

A SYSTEMATIC APPROACH FOR THE IDENTIFICATION OF PROCESS REFERENCE MODELS

Alta van der Merwe^{1,2}, Paula Kotzé^{1,2}

¹Meraka Institute, PO Box 395, Pretoria, South Africa, 0001

²University of South Africa, PO Box 392, Pretoria, South Africa, 0003
vdmeraj@gmail.com, paula.kotze@meraka.org.za

ABSTRACT

Process models are used in different application domains to capture knowledge on the process flow. Process reference models (PRM) are used to capture reusable process models, which should simplify the identification of process models and make it economically viable. In the identification of core elements within the process reference model, the focus is often on the end-product and not on the procedure used to identify the elements. As often proved in development of projects, there is a danger to emphasize the end-product without following a solid identification procedure. In this paper, the focus is on the identification of process reference models, where the authors suggest a method to follow for identification of process reference models, apply the method to a case study and reflect on the experiences during requirements elicitation for identification of process reference models at different institutions.

KEY WORDS

Identification of process reference models; requirements elicitation; process modeling and analysis.

1. Introduction

Understanding the requirements of a problem is one of the most difficult tasks that a software engineer faces. Requirements engineering is the field that assists software engineers in understanding the problem that they aim to solve [1]. Nuseibeh and Easterbrook [2] identified, as part of a paper on requirements engineering, the establishment of reusable requirements models as an important activity that requires future research efforts. The goal of these reusable requirements models will be to act as reference models for specifying requirements in such a way that the effort of developing requirements models from scratch is reduced.

During requirements elicitation the software engineer is involved in the interpretation, analysis, modelling and validation of information. For modelling, one of the tools used is the process model, which captures process knowledge used in a specific application domain. Curtis et. al [3] define a process model as an abstract description of an actual or proposed process that represents selected

process elements that are considered important to the purpose of the model and can be enacted by a human or machine.

In different application domains there are interest in the capturing and reusing of best practices for certain domains in the form of generic conceptual models or reference models. These models are often labelled with the term 'best practice', which are reference models that capture reusable state-of-the-art practices [4]. For process model structures, the term process reference models are used to refer to reusable process models, which are generic in nature and therefore reusable for future analysis and design efforts [5].

There are a number of research initiatives into process reference models. One example is the Next Generation Process reference Models project [6], where the goal is defined as 'to develop techniques for increasing the productivity of business process analysts by allowing them to reuse as much as possible existing models rather than systematically designing new ones from scratch'. However, with regard to the identification of process reference models, the literature does not provide the researcher with clear methods and techniques to identify and classify the models as generic and therefore reusable. The focus of our research was on the approach used in the identification of process models that are reusable and therefore can be classified as process reference models. The intended audience for our approach is developers involved in the identification of process reference model structures in a specific domain.

In this paper we provide some information on process reference models in section 2, followed by the suggested approach for identification of process reference models in section 3. The Higher Educational Institution (HEI) application domain is used as case study environment and in section 4 we reflect on the use of the suggested approach in this domain. In section 5 we reflect on the data-gathering from an institutional, human resource and personal experience perspective.

2. Process Reference Models

Process reference models is a relatively new concept, with first publications on the idea to improve the development

of enterprise-specific models appearing in the late 90s [7]. A process reference model is a process model that is available for re-use [5]. According to Küster, Koehler and Rydina [8] process reference models:

- significantly speed up the design by providing reusable and high quality content
- lead to better and optimized process designs; and
- bridge the business and IT domain.

There are several accepted process reference models used in different application domains. SCOR (Supply Chain Operations Reference Model) is a well-known set of process reference models used for the supply-chain industry. IndustryPrint [9] provides business process models that reflect best practices for key business processes. With regard to software process reference models, SAP AG and IDS Scheer developed the SAP reference models [10]. Lastly, the MIT process handbook is an online knowledge base available with different business process references [11].

Jacobs [12] argues that there are three types of reference models, those used for enterprise architectures, those used as framework models for standardization and those for applying methods. Our focus is on models used to view the behaviour of processes for workflow application development, similar to those published in articles by Weske, Goesmann, Holten and Striemer [13] and Wu, Deng and Li [14].

