Emergent Concepts on Knowledge Intensive Processes

Gonzalo A. Aranda-Corral¹, Joaquín Borrego-Díaz², Juan Galán-Páez², and Antonio Jiménez-Mavillard³

¹ Universidad de Huelva, Department of Information Technology, Crta. Palos de La Frontera s/n. 21819 Palos de La Frontera, Spain

² Universidad de Sevilla, Department of Computer Science and Artificial Intelligence, Avda. Reina Mercedes s/n. 41012 Sevilla, Spain

³ Western University, Department of Modern Languages and Literatures, 1151 Richmond Street, London, Ontario, N6A 3K7, Canada

Abstract. An approach to refine and revise the general framework of KiP (Knowledge Intensive Process) is presented. The specific case of collaborative KiP is studied and the prominent role of collaborative KiPs in the general context of Business Processes is revealed. The approach is based on Formal Concept Analysis.

1 Introduction

Nowadays there exists a growing interest in deeply understanding business processes based on the use of Knowledge in an intensive way (Knowledge Intensive Process, KiP). This kind of processes are governed by Business Processes (BP) at several levels within knowledge companies. The task of obtaining an adequate integration of KiP into classic BP informational ecosystems represents a challenge that Knowledge Economy needs to solve [9], [12].

The design and extraction of patterns [18] for KiP is inherent to this challenge. BP patterns provide a number of benefits for BP models (BPM). As is it well known, obtaining BP patterns, allows to (see [11]): simplify work, encourage best practices, assist in BP analysis, show inefficiencies, remove redundancies and greatly aid to consolidate interfaces for a proper design of BPs as well as to facilitate their re-use. KiP are particularly complex BP (even more for the collaborative ones) and the study of their patterns is particularly challenging. An analysis of the requirements, characteristics and frameworks for KiPs is mandatory to state a successful formal basis for their modeling and extraction of patterns [8]. The formal basis is the first step in an attempt to answer questions as Which is the adequate pattern for a concrete KiP? Are we facing a new KiP? Which kind of KiPs are the best in a concrete BPM?

In order to formalize the detailed analysis of KiP and to study real-world applications and experiences, the aim of this paper is to show how to use Formal Concept Analysis (FCA) in the refinement of characterizations and requirements for managing and executing KiPs. Specifically we propose a formal refinement of the semantic relationships

^{*} Supported by TIC-6064 Excellence project (*Junta de Andalucía*) cofinanced with FEDER funds.

Table 1. Characteristics of KiPs [8]

Characteristic	Explanation		
C1: Knowledge driven	The status and availability of data and knowledge objects drive human decision mak		
C2: Collaboration-oriented	and directly influence the flow of process actions and events. Process creation, management and execution occurs in a collaborative multi-user en- vironment, where human-centered and process-related knowledge is co-created, shared		
	and transferred by and among process participants with different roles.		
C3: Unpredictable	The exact activity, event and knowledge flow depends on situation- and context-specific elements that may not be known a priori, may change during process execution, and may vary over different process cases.		
C4: Emergent	The actual course of actions gradually emerges during process execution and is deter- mined step by step, when more information is available.		
C5: Goal-oriented	The process evolves through a series of intermediate goals or milestones to be achieved.		
C6: Event-driven	Process progression is affected by the occurrence of different kinds of events that influ- ence knowledge of workers' decision making.		
C7: Constraint- and rule-driven	Process participants may be influenced by or may have to comply with constraints and rules that drive actions performance and decision making.		
C8: Non-repeatable	The process instance undertaken to deal with a specific case or situation which is hardly repeatable, i.e., different executions of the process vary from one to another.		

among characteristics and requirements of KiPs. An analysis of this kind can aid to clarify in which status KIPs are in BPM field. FCA has been successfully proved as an useful tool to analyze phenomenological reconstructions of Complex Systems [5], and BP in Knowledge-based companies are inherent complex systems where the understanding and classification of their elements strongly depends on the features used.

The motivation of this work is based on the fact that this approach clarifies a number of features associated with the evaluation and assessment of KiP, as for example: to refine evaluation and assessment frameworks by means of semantic methods, to compare requirements and characteristics of different KiP models and concrete cases, to decide which sets of requirements represent an innovation niche, etc.

Lastly, the application of FCA allows to revise KiPs associated with specific tools for Knowledge Management (KM). The resuls are applied to the analysis of OntoxicWiki's KiP for collaborative enrichment, extension and refinement of Documentation on ontologies [4] (see Sect. 5).

