Inf Technol Manag (2011) 12:97–109
DOI 10.1007/s10799-010-0076-z

Supporting process design for e-business via an integrated process repository

Harry Jiannan Wang · Harris Wu

Published online: 9 October 2010

Abstract Business process design is an integral part of e-business engineering. Given that e-business models usually involve a wide range of business processes across different business functions with complex activities, events, and documents, process design for e-business is a very challenging task. Although various process reference models (PRMs) have been developed to provide guidelines for process design, research on leveraging multiple PRMs to support process design for e-business has been scant. In this paper, we demonstrate that the diverse process design requirements in e-business are best satisfied by utilizing multiple PRMs via a case study. Then, we propose a collaborative approach grounded in knowledge management theory to integrating multiple process reference models to better support process design in e-business. We equip the integrated process repository with a set of novel features based on Web 2.0 technologies to enhance its utility, efficiency, and quality for process design support. A prototype system is developed and user experiments are conducted to evaluate the system.

Keywords Business process design · e-Business · Process reference model · Web 2.0 · Case-based design

H. J. Wang (✉)
Department of Accounting and MIS, University of Delaware, Newark, DE, USA
e-mail: hjwang@udel.edu

H. Wu
Department of Information Technology, Old Dominion University, Norfolk, VA, USA
e-mail: hwu@odu.edu

1 Introduction

Today, business process management (BPM) systems have become an integral part of modern enterprises' IT infrastructures as indicated by the fast growing BPM software market estimated to reach $6 billion in 2011 [36]. Specially for e-business, BPM is a critical successful factor because the efficiency and effectiveness of underlying business processes have become the major source of e-business companies' competitive advantages [3]. Process design is the foundation and most critical component of BPM, where new business processes are developed to meet specific business problems and/or existing processes are revised to improve their performance [21]. Current process design practice usually refers to process designs from some industry standards as the starting point instead of completely starting from scratch or only looking at as-is processes [43]. Over the past several decades, researchers and practitioners have packaged their process design knowledge and best business practices in different industries in the form of Process Reference Models (PRMs), such as SAP Process Reference Model [10], Oracle Best Practice Processes [37], MIT Process Handbook [27], RosettaNet Partner Interface Processes (PIPs) library [45], and various business process templates from different BPM software vendors.

These reference models help jumpstart a process design project and provide insights into discovering innovative business processes by offering massive best-practice-based process diagrams and descriptions from various business functions, e.g. sales, marketing, accounting, etc., in different industries, e.g. healthcare, telecommunications, and pharmaceuticals. Although these reference models often complement each other with different contents and scopes, e.g. SAP reference model is developed around ERP applications whereas RosettaNet PIPs are designed to facilitate B2B e-business, research on leveraging multiple reference models to support process design in e-business has been scant. Given the diversity of e-business in terms of the involved business processes, individual PRM can only provide limited process design support. Based on our case study (see Sect. 2 for details) of several well-known process reference models, we found several issues that causing this research gap. For example, it is not easy to find similar process designs across different reference models due to their diverse ways of classifying process designs. In addition, process designs in those reference models are in the form of text documents and process images and only simple keyword-based search is provided to explore the models. Thus, it is very difficult to find most relevant process information to the current process design need from multiple PRMs. These issues greatly hinder the usability of existing PRMs for supporting e-business process design. As such, a more efficient and innovative solution for integrating and managing process knowledge in multiple PRMs is an imperative.

The Internet has enabled mass collaboration for product innovation and knowledge discovery [8, 52]. A new trend on the Web is to utilize the “free brains and labor out there” among Internet users. Harnessing social intelligence, i.e., utilizing individual users’ knowledge and efforts, is one of the core competencies of “Web 2.0” [33]. Such phenomena can be seen everywhere now, such as collaborative spam e-mail filtering, Wikipedia, Netflix’s movie recommendation algorithm, and CNN’s attempt to create user-contributed news [52]. There remains a large potential for Web 2.0 to be utilized by process reference models to support business process design. For example, users can collaboratively evolve a unified process classification scheme by classifying process designs from different PRMs, which helps better integrate and organize process knowledge from different sources. Users can also recommend designs to others’ requirements, which help build connections among process designs from different reference models. Our fundamental belief is that by harnessing the collective power of a large group of people, knowledge from multiple process reference models can be better managed and utilized to support business process design.

In this paper, we propose a Web 2.0 approach to integrating multiple process reference models and support process design. We intend to make several contributions: first, we conduct a case study of major existing process reference models and identify the challenges of utilizing these models in supporting process design for e-business. Second, we develop a theoretical framework of process design as a collaborative knowledge creation process, which serves as the formal guidelines for the development of innovative process design support systems. Third, we propose a Web 2.0 approach to integrating multiple process reference models into a process repository to better support e-business process design. Last, we implement a prototype system to integrate several well-known process reference models and conduct experiments to evaluate the system’s utility, efficiency, and effectiveness in terms of supporting process design.

The rest of the paper proceeds as follows. Next, we first review the relevant literature. Then, we present a case study of several process reference models, leading to several research issues that motivate our research. After that, we propose a conceptual framework for collaborative process design based on knowledge management theory. We present our Web 2.0 approach to integrating and managing process reference models in Sect. 5 and evaluate our approach in both academic and industry settings in Sect. 6. Finally, we summarize the results and present future research plan at the end of the paper.

2 Literature review

Organizations continuously design and re-design their business processes to achieve operational excellence and meet new business requirements, i.e., process innovation [14, 31, 51]. Process designers usually refer to existing process design documents as the starting point instead of starting from scratch or as-is processes [29, 50]. Many process repositories, or so-called reference models, have been built over the past several decades by researchers and practitioners to package their process design knowledge and best practices in different industries [29, 42]. Many important business process management research issues in e-business have been identified, such as e-business workflows, inter-organizational workflow, and e-business process standards [3].

Business process design has been an important task in designing e-business models and e-business engineering [1, 5, 13, 40, 43, 49]. In the BPM lifecycle, process design is the most critical step, where existing processes are continuously revised for performance improvement and new processes are developed to cope with ever-changing business environments [43, 47, 53, 56, 62]. Many graphical representations of process models have also been established and applied in process design, such as UML Activity Diagrams [34], BPMN [35], and EPC (Event-driven Process Chain) [48]. These specifications define the semantics of the graphical process modeling symbols and related rules. In addition, various process specification languages and notations have been proposed to model dynamic and collaborative business processes for enterprise information systems [7, 38, 46]. Existing process design approaches can be classified as either mainly participative or analytical [44].

