SENTINEL: A Semantic Business Process Monitoring Tool

Carlos Pedrinaci
Knowledge Media Institute
The Open University
Walton Hall
Milton Keynes, MK7 6AA,
United Kingdom
c.pedrinaci@open.ac.uk

Dave Lambert
Knowledge Media Institute
The Open University
Walton Hall
Milton Keynes, MK7 6AA,
United Kingdom
d.j.lambert@open.ac.uk

Branimir Wetzstein
Institute of Architecture of
Application Systems
University of Stuttgart
Universitätsstraße 38
70569 Stuttgart, Germany
branimir.wetzstein@iaas.uni-stuttgart.de

Tammo van Lessen
Institute of Architecture of
Application Systems
University of Stuttgart
Universitätsstraße 38
70569 Stuttgart, Germany
tammo.van.lessen@iaas.uni-stuttgart.de

Luchesar Cekov
Ontotext Lab. / Sirma Group
135 Tsarigradsko Shose Blvd.
Office Express IT Center,
Sofia 1784, Bulgaria
luchesar.cekov@ontotext.com

Marin Dimitrov
Ontotext Lab. / Sirma Group
135 Tsarigradsko Shose Blvd.,
Office Express IT Center,
Sofia 1784, Bulgaria
marin.dimitrov@ontotext.com

ABSTRACT

Business Activity Monitoring (BAM) aims to support the real-time analysis of business processes in order to improve the speed and effectiveness of business operations. Providing a timely, integrated high-level view on the evolution and well-being of business activities within enterprises constitutes a highly valuable analytical tool for monitoring, managing and hopefully enhancing businesses. However, the degree of automation currently achieved cannot support the level of reactivity and adaptation demanded by businesses. We argue that the fundamental problem is that moving between the business level and the IT level is insufficiently automated and suggest an extensive use of semantic technologies as a solution. In particular, we present SENTINEL a Semantic Business Process Monitoring tool that advances the state of the art in BAM by making extensive use of semantic technologies in order to support the integration and derivation of business level knowledge out of low-level audit trails generated by IT systems.

Categories and Subject Descriptors

D.2.8 [**Software Engineering**]: Metrics—*Performance measures, Process metrics*; D.2.9 [**Software Engineering**]: Management—*Life cycle, Productivity*; H.4.1 [**Information Systems Applications**]: Office automation—*Workflow management*; I.2.4 [**Information Systems Applications**]: Knowledge Representation Formalisms and Methods

General Terms

Measurement, Management, Performance

Keywords

Business Activity Monitoring, Business Process Analysis, Semantic Business Process Management

1. INTRODUCTION

Business Process Management (BPM) intends to support "business processes using methods, techniques, and software to design, enact, control, and analyze operational processes involving humans, organizations, applications, documents and other sources of information" [41]. BPM acknowledges and aims to support the complete life-cycle of business processes which undoubtedly involves the analysis and reengineering of process models. However, BPM has made more evident the difficulties of obtaining automated solutions from high-level business models, and for analyzing the execution of processes from both a technical and a business perspective [20]. The fundamental problem is that moving between the Business Level and the IT Level is hardly automated. Deriving an IT implementation from a business model is particularly challenging and it requires an important and ephemeral human effort which is expensive and prone to errors. Conversely analysing automated processes from a business perspective, e.g., calculating the economic impact of a process or the performance of departments within an organisation, is again an expensive and difficult procedure which typically requires a human in the loop.

One of the distinguishing characteristics of BPM solutions with respect to traditional Workflow Management Systems is commonly referred to as Business Process Analysis (BPA) [41]. The main goals pursued by BPA are on the one hand the verification or validation of the execution with respect to prescribed or expected processes, and on the other hand the identification of potential improvements of business processes. The knowledge gained in this phase is thus employed for reengineering and fine tuning existing process definitions.

This area therefore comprises a wide-range of fields such as Business Activity Monitoring, Business Intelligence, Business Process Mining and Reverse Business Engineering. The importance of BPA is widely acknowledged and in fact all the main vendors provide their own solutions [45]. The quality and level of automation provided by these tools are rather similar and not surprisingly major efforts are devoted to presenting the information in a simple yet meaningful way better supporting humans in the analysis phase. As a consequence the state of the art in BPA represents yet another bottleneck in the management scalability of business processes.

We refer to Business Activity Monitoring (BAM) as the technology in charge of "providing real-time access to critical business performance indicators to improve the speed and effectiveness of business operations" [27]. For its very nature, within BAM the previously mentioned difficulties are even more outstanding. We have previously argued for the use of semantic technologies, namely ontologies and Problem-Solving Methods [38], as a means to enhance the state of the art in BPA [3]. In the light of this vision, we have defined Core Ontology for Business pRocess Analysis (COBRA) [33] which provides a core terminology where business practitioners can map domain-specific knowledge in order to analyse their business processes. We have also defined additional extensions for capturing semantically the logs produced by IT systems and for deriving knowledge in terms of COBRA. In this paper we present SENTINEL (SEmaNTic busINess procEsses monitoring tooL), a tool that advances the state of the art in BAM by making an extensive use of semantic technologies in order to support the integration and derivation of business level knowledge out of low-level audit trails generated by IT systems.

The remainder of the paper is organised as follows. First we analyse the main requirements BAM solutions need to fulfil paying particular attention to the difficulties currently encountered. We next describe SENTINEL, our semantic business processes monitoring tool. Finally, we discuss related work, present some conclusions from our work and define lines for future research.

2. ON THE NEEDS OF BUSINESS ACTIVITY MONITORING

Providing a timely, integrated high-level view on the evolution and well-being of business activities within enterprises constitutes a highly valuable analytical tool for monitoring, managing and hopefully enhancing businesses. The technical challenges it poses are, however, not so obviously addressed. In this section we review some of the main challenges that arise when building a BAM solution and present at the same time the requirements that have driven the development of SENTINEL which will be presented in Section 3.

2.1 Data Integration

Business Process Analysis (BPA) uses the logs captured by the underlying IT infrastructure such as Enterprise Resource Planning, Customer Relationship Management, and Workflow Management systems to derive information concerning the well-being of business processes [41]. Common practice within the industry is to build a Data Warehouse which consolidates all sorts of corporate information and enriches it with derived statistical data [9]. Constructing a Data Warehouse is however an expensive, delicate, and somewhat brittle process which is particularly sensitive to changes on the underlying IT infrastructure. Current approaches are based on a so-called Extract-Transform-Load phase which asynchronously takes data from a myriad of systems in, typically, highly heterogeneous formats and loads them into a Data Warehouse for further analysis. Not surprisingly one main challenge envisaged by BPA solutions regards gathering and integrating large amounts of heterogeneous yet interrelated data within a coherent whole.