Structure of the paper. The next section introduces the main characteristics and requirements of the KiPs considered in this paper. Sect. 3 succinctly reviews FCA. In Sect. 4 FCA based analysis of KiPs is described, showing a number of interesting features of KiPs from this semantic perspective. In Sect. 5 a first application of this analysis to some processes, assisted by a semantic knowledge externalization tool (OntoxicWiki) is presented. Lastly, related work and some insights on future work are given.

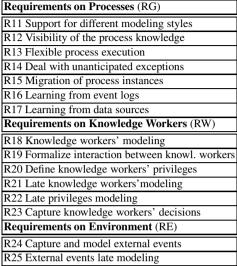
2 Characteristics and Requirements for KiPs

In collaborative KiPs three dimensions converge: the Knowledge Dimension, its collaborative nature, and its consideration as Business Process. All these three dimensions have to be modeled. A consensus on KiP definition is a key step in order to understand hidden mechanisms and patterns that operate in this kind of processes. A definition catching the complex nature of KiP could be the following (see [17] and also [8]):

A KiP is a process whose conduct and execution are heavily dependent on knowledge workers performing various interconnected knowledge intensive decision making tasks.

Table 2.	Requirements	for	KiPs	[8]
----------	--------------	-----	------	-----

Requirements on data (RD)	
R1 Data modeling	
R2 Late data modeling	
R3 Access to appropriate data	
R4 Synchronized access to shared data	
Requirements on Knowledge Actions (RK)	
R5 Represent data-driven actions	
R6 Late actions modeling	
Rules and constraints (RR)	
R7 Formalize rules and constraints	
R8 Late constraints formalization	
Requirements on Goals (RG)	
R9 Goals modeling	
R10 Late goal modeling	



KiPs are genuinely knowledge, information and data centric and require substantial flexibility at design- and run-time

A fine analysis of the main elements to consider in KiP analysis is given in [8]. The authors enumerate two sets of ingredients to describe and study KiP in order to provide a precise characterization. On the one side, it is mandatory to highlight the characteristics that could make KiP different of other BP (see Table 1). On the other side, a list of requirements retrieved for KiPs was considered (see Table 2). Both sets of ingredients were extracted from real-world application scenarios.

Requirements for KiPs are driven to achieve a sound representation and performance of the KiP instance models. In Table 2 a complete requirement list, due to Ciccio et al., [8], is shown. The aim of our work is to refine the analysis given in that article, by means of a systematic treatment of requirements and characteristics and other features which are essential in KiP analysis. In order to devise a robust (semantic based) refinement, a formal analysis is applied to concepts associated to Ciccio et al.'s framework. The analysis is carried out by means of formal concept reasoning.

3 Formal Concept Analysis

FCA mathematizes the philosophical understanding of a concept as a unit of thoughts composed of two parts: the extent and the intent. The extent covers all objects belonging to the concept, while the intent comprises all common attributes valid for all the objects under consideration [10].

A formal context M = (O, A, I) consists of two sets, O (objects) and A (attributes), and a relation $I \subseteq O \times A$. Finite contexts can be represented by a 1-0-table (identifying I with a boolean function on $O \times A$). Given $X \subseteq O$ and $Y \subseteq A$, it defines

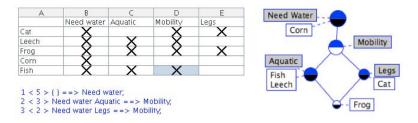


Fig. 1. Formal context of fishes, and its associated concept lattice

 $X' = \{a \in A \mid oIa \text{ for all } o \in X\} \text{ and } Y' = \{o \in O \mid oIa \text{ for all } a \in Y\}$

The main goal of FCA is the computation of the concept lattice associated with the context. A (formal) concept is a pair (X, Y) such that X' = Y and Y' = X. For example, the concept lattice from the formal context of fishes of Fig. 1, left (attributes are understood as *live in*) is depicted in Fig. 1, right. Each node is a concept, and its intension (or extension) can be formed by the set of attributes (or objects) included along the path to the top (or bottom). For example, the bottom concept ($\{eel\}, \{Coast, Sea, River\}$) is the concept *euryhaline fish*. CL contains every concept that can be extracted from the context. As well, concepts are defined but it is possible that no specific term (word) exists to denote it.