Participative approaches usually obtain process information using traditional data collection instruments such as interviews, meetings, and workshops [9, 20, 26, 48], whereas analytical approaches aim to apply formal theories and techniques to derive the process models [11, 44]. Process design practices conducted in industry is mostly participative [18, 44]. ARIS (Architecture of Integrated Information Systems) is one of the most well-known business process modeling framework adopted by industry practitioners, which has been incorporated in several commercial systems from companies like SAP and IDS Scheer [48]. In ARIS, process models are constructed by analyzing and grouping relevant business objects from five different views including function, organization, data, output, and control and most process information is gathered through interviews, workshops, document analysis, etc. In addition, although document analysis has been identified as a major means to get process information, no detailed systematic procedures on how to extract process model based on various existing process documents have been reported [48, 57]. In order to automate process design, several analytical process design approaches have been proposed, such as such as linear programming [2], cost optimization [55], computational experiments [15], data dependency analysis [44] and probability theory [11]. Several algorithms have also been developed to extract process models from structured event logs generated by transactional systems such as ERP, CRM, or workflow management systems [19, 54]. Web 2.0 technologies have also been leveraged to help collect and manage business process knowledge for process design innovation [22]. However, research on leveraging Web 2.0 technologies to integrating multiple process reference models for e-business process design has been scant.

Case-based design has been shown effective in process design/redesign and workflow management [25, 28]. Searching and exploring a case repository can benefit from social classification [60] and social tagging [12, 30]. Social tagging systems such as delicious allow users to assign free-formed keywords (tags) to any document on the Web, and share these tags. Any user can browse these tags, or search for items tagged with given keywords. Social tagging requires little effort from individual users and benefits them personally. Once individuals tag the documents for their own benefit, their knowledge is captured in tags and, in turn, benefits others. The integrated process repository illustrates the business value of social intelligence (the tags) in a knowledge creation setting [59].

While individuals can provide process design recommendations to each other, such "free help" will be limited, as process designers are often busy professionals. A recommender system can provide recommendations systematically at low cost [4, 16, 17, 25]. The recommender system can utilize a hybrid set of input including peer recommendations, classifications, tags, user profiles, and textual descriptions of problem situations, usage history, and user feedbacks. Analysis of links (hyperlinks or semantic links) among the documents as well as the user-document links can identify topics among the design documents and the high quality documents in each topic. Process design recommendation raises new challenges such as how to incorporate process structures in the recommendation algorithm.

3 A case study of process reference models

In this section, we present a case study of four process reference models (PRMs), i.e., MIT Process Handbook (PH) [27], SAP Process Reference Model (PRM) [10], Oracle Best Practice Processes (BPP) [37], and RosettaNet Partner Interface Process (PIP) Library [45]. We choose these reference models for several reasons: first, they are all well-known process reference models, which serve as good samples for our research; second, these reference models cover a broad range of business processes for e-business and complement each other in terms of contents; third, they represent reference models from different groups, i.e. academia (MIT PH), system vendor (SAP PRM and Oracle BPP), and standards organization (RosettaNet PIPs). Although there are many other process reference models, such as Supply Chain Operations Reference model (SCOR), PwC best practices (http://globalbestpractices.pwc.com/), and various business process templates from BPM software vendors [6], we believe that the four PRMs we investigate in this paper are representative enough to show the need for an integrated process design support system for e-business as we discuss next.

Many of existing PRMs are merely linked web pages, such as the MIT Process Handbook (Fig. 1), SAP PRM, and RosettaNet PIPs, while others are embedded in the packaged software, such as Oracle's BPP from its Business Process Analysis Suite (Fig. 2) [37]. A few PRMs can be accessed free of charge, while others are provided as commercial products or subscription services. We compare the four process reference models according to a set of features. The goal is to present the similarities and differences among those models in order to illustrate the need to integrate and enhance those models for better process design support for e-business.

- *Process Classification* (*PC*): all four reference models classify processes based on business functions. Function names and quantities are different among these PRMs as listed below.
  - MIT PH (10 functions): Procurement, Supply Chain Management, Marketing, Sales, Information