Once a Data Warehouse has been built and populated, Online Analytical Processing (OLAP) and Data Mining tools enable sophisticated data analysis that can help business analysts understand their businesses and even predict future trends. However, the semantics of the data being implicit, both OLAP and Data Mining techniques can hardly benefit from contextual knowledge about the organisation at analysis time, and strictly rely on human interpretation of the results [45]. This not only brings additional manual labour to an already complex and time consuming task, but it also prevents the automation of certain decision making procedures. As a result, enterprises often develop expensive domain-specific solutions which become an additional management overhead when changes within the enterprise need to be implemented.

The main difference between BAM and traditional real-time monitoring lies in the fact that BAM gathers data from a wide range of internal and external application systems in order to provide a richer view on business activities. As a consequence, BAM solutions need to address the technical challenges Data Warehousing currently faces and must do so virtually in real-time.

2.2 Monitoring Modes

Two kinds of monitoring are usually distinguished in existing tools [36], depending on whether the information is propagated actively, that is automatically by the infrastructure, or passively:

1. Active Monitoring is concerned with "real time" propagation of relevant data concerning the enactment of business processes, such as the status or the execution time. Active Monitoring is therefore limited to presenting in real time the data provided by the underlying IT infrastructure.

2. Passive Monitoring delivers information about process instances upon request. This information can be general details about the execution of processes or specific details concerning a particular process enactment, such as obtaining the current value of some variable. Passive Monitoring offers the possibility for obtaining information not provided by the Active Monitoring mode concerning ongoing processes. It also supports retrieving old information from monitoring logs if the user needs to see historic records.

Since the very purpose of BAM is to support the analysis of processes and the eventual adoption of corrective measures at runtime, BAM solutions need to support support both Active and Passive monitoring, allowing a seamless transition between both modes depending on the user interactions. For instance, a user should be able to request more information about some process if actively monitoring the execution triggers the interest for obtaining further details. In this sense BAM tools should seamlessly support a smooth exploration of the business process space, allowing the user to focus on certain aspects, to obtain further information with respect to certain details, etc. We shall deal with this in more detail in the next section.

2.3 Customisable and Dynamic Information Visualisation

Monitoring information is often structured around three different views [36]: (i) the Process View which is concerned with key performance indicators of business processes; (ii) the Resource View centred around the resources, human or mechanized, required for executing processes; and (iii) the Object View which focuses on business objects such as inquiries, orders or claims. These three views are populated with statistical information such as the minimum, the maximum, the average or the deviation of some parameter of interest.

These views are of major relevance to business analysis and management, and as a consequence they are typically supported by BAM tools. However, different users have different roles, interests, access rights, and preferences and these vary depending on the specific scenario, the focus of their analysis, etc. The user interface of a fully-fledged general purpose solution must therefore be characterised by its flexibility [29]. This includes for instance support querying, filtering and ordering the information according to user defined specifications [22]. Indeed, given the kinds of users addressed, the specification of these queries and filters should be supported in a simple way so that business analysts can browse the existing execution information effectively.

Similarly, different domains exhibit particular characteristics which impedes a "one size fits all" approach. The monitoring tool should therefore support users in defining their own visualisation templates to be populated with relevant monitoring information. The visualisation framework should be supported by a wide range of graphical representations such bar charts, line charts, pie charts, time series charts, etc. Additionally, the visualisation framework should support the presentation of user-defined information combining diverse statistical information about processes, etc. We shall next deal with the computation of this information in more detail.

2.4 Metrics Computation

In the business world the maxim "*if you can't measure it, you can't manage it*" is often used. Although it is too blunt a statement, it captures an important essence in current management approaches which try to maximise the aspects measured in order to evaluate, compare, and control the evolution of businesses. For instance the Balanced Scorecard is a popular "*set of measures that gives top managers a fast but comprehensive view of the business*" [24]. In a nutshell, the Balanced Scorecard defines four perspectives and suggests for each of them a set of aspects that managers should focus on. Assessing how well a company is doing is then a matter of calculating metrics and contrasting them with respect to pre-established goals for each of these key aspects. In the same vein but in more concrete terms, the Supply-Chain Council defines in the Supply-Chain Operations Reference-model a set of Supply-Chain targeted metrics such as "fill rate by order" or "total order fulfilment lead time" [39], which represent what is often referred to as Key Performance Indicators in the literature [24, 5].

BAM tools therefore need to be able to automatically compute metrics out of the body of information gathered at runtime by the execution and monitoring infrastructure. In the literature, metrics are often divided into two different kinds, namely general metrics and user-defined metrics [8]. The former are metrics which are generally useful for any process despite its specific domain or characteristics. Among these one can identify metrics like the "number or processes running", the "number of execution failures", statistical information about the execution time of processes, etc. The latter are metrics defined specifically by the user for a specific domain. The prototypical example is the "process cost" which is indeed domain and even process-specific and cannot therefore be predefined in advance. In order to cater for this a general purpose BAM tool must support users in defining their own metrics in a simple yet operational way, allowing them to create them and obtain automatically the results much like for general purpose ones.

2.5 Advanced Analysis

Effectively analysing business processes requires computing metrics that can help determining the health of business activities and thus the whole enterprise. However, this is not all there needs to be done. Aspects like client's satisfaction, whether a certain strategy will work out or not, how successful the research department is, or what would happen if we make a certain change in a process cannot "simply" be measured. Still, business practitioners need to take decisions on a daily basis. In some cases qualitative aspects (e.g., customer satisfaction) can artificially be transformer into quantities (e.g., happy = 3, very happy = 4). However, more complex cases like for instance what-if scenarios (i.e., what would happen if we make this change?) are not that easy to handle. Similarly, detecting, or better yet, anticipating process deviations with respect to expected behaviours can hardly be approached as a simple measurement problem.

In order to deal with these scenarios, BAM solutions need to apply advanced analysis techniques. In the literature we find techniques based on plain statistical analysis, data and process mining techniques, neural networks, case-based reasoning, or simply heuristic rules [8, 29, 17, 5, 45, 26, 4]. Some techniques focus on automating analysis to the greatest extent whereas others pay particular attention to obtaining results that can easily be explained and presented to the user. Regardless of the techniques though, there is a clear need for BAM solutions to take over as much burden as possible from the business analyst in the decision making process, and this necessarily requires the application of advanced analysis techniques. It is worth noting in this respect that the closer we get to strategic analysis, the more impact analysis results are likely to have, but the more complex analysis techniques are likely to be required in order to deal with qualitative aspects, approximations, and uncertainty.

2.6 Infrastructure Management

In [16] the authors identify three main levels for the analysis of e-businesses, namely the business level, the process level and the IT level. The first level is concerned about the value exchanges between the different actors involved (e.g., companies) and is therefore of particular relevance for business practitioners. The second level considers the process point of view (e.g., BPEL level) and is usually the focus of process architects. Finally, the third level is concerned about technical details such as the decomposition of a process into Web Services. An effective BAM solution must therefore provide the means for analyzing existing e-businesses at these three layers in a way that actions at the business level are propagated to the underlying IT support and conversely IT systems behaviour is escalated and reflected at the business level.

