosing to download, a task will be scheduled in our system that extracts the models selected by the user's filter and creates a zip file. As this may incorporate a significant amount of time, the user will receive an email with the download link.

Research on Process Models and Collections. The downloaded models unzip to a directory structure, and for each process model revision, a JSON file and an SVG file are provided. The JSON file contains the model's structure and attributes, and is used as the internal format of the process modeling tool, whereas the SVG file provides a ready-made vector graphic to display the diagram.

To support researchers in disseminating the directory structure and parse the JSON representation of diagrams, we also offer an open-source platform to process model collections research [1]. This platform provides import functionality for the BPM AI model collection, among others. Once the downloaded models have been imported, the platform offers utilities to filter, transform, and extract information from the process models, similar to the pipes and filter enterprise integration pattern. A mapping for EPC and BPMN models to a generic process model representation is provided, such that features of the jbpt[2] Java library can

[2] jbpt, cf. `http://code.google.com/p/jbpt/`

be used, e.g., workflow graph parsing, net unfolding generation, and to derive behavioral profiles.

4 Show Case

The demo addresses all researchers that focus on process model aspects, in all phases of the business process lifecycle. Models of the BPM AI are generally operative models, i.e., they do not contain technical details as required for enactment or performance evaluation.

In the demonstration, we present a show case that targets at the creation of a word cloud from activity labels in chosen process models. For a short screen cast of this show case, visit `http://vimeo.com/43098307`.

1. In the first part, we explain the capabilities of the user interface to select a subset of process models from the collection. We concisely present the characteristics of the filter criteria.
2. The process of requesting access to the process model collection by providing a research proposal, receiving a response, and downloading the collection is laid out comprehensively. We also discuss the structure of the directories and files included in the downloaded zip file.
3. Finally, we will use the process model collection research platform to extract activity labels from the given process models, independent of their modeling language, and feed them into a word cloud generator. This is intended to show, how researchers can leverage the knowledge of the BPM AI models with small effort.

   A word cloud visualizes the distribution of words in a large set by their size printed on a canvas; an example is depicted in Fig. 3.

Fig. 3: Example word cloud.

References

1. Rami-Habib Eid-Sabbagh, Matthias Kunze, Andreas Meyer, and Mathias Weske. A Platform for Research on Process Model Collections. In *BPMN 2012*, (to appear).
2. D. Fahland, C. Favre, J. Koehler, N. Lohmann, H. Völzer, and K. Wolf. Analysis on demand: Instantaneous soundness checking of industrial business process models. *Data Knowl. Eng.*, 70(5):448–466, 2011.
3. M. Kunze, A. Luebbe, M. Weidlich, and M. Weske. Towards Understanding Process Modeling – The Case of the BPM Academic Initiative. In *(BPMN 2011)*, volume 95 of *LNBIP*, pages 44–58. Springer, 2011.
4. M. Kunze and M. Weske. Signavio-Oryx Academic Initiative. In *BPM 2010 Demo*, volume 615 of *CEUR*, 2010.
5. A. Polyvyanyy, L. García-Bañuelos, and M. Dumas. Structuring acyclic process models. In *BPM'10*, volume 6336 of *LNCS*, pages 276–293. Springer, 2010.