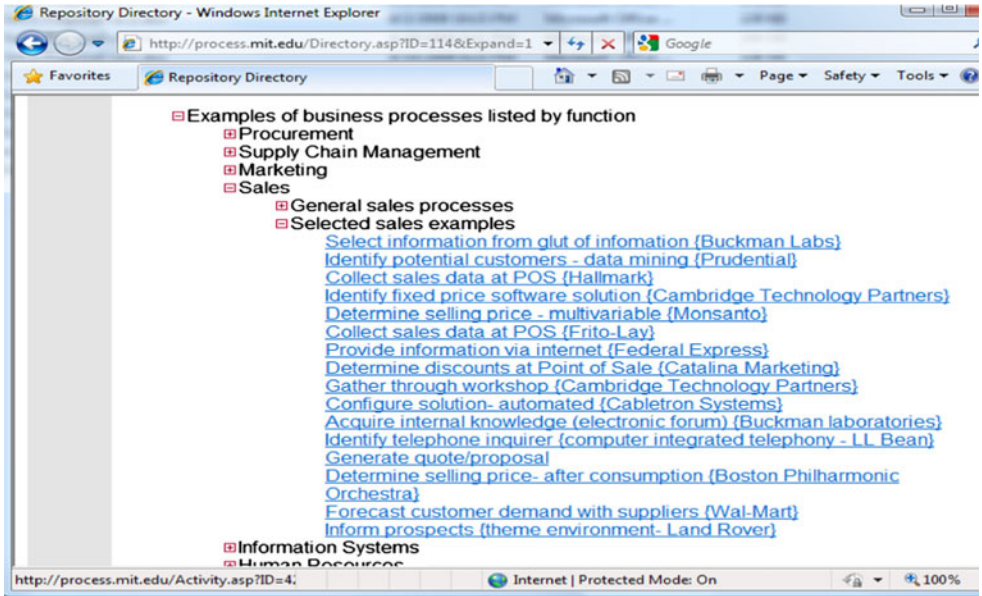


Fig. 1 MIT process handbook repository

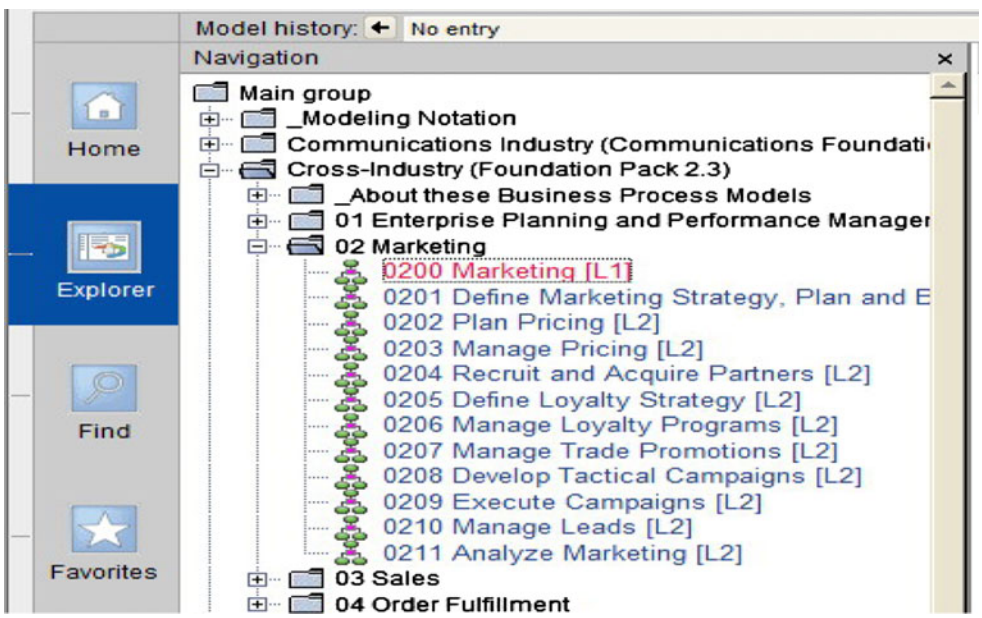


Fig. 2 Oracle business process repository

Systems, Human Resources, Strategic Planning, Finance/Accounting, Manufacturing/Logistics, Engineering.

- SAP PRM (9 functions): Sales, Production, Procurement, Accounting, Organization and Human Resources, Business Planning and Controlling, Capital Asset Management, Finance Management, Customer Service.
- Oracle BPP (20 functions): Enterprise Planning and Performance Management, Marketing, Sales, Order Fulfillment, Customer Service, Field Service and Depot Repair, Supply Chain Planning, Product Management, Production, Procurement, Materials Management and Logistics, Transportation Management, Project Management, Financial Control and Reporting, Cash and Treasury Management, Asset Lifecycle Management, Enterprise Information Management, Recruiting, Workforce Deployment, Compensation Management.
- RosettaNet (RN) PIPs (7 functions): Partner Product and Service Review, Product Information, Order Management, Inventory Management, Marketing, Service and Support, Manufacturing.

- *Classification Granularity* (*CG*): SAP PRM, Oracle BPP, and RN PIPs only classify the processes as a whole, i.e. process-level classification. Besides process-level classification, MIT PH also classifies the tasks of each process according to their types, i.e. task-level classification.

- *Graphical Process Representation* (*GPR*): MIT PH only describes processes using texts. SAP PRM represents processes as Event-driven Process Chains (EPCs). Oracle BPP and RN PIPs use proprietary flowcharts to depict their processes.
- *Technical Details* (*TD*): MIT PH focuses on business process knowledge and is system and implementation independent. SAP PRM is designed for SAP ERP system, where some of the pre-configured processes are ready for execution. RN PIPs provides lots technical documents specifying the format of business objects and messages. Oracle BPP has very limited technical details in term of process implementation.
- *Business Cases* (*BC*): For each general process description in MIT PH, examples of processes implemented by specific companies are provided to stimulate new ideas on process design. Only a few real-world cases are presented in [10] for SAP PRM. RN PIPs and Oracle BPP only provide standard process templates without specific cases.
- *Searchability* (*SE*): only limited keyword-based search is supported by all four models.
- *Scope* (*SC*): we informally define scope as the number of process templates provided in the reference model. MIT PH provides 19 general process descriptions. For SAP PRM, we conducted a workflow template search in one instance of SAP ECC IDES (ERP Central Component Internet Demonstration and Evaluation System) and got 1,571 returned results. There are 209 process diagrams in Oracle BPP. RN PIPs currently list 99 PIPs in their library.

Based on the comparison result, we present next a use case to illustrate the needs and challenges of integrating multiple reference models for better process design support. Assume we have a business analyst who is in charge of designing a purchasing process for an e-business. The following are some of the requirements:

R1: Find an example on general purchasing process to get a basic understanding.
R2: Find cases on how other companies have implemented purchasing processes.
R3: Find out whether the company's ERP system supports such a process.
R4: Explore information on automating purchasing processes with business partners.

Table 1 shows the results of comparing the four reference models in terms of supporting the process design requirements listed above. For R1, MIT PH provides a very easy-to-understand general procurement process consists of the following key steps: Identify potential sources → Identify own needs → Select supplier → Place order → Receive → Pay → Manage suppliers. SAP PRM, Oracle BPP, and RN PIPs provide several detailed sub processes focusing on certain key tasks in the purchasing process without showing the overall process. For R2, MIT PH has 13 business cases from companies, such as Motorola, Kodak, and Harley Davidson, to show specific implementations of procurement examples, whereas SAP PRM and RN PIPs do not have any associated business cases. R3 is a very special requirement that only partially supported by SAP PRM, because tasks in SAP PRM are associated to SAP ERP transactions with documents and objects which provide useful information for R3 even when the ERP system is not from SAP. Given its non-technical nature, MIT PH does not provide support for R4. In contrast, RN PIPs have detailed implementation guides for coordinating purchasing processes electronically with business partners. This use case illustrates that no individual reference model can best support the sample process design requirements. In a real-world setting, there are often more diverse process design requirements for e-business.

Table 1 Results of supporting process design requirements

	R1	R2	R3	R4
MIT PH	Y	Y	N	N
SAP PRM	T	N	T	T
Oracle BPP	T	N	N	N
RN PIPs	T	N	N	Y

Y Yes, *N* No, *T* To some extent

We summarize the comparison result of the four PRMs and the use case discussed above into a set of research issues that motivate our research.

Research issue 1 Better support for e-business process design can be achieved by integrating multiple process reference models. Table 2 shows that different process reference models complement each other in terms of contents and scopes. Therefore, the union of multiple process reference models provides a much larger process knowledge base than each individual ones. Table 1 further illustrates that given the diverse process design requirements in e-business, a certain process design need might only be best supported by more than one reference model. In addition, different reference models usually provide their unique perspectives even for the same business process. These different perspectives often help process designers gain more insights on the process, such as additional tasks, actors, and data items. More research efforts on leveraging multiple process reference models for better process design for e-business are needed.