As mentioned previously, our concern was the identification of process model structures. We therefore suggested an approach where the most important factor is that the process models should be confirmed to exist within different environments, in order to claim that they are generic within different environments and can therefore be stored for future reuse in that specific environment. In designing an approach of this nature, the authors were involved in the discipline of method engineering. From a methodological perspective, method engineering refers to the process of designing, constructing and merging methods and techniques to support information systems development [15]. The development of the approach can be mapped to a Type I method in method engineering. The goal of a Type I method is to bring order to an environment using data modeling or process modeling [29].

The approach suggested is described in more detail in section 3.

3. An Approach for the Identification of Process Reference Models

The suggested approach consists of five phases (Figure 1) that the developer may follow during the identification of process reference models, including:

1. the *definition* of scope.

2. the *identification* of a procedure to derive the process model structure,
3. the *data-gathering* at different institutions,
4. the *comparison* of the results, and
5. a *verification* of the results obtained in (4).

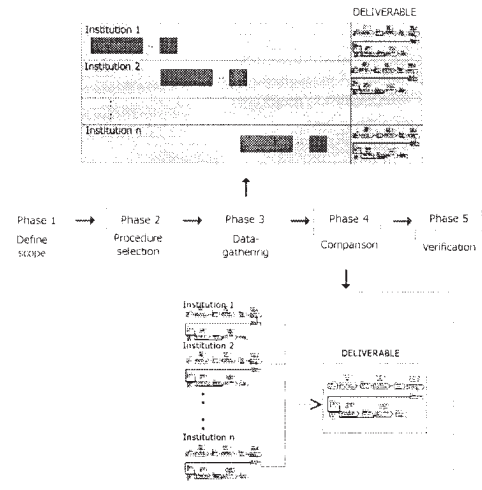


Figure 1. Identification of process reference model structures

For each of the phases there are different issues that the development team should consider. During Phase 1, *Define scope*, the goal is to define role players and identify key persons responsible for the development team. A feasibility study should be done with regard to time, human resources and financial implications. Furthermore during this phase the institutions or enterprise that you will use in data gathering should be identified. The most important task is to get approval from management for doing the project.

During Phase 2, *Procedure selection*, the goal is to select a procedure with which the team could do requirements elicitation. The team should select a procedure that supports a model with a diagrammatic presentation to represent the processes and the flow between them. All the tools used should focus on process modeling and the method focuses on identification of existing processes with the goal to create process reference models for organizational activities such as process re-engineering. It is important that the procedure should support the definition of deliverables after each phase and is not too costly. The procedure selection should consider the use of existing procedures [16] or should consider the possibility of deriving a new procedure.

If the team consider the development of a new procedure, they should ensure that the procedure include steps that are cyclic to revisit different units in identification of processes within the unit, use a mathematical notation for the definition and categorization of the sets of processes, include steps that eliminate any process duplication, provide mechanisms for definition of deliverables at the end of each phase,

supports a standard process model notation, include an identification process of quality control mechanisms, and use accepted guidelines to construct the process model.

During Phase 3, *Data-gathering*, the goal is to identify the different processes in the organizations. An incremental model [29] is suggested where each institution is visited and the procedure selected in Phase 2 is used to derive the process models for the specific institution. The team involved in the requirements elicitation should consider the data collection techniques that are applicable for the specific institution, the identification of key persons in the different units, the inclusion of phase checking to validate the deliverable for each phase, the use of CASE-tools to support the data-gathering technique if available for the selected procedure. During data-gathering it is essential to document all interviews and documents used during the activity and the analyst should be careful of using pre-knowledge after the first increment of data-gathering at the second and third institution. If financially feasible, the use of an alternative analyst is suggested.

For Phase 4, *Comparison*, the goal is to use the different process model structures derived in Phase 3 and to identify a set of unique process models that is referred to as the process reference models. The guidelines to use in this step include that the at most a process model consists of 10 to 15 processes to complete one single function. A comparison table should be sufficient to identify the core or generic processes on a level for a specific scope. The team should remember that a high level of skills may be required. Some interpretation may be needed where similar processes are called different names in different environments. On higher levels the 'set' of processes that are generic on one level should be larger while on the lower levels, where processes become atomic, one ends up with single generic sub-processes. This is natural because the identification of sub-processes for lower levels consists of decomposition of processes on higher levels.

