Class Notes: BPM Research and Education

Jan Recker

Jan Recker [j.recker@qut.edu.au]

BPM Technology Research – the Case of Apromore

Jan Recker with Marcello La Rosa

Over the last Columns, I have dipped into various aspects of the academic life as it surrounds the management of business processes – I have shared views on the debates between industry and academia [1], sketched out current trends in expanding BPM to new frontiers [2, 3], and reviewed efforts in BPM classroom settings [4].

My intention for this Column was to examine the technological side of BPM research and education a bit closer. To that end, I have teamed up with two colleagues, Dr Marcello La Rosa, who heads an exciting initiative on the development of advanced BPM technologies, in the form of an Advanced Process Model Repository – Apromore, and Eike Bernhard, a doctoral student researching the impact of process modeling technologies on organizations. Our intent is to use the example of Apromore to showcase how BPM technologies are incepted, developed, built and applied.

The Starting Point

All started with the realization that organizations, over recent years (if not decades) have significantly expanded their BPM capabilities reaching higher levels of BPM maturity, and consequently, did a lot more process modeling work. Many process-related initiatives almost inevitably lead to organizations not only having some of their processes documented and modeled, but very often most if not all of them existing in different versions, variants, as as-is models, to-be models, should-be models or even could-be models. In fact, enterprises tend to have hundreds if not thousands of process models by now. It is quite common that such sets of industry-strength business process models include thousands of process activities plus further related information about process data, applications, risks and so forth. These models are increasingly published over an Intranet to a large number of stakeholders with varying skills and responsibilities.

While these efforts are laudable, it may not come as a surprise that many organizations find it difficult to keep track of such large amounts of complex process models – an issue that is referred to as the challenge of "model (lifecycle) management" [5], "model maintenance" [6] or "modeling in the large" [7]. The problem is amplified by overlapping content across models, poor version management, process models that are used simultaneously for different purposes, the use of different modeling notations such as EPCs, BPMN, etc. Do we need to go on?

A Challenge and an Opportunity

The requirement to deal with an increasing number of business process models within organizations forms both a challenge and an opportunity. The challenge is how to sensibly deal

with such large volumes of models and how to scale up approaches that we found to work with individual models or even with a small set of models. Especially, it becomes essential to keep track of the various models, as they may refer to each other, mutually overlap, supersede one another, and evolve over time.

Let us illustrate this challenge with some examples from our work: When we talk to organizations undertaking process modeling initiatives, they often report on their struggle with two key dimensions – model creation and model management. The first dimension relates to populating an organizational repository with process models as documentations. In other words: While populating the repository, how can the development of multiple different forms of process documentation across business units be prevented and how can duplicate work be avoided? One of our industry partners in charge of enterprise-wide modeling efforts put it this way:

> "*We have many areas which have run projects and modeled their own processes 'from scratch' without looking at what already exists.* […] *I find that users create duplicate activities and indeed fragments, very quickly, and often across multiple business functions which each have their own quirks and resourcing issues.*"

The most significant challenge in this regard is how process models – or at least significant parts of them – can be re-used for different purposes. This is especially important since modeling work entails substantial costs.

The second dimension of struggle, model management, concerns problems organizations face after having identified and documented their processes in large model collections. One struggle here, for example, concerns standardization: how can the number of existing model variations be reduced? In a banking corporation, for example, many analogies between private and corporate banking processes exist that could be or should be identified from the models – beyond what the customer sees. Being able to identify these common parts across sets of process models goes a long way not only in standardizing processes across variants or product lines, but may also help to identify company-internal best practices and mapping out where these best practices could be re-used in the organization.

Another struggle deals with the matter of simplification: use, review, analysis and improvement of processes benefits from a smaller and simpler model collection. Many models in a repository are typically not that relevant because they may be drafts, old versions, overhauled as-is models, under-utilized variants and so forth.