Table 2 Comparison of four process reference models

	PC	CG	GPR	TD	BC	SE	SC
MIT PH	Function (10)	Task	No	Low	Many	Keyword	Limited
SAP PRM	Function (9)	Process	Yes	High	Some	Keyword	Large
Oracle BPP	Function (20)	Process	Yes	Low	Few	Keyword	High
RN PIPs	Function (7)	Process	Yes	High	Few	Keyword	Medium

Research issue 2 There is no standard classification scheme used by existing process reference models to categorize their process designs. SAP, Oracle, MIT and RosettaNet classify their processes into different business functions, such as sales, marketing, accounting, etc. However, the number of business functions and function names are not consistent. For example, sales processing is categorized in SAP's reference model under "Order to Cash" category, while a similar process is categorized in Oracle's model under "Sales". Other classification schemes are also used. For example, MIT also classifies their processes using a "process compass" with four dimensions, whereas PwC uses the process classification framework developed by APQC (http://www.apqc.org) to categorize its best practices into 13 relatively general business processes. We argue that a standard and multi-facet process classification framework is necessary in order to efficiently combine and organize process knowledge from different sources.

Research issue 3 There is very limited support for efficient searching, discovering, and sharing process knowledge in existing process reference models. The process reference models we studied are provided mostly in the form of text documents and process maps classified into a tree-like hierarchy according to their classification schemes as either linked web pages, e.g. reference models from SAP, MIT, RosettaNet, and PWC or packaged software components, e.g. reference models from Oracle and Tibco. In order to find a certain process design case, the users usually have to understand how the case is classified or use keyword-based search to locate the case, which often does not return the most relevant results based on our testing on the web-based reference models from SAP, MIT and RosettaNet. Thus, it is not a trivial task to find most relevant process information to the current process design need. Therefore, a more efficient and innovative mechanism for exploring process knowledge in various reference models is necessary.

Although process design can benefit from integrating multiple reference models, integrating multiple reference models is not easy due to some key challenges listed below:

- Lack of an integrated process classification scheme. The current classification schemes of different models are not consistent and often overlapping.
- Proprietary modeling notations. MIT PH lacks of graphical process representation, where other two models use their proprietary modeling notations. An industry standard notation, e.g. BPMN, should be adopted.
- Lack of process semantics. Different terms with same meaning are used in different models, such as purchasing in SAP PRM and procurement in MIT PH and Oracle BPP, resulting in ambiguous process specification and poor model interoperability.

These challenges lead to the following research issue:

*Research issue 4 Integrating multiple process reference models is a non-trivial task, which requires a systematic approach.*The research issues identified in this section greatly hinder the usability of existing PRMs in term of e-business process design support, which motivate us to develop an innovative integrated process repository to provide a solution. Next, we study process design from a collaborative knowledge management perspective to provide a theoretical foundation for the design of our process repository.

4 A knowledge management perspective of process design

From a knowledge management perspective, process design is a process of creating process knowledge. Case-based process design, in particular, is a cyclic, social process of knowledge creation through reusing existing and creating new process designs. As business processes are evolving, process design knowledge is perishable [39]. The success of case-based process design depends on a knowledge creation cycle among individuals and organizations that continuously grows the knowledge base of designs.

In Nonaka's classic "spiral" model for organizational (social) knowledge creation [32] (See Fig. 3), knowledge is created through a cycle of four intertwining modes of conversion between tacit (unexpressed) and explicit (expressed) knowledge: internalization, externalization, socialization and combination. Internalization refers to the conversion of explicit knowledge into tacit knowledge, which corresponds to the traditional notions of learning, understanding or sense-making. Externalization refers to the expression of tacit knowledge as explicit knowledge,

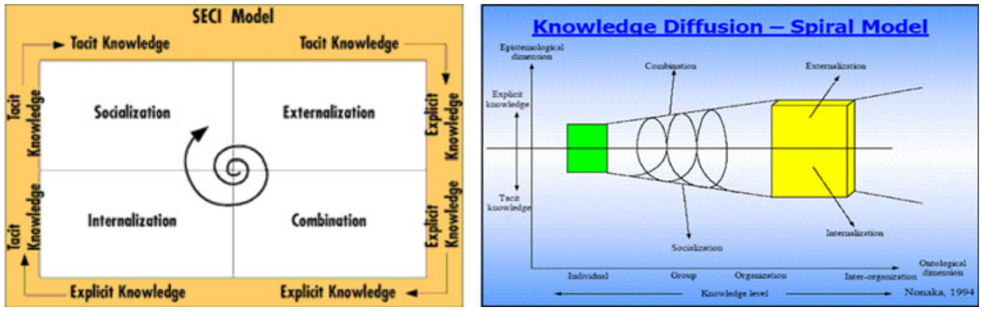


Fig. 3 Spiral model for organizational knowledge creation [32]

Table 3 Key features of proposed integrated process repository

Key features of proposed process repository	Knowledge creation mode(s) supported	Research issue(s) addressed
Collaborative classification and tagging	E, S, C	1, 2, 4
Advanced search and exploration	I	3
Design classification and recommendation	C	1, 2, 3
Feedback and reputation management	S, C	3, 4

I Internalization, *E* Externalization, *S* Socialization, *C* Combination

which corresponds to the traditional notion of codification. Socialization refers to creating tacit knowledge through social interactions and shared experience. Combination refers to creating explicit knowledge from other explicit knowledge, through merging, categorizing, sorting, or re-contextualizing. The cycle of these conversions expands along the social dimension from individuals to groups, organizations, inter-organizations, and to the whole society.

The process design process can be studied using Nonaka's spiral model of knowledge creation. Process designers "internalize" by exploring and learning from past designs. They "externalize" by documenting their design knowledge. They "socialize" through collaborating in designs or other means of interaction. They "combine" by categorizing, reusing and assembling past designs. Process design is carried out on different social levels: individual, group (such as small teams), organization (such as a consulting firm), or inter-organizations (such as a community of process design professionals, or during negotiation of an inter-organizational process). Through a cycle of these knowledge conversions, the organization (or community) builds up its process knowledge base from individuals' knowledge.

Using Nonaka's framework, the key to process model integration is Combination, i.e. merging, categorizing, sorting, or re-contextualizing process knowledge embedded in process models. Besides combination, a process model repository needs to support all other forms of knowledge creation and foster spiral knowledge creation cycles. Motivated by our theoretical framework of process knowledge creation, we specify several key requirements of an online integrated process repository, which help address the research issues presented in the previous section as shown Table 3. These features collectively support all four modes of process knowledge creation. We show next how we leverage Web 2.0 technologies to realize these features in our integrated process repository.