It is therefore particularly appealing a feature, if not necessary, to include management functionality allowing the user to control processes execution within the Workflow Management systems, work allocation through Enterprise Resource Management, etc. After all, BAM aims to support business analysts in adopting decisions with respect to deployed business processes. Being able to propagate complex decisions adopted at strategic levels down to the IT layer appears to be a necessary step towards supporting the level of adaptation demanded by businesses nowadays. Additionally, this is a necessary step towards fully automated reactive solutions that can take certain decisions on behalf of users when critical conditions are met.

3. SENTINEL

In order to achieve the level of automation and genericity required by businesses we advocate extensive use of semantic technologies. In the remainder of this section we will describe SENTINEL, a tool that advances the state of the art in BAM by making an extensive use of semantic technologies in order to support the integration and derivation of business level knowledge out of low-level audit trails generated by IT systems. We first provide the overall rationale underlying the tool paying particular attention to the use of semantics and then focus on the technical aspects of SENTINEL, starting with an overview of its architecture and following with more technical details of each of the components.

3.1 Semantic Business Activity Monitoring

Semantic BPM, that is, the extension of BPM with Semantic Web and Semantic Web Services technologies has been proposed as a means for increasing the automation of tasks through the provisioning of semantic descriptions of the artefacts involved in the life-cycle of business processes [20]. This vision is pursued within the SUPER project[1] which has already produced an extensive set of ontologies and tools within an overall framework that spans from methodological aspects of SBPM to the deep technical details required to orchestrate a set of Web Services which allows a business process to achieve the desired business goals. The tool presented in this paper is part of this framework and as such is strongly based on the use of semantic technologies to enhance the state of the art in BAM.

From our brief characterisation of some of the main requirements BAM solutions need to address, it can be distilled that BAM is the meeting point between Data Warehousing, Business Intelligence, Process Monitoring and Process Mining. It is therefore based on the integration and application of diverse technologies which are already challenging on their own. The quality and level of monitoring provided by state-of-the-art monitoring tools are rather similar and not surprisingly major efforts are devoted to presenting the information in a simple yet meaningful way better supporting humans in the interpretation of monitoring information [29].

Often companies invest in very expensive customised solutions that integrate domain-specific details in order to increase the level of automation. After long periods of consultancy and development, quite advanced solutions can be obtained but their benefit in mid and long term is not so clear. The business world is characterised by ever changing conditions, and one key to success is precisely the capacity to adapt and react to these changes. Customised applications typically make certain assumptions which after a while do not hold anymore. As a consequence, companies need to engage into expensive development processes in order to readapt the software. What is needed instead are general purpose solutions that can handle heterogeneity and evolution and still support advanced BAM facilities.

Despite the advances so far, there is still a long way to go to achieve the level of adaptability in process-aware systems that current businesses require. The reason for this is mainly that the semantics of the data manipulated concerning some specific business domain, are only present in the head of the business analyst and are not available for automated processing by machines [20]. Automating these tasks in a domain-independent manner requires capturing both static knowledge, like for example a company's processes and organisational structure, and procedural knowledge such as how to detect that a process will miss a deadline.

Conceptualising static knowledge in a way that can support automated reasoning by machines is well supported by means of ontologies [19]. On the other hand, research in Knowledge Engineering has shown that Problem-Solving Methods (PSM) are an appropriate way for encapsulating procedural knowledge in a reusable way [37, 38]. PSMs are reusable knowledge-based components able to support the development of highly complex systems by integrating diverse task-specific but domain-independent expertise for solving knowledge intensive tasks using ontologies as the lingua franca [37, 38]. Their genericity stems from the formalization of the relevant concepts for performing a specific task in an ontological form constructing in this way a formal interface to the task-specific expertise. This interface can then be used for applying the problem-solving expertise over domain specific data by defining mappings that bridge the gap between both conceptualizations. Additionally, there is often a conceptual separation between the task to solve and the method used which supports the application of diverse

[1] http://www.ip-super.org

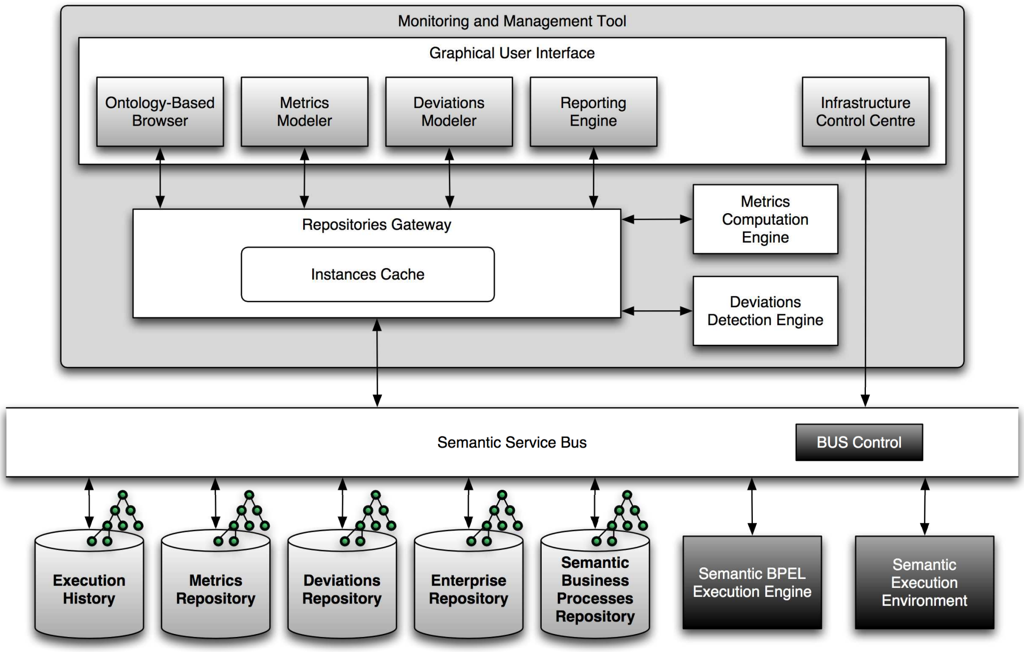

Figure 1: Monitoring and Management Tool Architecture.

techniques on a per case basis.

3.2 Architecture

Figure 1 presents the overall architecture of SENTINEL including external components that interact with the tool represented as black boxes. In particular, the Semantic BPEL Execution Engine [43] is an extension of the open source BPEL engine Apache ODE that implements the operational semantics of BPEL4SWS [30]. In addition to BPEL 2.0, it provides means to interact with semantic Web services, to perform data mediation (based on semantic annotation of the data types) and to utilise logical expressions for describing conditional flows. Semantic Web services execution is supported by implementations of the Semantic Execution Environment [31], a reference architecture which is under standardisation within OASIS. The Semantic Execution Environment architecture is derived from WSMX [13] and IRS-III [12] and uses a comprehensive conceptualisation of services stemming from WSMO [14] in order to support the discovery, mediation, choreography and orchestration of semantic Web services.