The last phase, Phase 5 *Verification*, consists of verification at an institution not included in the data-gathering activities. Remember that this act as a triangulation exercise where results obtained from previous institutions is confirmed. The analyst is not involved in data-gathering within the organization but rather discusses the results with key persons at the institution. The profile of the selected institution should match the profile selected for generic process model structure identification.

4. Case Study Description

The application domain that we used as a case study environment in order to derive process reference models was the higher education institution (HEI) application domain. In section 4.1 to 4.5 the reflection on the use of the different phases in the approach are discussed.

4.1 Phase 1: Define Scope

During the first phase, the institutions for data gathering and verification were identified. The three institutions included a distance education university (DEU), a residential university (RU1) and a residential technikon (RT). A fourth residential university (RU2) was selected at which to carry out data verification. For all four institutions management on the highest level were approached for approval. There were no financial implications for the project since the project was part of a PhD study and therefore sponsored by the authors' institution. The focus was on the processes involved in supporting the student, from enrolment to course presentation. The goal was to identify the high-level process reference model and also to do a more in-depth analysis of one of the high-level processes, in order to comment on the generic nature of these sub-processes.

4.2 Phase 2: Procedure Selection

For the second phase the goal is to select a procedure with which the team could do requirements elicitation. For this study, a requirements elicitation procedure called the REPPMS (Requirements Elicitation Procedure for Process Model (PM) Structures) with the goal to identify a set of unique process reference models was developed and used. [16-18]. The procedure consisted of five different phases (Figure 2), where each phase may be visited more than once.

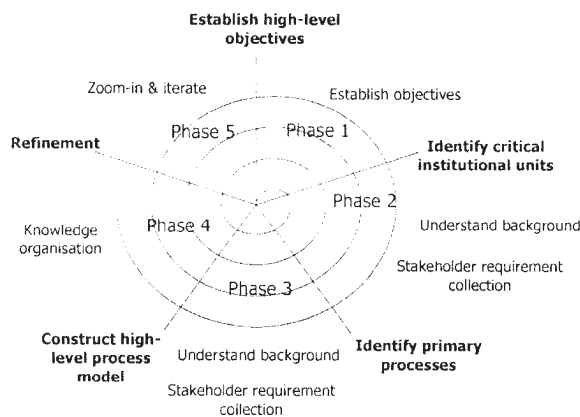


Figure 2. Phases in the requirements elicitation procedure

4.3 Phase 3: Data-gathering

The goal of the data-gathering phase is to identify the different processes in the organizations. We used the REPPMS procedure mentioned in section 4.2 for data-gathering at the three different institutions. All the units (departments, institutes, bureaux, etc.) were included in the initial data sample for each institution. The main data collection technique used was the interviews. A key person was identified in each unit for discussing the processes within the unit and the way that they interact

with one another. Data collection also included non-participant observation and participant observation.

An observation during the application of the procedure at the DEU was that it is impossible to separate the processes and the persons involved in them. People and processes form a unit and both should be included in any analysis effort, even if one knows what the processes are. The deliverable of the requirements elicitation at the DEU was the high-level process model and the set of sub-processes for all the processes in the high-level process model.

At the RUI we were confronted with respondents who were very sceptical about the work and unsure of the reasons for being interviewed. It was necessary to make use of innovative approaches to set the respondents at ease and to assure them that there was no threat in the questions directed at them. The strategy used was to return to the theory and read what has been written on interview techniques.

At the last institution, the residential technikon (RT) there were, similar to the RUI, some negative respondents who were unsure about the purpose of the research and therefore questioned the reason for the interview. During this research iteration we were much more comfortable with the application of the procedure and owing familiarity with the activities at educational institutions, preferred not to become involved in lengthy discussions. We were also more skilled at putting the respondents at ease and began each conversation with the knowledge that we needed to make the person comfortable and to ensure that this was not a judgement of work done, but rather a data-gathering procedure for the purpose of reflecting on the activities in the unit.