5 A web 2.0 approach to integrating and managing process reference models

5.1 PRM integration procedure and management mechanisms

Based on our theoretical framework of process knowledge creation, our process model integration approach is a collaborative process initiated by experts, evolved by a community of users, and aided by automated procedures as shown in Fig. 4. Key to this process is an integrated process repository with Web 2.0 technologies such as social classification, tagging and recommendation.

The integrated process repository starts from multiple process reference sources with different process models. Experts perform the initial schema mapping and merging,

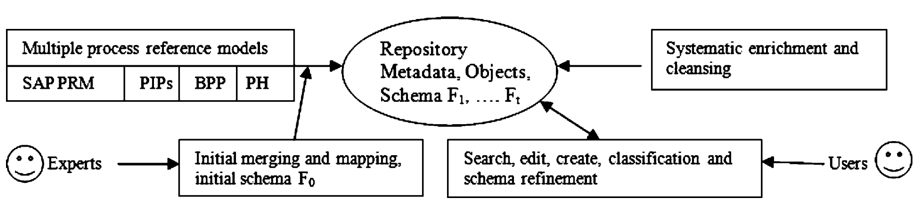


Fig. 4 Process reference model integration procedure

and create a baseline integrated schema of process models. Delphi approach may be used by a panel of experts. The baseline schema may be far from perfect, and processes may not be fully classified into the integrated model. However, the repository will be evolved by a much larger user community. Users may search and edit existing process designs, or create new ones. Users may classify (or re-classify) process designs into the integrated model. Users may refine and evolve the integrated schema created initially by experts. In addition to manual efforts, automated schema enrichment and cleansing procedures will be developed utilizing data mining techniques and semantic reference libraries.

More specifically, we operationalize the procedure via a multi-facet process classification framework. As noted in the Research Issue 2, existing process reference models classify process designs according to different classification schemes (facets) [41]. In addition, even if the same facet is chosen, e.g. business function, the categories within the facet defined by different reference models usually do not agree with one another. We propose a collaborative multi-facet process classification framework to address this issue and help with the integration of multiple process reference models. In particular, there are three key tasks in this framework as discussed below.

(1) Consolidating classification categories within a single facet from different sources. SAP, Oracle, RossettaNet and MIT all classify their process designs according to business functions, such as sales, marketing, production, etc., but the number and names of business functions used are quite different. For example, the numbers of top-level business functions defined in each reference models are as follows: SAP (9), Oracle (20), RossettaNet (7), MIT (10). We manually consolidated the business function categories from those four datasets using the approach informally described as follows: (a) we take the set of categories $C = \{c_1, c_2, \ldots, c_i\}$ from one dataset as a starting point, then (b) we go over each category from another dataset $c'_j \in C'$, $C' = \{c'_1, c'_2, \ldots, c'_j\}$ and determine whether there is a category $c_i \in C$ that is semantically equivalent to c'_j. If yes, c_i c'_j and are merged into c_i, otherwise add c'_j to C. (c) repeat step (b) for all remaining datasets and the resulting C is the set of consolidated categories. By consolidating the business function categories from the aforementioned four datasets, we get a total number of 14 categories for business function facet. We are in the process of consolidating subcategories for more detailed process classification based on business functions.

(2) Developing baseline multiple classification facets with initial categories. A process design usually has multiple perspectives. Therefore, a multi-facet process classification is needed to represent those perspectives and support exploration of process designs based on multiple perspectives. We have identified several process classification facets as the baseline as discussed below:

- Business function: we have discussed this in task 1.
- Industry: business processes are industry dependent. For example, processes from the same business function, such as R&D, could be very different between telecommunication companies and pharmaceuticals companies. We use the business categories from the Open Directory Project (http://www.dmoz.org/Business/) as the baseline industry categories for this facet, which has 27 categories in total.
- APQC's process classification framework: This framework has been widely used to benchmark process performance, where processes are classified into operating processes with 7 subcategories and management processes with 6 subcategories.
- Supply Chain Operations Reference-model (SCOR) (http://www.supply-chain.org/): SCOR is another well-known classification scheme for supply chain processes, which is based on five distinct management processes: Plan, Source, Make, Deliver, and Return.

Note that additional facets can always be added to provide additional classification perspectives. For the same process from different process reference models, such as the procurement process from SAP, Oracle, and MIT PH, they are classified in the same category for at least one facet, such as business function facet for the procurement process, and may be classified in different categories for other facet, such as the creator facet and industry facet.

(3) Utilizing social tagging and classification in maintaining and evolving the multi-facet process classification framework. Tasks 1 and 2 showed that a multi-facet process classification takes a lot of efforts to build and maintain. And yet, predefined categories may not be sufficient to classify some process designs. Social tagging and social classification are new fields that have attracted significant research interests [59, 60]. Social tagging allows metadata to be created collaboratively by data consumers versus data producers. In our approach, users can develop new facets and categories through social classification. Users can also add free-form tags to process design documents to add additional information for process design exploration and reuse. The tags can be used later to classify

process designs into facets and categories. In this way, our process classification framework can be maintained by leveraging social intelligence.

Compared to keyword-based searches found in most existing process reference models, our multi-facet classification allows users to contribute structural knowledge, in addition to knowledge that can be embedded in tags or textual description. The classification system allows users retrieve information through recognition of category names instead of recall of query keywords, where recognition is far easier than recall. Based on the initial multi-facet process classification framework, the process repository supports "guided classification", using novel interfaces to guide users to manually categorize process design documents into the framework. However, for a large growing process design repository our proposed integrated process repository, human efforts alone will not be sufficient to keep the design documents "fully" classified. Our integrated process repository also classifies process design documents automatically based on their contents and metadata. Feedback and reputation management are also important to any online knowledge communities. User feedbacks help improve a design, or help determine a design's quality and reusability. A reputation system helps motivate a user to contribute to the design repository. WPR utilizes user feedbacks, reuse statistics and user interactions to manage user reputation. With guided classification, automated classification and feedback mechanisms, WPR allow users collaboratively evolve a rich, multi-facet process classification.