The monitoring tool connects to external components such as the execution engines through the Semantic Service Bus (SSB) an extension of Apache ServiceMix[2]. The SSB is a JBI-based Enterprise Service Bus implementation which integrates various components via normalised messages that are routed through a message-oriented middleware [21]. In addition to message routing and endpoint virtualisation, JBI employs a sophisticated deployment model that allows for deploying a complex integration scenario at once. Service engines (e.g. BPEL engine, SEE) and binding components (e.g. HTTP or JMS endpoints) are provided with lifecycle management facilities that allow for starting, stopping, removing components and/or artefacts that are deployed to them.

[2]http://servicemix.apache.org/

Although the tool is here depicted in the context of the overall SUPER architecture it is worth noting that the technologies used have either been devised as general purpose solutions, or have directly been based on standards (e.g., BPEL). The results described herein are therefore essentially reusable in more general enterprise setups. This applies for instance to environments where Enterprise Resource Planning or Customer Relationship Management systems are deployed. In these cases, the Semantic Service Bus can support populating monitoring information and this information should seamlessly be interpreted by the tool provided that it is based on the core conceptualisations used. We shall give more details about the conceptualisations and their suitability for capturing logs from a wide-range of BPM systems in Section 3.2.1.

The architecture of the monitoring tool is basically decomposed into three main parts. At the lowest level there is the support for accessing repositories capturing knowledge about the enterprise, its processes, previous process executions, or even metrics and deviations. This knowledge is used by the other two components, namely the analysis engines (i.e., Metrics Computation Engine and Deviations Detection Engine) and the advanced graphical user interface.

In order to support effectively monitoring business processes, the monitoring tool is listening to the SSB in order to capture monitoring messages populated by the execution infrastructure. Upon reception, events are processed by the tool to populate the user interface, to derive new knowledge, and to trigger additional computations thanks to the

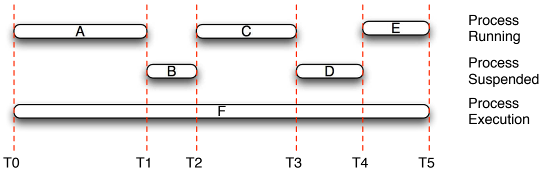

Figure 2: Process Instance life-cycle.

forward-chaining engine part of IRS-III [28, 12]. In particular, the tool includes a set of forward-chaining rules that detect certain conditions (e.g., the instantiation of a process, its completion, etc) and derive additional knowledge which stored in the repositories. For example, any time a new event is received, the status of the process instance is updated and a new interval capturing the different states of the process instance during its life-cycle is created (see Figure 2). In the remainder of this section we shall go in detail through the main parts of the monitoring tool.

3.2.1 Knowledge Bases

The main underlying characteristic of SENTINEL is the use of semantic technologies as the key pillar for achieving a domain-independent yet highly automated and advanced monitoring tool. The tool is currently supported by a set of repositories as illustrated in Figure 1. The Execution History repository captures audit trails and additional information for supporting business process analysis. The Enterprise Repository captures domain-specific organisational information (e.g., departments, resources). The Semantic Business Processes Repository captures process definitions, whereas the remain two capture metric and deviations definitions as well as their results.

These repositories are expressed in terms of a set of domain-independent yet BPM-specific ontologies. Ontologies provide the means for describing the concepts of a domain of interest and the relationships between them in a way that amenable to automated reasoning. Within SUPER, major efforts have been devoted to producing a comprehensive set of ontologies about the BPM domain. The ontologies defined so far range from ontologies for supporting the definition of organisational structures [2], business processes [1] or even business goals to ontologies for capturing execution logs and supporting business process analysis [33]. These ontologies define the *lingua franca* that concrete organisations can use in order to, on the one hand, simplify the modelling of their enterprise and, on the other hand, open up the way for applying the whole generic toolset over their concrete domain.

Although describing all the ontologies is outside of the scope of this paper, it is worth however describing those aimed at supporting Business Process Analysis. Within the set of ontologies, Core Ontology of Business pRocess Analysis (COBRA) [33], which is the metamodel used by the Execution History, represents the main conceptualisation SENTINEL is based upon. COBRA, depicted in Figure 3, provides a pluggable framework based on the core conceptualisations required for supporting BPA and defines the appropriate hooks for further extensions in order to cope with the wide-range of aspects involved in analysing business processes.

COBRA builds upon Time Ontology that provides a temporal reference by means of which one can determine temporal relations between elements. COBRA provides a lightweight foundational basis in order to ensure a coherent view among additional extensions. It defines concepts such as *Process*, *Activity*, *Process Instance*, *Role*, *Resource*, *Organisation* or *Person* which are to be refined within specific ontologies as defined within SUPER, or other approaches like the Enterprise Ontology [40] or TOVE [15].

COBRA has been extended with a reference Events Ontology (EVO) [33] that provides a set of definitions suitable to capture monitoring logs from a large variety of systems and ready to be integrated within our core ontology for analysing business processes. EVO is the monitoring format employed by the SUPER execution infrastructure and therefore that of the Execution History. Still, it is based on existing syntactic formats, e.g., MXML [42] or the Audit Trail Format by the Workflow Management Coalition [29] which therefore confers on it the ability to capture logs generated by a plethora of systems. As prescribed by COBRA, EVO is centred around a state model that accounts for the status of processes and activities, see Figure 4. The state model has been captured ontologically and enhanced with additional relations. For instance it is possible to determine whether an Activity Instance has been allocated–*isAllocated*–which is true for those that are either in state Running, Suspended, or Assigned. It is also possible to determine whether a Process is active–*isActive*–which is equivalent to Running, or inactive–*isInactive*–which is true for the rest of the states, etc.

Concerning the analysis themselves such as metrics, the previously presented version of COBRA [33] solely captures the concept *Analysis Result*, a Temporal Entity, which has two disjoint sub-concepts: *Qualitative Analysis Result* and *Quantitative Analysis Result*. As part of our work on metrics definition and computation, we have slightly extended COBRA itself. First, COBRA includes support for units of measure and their manipulation. Secondly, we have introduced the concept *Analysis*, which is refined into *Qualitative Analysis* (e.g., "is the process critical?") and *Quantitative Analysis* (e.g., "process execution time") based on the type of Analysis Result they produce. This provides us the means for maintaining a library of Analysis specifications (e.g., metrics, time series, etc.), and it allows us to distinguish between the analysis themselves and the actual results. Indeed, the relationship between Analysis and Analysis Result has also been captured, in such a way that every Analysis Result is a result for a particular Analysis, and every Analysis may have several Analysis Results. Hence we can obtain all the results for a particular analysis, track its evolution over time, apply time series analysis, etc.