Knowledge Bases (KB) in FCA are formed by *implications between attributes*. An implication is a pair of sets of attributes, written as $Y_1 \to Y_2$. It is true with respect to M = (O, A, I) according to the following definition. A subset $T \subseteq A$ respects $Y_1 \to Y_2$ if $Y_1 \not\subseteq T$ or $Y_2 \subseteq T$. $Y_1 \to Y_2$ is said to hold in M ($M \models Y_1 \to Y_2$ or $Y_1 \to Y_2$ is an implication of M) if for all $o \in O$, the set $\{o\}'$ respects $Y_1 \to Y_2$.

Definition 31. Let \mathcal{L} be a set of implications and L be an implication.

- 1. L follows from $\mathcal{L}(\mathcal{L} \models L)$ if each subset of A respecting \mathcal{L} also respects L.
- 2. *L* is complete if every implication of the context follows from *L*.
- *3.* \mathcal{L} is non-redundant if for each $L \in \mathcal{L}$, $\mathcal{L} \setminus \{L\} \not\models L$.
- 4. \mathcal{L} is a (implication) basis for M if \mathcal{L} is complete and non-redundant.

A particular basis is the *Duquenne-Guigues* or so called *Stem* Basis (SB) [13]. The SB for the context of Fig. 1 is shown (down). In this paper no specific property of the SB is used, so it can be replaced by any other basis. In order to reason with implications, a production system can be used (see e.g. [2]).

Theorem 1. Let S be a basis for M and $\{A_1, \ldots, A_n\} \cup Y \subseteq A$. The following statements are equivalent:

1. $S \cup \{A_1, \ldots, A_n\} \vdash_p Y (\vdash_p \text{ is the entailment by means of a production system}).$

2. $\mathcal{S} \models \{A_1, \dots, A_n\} \to Y$

3. $M \models \{A_1, \dots, A_n\} \rightarrow Y$.

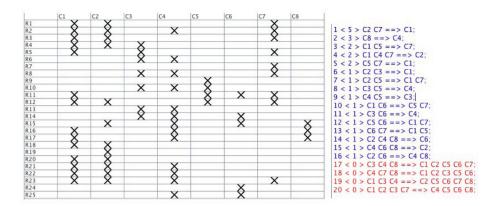


Fig. 2. Formal Context extracted from [8] and its associated STEM basis

In FCA, association rules are also implications between sets of attributes. Confidence and support are defined as usual in data mining. The *Stem Kernel Basis* (SKB) is the subset of the SB formed by the implications with nonzero support. To simplify, we assume that \mathcal{L}_M is a concrete basis for M (it is not necessarily the Stem basis). Likewise, the kernel of \mathcal{L}_M is denoted by \mathcal{L}_M^s . As it was above-mentioned, in the specific framework of CS observability, the set of implications with nonzero support gives some insights in a number of applications on both, micro and macro levels (cf. [1,3]).

4 FCA-Based Analysis of KiPs

In this section a number of results on the nature of (collaborative) KiP, obtained by means of FCA tools, are presented. The following subsections show a number of consequences, obtained from the semantic analysis performed on the concepts involved in Ciccio et al.'s framework, about the analysis performed about the refinement of the set of the characteristics, requirements, characterization of kind of KiPs and formal relationships. A brief analysis of the basis gives some insights about the KiP framework. Due to the lack of space, different applications have to be selected in each subsection. The analysis starts from an excellent and deep review [8]. A natural assumption from this paper is that the deep analysis made on it, is consequence of the study of a great number of KiPs, tools an BPM.

The relationship among KiP requirements (as objects) and characteristics (as features) is described in the formal context M_1 depicted in Fig. 2 (left). The Stem basis \mathcal{L}_C associated to the context, is also shown in Fig. 2 (right). Likewise, the dual context M_2 is considered (that its, that one built by using characteristics as objects and requirements as attributes). M_1 is very useful to understand how characteristics are related in the basis on requirements, while M_2 is very useful to relate requirements, showing if they are independent, if there exists subsumption among them, etc.