Design recommendations address the information overload issue in a large design repository. Exploring and reusing past solutions as a problem-solving approach is called case-based reasoning [61]. Case-based reasoning has been applied in many fields, such as software design and reuse [58]. We apply case-based reasoning in our integrated process repository to allow users to post their design requirements, request for design recommendations and recommend relevant existing designs to each other's problem situations. As a recommender suggests an existing business process design to a problem situation, the system will ask the recommender to specify the facets along which they are related to each other. In other words, the system will ask recommender to categorize the past design into categories related to the current situation. The recommender who has suggested design cases also in effect suggest perspectives (facets) along which the design space can be further explored. In addition, such a recommendation mechanism enriches the process repository by classifying existing design cases and help design reuse in the future. In addition to user recommendations, our process repository will also provide systematic recommendations based on a combination of inputs by leveraging research in recommender systems.

5.2 The prototype system

Based on the PRM integration procedure and the related process design management mechanisms, we developed a prototype system, whose major components are shown in Fig. 5. All system components are implemented on the server side based on an open source content management platform. By placing no software requirements except a browser on the client side, the system provides a common platform for collaborative process knowledge management across organizational boundaries. Figure 6 shows the screenshots of the system. On the left, you can see the main page of the repository, where advanced search and navigation function based on multi-facet process classification framework is provided. On the right, you can see the detail page for each uploaded process design document. On this page, the users can classify process design into different facets and categories, add custom tags and provide feedback to the design and classification. As discussed in Sect. 5.1, we have developed the baseline classification facets with highest level of categories from the analysis of several existing process reference models and have created those facets and categories in the prototype system. Some subcategories have also been created. We will rely on the users of the system to further add, modify, and evolve the classification categories as we discuss in the next section.

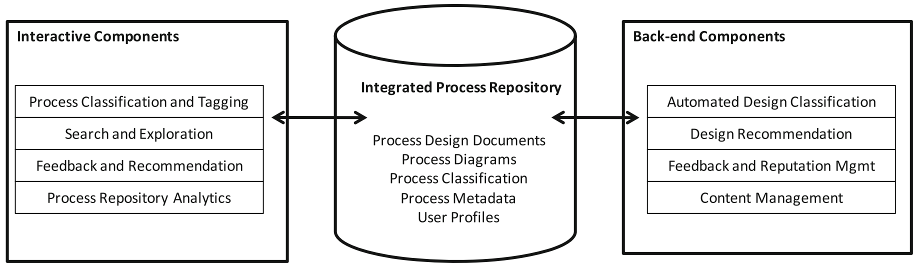
6 Evaluation

In this section, we evaluate our proposed system by deploying the system in both academic and business settings and conducting user experiments to evaluate the systems' usability and effectiveness. Below are some preliminary results.

6.1 Evaluation in an academic setting

The integrated process repository was deployed at a large tier-1 public university and used to teach process design in a graduate class with 30 students. The repository was seeded with 277 process designs from SAP and Oracle reference models. The process designs are diagrams with titles and brief textual descriptions. The students were required to design business processes for procurement, marketing, sales, and customer service for e-business in a controlled user experiment by leveraging the process designs in the repository. Initially, the process designs are unclassified, i.e., un-integrated, and the students were asked to only use keyword search and simple page-by-page

Fig. 5 Major components of the prototype system



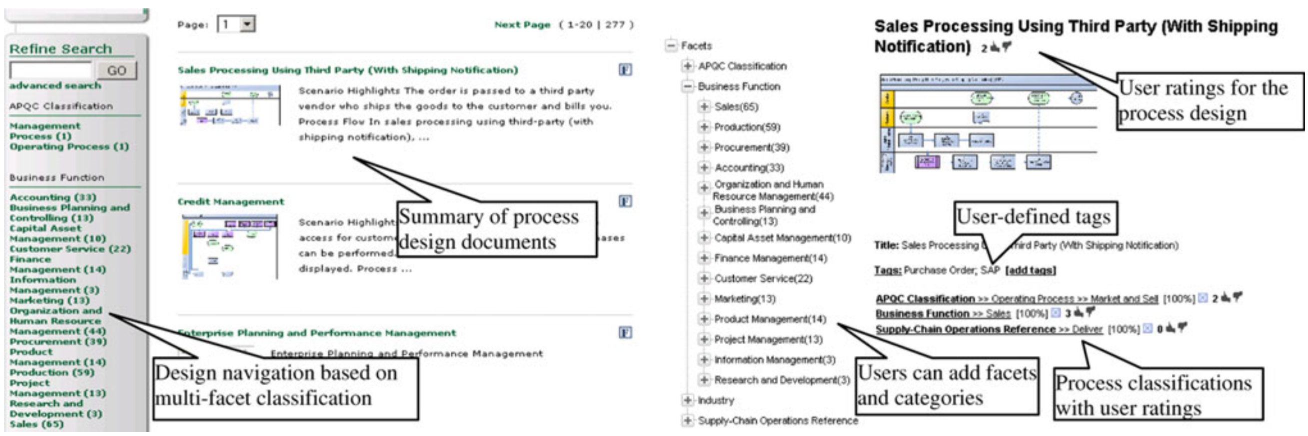


Fig. 6 Screenshots of integrated process repository

browsing to find relevant process designs. Then, the students collectively classify the process designs into different instructor-created facets, such as APQC's process classification scheme, business function, and Supply Chain Operations Reference classification. Students were also allowed to create new facets, categories and sub categories. Besides classification, the students can add free-form tags to process design and rate other students' or experts' classifications by a simple voting mechanism: either a positive vote (thumb-up) or a negative vote (thumb-down) as shown in Fig. 6. Once the process designs are classified, we asked the students to use the integrated process repository to try to find relevant process designs. The students' click stream data was collected by the system for analysis. A survey with 14 questions was used to help us evaluate whether students found the repository easy to use and whether the repository led to effective reuse and higher quality process designs.

Over 97% of the process designs are classified into at least one category in the multi-facet process classification schema to help integrate process designs from different sources. According to the survey, on average 96% of the students agreed that the system is easy to learn and use. 80% of the students agreed that the integrated process repository enabled by the multi-facet process classification is more efficient for finding relevant process designs than keyword-search-based unclassified process reference models. In summary, our academic evaluation proved our integrated process repository to be user friendly and effective for supporting process design.

6.2 Evaluation in an industry community setting

The integrated process repository was adapted and deployed at a software implementation community website for a popular Enterprise Resource Planning (ERP) software package. The community included implementation partners of the ERP software vendor. The process repository mainly consists of e-Business engineering specifications for implementation (including customization and integration) of the ERP software package, classified in multiple dimensions including the dimensions used in popular process reference models. The repository was seeded with process designs in the core ERP software package, and evolved with the process designs from custom implementations contributed by implementation partners. To encourage participation, the repository adopted an exchange model to allow partners to charge each other for reusing their process designs. Active participation allows implementation partners to achieve higher partnership levels without paying associated tiered partnership fees. The process repository also contained a recommendation feature based on customer characteristics, partnership level and user feedback. The recommendation feature encouraged participation and helped implementation partners to locate best practices from their peers.