As part of this extension, COBRA now also includes *Query*, the Qualitative Analysis *par excellence*. The concept Query supports capturing ontological queries in a similar vein to stored procedures in Databases. Capturing Queries ontologically provides the means for maintaining a library of typically useful queries parametrised in a way such that business analysts can directly reuse and apply them over their specific domain. This addition is therefore of particular relevance for better supporting business practitioners–typically

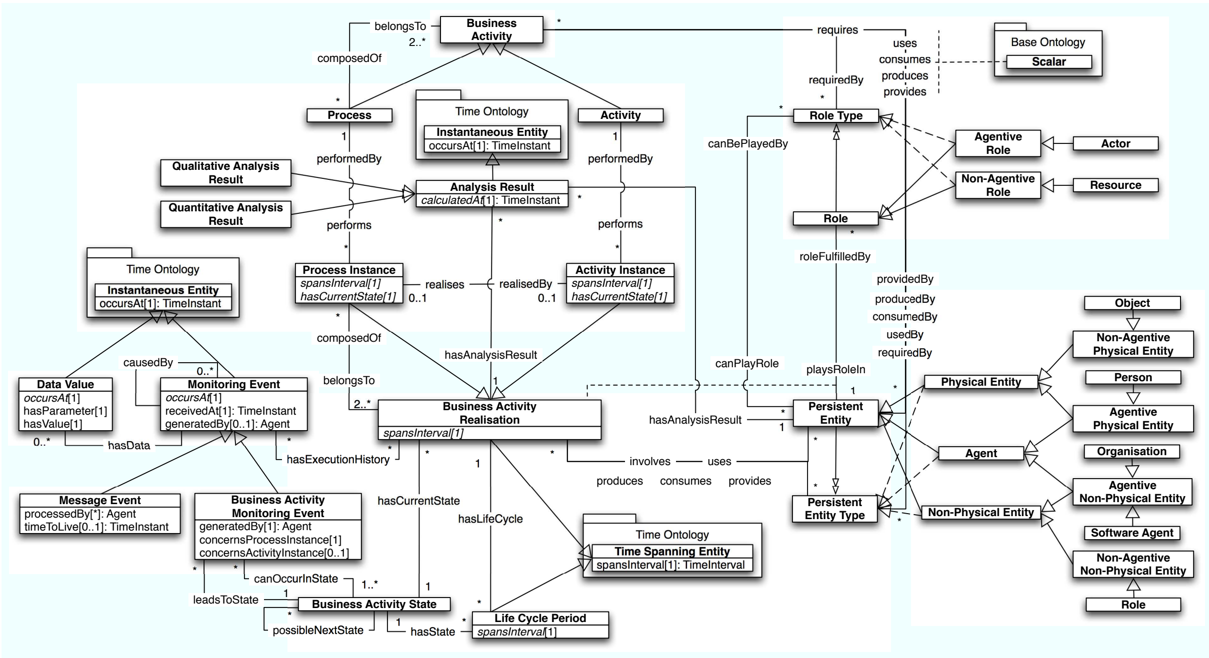

Figure 3: Core Ontology for Business pRocess Analysis.

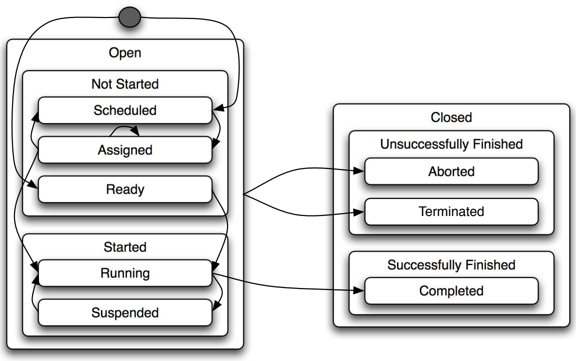

Figure 4: Events Ontology state model.

not highly-skilled technically–in analysing processes.

The ontologies briefly described herein represent the core conceptualisations underlying SENTINEL. They support the tool in capturing monitoring logs from a plethora of systems and bringing them to the business level where they can appropriately support BPA through the use of additional enterprise modelling conceptualisations. The whole body of knowledge about enterprises and their processes executions is captured within the Enterprise Repository, see Figure 1. It is accessed by the monitoring tool on demand through a repositories gateway that distributes the queries as appropriate and also caches information for performance reasons. The caching mechanism that has been developed includes specific caches for certain concepts based on the variability of the instances. For example, monitoring events once stored will not be modified anymore and FIFO is a suitable cache eviction algorithm. On the other hand, information about Process Instances will be highly dynamic and therefore little can be gained from caching mechanisms.

3.2.2 Graphical User Interface

SENTINEL is build upon WSMO Studio [10], an open source Eclipse-based semantic Web services modelling environment. WSMO Studio provides support for modelling and browsing WSMO entities such as ontologies, goals and WSML logical expressions (for pre-conditions, post-conditions, assumptions and effects). An important feature of Eclipse is that its component model presents a declarative specification of ways to extend an application, called extension points [18]. New plug-ins may extend existing plug-ins and may be easily incorporated into an application. Among the plug-ins developed so far the support for accessing ontology repositories and for modelling processes based on the set of ontologies that have been produced within SUPER so far [11] are of particular relevance.

WSMO Studio provides an advanced modelling environment able to model and browse existing ontologies and knowledge bases, as well as to graphically represent and edit semantic business processes. It therefore constitutes a particularly suitable framework for supporting the monitoring tool. SENTINEL has been developed as a plug-in for WSMO Studio in order to benefit form all the previously mentioned features, including those directly provided by Eclipse. The tool currently contemplates three views, namely the Execution History view, the Statistics view, and the Control Centre view, in addition to those provided by WSMO Studio itself.