The currency of their process models is also a challenge for organizations. If models are created for a certain purpose and never touched again after that, they are of limited value. In fact, most organizations report that they have "*quite a few models that are old and out of date*". To ensure continued use, they have to be updated regularly to reflect actual rather than historic practices. This now links back to standardization – the existence of dozens of model variations severely complicates change implementation.

Addressing these challenges does not only entail significant savings in terms of less modeling work and more assistance to the modeling teams in an organization, but also provides amplified opportunities to facilitate process improvement as well as cross-departmental communication and coordination. These opportunities are only possible if *innovative process model management technology* is made available, which can unleash the latent knowledge and power that lie in process model repositories.

And indeed, our initial explorations of this topic highlight a number of facts about "current practices". We learned, for example, that current users are not entirely satisfied with current proprietary BPM repositories. One of the related problems stems from the fact that organizations often use multiple modeling notations, especially in different task settings; for instance, when they wish to communicate processes to management and end users versus when the same processes

www.bptrends.com   2

need to be modeled for automation. Support for "*swapping between model languages [would be] excellent*."

A repository with advanced features could go beyond the obvious and traditional management of model collections, but actually affect the business. One of our key informants noted that the emerging implications would yield potential benefits

> "*not only in terms of managing the repository but as a means to simplify the business and potentially reduce costs.* […]. [It yields] a *shift in focus from how it helps you manage your repository of processes, to how it helps you manage your business. The business benefit of reducing the number and complexity of processes that the business maintains and runs.*"

So, opportunities abound. The key research that is being undertaken on the Apromore platform is thus driven by the aim to develop an initial solution to the above challenges as well as to enable new opportunities around the management of process model collections in terms of identifying and leveraging latent knowledge and possibilities from process model collections that may have remained unknown until now.

What does Apromore do?

Apromore [8] is a business process model repository. Like any other repository, it can be used to import, export and store process models, and control user access to these models based on authorization rules. Beyond these typical repository amenities, Apromore offers sophisticated, state-of-the-art features to facilitate the management of large process model collections. These features are especially useful when dealing with large volumes of process models that are interrelated on different levels of abstractions and in various relationships (predecessor, successor, resource dependency) on the same level. Apromore's features can be classified according to four broad service areas (see Figure 1):
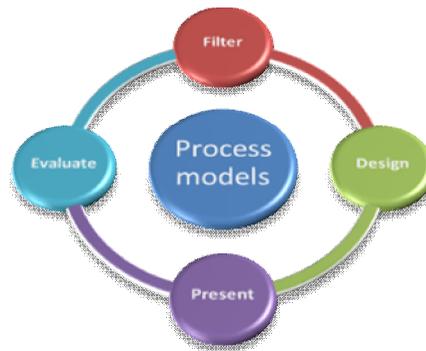

Figure 1: The Four Service Areas of Apromore

1) **Evaluation**, offering capabilities for checking the quality of process models based on various notions such as syntactic quality and pragmatic quality (including understandability and maintainability).

2) **Filtering**, offering capabilities to filter the repository based on different criteria, for example selecting all process models similar to a search model, or all process models that contain a given path of activities.

www.bptrends.com   3

3) **Intelligent Design**, supporting the creation or completion of process models by reusing existing content, for example creating a new process model by merging a set of similar process models.

4) **Presentation**, simplifying the understanding of large process models or collections thereof, by, for example, abstracting from details or highlighting certain model elements.

It's very common that within the same organization more than one process modeling language is used. To cope with this problem, Apromore supports a variety of languages, including BPMN, eEPCs, YAWL and Workflow nets. When imported, a process model is converted into Apromore's canonical process format. This format is used for internal representation (though a copy in the native format is also kept). The advantage of using this canonical format is that any operation can be performed over models represented in different languages. For example, one can compare an EPC with a BPMN process model and then export the results in either of these two languages, or in a third language, such as YAWL.

The Building Blocks

Apromore is implemented according to a service-oriented architecture (see Figure 2) and deployed over the internet as a Software-as-a-Service. The core service is the Repository Manager. This service exposes all the repository features via Web service operations for integration with third-party applications, e.g. an external BPM System.