4.4 Phase 4: Comparison

For Phase 4, the comparison phase, the goal is to use the different process model structures derived in Phase 3 and to identify a set of unique process models that is referred to as the process reference models. For this phase comparison tables were used to compare the results from the different HEIs and to report on the generic nature of the processes on different levels of the HEI. In more complex process model structures, a more formal measurement method than comparison tables may be necessary, but in this study, where the structures included not more than 10 to 15 processes per level, the use of comparison tables was sufficient.

After completion of the comparison between the process models retrieved from the different institutions, the high-level process reference model representation consisted of 8 high-level processes. The eight processes and the flow between the processes are illustrated in Figure 3.

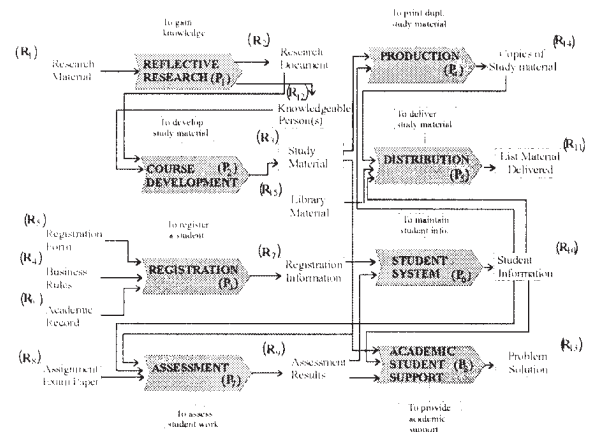


Figure 3. High-level process reference model

In the representation of the HEI structure the flow between processes is supported by thirteen resources. The input for the first process, REFLECTIVE RESEARCH, is the research material used to conduct the research. This includes prescribed books, journals, publications, web-resources, etc. The output for this procedure is a staff member who can be seen as knowledgeable on the research topic and / or a written report on the findings of the research activity. Both these can be input resources for COURSE DEVELOPMENT where the output is a piece(s) of study material, including tutorial letters, study guides, examination papers, video, audio etc. These source documents needed for duplication are sent to the PRODUCTION process where the printing is started based on the number of students in the course (retrieved from STUDENT SYSTEM). The DISTRIBUTION process sends course material to students based on student information retrieved from STUDENT SYSTEM. Material could also be distributed from other resources, e.g. from the library (books). REGISTRATION is done using an application form received from the student, his academic record and the rules of the institution for registration. The data is captured and stored on the STUDENT SYSTEM. ASSESSMENT is done based on the assignment / examination paper received by the students (once again the student information is retrieved from STUDENT SYSTEM). For ACADEMIC STUDENT SUPPORT the lecturer needs the student information (if it is relevant to marks obtained), the course material (if it is course related) or / and the assessment results to assist successfully in answering queries.

Most of the processes on the highest level were duplicated at all three institutions. Although, both REI and RT argued that the DISTRIBUTION and PRODUCTION processes are not primary processes at their institutions. This may be the case if you assume that the lecturer is responsible for production and distribution of course material. But this is not always the case. For web material there may be a developer responsible for the publishing of the material on the web. This is once more a responsibility shift. Also, the distribution of audio

material may not necessarily be the responsibility of the lecturer. After discussions we agreed that the processes should be represented on the high-level process model.

The REGISTRATION process was used in Phase 5 of the requirements elicitation procedure as an example of refinement of a high-level process. From this refinement it was possible to deduce that there are process reference modes on lower levels of the high-level process reference model. For example, the counter registration scenario, for a new and existing student, in the REGISTRATION process, has four generic sub-processes that were identified. These sub-processes include Application Process, Payment Verification, Academic Verification and Course Material Distribution. Note that the sequence of execution on second level and the sub-processes on third level differed in the sub-process model structure. The indication that there are generic process models on lower levels contributes to the knowledge of the elements of the process reference model on different levels.