The Execution History view (see Figure 5) provides real-time visualisation of monitoring events as they are populated by the execution infrastructure. Events, as previously introduced, are directly expressed in terms of EVO, and

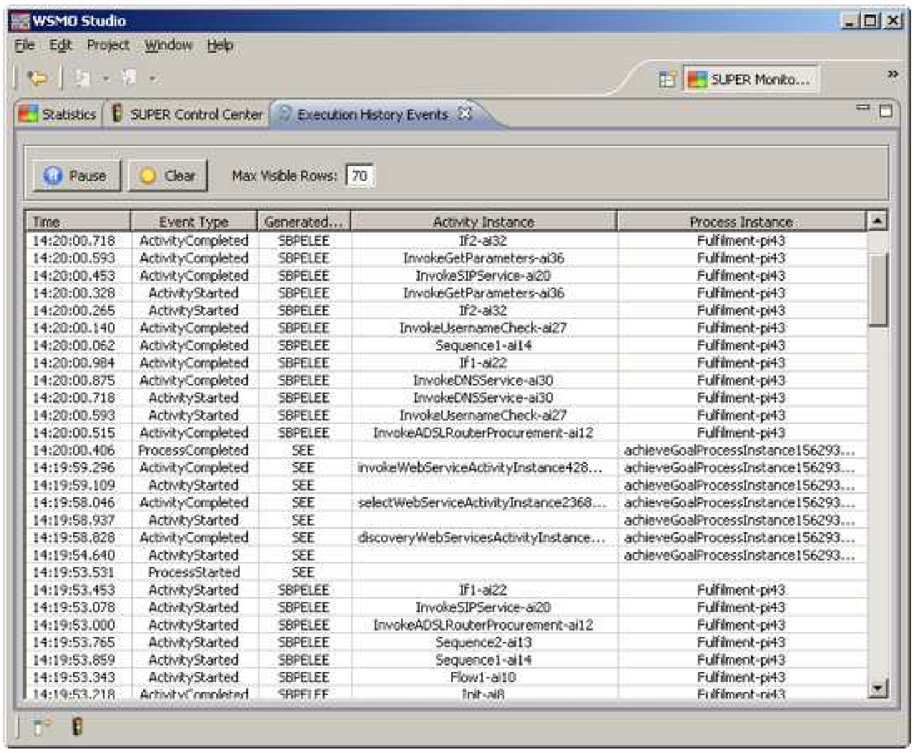

Figure 5: Execution History view.

as a consequence are directly amenable to reasoning, and are tightly integrated with the rest of the ontologies defined within SUPER. This allows us for instance to display, query and filter the information in generic terms and it supports, if the user requires so, to access additional information captured within the different repositories. For instance, if an event refers to a concrete Process Instance we can access the Process Instance definition and visualise graphically its definition using a BPMN representation. Similarly, information about the Agents (e.g., people, systems) involved and the data exchanged at runtime can be represented.

The Statistics view presents a summary of relevant statistical information concerning processes and their executions. This view is populated with metrics and so-called Key Performance Indicators concerning business activities and their execution. Additionally the tool has support for including a chart-based representation of these results in order to easily convey an overall view on the metrics computed. The current version of the tool includes a predefined set of general purpose metrics (see Section) and the corresponding prefixed report. The information currently presented are the average, minimum and maximum execution time of a given process in seconds, the number of process instances of that process (running or not), and the number of successfully and unsuccessfully finished process instances. There is however work underway towards providing complete flexibility with respect to the metrics to be computed, including the definition of domain-specific ones, and how to present them. Further details will be presented in Section 3.2.3

Besides the monitoring functionality, complex enterprise integration scenarios demand also on support for managing the whole infrastructure. These management functionality is exposed via the Java Management Extensions (JMX) [23]. JMX employs so called MBeans that define the interface to the supported management functionality. This comprises reading and writing of attributes, invoking management functionality as well as notification mechanisms, i.e. the observed middleware can notify attached management tools when certain conditions arrive or attributes change. The Control Centre utilizes this functionality to enable system administrators to manage the SUPER infrastructure as a whole. As all SUPER components involved in process/service execution are rendered as JBI components, they are automatically registered in the SSB's JMX MBean server. Once connected to the SSB, the Control Centre allows for listing all SUPER components and artefacts deployed to them and provides access to their lifecycle management, i.e. administrators have the ability to stop certain components or artefacts.

3.2.3 *Analysis Engines*

We previously introduced some of the main requirements that BAM solutions need to address from an analysis perspective. The current version of SENTINEL contemplates two modules in this respect, see Figure 1 . On the one hand, Metrics Computation Engine is in charge of supporting the automated computation of general purpose as well as user-defined metrics. On the other hand, Deviations Detection Engine aims to support the automated detection of process deviations. Although their level of maturity is quite heterogeneous, important efforts have been devoted to devising their underlying structure.

Our approach is based on previous research on Problem-Solving Methods [37, 28]. In particular, we build upon the Task Method Domain Application (TMDA) framework [28] for Knowledge-Based Systems reuse and development. In a nutshell, TMDA prescribes constructing Knowledge-Based Systems based on the definition of task ontologies that define classes of applications (e.g., diagnosis, classification), method ontologies that capture the knowledge requirements for specific methods (e.g., heuristic classification), domain ontologies which provide reusable task-independent models, and application ontologies for integrating domain models with domain-independent problem solvers.

TMDA has been applied to the construction of systems that tackle diverse knowledge-intensive tasks. This includes parametric design, planning, and classification. It has therefore proven its genericity and versatility which makes it a particularly good candidate for structuring our development. Despite the relative simplicity of the metrics computation endeavour with respect to more complex knowledge-intensive tasks as typically tackled within Problem-Solving Methods research [28, 37], this approach gives us the appropriate genericity and support for interchanging methods as the need arises. Furthermore, this represents a step towards the creation of library of tasks for Business Process Analysis and their corresponding methods which we plan to base on previous research in Problem-Solving Methods such as diagnosis, classification and configuration design [3].

Metric computation is defined within the TMDA framework as a kind of task that takes a Metric definition as input and returns a Quantitative Analysis Result with the actual value for the Metric at that particular point in time, see Figure 6. A key aspect with respect to metrics computation concerns the support included for defining the metrics themselves. Metrics Ontology provides us with the capacity for specifying and computing metrics, as necessary for analysing and managing business processes, in a domain-independent way.

On the basis of our conceptualisation we can capture kinds of metrics, e.g,. "process instance execution time", as well as specific metrics to be computed, e.g., "process instance X execution time". The former are defined as concepts, whereas the latter are modelled as instances. In this way we can provide libraries of metrics such as general purpose ones, or specific for some domain like Supply-Chain, and at analysis time the analyst can specify which of these metrics should be computed over which entities by instantiating them. This provides a convenient way for organising metric definitions and seamlessly supports the comparison of results by kind of metric, e.g., "which is the process which takes longer", as well as it allows tracking their evolution over time.

Central to Metrics Ontology is the concept *Metric* which is defined as a Quantitative Analysis (see Section 3.2.1). Metrics are specified by a set of input roles that point to domain-specific knowledge [28]. We refine Metrics into two disjoint kinds, Function Metrics and Aggregation Metrics. A Function Metric is a metric that can be evaluated over a fixed number of inputs. For example, the Metric *Process Instance Execution Time* is a Function Metric which takes as input one Process Instance. Conversely, Aggregation Metrics (e.g., "average process execution time") take an arbitrary number of individuals of the same kind (e.g., a set of Process Instances) as input. Therefore, Aggregation Metrics are computed over a population in order to obtain an overall perception of some aspect of interest such as the average execution time of some particular process. The population to be processed can be defined intensionally as an (domain-specific) ontological query so that the metric computation can focus on certain processes, or resources of interest. In this respect the use of semantic technologies play a key role towards supporting business analysts in the analysis of processes, allowing them to use their domain-specific terminology and still use a generic machinery to process the information in a seamless way.