Apromore offers a standard client to access these features. This client is the Portal, a Web application able to render on screen all features internally provided by Apromore through dedicated user interfaces. In turn, the Portal communicates with the Editor to allow users to visual and modify the process models in the repository, and create new ones.

The Canonizer is responsible for canonizing process models (i.e. converting them into Apromore's canonical process format) as they are imported into the repository, or to de-canonize these process models back to a native format (e.g. BPMN 2.0 XML, XPDL, EPML) as they are exported from the repository. The Canonizer is equipped with a plug-in interface such that new formats can be supported simply by dumping new plug-ins into Apromore.

The Toolbox is a facade over the advanced features that can be performed on the stored process model collections, such as filtering the models according to their similarity or merging them in a new model. Similarly to the Canonizer, this service offers a plug-in interface for new features to be added to Apromore on the fly.

Access to the models stored in Apromore is achieved via the Data Access service which encapsulates data-centric operations for reading/writing data upon requests made by the other services. Finally, the Security Manager controls security aspects such as user authentication and authorization.

From a technology perspective, all services are implemented as Java Spring components while JAX-B is used for the internal representation of Java objects. WSDL is used to describe the Web service interface of the Repository Manager, while the Web service messages are exchanged through SOAP. The portal is implemented using the ZK Enterprise Framework[1] whereas the plugin interface is realized via OSGi on top of the Eclipse Virgo application server.[2] The backend database is MySQL. The editor is realized by reusing components of the Signavio Core Components project.[3]

---

[1] http://www.zkoss.org

[2] http://www.eclipse.org/virgo/

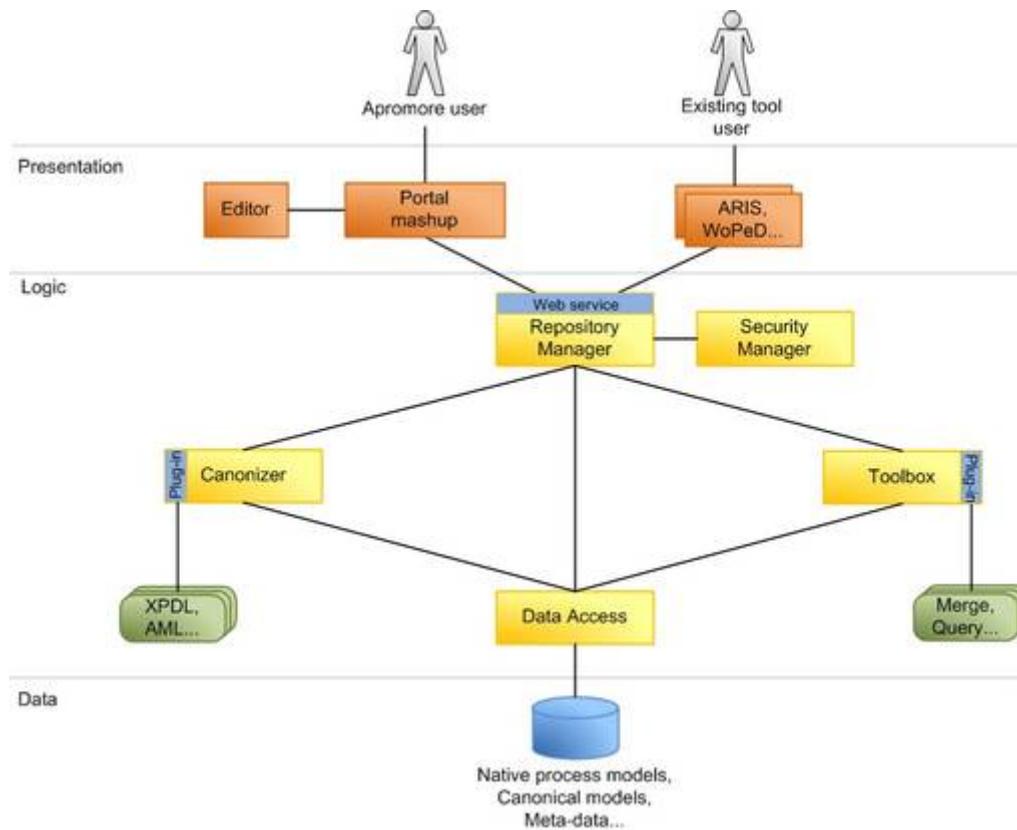[3] http://code.google.com/p/signavio-core-components/