4.5 Phase 5: Verification

Phase 5, the last phase consists of verification at an institution not included in the data-gathering activities. In the case study we used a residential institution (RU2) that was not used in Phase 3, during data gathering for this phase.

Leading questions on the results obtained at the three institutions were used as a guideline in discussions at the RU2. The feedback of the group was recorded using field notes. The respondents at RU2 agreed that the structures do represent their own process flow and that the high-level PRM could be valuable in their own process re-engineering efforts and requested copies of the structures.

In summary, after the data gathering at three institutions and verification at RU2, it was recorded that on lower levels there is a set of sub-processes that forms the core of the level viewed. For example, a pattern was recorded in the sub-processes listed on the second level for the REGISTRATION process, which confirm the existence of generic processes not only on the highest level, but also on lower levels.

5. Reflection of the Identification of Process Reference Models

In reflecting on the use of the requirements elicitation process to identify process reference model structures, it is possible to comment on issues within the institution concerned, human resources and personal experiences.

5.1 Reflection on Institutional Level

The following experiences were significant at institutional level

- At unit level within institutions there were often good descriptions of processes within the unit, but the

institutional process model structure was poorly described.

- The use of the web as an information tool is expanding in institutions. With the growth in users of web-technology, HEIs realized the worth of web pages in attracting students to their own institution. An intranet also plays a bigger role in 'sharing' information within the institution, making it a valuable resource not only for researchers, but also as a communication tool amongst staff. The web pages were used extensively during requirements elicitation, after we discovered how much information is available regarding the web on the structures of the different units in the institutions.
- A valuable set of documentation was created during the requirements elicitation. The high-level process model was defined and valuable information on the institutional units, resources and flow between processes was documented. Often developers neglect the documentation of ideas and of 'what is going on' and this is identified as one of the reasons why systems fail [1, 19].

5.2 Comments on Human Resource Level

Much is written on data-gathering during interaction with human subjects, e.g. how one should conduct interviews and what to ask and what not to ask [20, 21]. From our experience of conducting interviews during this research, we found the following to be significant:

- At the DEU which was the environment in which the researchers was staff, people were very approachable. We believe this was because they did not feel we was from the 'outside' and in most cases we did not even have to explain why we were involved in the research effort.
- At the DEU people felt threatened the moment they were asked why something was done in a certain way. We had to be very careful not to sound as if we were criticizing the way in which it was done, especially in cases where alternative procedures, such as the use of automatic electronic processes could easily replace the worker.
- At institutional level, being outsiders at the RU and RT meant that people were more aware when answering questions. However, because the respondents were more clinical in answering questions on the working of units, we believe this created a balance since staff at our own institution could sometimes have been rather emotional.

5.3 Reflection on Personal Experiences

The following are personal recommendations on conducting a research effort where data is gathered at different institutions:

- Identify a key person at the institution. This person should be in upper management and support your research. Use this key person to refer you to other key persons in subunits. Use the name of the key person in conversations. People being interviewed are more positive if they realize management supports the research being conducted.
- Remember that people may easily feel threatened during interviews, especially if you ask them for reasons why something is done in a certain way. This substantiates findings on what BPR lacked in the early 1990s [22], causing later procedures to focus more on change management [23].
- Listen carefully to what is said. Sometimes respondents say what they think you want to hear.
- Consider using staff on the development team who are familiar with the environment in which the requirements elicitation procedure will be used.

At RUT and RT, being outsiders gave us a more clinical perspective on the processes. A disadvantage was that we had to be careful not to use our pre-knowledge and not to 'decide' how things were working. Once again, documenting findings and going through a step-wise refinement assisted us in this procedure. This is confirmed by Bergey et al. [24] who emphasized that one of the reasons why re-engineering efforts fail is the lack of a strategy.

6. Conclusion

In this paper we reported on an approach for identification of process reference models. For small environments this is relatively easy, but for a bigger application domain it is a complex task to identify the generic process model structures. The application domain that we used as a case study environment was the higher education institution application domain, where we identified a high-level process model as well as lower level process models for the Registration sub-process. Since the processes repeat within the different institutions, there is strong evidence that the approach was successful for identification of generic processes across different institutions and since the processes are repeating, the processes could be classified as a reference models for HEI institutions. It is therefore possible to derive that the approach did identify process reference models for this specific application domain. In order to generalize the findings, it is necessary to repeat the approach in a different application domain.