The ERP software vendor incorporated the process repository into its website for implementation partners, in an effort to improve reuse of process designs and sharing of the best practices in its implementation community. Effective and efficient software implementations can bring a customer millions of dollars in returns, improve implementation consultancies' profit margins, and therefore sustain software vendors' growth.

ERP software implementations solve business problems typically by reengineering the business process in addition to customizing and integrating an off-the-shelf software package. Successful ERP package implementations rely heavily on reusing prior implementation experience. In fact, when facing a problem situation, implementation professionals frequently practice case-based reasoning: (1) explore the past cases (experience and best practice in the field), (2) find a similar case, and (3) reuse the solution for the past case in the new problem situation. When the same software package is implemented at many sites, such case-based reasoning can significantly improve quality, reduce costs, and lower design risks. Effective knowledge management is the key to quality improvements and cost or risk reductions in ERP implementations [24].

The implementation partners have contributed several thousands of documents into the repository. The documents were classified into over 20 categories under 4 facets. We interviewed consultants from several different implementation partners to evaluate the effectiveness of the process repository. Interviewees reported that by searching along multiple dimensions and different process reference models, they identified relevant e-Business scenarios and solutions that would have been otherwise difficult to find. For example, one consultant worked on report definitions for sales agents. The integrated repository has a functional dimension, which contains a Sales category, as well as an output dimension, which contains a Reports category. The consultant found several sales-oriented reports that saved her weeks of work. The process repository allows users to compare the ERP software package against its competitors in fine levels of detail, as the process designs embedded in the software package are classified along multiple dimensions, including certain dimensions in Oracle and SAP's process models. The sales and marketing teams of the software vendor have greatly benefitted from the process repository, which shows the capability and potential of the ERP software package.

Interesting Web 2.0 phenomena emerged from the prototype deployment. The exchange feature in the process repository provided a financial incentive for implementation partners to contribute their knowledge. While the process design details remained proprietary in many cases, a large number of engineering specifications such as the process input and output definitions were contributed to the process repository. The process repository seems to have greatly improved the willingness of the partners to share their knowledge with past implementations. Apparently the partners competed for reputation in the implementation community, as well as the partnership levels which help them close implementation deals. Based on interviews with the consultants, the increased knowledge contribution and collaboration have a significant impact on improving the process design quality across the whole implementation community of the ERP software package.

6.3 Discussion and limitations

Both academic and industry evaluation proved our integrated process repository to be user friendly and effective for supporting process design. While the integrated process repository approach is expected to support medium to large size of user communities, the evaluation has been limited to relatively small communities with no more than hundreds of users. The relatively small community size contributed to the validity and reliability of the schema. For example, we have not found any users purposely "spamming", "fighting" or otherwise deteriorating the repository. For larger user communities, the system will be more dependent on automated schema cleansing efforts to maintain quality and minimize information overload. The user load was relatively light in those evaluation environments, and the system has not been fully tested for scalability.

7 Conclusions and future work

In this paper, we first conducted a case study to show that better process design for e-business can be achieved by integrating multiple process reference models. Then, we proposed a collaborative knowledge management approach to developing such an integrated process repository. The integrated process repository incorporated many Web 2.0 technologies to help organize and maintain process design knowledge. We evaluated our approach by deploying the prototype system in academic and real-business settings and conducted user experiments to evaluate the system's utility, efficiency, and quality. To the best of our knowledge, our integrated process repository is the first process repository that leverages Web 2.0 technologies to combine and manage process knowledge from multiple sources to support process design for e-business.

Our research presented in this paper also has its limitations, which we plan to address in our future research. First, we are continuously enhancing the system's features, such as browser-based collaborative process modeling and social network based process design recommendation. In particular, we are integrating an open source collaborative process modeling system named Oryx BPMN modeler to our integrated process repository. Second, we plan to enrich the contents of the repository with process designs from other sources, such as PwC best practices, Supply Chain Operations Reference Model, and IBM Blueworks. In addition, we also want to incorporated process designs from less structured data sources, such as email exchanges, professional BPM forums, and search engine query results. Third, we are further evaluating our approach by conducting additional user experiments in broader industry settings such as supply chain management [23]. Finally, we are looking at the social issues in using the integrated process repository, such as incentives for community involvement, privacy, and data confidentiality.

Acknowledgments The project is supported in part by the United States National Science Foundation, Awards #0725277 and #0713290, and a research grant from Delaware Design Institute at the University of Delaware.