Figure 2: Apromore's architecture

A Closer Look Inside

At the time of writing, Apromore caters for various advanced features, each exposed as a plugin. These include language interchange, exact and approximate clones detection, fragment-based version control, fragment-based access control and locking, change propagation, structuring, similarity search and process merging. Over time, ongoing research and development will provide further features, each being made available as a plugin that can easily be added to the existing features. Below, we illustrate some of the existing features in more detail.

Exact Clone Detection and Change Propagation
Apromore can identify fragments that are shared by different process models across the whole repository in a matter of milliseconds. These fragments are called "exact clones". If a clone is updated from within a given process model, all process models sharing that fragment are notified of the change. Then the change is propagated according to the propagation policy specified for each process model. If an "instant propagation policy" is used for a process model, this model will be automatically updated as soon as one of its fragments gets changed. If a "controlled propagation policy" is used, the owner of the process model can review the available changes and apply only the required ones.

You can use this method to avoid creating "unwanted" inconsistencies across the process models of your repository. Moreover, identifying exact clones offers opportunities for implementing these clones with shared services, thus it can help you rationalize the IT infrastructure implementing your business processes.

Similarity Search and Process Merging
Apromore can find all process models that are similar to a given process model – the "search model" – very efficiently. The similarity between two process models is determined by a "similarity

threshold" that takes into account both the textual similarity between labels of process elements and the graph similarity between process model structures. Roughly, a threshold of 50% means that two models share 50% of their structures and have labels which are 50% similar to each other. The similarity threshold is used to rank the search results, so that you can easily identify the most similar process models to the search model.

Once two or more models have been found that are sufficiently similar to each other, you can consolidate these models by *merging* them in a "configurable process model". A configurable process model [9] is essentially the union of all the elements of the similar models. Still, you don't want to lose track of the differences that existed between these models. In order to avod this problem, configurable process models feature variation points to distinguish commonalities from variant-specific branches, and annotations to mark where each element in the configurable model originates from.

These variation points can be used to create projections of the configurable process model that only show certain elements through an operation called "configuration". For example, one may project the configurable process model to one of its input models, or to a model showing only the commonalities between all input models used for the merge. The idea is that you can maintain only the configurable process model, and create projections of this model on demand, depending on your target audience.

This is only a glimpse at some of Apromore's advanced features. If you want to know more about Apromore, check out http://apromore.org.

Finally, an Invitation

We expect Apromore will generate significant impact on the process modeling practice within organizations, as well as stimulate further research into the discipline of process modeling. The direct beneficiaries of Apromore are clearly those organizations that already apply process modeling on a large scale and wish to extract greater value from their process models; specifically those who are involved in process improvement plans. Their success will be enhanced to comply with regulations and innovate in their operations using process modeling. Moreover, technology vendors may find it interesting to extend their offerings by tapping into Apromore advanced features.

Our BPM research group is collaborating with Suncorp, one of Australia's top 25 listed companies, on a project supported by the Australian Research Council, to further develop Apromore, and explore its potential to solve some key BPM problems. In particular, this project is supported by Suncorp Insurance division. Suncorp Insurance offers more than 30 products for personal, motor vehicle and commercial insurance by controlling over 15 insurance brands, which are the result of a series of mergers and acquisitions the company has recently gone through. This has led to a repository of more than 3,000 process models, managed by various teams of analysts. This collaboration with Suncorp provides us with the opportunity to work with real-life business process models, and the possibility to conduct case studies to validate the outcomes of our research.

We are eager to extend our network of collaborations and secure further participation from industry. For this purpose, we established the **Apromore Consortium**, which seeks to gather organizations which deal with the problem of managing large process model collections. If you have a related challenge or opportunity, do let us know by contacting one of the authors.