The concept lattice associated to M_1 has 31 concepts. In this case, a concept can be viewed as a set of characteristics with common attributes, which are the common requirements for a set of characteristics. The analysis of \mathcal{L}_C gives a number of interesting insights of the framework. Without being exhaustive, the main consequences are:

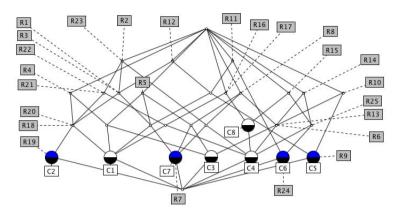


Fig. 3. Concept lattice considering requirements as attributes

4.1 Any Non-repeatable KiP Has Emergent Nature

This is the interpretation of the fact

$$\mathcal{L}_C \models C8 \to C4$$

Thus, from the point of view of [8] framework, the overall process of non repeatable KiP cannot be univocally determined by step-by-step elements.

4.2 Essential Requirements for Characteristics of KiPs

By analyzing the concept lattice associated to M_2 (see Fig. 3), it is possible to associate to each characteristic an essential requirement characterizing it. For example, from the point of view of the collaborative dimension, the following facts hold:

- C1 is characterized by R2 and R4: late data modeling and data access to shared data is essential for knowledge-driven KiP.
- C2 is characterized by R19: That is, every KiP accomplishing R19 (the process formalizes the interaction between knowledge workers) is collaboration-oriented.
- C3 is characterized by R6 and R4: KiP which accomplishes synchronized access to shared data and late actions modeling are essentially unpredictable.
- C4 is characterized by R2 and R6: late data modeling and late action modeling are essential requirements of emergent Kips.
- C5 is characterized by R10 and R12: KiP satisfying the visibility of process knowledge and late goal modeling are goal-oriented KiPs.
- C6 is characterized by R24: The capture and modeling of external events is the key feature of event-driven KiPs.
- C7 is characterized by R7: Constraint- & rule-driven KiPs formalize rules and constraints. Another more interesting characterization is " $R2 \wedge R8$ ": late data modeling and late constraints formalization.

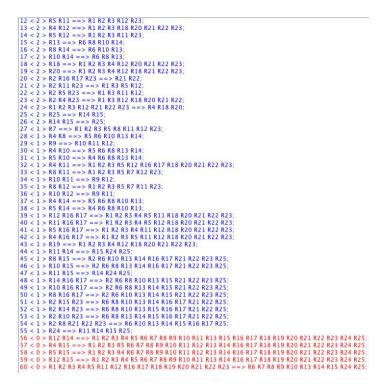


Fig. 4. STEM basis for M_2

- C8 is characterized by R15 and R16 (or R15 and R17). That its, non-repeatable KiPs usually provide the migration of process instances -R15- (that would that facilitates the instantiation and number of agents working in the process) and also helps to learn from data (event logs, as R16 or data sources, as R17).

4.3 Requirements Refinement

The analysis of the stem basis for M_2 , \mathcal{L}_R (see Fig. 4), shows a number of redundant requirements. A set of requirements $S \subseteq \{R_1, \ldots, R_25\}$ is a **minimal descriptional requirement system (mrds)**. It is possible to describe any requirement by conjunction of requirements of S, and any proper subset of S does not. In logic terms,

$$\mathcal{L}_R \cup S \models R_i \text{ for any } 1 \leq i \leq n$$

and $\mathcal{L}_R \cup (S \setminus \{R\}) \not\models R$ for any $R \in S$. The following set of requirements is a mrds:

$$\{R_2, R_4, R_6, R_7, R_{10}, R_{12}, R_{15}, R_{16}, R_{19}, R_{24}\}$$

A fine analysis of the mrds is interesting. On the one side, it contains requirements of each kind, supporting the completeness and non redundancy of the classification of the requirements from [8]. On the other side, any requirement is subsumed by the conjunction of at most two requirements of the mrds. Thus the mrds is a nice set of requirements to accomplish in order to design complete KiPs models and software.

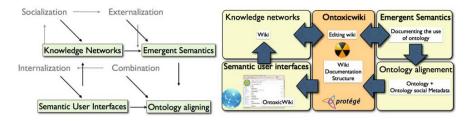


Fig. 5. Projection of Nonaka & Takeuchi's cycle (left) and its interpretation in Ontoxicwiki[4]

4.4 The Conceptual Nature of Collaborative KiPs

Collaboration oriented (C2) KiPs are important KiPs to consider in current BPMs. The recognizing of the collaborative nature of actions/process is important in order to design proper models as well as to detect patterns. In the concept lattice, C2 represents the intent of a concept that comprises the 48% of requirements, which gives us a sense of the relative complexity of collaborative KiPs within KiP. Moreover, the extent provides a fine definition in terms of requirements by means of FCA of collaborative KiPs:

 R_1, R_2, R_3, R_4 (that is, every requirement on data), R_{12}, R_{15} , $R_{18}, R_{19}R_{20}, R_{21}, R_{22}, R_{23}$ (that is, every requirement on knowledge workers)

Therefore, data and knowledge workers are essential elements in collaborative KiPs, allowing to (*a priori*) recognize collaboration in KiPs.

