On Integrating Data Mining into Business Processes

Dennis Wegener[1] and Stefan Rüping[1]

Fraunhofer IAIS, Schloss Birlinghoven, 53754 Sankt Augustin, Germany,
{dennis.wegener,stefan.rueping}@iais.fraunhofer.de

Abstract. Integrating data mining into business processes becomes crucial for business today. Modern business process management frameworks provide great support for flexible design, deployment and management of business processes. However, integrating complex data mining services into such frameworks is not trivial due to unclear definitions of user roles and missing flexible data mining services as well as missing standards and methods for the deployment of data mining solutions. This work contributes an integrated view on the definition of user roles for business, IT and data mining and discusses the integration of data mining in business processes and its evaluation in the context of BPR.

Key words: Data Mining, Business Processes, Integration, BPM, CRISP

1 Introduction

The capability of data analysis is a crucial factor of success for business today. Main parts of the business world are based on IT and deal with huge amounts of electronic data. Conventional approaches of modelling data mining in business contexts are limited in addressing a major current trend: Data Mining becomes more and more an integral part of executing a business. Tasks like placing advertisements, recommending products, or detecting fraud have become standard application fields of data mining, and have a serious implication for business profits. A frequent re-engineering of business processes is a consequence of this development. For example, as soon as data mining discovers that one advertisement channel performs much better for a certain group of customers, the delivery mode of advertisements is changed. When fraud detection discovers that a certain mode of payment is associated with a high fraud rate, this mode of payment is no longer allowed. In effect, every change in the deployed process immediately changes the incoming data: users cannot be analyzed for their response to one type of advertisement, if they are targeted with another type of campaign. A fraud detection rule makes one type of fraud disappear (but may in turn cause fraudster to shift to other types of fraud). For a process modelling of data mining this means that the steps of defining the data mining goal and understanding the data on the one hand, and deployment on the other hand, are much more connected and dependent on each other than it appears in traditional approaches such as CRISP [1].

From an organizational perspective, the growing integration of data mining also calls for a much stronger role of a data analyst, who needs not only to transport business requirements to the data mining, but also to bring data mining input into a business reengineering process. In contrast to traditional role definitions, which do not address the process of data mining specifically, we argue that the definition and execution of any data mining process which brings significant innovation to the business needs to have a balanced view of the business perspective and the data mining perspective. It is necessary to not only answer the question of what should be done by data mining (e.g., detect fraud), but also of what can be done to allow for a better data mining (e.g., record additional information), and of what significant new input data mining can be expected to offer (e.g., from the use of a novel data mining approach).

Current research progress in business process management (BPM) is based on defining and managing business processes in BPEL- and BPMN-based [2, 3] service-oriented environments. Such modern BPM environments provide flexible design, management and deployment of business processes. Given that these powerful BPM environments and the CRISP model exist, one could assume that it is very straightforward to efficiently integrate data mining in business processes. But is this really the case? In this paper we make the point that in practice still many redundancies and inefficiencies exist. We believe that the integration of data mining into business processes in such environments can be simplified by developing an integrated approach. We aim at a concept for the integration that facilitates the modelling of the data mining process within the business process as well as the technical deployment into the business IT environment. This includes an integrated view on the interfaces between data mining and the business, as well as additional specifications, standards and definitions.

This work contributes an integrated role model for business, IT and data mining roles, an analysis on how the integration of data mining matches with business process reengineering best practices, and an approach for the evaluation of the integration of data mining processes into business processes.

The remainder of the paper is as follows: Sec. 2 introduces the field of BPM and gives details on BPR and role definitions that are related to the problem of integrating data mining in business processes. Sec. 3 gives an overview on data mining, including data mining problems, methods and the CRISP-DM process. In Sec. 4 we go into details on integrating data mining in business processes. Sec. 5 discusses related work. Sec. 6 concludes and gives an outlook on future work.