References

[1] R.S. Pressman, *Software Engineering, A Practitioner's Approach* (New York, McGraw-Hill International, 2005).
 [2] B.A. Nuseibeh and S.M. Easterbrook. Requirements Engineering: A Roadmap. in *22nd International Conference on Software Engineering* Limerick, Ireland: 2000. 35-46.

[3] B. Curtis, M.I. Kellner, and J. Over, Process Modeling, *Communication of the ACM*, 35(9),1992, 75-90.
 [4] BMP Center. *Configurable process models*. 2005 [cited 2007 November]; Available from: <http://is.tu.tue.nl/staff/wvdaals/BPMcenter/configurable%20process%20models.htm>.
 [5] B. Ramesh and M. Jarke, Towards reference models for requirements traceability, *IEEE Transactions on Software Engineering*, 27(1),2001, 58-93.
 [6] QUT. *Next Generation Reference Process Models*. 2006 [cited 2007 April]; Available from: <http://www.bpm.fit.qut.edu.au/projects/nextgenref/>.
 [7] K. Mertins and P. Bernus, *Reference Models*, in *Handbook on Architectures of Information Systems*, G. Schmidt, Editor. 1998, Springer: Berlin. p. 615-617.
 [8] J.M. Kuster, J. Koehler, and K. Ryndina. Improving Business Process Models with Reference Models in BDD. in *BMP 2006 Workshops*. 2006. 35-44.
 [9] Deloitte Consulting, *IndustryPrint*. 2004.
 [10] SAP AG, *SAP Solution Manager*. 2004.
 [11] T.W. Malone, K. Crowston, and G.A. Herman, *Organizing Business Knowledge* MIT, 2003).
 [12] D. Jacobs, *Towards a Business Process Model Warehouse Framework*, in *School of Computing*. 2008, University of South Africa: Pretoria. p. 188.
 [13] M. Weske, et al. A reference model for workflow application development processes. in *The International Joint Conference on Work Activities coordination and Collaboration*. 1999. 1-10.
 [14] Z. Wu, S. Deng, and Y. Li. Introducing EAI and service components into process management. in *The 2004 IEEE International Conference on Service Computing (SCC '04)*. 2004. 271 - 276.
 [15] D. Avison and G. Fitzgerald, *Information Systems Development: Methodologies, Techniques and Tools* (Berkshire, UK, McGraw-Hill, 2003).
 [16] A. Van der Merwe, *Towards a Reusable Process Model Structure for Higher Education Institutions*, in *School of Computing*. 2005, University of South Africa: Pretoria. p. 353.
 [17] A. Van der Merwe, J. Cronjé, and P. Kotzé. The Functionality of a Requirements Elicitation Procedure Developed for Process Modelling within the Higher Education Application Domain. in *SACLA 2004*. Durban, South Africa: 2004. 895-917.
 [18] A. Van der Merwe, P. Kotzé, and J. Cronjé. Selecting a Qualitative Research Approach for Information Systems Research in *Southern African Computer Lecturers Association*. Kasane, Botswana: 2005
 [19] J.L. Whitten, L.D. Bentley, and K.C. Dittman, *Systems analysis and design methods* (New York, McGraw-Hill Higher Education, 2000).
 [20] I. Seidman, *Interviewing as Qualitative Research: A Guide for Researchers in Education and the Social Sciences* (New York, Columbia U Teachers' College, 1991).
 [21] P. Leedy, *Practical Research, Planning and Design* (Columbus, Ohio, Merrill, 1993).
 [22] T. Davenport, Reengineering Perceptions, *Information Week*, (518),1995.
 [23] K. Bruno, B. Vrana, and L. Welz. Practical Process Engineering for Higher Education. in *CAUSE98*. Seattle, Washington: 1998
 [24] J. Bergey, et al., Why Reengineering Projects Fail, CMU/SEI-99-TR-010, Carnegie Mellon Software Engineering Institution.