5 Case of Study: OntoxicWiki as a Collaborative Tool for KiP

OntoxicWiki [4] was designed to provide a semantic bridge between the knowledge activities of the projection of Nonaka and Takeuchi's cycle (NTC) for Knowledge externalization [15], enhancing both Web2.0 and SW solutions in this context (fig. 5). The tool is designed to satisfy several needs which arise when NTC is adapted for a semantic framework. Specifically, OntoxicWiki aims to bridge the gap between user and ontology. The main objective of this application is to represent ontologies in an intuitive and easy understandable way for any user by providing them with an environment from which the can repair and document ontologies socially, concretely with wiki technologies and associated patterns collaborative methods. KiP activities associated to OntoxicWiki are intimately related with the semantic specialization of NTC, which shows four needs for creating truly SW2.0 communities: *emergent semantics, semantic user interfaces, knowledge networks and ontology alignment* (see fig. 5):

Therefore, KiPs associated to Ontoxicwiki use fit on CKW life cycle from [14]. As KiP tool, OntoxicWiki exploits the social nature of Wiki technologies with the formatting, by means meta-data, of knowledge on use of concepts by knowledge workers. The knowledge processing cycle associated to OntoxicWiki ecosystem is depicted in Fig. 6 (specialized to a Pharmaceutical lab). To study its collaborative nature, authors have analyzed the fourteen requirements associated to collaborative KiP according to FCA interpretation of Ciccio et al.'s framework.

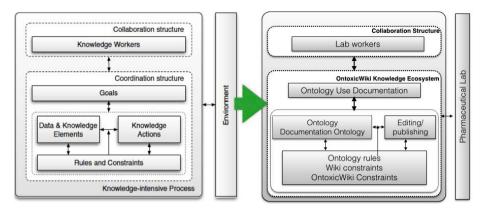


Fig. 6. Ontoxicwiki assisted process as a KiP

The requirements on data are accomplished by the Reparation of Ontology Process (ROP) and the Use Documentation Process (UDP). Both process also accomplish the full set of requirements on knowledge workers. However R_{12} and R_{15} are not satisfied by ROP and UDP. Therefore, an interesting consequence of the FCA-based analysis applied to OntoxicWiki's KiP suggest the inclusion of tools for visualizing the knowledge process (which Wiki technologies hide on edition logs and it is affected by user permissions), allowing the fair migration of process instances (mainly, the basic operations of ontology class edition and documentation).

6 Related Work

An approach to refine and revise the general framework of KiP has been presented. The specific case of collaborative KiP is has been studied, and the prominent role of collaborative KiPs in the general context was revealed. Other approach to recognize collaboration by means of FCA applications can be found in the scientific literature. In [6] FCA is applied for detecting and recognizing recurring collaborations among software artifacts. Our approach is at general level and it covers general KiPs in BP. An ambitious program would be the application of this approach to the BPs archived in the MIT process handbook (http://ccs.mit.edu/ph). In this case, it is convenient the use of the association rules basis (called *Luxenburger basis*) instead of Stem Basis.

Our approach is bottom-up, but an top-down approach to the semantic analysis of KiP can be achieved by means of ontologies. In [7] an ontology for KiP, KIPO, is presented. This approach is useful due to the fact that ontologies provide reuse of KiPs, and current research in the literature points to the lack of approaches of this kind. Requirements enumerated in that paper highlight the role of tacit knowledge in KiPs. Authors propose an orthogonal approach to the one devised in this paper: KIPO declares internal elements of KiPs whilst our approach only declares *exogenous* requirements for KiPs.

7 Future Work

With respect to the preliminary analysis of KiPs associated to the use of OntoxicWiki, it is interesting to remark that OntoxicWiki outputs documented evidence of KiPs

(ontology versions). Therefore it is possible to enhance OntoxicWiki in order to accomplish requirements on learning from log and data (R_{16}, R_{17}) . With respect to the documentation of KiPs, a similar analysis to the one made in this paper can facilitate the reuse as well as to detect patterns in documentation of KiPs [16].