2 Business Process Management

2.1 Business Processes and User Groups

A business process is a series of steps designed to produce a product or a service which includes all the activities that deliver particular results for a given customer [4]. Business process management (BPM) is a discipline combining

software capabilities and business expertise to accelerate business process improvement and to facilitate business innovation [5]. Integrating data mining in business processes involves different groups of users with different responsibilities, knowledge and background. Naively speaking, this will include the roles of business experts, who are in charge of the business but do not have knowledge in the implementation and in data mining, data miners, who are experts in data mining and who are in charge of designing and implementing the data mining process, and the IT experts, who manage the hardware and software resources of the business and who are in charge of implementing and integrating the business process in the IT environment.

In literature, a lot of classifications of roles in the context of BPM is available. In [6], a detailed study on employee competencies is given based on the following classification of roles and their relations: The *Project Leader*, the *Process Consultant*, the *Process Coordinator*, the *Process Owner*, the *Process Controller*, and the *Process Staff*. However, this definition lacks in describing the IT roles that are involved in the scenario and thus is not suitable for discussion the integration problems. In [5], four key groups that have to work together in BPM projects are defined, namely the *Workforce*, the *IT*, the *Management*, and the *Business Analysts*, including a lot of sub-groups among these.

2.2 Business Process Reengineering and its Evaluation

Current research on business process management (BPM) outlines the importance of business process reengineering (BPR) and its role in improving business processes [7], e.g., in terms of improving efficiency and effectiveness of an existing business. BPR regards business processes from a "clean slate" perspective to determine how to best construct these processes to improve how the business is conducted. Reengineering is the fundamental rethinking and radical redesign of business processes to achieve dramatic improvements in critical, contemporary measures of performance [8]. In literature, BPR is often evaluated according to a set of dimensions in the effects of redesign measures, e.g. time, cost, quality and flexibility [9], cost, quality, service and speed [8] or cycle time, cost, quality, asset utilization and revenue generated [4]. Ideally, a redesign or modification of a business process decreases the time required to handle incidents, it decreases the required cost of executing the business process, it improves the quality of the service that is delivered and it improves the ability of the business process to react flexible to variation [7]. However, a property of such an evaluation is that trade-off effects become visible, which means that in general, improving upon one dimension may have a weakening effect on another.

In [9], a set of best practice heuristic rules on business process (re)design is evaluated according to the metric cost, time, flexibility and quality. In total, 29 rules are described which are classified according to their orientation towards customers, business process operation, business process behaviour, organization, information, technology and external environment. Later, we will discuss which rules are involved in the integration of data mining in business processes.

3 Data Mining

3.1 Data Mining Problems, Goals and Methods

Data mining, also called knowledge discovery in databases (KDD), is the process of extracting (unknown) patterns from data. In general, a data mining process includes several iterations of single data mining steps (algorithm executions). The goals of the data mining process are defined by the intended use of the system from the user perspective and can be classified into two types: *verification*, where the system is limited to verifying the user's hypothesis, and *discovery*, where the system autonomously finds new patterns. The discovery goal can be further subdivided into *prediction*, where the system finds patterns for predicting the future behavior of some entities, and *description*, where the system finds patterns for presentation to a user in a human-understandable form [10].

A variety of data-mining methods exists which help in achieving the goal of prediction and description. According to [11], data mining methods commonly involve the following classes of tasks: Inferring rudimentary rules, statistical modeling, constructing decision trees, constructing rules, mining association rules, linear models, instance-based learning, and clustering. For each of these methods a variety of data mining algorithms exist that incorporate these methods. We refer to [11] for an updated reference on data mining methods and algorithms.