Acknowledgments

In this Column we discuss ongoing research work at the Queensland University of Technology (QUT), as well as in a network of academics around the globe. Several individuals from various

universities besides QUT are contributing to the Apromore initiative. These key contributors are listed at http://apromore.org/about, and we herewith appreciate and acknowledge their work, input and views. As usual, our interpretation of their work should be considered as the authors' personal views. We also thank Eike Bernhard for his insights into model management practices and for his input to writing this paper.

Please feel free to contact us with your suggestions, feedback and comments, or for a copy of articles related to the topics above.

About the Authors

Jan Recker is a Full Professor and the Woolworths Chair of Retail Innovation at the Queensland University of Technology in Brisbane, Australia. Jan's research interests focus on the use of process design in organizational practice, business transformations and organizational innovations. Jan has authored and edited several books, including one on BPMN and one on Green BPM. He also co-authored over 100 academic papers in journals and conferences and presented his research all over the globe. He holds a PhD in Information Systems from Queensland University of Technology and a MS in Information Systems from the University of Muenster, Germany.
His research and publications can be accessed at http://www.janrecker.com. The best way to contact Jan is via email (j.recker@qut.edu.au). You can subscribe to his tweets at www.twitter.com/janrecker.

Marcello La Rosa is an Associate Professor and the Academic Director for corporate programs and partnerships with the Information Systems School at the Queensland University of Technology, Brisbane, Australia. He is also a Researcher at NICTA and an IS Research Fellow at the University of Liechtenstein. He has over 60 publications on BPM topics in scientific journals, conference proceedings and books. His research interests include the management of large process model collections, process configuration and automation. Marcello has taught BPM to students and practitioners in Australia for over five years. He holds a PhD in Computer Science from Queensland University of Technology and a MS in Computer Engineering from Politecnico di Torino, Italy.
His research and publications can be accessed at http://www.marcellolarosa.com. The best way to contact Marcello is via email (m.larosa@qut.edu.au).

References

1. Recker, J., *Of Ivory Towers and Boundary Spanners.* BPTrends, 2011. **8**(5): p. 1-7.
2. Recker, J., *Green, Greener, BPM?* BPTrends, 2011. **8**(7): p. 1-8.
3. Recker, J., *X-aware Business Process Management.* BPTrends, 2011. **8**(12): p. 1-7.
4. Recker, J., *How Was School Today?* BPTrends, 2012. **9**(3): p. 1-6.
5. Reijers, H.A., R.S. Mans, and R.A. van der Toorn, *Improved Model Management with Aggregated Business Process Models.* Data & Knowledge Engineering, 2009. **68**(2): p. 221-243.
6. Fischer, R., S. Aier, and R. Winter, *A Federated Approach to Enterprise Architecture Model Maintenance*, in *2nd International Workshop on Enterprise Modeling and Information Systems Architectures*, M. Reichert, S. Strecker, and K. Turowski, Editors. 2007, Gesellschaft fuer Informatik: St. Goar, Germany. p. 9-22.
7. Raduescu, C., et al. *A Framework of Issues in Large Process Modeling Projects*. in *14th European Conference on Information Systems*. 2006. Goeteborg, Sweden: Association for Information Systems.
8. La Rosa, M., et al., *APROMORE: An Advanced Process Model Repository.* Expert Systems with Applications, 2011. **38**(6): p. 7029-7040.
9. La Rosa, M. and M. Dumas, *Configurable Process Models: How to Adopt Standard Practices in Your Own Way* BPTrends, 2008. **5**(11): p. 1-10.

BPTrends Linkedin Discussion Group
We created a BPTrends Discussion Group on Linkedin to allow our members, readers and friends to freely exchange ideas on a wide variety of BPM related topics. We encourage you to initiate a new discussion on this publication, or on other BPM related topics of interest to you, or to contribute to existing discussions. Go to Linkedin and join the **BPTrends Discussion Group.**