Lastly, a further work is the analysis and characterization of KiPs which autonomous agents execute in Knowledge based tasks. In the case presented in [3], dialogue and argument-based process follow KiP patterns which can be reused in other BPM.

References

- Aranda-Corral, G.A., Borrego-Díaz, J., Galán-Páez, J.: Qualitative reasoning on complex systems from observations. In: Pan, J.-S., Polycarpou, M.M., Woźniak, M., de Carvalho, A.C.P.L.F., Quintián, H., Corchado, E. (eds.) HAIS 2013. LNCS, vol. 8073, pp. 202–211. Springer, Heidelberg (2013)
- Aranda-Corral, G.A., Borrego-Díaz, J., Galán-Páez, J.: Complex Concept Lattices for Simulating Human Prediction in Sport. J. Syst. Sci. and Complexity 26(1), 117–136 (2013)
- Aranda-Corral, G.A., Borrego-Díaz, J., Giráldez-Cru, J.: Agent-mediated shared conceptualizations in tagging services. J. Multimedia Tools and Applications 65(1), 5–28 (2013)
- Aranda-Corral, G.A., Borrego-Díaz, J., Jiménez-Mavillard, A.: Social Ontology Documentation for Knowledge Externalization. Comm. in Comp. and Inf. Sci. 108, 137–148 (2010)
- Aranda-Corral, G.A., Borrego-Díaz, J., Galán-Páez, J.: On the Phenomenological Reconstruction of Complex Systems–The Scale-Free Conceptualization Hypothesis. Systems Research and Behavioral Science 30(6), 716–734 (2013)
- 6. Arevalo, G., Buchli, F., Nierstrasz, O.: Detecting Implicit Collaboration Patterns. In: Proceedings of WCRE (2004)
- Carvalho, J.E.S., Santoro, F.M., Baião, F.A., Pimentel, M.A.: KiPO: the knowledge-intensive process ontology. Software & Systems Modeling, 1–31 (2014)
- Di Ciccio, C., Marella, A., Russo, A.: Knowledge-Intensive Processes–Characteristics, Requirements and Analysis of Contemporary Approaches. Journal on Data Semantics (2014)
- Fiechter, C.A., Marjanovic, O., Boppert, J.F., Kern, E.-M.: Knowledge management can be lean: Improving knowledge intensive business processes. In: Howlett, R.J. (ed.) Innovation through Knowledge Transfer 2010. SIST, vol. 9, pp. 31–40. Springer, Heidelberg (2011)
- 10. Ganter, B., Wille, R.: Formal Concept Analysis. Mathematical Foundations. Springer (1999)
- 11. Glushko, R.J., McGrath, T.: Document engineering. MIT (2008)
- Gronau, N., Weber, E.: Management of knowledge intensive business processes. In: Desel, J., Pernici, B., Weske, M. (eds.) BPM 2004. LNCS, vol. 3080, pp. 163–178. Springer, Heidelberg (2004)
- 13. Guigues, J.-L., Duquenne, V.: Familles minimales d'implications informatives resultant d'un tableau de donnees binaires. Math. Sci. Humaines 95, 5–18 (1986).
- Mundbrod, N., Kolb, J., Reichert, M.: Towards a system support of collaborative knowledge work. In: La Rosa, M., Soffer, P. (eds.) BPM Workshops 2012. LNBIP, vol. 132, pp. 31–42. Springer, Heidelberg (2013)
- 15. Nonaka, I., Takeuchi, H.: The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation. Oxford Univ. Press (1995)
- Scheithauer, G., Hellmann, S.: Analysis and documentation of knowledge-intensive processes. In: La Rosa, M., Soffer, P. (eds.) BPM Workshops 2012. LNBIP, vol. 132, pp. 3–11. Springer, Heidelberg (2013)
- Vaculin, R., Hull, R., Heath, T., Cochran, C., Nigam, A., Sukaviriya, P.: Declarative business artifact centric modeling of decision and knowledge intensive business processes. In: 15th IEEE International Conference on Enterprise Distributed Object Computing, EDOC 2011 (2011)
- Verginadis, Y., Papageorgiou, N., Apostolou, D., Mentzas, G.: A review of patterns in collaborative work. In: Proc. 16th ACM Int. Conf. Supporting Group Work (GROUP 2010), pp. 283–292. ACM (2010)