3.2 The Data Mining Process

The data mining process is an interactive and iterative process that involves numerous steps with many decisions made by the data miner. CRISP-DM [1] is a standard process model for data mining that depicts corresponding phases of a project, their respective tasks, and relationships between these tasks. According to CRISP-DM, the lifecycle of a data mining project consists of the following six different phases: *Business Understanding* - understanding the project objectives and requirements from a business perspective and converting this knowledge into a data mining problem definition; *Data Understanding* - getting to know the data and to identify data quality problems; *Data Preparation* - construct the final dataset from the initial raw data as input for the modelling; *Modeling* - various modeling techniques are selected and applied, including the calibration of their specific settings; *Evaluation* - assess how well the built model achieves the business objectives; *Deployment* - the results of the data mining and the knowledge gained are delivered to the user, reaching from generating a simple report up to a complex implementation of a repeatable data mining process. The process in general is iterative, but also foresees stepping back between certain phases to adjust some of the decisions made. From the data mining perspective, the (business) user is mainly involved in the phases Business Understanding and Deployment, while the other phases are mostly performed only by the data miner. In terms of integration, it has to be distinguished between the (technical) deployment of the data mining solution as a whole, which might be done only once for a given business process, and the deployment of new data mining models, which might be done frequently.

3.3 Evaluating Data Mining

At the end of the modeling step of the data mining process, the data mining model is evaluated in order to determine its quality. Model-evaluation criteria are quantitative statements of how well a particular data mining model meets the goals of the KDD process [10] and the business objectives [12]. The criteria differ according to the data mining goal. For each of the goals verification, prediction and description a number of performance measures exist for the evaluation. E.g., predictive models are often evaluated by the empirical prediction accuracy on some test data and descriptive models can be evaluated along the dimensions of predictive accuracy, novelty, utility, and understandability of the fitted model [10]. Other measures include, e.g., precision, recall and f-measure. In addition, techniques like cross-validation or bootstrap [11] are applied in order to ensure the representativeness of the evaluation. Despite these metrics for evaluating the result of the data mining process, other important factors like the time spent for the process, the resources that were used, etc., can be included in the evaluation. In the evaluation phase of the data mining process, the model as well as the way it was constructed is evaluated according to the business objectives.

4 Integration of BPM and Data Mining

Data mining is often integrated with a business process to provide value [12]. However, data mining needs a lot of domain knowledge and thus is difficult to handle for non-experts. On the other hand, continuous data mining and flexible adaption of the data mining and the business process plays a crucial role in the context of modern data mining solutions as part of business applications, e.g. online advertisement or fraud detection. From the business point of view, the data mining results have to be frequently evaluated according to relevance for the business and respective actions have to be taken, e.g. adapting the business process due to new emerging fraud patterns or adapting the data mining to provide more detailed results. This means, that the CRISP-DM phases deployment and business understanding are much more connected than before. Thus, there is a clear need for approaches for the integration of data mining into business processes which need to be compatible with modern business process management environments. In particular, they need to cover flexibility w.r.t. adapting the data mining process as well as the business process itself in an online environment, which they do not provide so far. While CRISP-DM includes longer process cycles, modern data mining needs for fast development and quick experiment setup and evaluation (also on live systems). Motivating examples for integrating a data mining process into a business process in order to improve the business are, e.g., the EC-funded projects RACWeB[1] and iWebCare [13]. RACWeB aimed at improving the Western Balkans and EU countries' customs efficiency and transparency in risk assessment by enhancing the identification

---

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

of risk profiles through the utilisation of data mining techniques. In iWebCare, a flexible fraud detection platform was developed in order to ensure quality and accuracy and minimise loss of health care funds in the Healthcare business. The basic idea behind both projects was to develop a web service platform where participating organizations can upload datasets and semi-automatically select an appropriate data mining process. Based on the analysis of the business processes of participating organizations, it was possible to develop generic data mining solutions that can be re-used in similar business processes. Because of the detailed business process modelling it is not necessary to follow the CRISP-DM process step by step again a second time, as much of the relevant business knowledge is already contained in the business process model. In other words, redundancies exist between the business perspective and the data mining perspective.
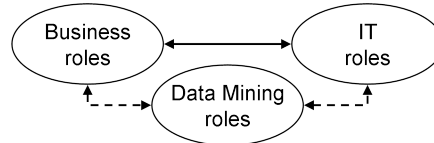
4.1 An Integrated Role Model

Fig. 1. Interaction of the different groups of roles - existing interaction between business and IT roles and new interactions with data mining roles