In order to support the automated computation of metrics, which is indeed metric dependent, each metric has a computation expression which is defined as a unary procedure. In this respect it is worth noting that the language used to define the metrics themselves as well as to develop Metrics Computation Engine is Operational Conceptual Modelling Language (OCML) [28]. OCML seamlessly supports the integration of static and dynamic knowledge paving the way for a rapid prototyping of a fully operational solution[3]. It is worth noting however that OCML provides support for importing and exporting data represented in other languages such as OWL and WSML–the language used within SUPER–and therefore allows the wider application of our techniques over data represented in Semantic Web and semantic Web services formalisms.

Work on Deviations Detection Engine is less mature than that for metrics computation. The engine will be based on the application of PSMs ideas in order to provide a generically applicable engine as well as a set of interchangeable methods that can be selected an applied at runtime depending on their suitability [3]. So far however, the only method implemented is a threshold-based one whereby a business activity execution is considered to be deviating if its execution time is beyond that of the average plus the standard deviation of previous executions. Future efforts will be devoted to the development of more complex methods reusing previous research in PSMs, such as Classification problem-solving and Diagnosis.

4. RELATED WORK

In [29] the author provides a comprehensive analysis on current techniques in process monitoring and control within organisations. The author covers the topic by addressing four relevant perspectives to process controlling: the data perspective, the usage perspective, the tool perspective and the method perspective. The data perspective is concerned with collecting, storing and representing audit trail information. The author describes existing techniques and proposes an audit trail format that we took into account while devising our conceptualisations (i.e., COBRA and EVO). Our approach, although similar in many respects, provides a more formal conceptualisation that is amenable to automated reasoning. This conceptualisation, which underlies SENTINEL, provides us with the capacity to apply advanced knowledge-based techniques in a domain-independent

[3]The ontologies described herein can be found at http://www.cpedrinaci.net

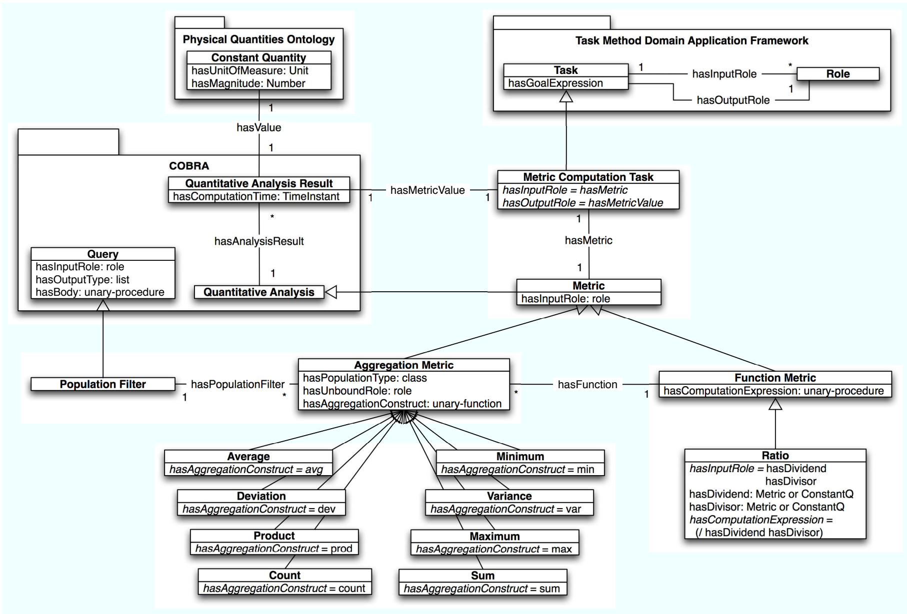

Figure 6: Metrics Ontology.

manner, as opposed to current practices.

The usage perspective concerns how process controlling management is approached. Some of the state of the art solutions focus on exception handling [7], whereas others focus on the global management of processes [25]. Our work so far has focussed mostly on gathering knowledge about processes in a way that can support automated reasoning as well as it can help business analysts in the management of processes. Although, automated process control is so far not supported by the tool, the very goal of our extensive conceptualisation work has precisely been carried out in order to better support machines in automatically controlling processes [3]. The work carried out so far represent substantial steps in this direction.

The tool perspective is concerned with the architecture and tools that have been developed for process monitoring and control so far. Among these the most relevant to us are for instance the work on PISA (Process Information System based on Access) which precedes the work presented in [29], and the work on the Business Process Intelligence tool suite [35, 17, 8]. While the former does not make any use of semantic technologies to enhance the monitoring capabilities the latter uses lightweight taxonomies. Their work is however more focussed on the integration of mining techniques and support for explanations. In this respect we believe both approaches are complementary and consider that a more extensive use of knowledge-based techniques could indeed enhance their results. In a similar way, future work on SENTINEL will indeed be inspired by this research.

On a broader sense, BAM functionality as part of the BPM system is already supported in several products, e.g., [44] and [32]. Typically the BAM solution is tightly integrated with the BPEL engine which is part of the BPM system. Metrics are defined and calculated based on the events which are published by that BPEL engine. The event publishing mechanism is thereby proprietary. Another approach is to extend the BPEL process with event publishing activities, which invoke operations on a monitoring tool. This approach is utilized in [34] and [6]. The benefit of this approach is that event publishing does not depend on a proprietary mechanism of a BPEL engine. The main disadvantage however is that the BPEL process is more difficult to read and maintain, as it contains new activities, which deal with technical issues, and not just business logic. Our work relies on event publishing mechanism by the execution engines (e.g., BPEL engine). The difference to the existing tools is that in our approach ontologies are used for the description of events and the data they contain. In this way we better support the integration of proprietary formats via ontological mappings, and still allow inferring implicit knowledge at runtime.

5. CONCLUSIONS

BPM systems aim to support the whole life-cycle of business processes. However, BPM has made more evident the current lack of automation that would support a smooth transition between the business world and the IT world. Moving back and forth between these two aspects is a bottleneck that reduces the capability of enterprise to adapt to ever changing business scenarios. As a consequence there is a growing need for integrating semantics within the BPM domain. A crucial branch of BPM where semantics have a clear and direct impact is Business Process Analysis, where Business Intelligence and Business Activity Monitoring are appearing as a key enablers for increasing value and performance [45].

In this paper we have described SENTINEL, a semantic business process monitoring tool that aims to support the integration and derivation of business level knowledge out of low-level audit trails generated by IT systems. The fundamental specificity of our tool lies on the extensive use of semantic technologies, namely ontologies and PSMs. Ontologies define the *lingua franca* employed internally by the tool in a formal way so that tasks like integrating heterogeneous sources of information, using domain-specific terminologies, querying or filtering can better be supported. This is achieved thanks to, on the one hand, the possibility to apply automated reasoning, and on the other the use of refined conceptualisations that are closer to human understanding.