References

1. Aissi S, Malu P, Srinivasan K (2002) E-business process modeling the next big step. IEEE Comput 35(5):55–62
2. Aldowaisan TA, Gaafar LK (1999) Business process reengineering an approach for process mapping. Omega 27(5):515–524
3. Basu A, Kumar A (2002) Research commentary: workflow management issues in e-Business. Inf Syst Res 13(1):1–14
4. Bridge D, Goker MH, McGinty L, Smyth B (2005) Case-based recommender systems. Knowl Eng Rev 20:315–320
5. Byrd T, Davidson N (2006) An empirical examination of a process-oriented IT business success model. Inf Technol Manag 7(2):55–69
6. Cantara M et al (2009) Hype cycle for business process management. Gartner Research, Stamford
7. Capozucca A, Guelfi N (2010) Modelling dependable collaborative time-constrained business processes. Enterp Inf Syst 4(2):153–214
8. Chesbrough H (2005) Open innovation: the new imperative for creating and profiting from technology. Harvard business school press, Cambridge
9. Cobb CG (2004) Enterprise process mapping: integrating systems for compliance and business excellence. ASQ Quality Press, Milwaukee
10. Curran T, Keller G, Ladd A (1998) SAP R/3 business blueprint: understanding the business process reference model. Prentice Hall PTR, Upper Saddle River
11. Datta A (1998) Automating the discovery of AS-IS business process models. Probabilistic and algorithmic approaches. Inf Syst Res 9(3):275–301
12. Golder, Scott AB, Huberman (2006) Usage patterns of collaborative tagging systems. J Inf Sci 32(2):198–208
13. Gordijn J, Akkermans H, Van Vliet J (2001) Designing and evaluating e-Business models. IEEE Intell Syst 16(4):11–17
14. Harmon P (2007) Business process change: a guide for business managers and BPM and six sigma professionals. Elsevier, Amsterdam
15. Hofacker I, Vetschera R (2001) Algorithmical approaches to business process design. Comput Oper Res 28(13):1253–1275
16. Hornung T, Koschmider A, Oberweis AA (2007) Recommender system for business process models. In: Proceedings of 17th annual workshop on information technologies & systems
17. Hu D, Zhao JL (2008) Expert recommendation via semantic social networks. In: Proceedings of international conference on information systems. Paris, France
18. Hunt VD (1996) Process mapping: how to reengineer your business processes. Wiley, New York
19. Hwang S, Yang W (2002) On the discovery of process models from their instances. Decis Support Syst 34(1):41–57
20. Kettinger WJ, Teng JTC, Guha S (1997) Business process change: a study of methodologies, techniques, and tools. MIS Q 21(1):55–80
21. Lee J, Wyner GM, Pentland BT (2008) Process grammar as a tool for business process design. MIS Q 32(4):757–778
22. Lee S, Olson D, Lee S (2009) Open process and open-source enterprise systems. Enterp Inf Syst 3(2):201–209
23. Li L (2006) The effects of information technology implementation on supply chain collaboration. Int J Internet Enterp Manag 4(2):118–134
24. Li L, Zhao X (2006) Enhancing competitive edge through knowledge management in implementing ERP systems. Syst Res Behav Sci 23(2):129–140
25. Madhusudan T, Zhao JL, Marshall B (2004) A case-based reasoning framework for workflow model management. Data Knowl Eng 50(1):87–115
26. Madison D (2005) Process mapping, process improvement and process management. Paton Press, Chico
27. Malone TW, Crowston K, Herman GA (2003) Organizing business knowledge: the mit process handbook. MIT Press, Cambridge
28. Mansar S, Marir F, Reijers H (2003) Case-based reasoning as a technique for knowledge management in business process redesign. Electronic J Knowl Manag 1(2):113–124
29. Mansar SL, Reijers HA (2007) Best practices in business process redesign use and impact. Bus Process Manag J 13(2):193–213
30. Marlow C, Naaman M, Boyd D, Davis M (2006) HT06, tagging paper, taxonomy, Flickr, academic article, to read, in Proceedings of ACM conference on hypertext and hypermedia. p 31–40
31. MØller C (2007) Process innovation laboratory a new approach to business process innovation based on enterprise information systems. Enterp Inf Syst 1(1):113–128
32. Nonaka I (2005) A dynamic theory of organizational knowledge creation. Knowl Manag: Crit Perspect Bus Manag 5(1):14–37
33. O'Reilly (2005) What is web 2.0, O'Reilly
34. OMG (2005) UML superstructure specification, v2.0, http://www.omg.org/cgi-bin/doc?formal/05-07-04
35. OMG (2006) Business process modeling notation specification, http://www.bpmn.org/Documents/
36. Oracle (2008) State of the business process management market 2008
37. Oracle (2009) Oracle application integration architecture foundation pack for oracle applications
38. Park J, Kim K (2009) Hyperlinking the work for self-management of flexible workflows. Commun ACM 52(6):113–117
39. Peffers K (2002) Perishable research and the need for a new kind of IS journal. J Inf Technol Theory and Appl 4(1)
40. Popovaa V, Sharpanskykhb A (2008) Process-oriented organisation modelling and analysis. Enterp Inf Syst 2(2):157–176
41. Prieto-Diaz R (1991) Implementing faceted classification for software reuse. Commun ACM 34(5):88–97
42. Reijers H, Liman Mansar S (2005) Best practices in business process redesign an overview and qualitative evaluation of successful redesign heuristics. Omega 33(4):283–306
43. Reijers H, Mansar S, Rosemann M (2008) From the editors: introduction and a compass for business process design. Inf Syst Manag 25(4):299–301

44. Reijers HA, Limam S, van der Aalst WMP (2003) Product-based workflow design. J Manag Inf Syst 20(1):229–262
45. RosettaNet (2007) Overview: cluster, segments, and PIPs
46. Ryu K, Yucesan E (2007) CPM: a collaborative process modeling for cooperative manufacturers. Adv Eng Inform 21:231–239
47. SAP (2009) Process Management Lifecycle (PML) http://www.sdn.sap.com/irj/bpx/cycle
48. Scheer AW (2000) ARIS—Business process modeling, 3rd edn. Springer, Berlin
49. Segev A, Patankare A, Zhao JL (2003) e-Business process interleaving managerial and technological implications. Inf Syst e-Bus Manag 1(4):331–352
50. Snabe JH, Rosenberg A, Møller C, and Scavillo M (2008) Business process management: the sap roadmap. SAP PRESS, Dedham, MA
51. Tallon P (2008) Inside the adaptive enterprise an information technology capabilities perspective on business process agility. Inf Technol Manag 9(1):21–36
52. Tapscott D, Williams AD (2006) Wikinomics: how mass collaboration changes everything. Portfolio, New York
53. van der Aalst W, Hofstede A, Weske M (2003) Business process management: a survey. Lect Notes Comput Sci 2678:1–12
54. van der Aalst W, Weijters T, Maruster L (2004) Workflow mining discovering process models from event logs. IEEE Trans Knowl Data Eng 16(9):1128–1142
55. van der Aalst WMP (2000) Reengineering knock-out processes. Decis Support Syst 30(4):451–468
56. Weske M (2007) Business process management: concepts, languages, architectures. Springer, Heidelberg
57. Weske M, Goesmann T, Holten R, Striemer R (1999) A reference model for workflow application development processes. In: Proceedings of international conference on work activities coordination and collaboration. pp 1–10
58. Wu H, Cao L (2008) Supporting case-based design for packaged software implementations. In: Proceedings of 12th international conference on computer supported cooperative work in design. Xian, China, pp 306–311
59. Wu H, Zubair M, Maly K (2006) Harvesting social knowledge from folksonomies. In: Proceedings of the seventeenth conference on hypertext and hypermedia. ACM, New York, USA, pp 111–114
60. Wu H, Zubair M, Maly K (2007) Collaborative classification of growing collections with evolving facets, in the eighteenth conference on Hypertext and hypermedia
61. Xu L (1995) Case-based reasoning for AIDS initial assessment. Knowl Based Syst 8(1):32–38
62. zur Muhlen M (2004) Organizational management in workflow applications—issues and perspectives. Inf Technol Manag 5(3): 271–291