Data mining is typically performed by a specialist with lots of expertise in data mining (the data miner) who is responsible for the steps of the data mining project as, e.g., described by CRISP-DM. The client (from the data miner's perspective) is involved in the first phase (Business Understanding) and the last phase (Deployment); all phases in-between are solely carried out by the data miner. The data miner is not yet described by one of the roles in BPM. He is not responsible for the full process, but as data mining consists of several steps, also not just for a single task. There are tasks and responsibilities of the data miner which are not yet described in the area of BPM. From the data mining point of view, a lot of domain knowledge is needed to set up a high quality process. Typically, the setup of a data mining process requires a lot of interaction with other roles, e.g. the business user and the IT experts. So, important in the context of integrating data mining in business processes are the groups of business roles and of IT roles. In addition to the existing interaction between these roles, new interactions with the group of data mining roles appear (see Fig. 1).

Persons who are responsible for the data mining can also be classified into different roles, e.g. as follows: *Data Mining Senior* - responsible for the conceptual design; *Data Mining Experimenter* - responsible for the experiments; and *Data Mining Architect* - responsible for the integration, deployment and integration. The typical approach is to let the data mining senior define the problem

and the general procedure, to let the experimenter conduct experiments with a data mining toolkit, and to let the data mining architect implement, deploy and integrate the solution into a bigger context.

4.2 Integration of Data Mining Processes into Business Processes

For the integration, we assume that there is an existing business process. The goal is to integrate a data mining process into this business process in order to improve the business. A data mining process cannot be seen as normal business process. The necessary domain knowledge for the setup of data mining and the resources needed for the implementation justify the special role for the data miner.

CRISP-DM [1] describes a standard process model for data mining processes, but lacks in defining detailed concepts, definitions and standards for a deployment of the data mining solution as well as the integration of the whole data mining process. Often, only the integration of the data mining results is discussed, but not the integration of the data mining process as a whole. Business processes and data mining processes interface at the CRISP-DM phases Business Understanding and Deployment. The other CRISP-DM phases are mainly data mining specific with small or no involvement of the user. As there is already a business process defined, many of the decisions that need to be taken in the CRISP-DM phases Business Understanding and Deployment are already fixed in advance. Such decisions include, e.g., which task to replace or to add to the original business process, which data mining problem to solve, the definition of the data mining constraints in terms of quality, time, etc., and how to integrate things technically in the IT environment. On the other hand, results of data mining might suggest taking decisions in a different way. Hence, there exist redundancies when not carefully coordinating CRISP-DM and the revision of the business process.

Workflows involving data mining introduce additional dependencies among tasks as well as a combination of automated and manual tasks [12]. Thus, data mining cannot be described by a single task or box inside a business process. As described above, data mining itself can be seen a complex process. In addition, the data mining services (modelling part) can be quite complex or can have totally different functionality. There exist solutions for distributed and grid based data mining, including highly parameterized generic services as GridR [14], where the algorithm itself could be passed as a parameter.

Integrating data mining in business processes means to change the business processes. Thus, the integration has to follow the rules, standards and practices for BPR. In the following, we give details on which rules from [9] apply or explicitly apply not for the integration:

– Task elimination: 'eliminate unnecessary tasks from a business process' - a typical goal of integrating data mining is to replace (manual) tasks. However, new tasks for the data mining itself appear.

– Order-based work: 'consider removing batch-processing and periodic activities from a business process' - the model building step will not be executed once per order but rather as batch job (a periodic activity that is started according to some conditions).
– Parallelism: 'consider whether tasks may be executed in parallel' - if data mining replaces a manual task, it can, once it is deployed in a business process, be executed automatically. Thus, this might give an option for parallelization of preceding or subsequent tasks.
– Split responsibilities: 'avoid assignment of task responsibilities to people from different functional units' - the data miner will be involved in the data mining tasks, so responsibilities are distributed to a larger number of persons. However, we aim at a minimal involvement.
– Numerical involvement: 'minimize the number of departments, groups and persons involved in a business process' - see split responsibilities.
– Extra resources: 'if capacity is not sufficient, consider increasing the number of resources' - by the integration of data mining, resources in terms of people (the data miner) and hardware (computing and storage) are extended. If data mining replaces a manual task, previously used human resources become free.
– Control addition: 'check the completeness and correctness of incoming materials and check the output before it is send to customers' - by data mining, completeness and correctness of data can be checked more extensively.
– Task automation: 'consider automating tasks' - with the help of data mining, some manual tasks could be replaced.
– Integral technology: 'try to elevate physical constraints in a business process by applying new technology' - data mining represents a new technology that can help improving the business process.
– Outsourcing: 'consider outsourcing a business process in whole or parts of it' - parts of the business process are outsourced in the way that a new group, the data miner, is responsible.

The details on the integration of the data mining into business processes also depend on the data mining problem to solve. E.g., we distinguish between supervised and unsupervised data mining, as different characteristics raise different requirements in terms of modeling. In the context of supervised learning, labeled datasets are needed for training a model (which, e.g., specify if a dataset is fraud or not), whereas in the context of unsupervised learning no labels are needed. Thus, a different set of components is needed for the integration.

Instead of deciding on the kind of integration each time anew and of integrating these steps manually, we propose to define a standard set of data mining services that can interface with modern BPM environments. For the adaption of the business process models, templates could be developed that define on how exactly the data mining can be integrated. Our approach is generic, but works especially well in certain scenarios. These include scenarios with an objective, automatically computable measure of quality, and with standardised structured data with limited volume. This allows for automating many decisions in the data mining process and working with existing standardised solutions.

4.3 Evaluating the Integration

In order to determine if the integration of data mining into a business process makes sense and is valuable, we have to prove that the new integrated solution is better. In addition, as data mining has special requirements for IT [12], there is the need for a metric that helps to decide on if and how to adapt the IT environment. As integrating data mining in business processes means to apply the rules and standards of BPR, we propose to evaluate the integration including the metrics for BPR, taking into account the evaluation of data mining related BPR best practices as described above. The evaluation has to be two-fold. Both, the new operational business processes as well as the effort for the integration itself have to be evaluated.

For the former, there is a need to find a metric that covers both, available performance measures of business process redesign and data mining processes. In principle, the integration of a data mining process can be evaluated in the same way as standard BPR according to the key performance indicators time, cost, quality and flexibility that have been described earlier. The measures can be mapped as follows: The (IT-) resources that are needed by the data mining map to the cost measure; The runtime of the data mining process maps to the time measure; The model-evaluation criteria that describe the quality of the data mining result map to the quality measure; The level of generality of the integrated data mining solution maps to the flexibility measure. In addition, this can incorporate, e.g., how many data mining processes (e.g. classifications) can be performed simultaneously. It has to be noted that in the context of data mining not all of these measures are always fix. E.g., the time needed for learning a model is hard to predict. In addition, there is a non-trivial correlation between resources denoted to the data mining process and the quality of the result.

For the evaluation of the integration the time needed for the redesign and integration can be measured. E.g., one solution can be realized very fast, but the quality of the data mining result is not best. There is a solution with a better result quality (e.g. based on distributed computing), but the realization takes a lot longer. So, there is a trade-off that has to be taken into account. Thus, the speed-up of integration and development is also an important measure that has to be considered.

In modern frameworks incorporating BPEL and BPMN, a simple design, setup and deployment of business processes is possible. But, it is not clear how to model an integrated process in detail, meaning how much detail of the data mining process has to be visible in order to integrate it conceptually and technically. In the end, all technical steps have to be integrated. But, for instance it has to be answered which parts of the data mining process have to be integrated into the business process at minimum and which can be aggregated or hidden from the user. Thus, the complexity of the modeling of the business process can serve as metric. This could include, e.g., the number of additional components and tasks that have to be added to the original business process, the number of services that have to be implemented or used by the data miner, etc.

5 Related Work

In [10] the integration of data mining with other systems, e.g. database management systems or spreadsheet and visualization tools, is described as a research and application challenge. As applications today are getting more and more complex due to larger amounts of data or the distribution of data or resources, the problem of integration is also getting harder. Most important benefits of data mining for the business are achieved when the data mining results are deployed into a business process in a repeatable manner, which, e.g., involves the ability to rebuild, assess and apply models automatically [12]. [15] gives a survey on data mining techniques and business applications in the context of mining data from the business. There even exists some literature on integrating data mining in business processes from the end of the 90s [16]. They describe their experience on integration of data mining and observe open issues, but do not incorporate modern concepts and techniques for BPM yet.

According to CRISP-DM, "in many cases it will be the customer, not the data analyst, who will carry out the deployment steps". "Even if the analyst will not carry out the deployment effort it is important for the customer to understand up front what actions will need to be carried out in order to actually make use of the created models" [1]. Thus, the CRISP-DM model lacks in the deployment phase [17] and misses phases important for engineering projects [18]. There is no specification or support for standards on how to deploy the data mining results into the business. In addition, problems on integrating the data mining process into a business process in praxis are not addressed.

In [17] a methodology for the implementation of data mining into operational business processes is proposed, consisting of the phases Exploratory Data Mining, Deployment of IT into the Business Process, and Operational Data Mining. The first phase consists of performing a CRISP-DM process and evaluating the readiness for an integration of the data mining solution. In the second phase, the data mining solution is deployed into the business environment, including changes in design and implementation of the business process (which means to perform BPR) and associated applications. The third phase describes the steps of the operational solution, including all CRISP-DM phases but business understanding and an additional phase for model updating. They do not specify a general concept for the technical integration and the re-design of the process and its evaluation. In their example of integrating data mining into a direct marketing business process, the data mining is nearly modeled as black box.

Based on experience in software engineering, [18] proposes a model for data mining engineering that includes engineering-related phases which are missing in CRISP-DM. They identify as open issue, that available process models specify what to do, but not how to do it. This is a point that we start addressing with this work. [19] also identifies the lack of guidance towards implementing particular tasks of data mining methodologies and introduces a framework for the implementation of the business understanding phase of data mining projects. However, they do not focus on business integration and deployment of data mining solutions into business processes.

Scheduling and workflow environments are slowly being addressed by business process environments [12], e.g., with frameworks like BPEL, that allow and facilitate the inclusion of (external) web-services within business processes. JDM [12] provides web service interfaces for such a kind of integration on API level that can be used in modern BPEL based frameworks. In addition, a use case is presented on how to integrate data mining based on JDM into a business process based on BPEL in a Bank scenario. The authors use a role model with the groups Business Analysts, Data Analysts and IT. However, they do not specify in general how to adapt a business processes in detail, how to evaluate the integration and which requirements for the modelling of the business process when including data mining exist.

6 Conclusion

Both areas of BPM and data mining are very complex and require a lot of human resources. In response, many approaches for BPM and data mining have been proposed in the literature. In many practical applications, e.g. integrating fraud detection in an e-commerce system, these approaches work out fine. However, in this paper we showed that following the CRISP-DM model in parallel to standard BPR approaches results in many redundancies. In the context of integrating data mining into business processes we have to deal with two different processes that work together at two defined steps - the Business Understanding and the Deployment phase according to CRISP. Business processes can be quite different, but the adaption and redesign for the integration of data mining is often similar. Hence, there is a clear need for a coordinated approach. From the data mining point of view, the deployment in terms of business process design is similar (e.g., exchange one BP-step by another), but with respect to deployment on the IT infrastructure it can differ a lot. Thus, additional DM-process steps (w.r.t. deployment) and a standard set of services are needed. This motivates the development of flexible services such as those from the context of GridR.

This work contributed an integrated role model for business, IT and data mining roles, an analysis on how the integration of data mining matches with business process reengineering best practices, and an approach for the evaluation of the integration of data mining processes into business processes. It represents a further step towards an integrated environment which allows for flexible adding, exchanging or adapting data mining and business process tasks and fast experimenting, even on live systems, which is needed by modern data mining solutions in business applications.

As future work we foresee improvement on the concepts and techniques for an easy deployment of flexible data mining solutions into business processes in the context of modern BPM frameworks based on BPEL and BPMN, based on a uniform concept for both modelling and technical integration. This will include work on data mining services that support an easy integration as well as process templates for the business processes.

References

1. Shearer, C.: The CRISP-DM model: the new blueprint for data mining. Journal of Data Warehousing, Vol. 5 , Nr. 4, pp. 13–22 (2000)
2. Jordan, D., Evdemon, J.: Web Services Business Process Execution Language Version 2.0. Technical report, OASIS Standard (2007)
3. White, S. A., Miers, D.: BPMN Modeling and Reference Guide Understanding and Using BPMN. Future Strategies Inc., Lighthouse Pt, FL (2008)
4. Mayer, R.J., Dewitte, P.S.: Delivering Results: Evolving BPR from art to engineering. In: Elzinga, D.J., Gulledge, T.R., Lee, C.Y. (eds.): Business process engineering: advancing the state of the art (1998)
5. Peisl, R.: The Process Architect: The Smart Role in Business Process Management. IBM RedPaper (2009)
6. Eicker, S., Kochbeck, J., Schuler, P. M.: Employee Competencies for Business Process Management. In: Abramowicz, W., Fensel, D. (eds.): Proc. of 11th International Conference on Business Information Systems. Springer, Berlin, pp. 251–262. (2008)
7. Bessai, K., Claudepierre, B., Saidani, O., Nurcan, S.: Context-aware Business Process Evaluation and Redesign. In: Int. Workshop on Business Process Management, Design and Support, at Int. Conference on Advanced Information Systems, Montpellier, France (2008)
8. Hammer. M, Champy, J.: Reengineering the Corporation: A Manifesto for Business Revolution, Harper Collins, London (1993)
9. Reijers, H.A., Liman Mansar, S.: Best practices in business process redesign: an overview and qualitative evaluation of successful redesign heuristics. Omega - The International Journal of Management Science 33 (4), pp. 283–306 (2005)
10. Fayyad, U., Piatetsky-Shapiro G., Smyth, P.: From Data Mining to Knowledge Discovery in Databases. AI Magazine 17, pp. 37–54 (1996)
11. Witten, I.H., Frank, E.: Data Mining: Practical machine learning tools and techniques. 2nd Edition, Morgan Kaufmann, San Francisco (2005)
12. Hornick, M.F., Marcad, E., Venkayala, S.: Java Data Mining: Strategy, Standard, and Practice. Morgan Kaufmann, San Francisco (2006)
13. Tagaris, A., Konnis, G., Benetou, X., Dimakopoulos, T., Kassis, K., Athanasiadis, N., Rüping, S., Grosskreutz, H., Koutsouris, D.: Integrated Web Services Platform for the facilitation of fraud detection in health care e-government services. In: Proc. ITAB 2009, Lacarna, Cyprus (2009)
14. Wegener, D., Sengstag, T., Sfakianakis, S., Rüping, S., Assi, A.: GridR: An R-based tool for scientific data analysis in grid environments. Future Generation Computer Systems 25 (4), pp. 481–488 (2009)
15. Bose, I., Mahapatra, R.K.: Business data mining - a machine learning perspective. Information and Management, Vol. 39, No. 3, pp.211-225 (2001)
16. Holsheimer, M.: Data mining by business users: integrating data mining in business processes. In: J. Han (ed.) Tutorial Notes of the 5th ACM International Conference on Knowledge Discovery and Data Mining. ACM, New York, pp. 266–291 (1999)
17. Rupnik, R., Jaklic, J.: The Deployment of Data Mining into Operational Business Processes. In: Ponce, J., Karahoca, A. (eds.) Data Mining and Knowledge Discovery in Real Life Applications, I-Tech, Vienna, Austria (2009)
18. Marbn, O., Segovia, J., Menasalvas, E., Fernndez-Baizn, C.: Toward data mining engineering: A software engineering approach. Information Systems 34 (1) (2009)
19. Sharma, S., Osei-Bryson, K.: Framework for formal implementation of the business understanding phase of data mining projects. Expert Systems with Applications 36 (2) (2009)