Despite its relative lack of maturity, SENTINEL has already proven some of the benefits of using semantics within BAM. In particular, it has shown the benefits that bringing low-level syntactic audit trails to the semantic level can bring to the analysis of process executions. However, there are aspects that need to be improved and indeed many others where we could further exploit the set of comprehensive conceptualisations defined so far. Among the aspects we plan to improve next we contemplate two main areas: the user interface and the analysis engines. With respect to improving the user interface, we want to include an ontology-based facility for better supporting users in the definition of ontological queries, metrics and deviations. From an analysis perspective we plan to enhance the computation of metrics with support for their goal-oriented computation, and we plan to utilize a Heuristic Classification PSM for supporting the detection of process deviations.

Acknowledgments

This work was funded by the European project SUPER (FP6-026850).

6. REFERENCES

[1] W. Abramowicz, A. Filipowska, M. Kaczmarek, and T. Kaczmarek. Semantically enhanced business process modelling notation. In *Proceedings of SBPM 2007 Semantic Business Process and Product Lifecycle Management in conjunction with the 3rd European Semantic Web Conference (ESWC 2007)*, Innsbruck, Austria, June 2007.

[2] W. Abramowicz, A. Filipowska, M. Kaczmarek, C. Pedrinaci, M. Starzecka, and A. Walczak. Organization structure description for the needs of semantic business process management. In *3rd international Workshop on Semantic Business Process Management colocated with 5th European Semantic Web Conference*, 2008.

[3] A. K. Alves de Medeiros, C. Pedrinaci, W. van der Aalst, J. Domingue, M. Song, A. Rozinat, B. Norton, and L. Cabral. An Outlook on Semantic Business Process Mining and Monitoring. In *Proceedings of International IFIP Workshop On Semantic Web & Web Semantics (SWWS 2007)*, 2007.

[4] A. K. Alves de Medeiros, W. Van der Aalst, and C. Pedrinaci. Semantic Process Mining Tools: Core Building Blocks. In *16th European Conference on Information Systems*, Galway, Ireland, June 2008.

[5] B. Azvine, Z. Cui, D. D. Nauck, and B. Majeed. Real Time Business Intelligence for the Adaptive Enterprise. In *The 8th IEEE International Conference on E-Commerce Technology*, Los Alamitos, CA, USA, 2006. IEEE Computer Society.

[6] L. Baresi and S. Guinea. Towards Dynamic Monitoring of WS-BPEL Processes. In *Service-Oriented Computing - ICSOC 2005*, LNCS, pages 269–282. Springer, 2005.

[7] F. Casati, S. Castano, M. Fugini, I. Mirbel, and B. Pernici. Using patterns to design rules in workflows. *IEEE Trans. Softw. Eng.*, 26(8):760–785, 2000.

[8] M. Castellanos, F. Casati, U. Dayal, and M.-C. Shan. A comprehensive and automated approach to intelligent business processes execution analysis. *Distributed and Parallel Databases*, 16(3):239–273, 2004.

[9] S. Chaudhuri, U. Dayal, and V. Ganti. Database Technology for Decision Support Systems. *Computer*, 34(12):48–55, December 2001.

[10] M. Dimitrov, A. Simov, M. Konstantinov, and V. Momtchev. WSMO Studio - a Semantic Web Services Modelling Environment for WSMO (System Description). In E. Franconi, M. Kifer, and W. May, editors, *Proceedings of the 4th European Semantic Web Conference (ESWC)*, number 4519 in LNCS, pages 749–758, Innsbruck, Austria, June 2007.

[11] M. Dimitrov, A. Simov, S. Stein, and M. Konstantinov. A BPMO Based Semantic Business Process Modelling Environment. In M. Hepp, K. Hinkelmann, D. Karagiannis, R. Klein, and N. Stojanovic, editors, *SBPM*, volume 251 of *CEUR Workshop Proceedings*, Innsbruck, Austria, June 2007. CEUR-WS.org.

[12] J. Domingue, L. Cabral, S. Galizia, V. Tanasescu, A. Gugliotta, B. Norton, and C. Pedrinaci. IRS-III: A broker-based approach to semantic Web services. *Web Semantics: Science, Services and Agents on the World Wide Web*, 6(2):109–132, 2008.

[13] D. Fensel, M. Kerrigan, and M. Zaremba, editors. *Implementing Semantic Web Services: The SESA Framework*. Springer, 2008.

[14] D. Fensel, H. Lausen, A. Polleres, J. de Bruijn, M. Stollberg, D. Roman, and J. Domingue. *Enabling Semantic Web Services: The Web Service Modeling Ontology*. Springer, 2007.

[15] M. S. Fox. The TOVE Project Towards a Common-Sense Model of the Enterprise. In *IEA/AIE '92: Proceedings of the 5th international conference on Industrial and engineering applications of artificial intelligence and expert systems*, pages 25–34, London, UK, 1992. Springer-Verlag.

[16] J. Gordijn and H. Akkermans. Designing and evaluating e-business models. *IEEE Intelligent Systems*, 16(4):11–17, 2001.

[17] D. Grigori, F. Casati, M. Castellanos, U. Dayal, M. Sayal, and M.-C. Shan. Business Process Intelligence. *Computers in Industry*, 53(3):321–343, Apr. 2004.

[18] O. Gruber, B. J. Hargrave, J. McAffer, P. Rapicault, and T. Watson. The Eclipse 3.0 platform: Adopting OSGi technology. *IBM Systems Journal*, 44(2):289–299, 2005.

[19] T. R. Gruber. A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5(2):199–220, 1993.

[20] M. Hepp, F. Leymann, J. Domingue, A. Wahler, and D. Fensel. Semantic Business Process Management: A Vision Towards Using Semantic Web Services for Business Process Management. In F. C. M. Lau,

H. Lei, X. Meng, and M. Wang, editors, *ICEBE*, pages 535–540. IEEE Computer Society, 2005.
[21] G. Hohpe and B. Woolf. *Enterprise Integration Patterns: Designing, Building, and Deploying Messaging Solutions*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2003.
[22] W. Hur, H. Bae, and S.-H. Kang. Customizable Workflow Monitoring. *Concurrent Engineering*, 11(4):313–325, 2003.
[23] JSR 3 Expert Group. Java management extensions (JMX). JSR 3, Java Community Process, 2000. Available at http://www.jcp.org/en/jsr/detail?id=3.
[24] R. S. Kaplan and D. P. Norton. The Balanced Scorecard - Measures that Drive Performance. *Harvard Business Review*, January/February 1992.
[25] P. Kueng, P. Kueng, A. Meier, A. Meier, and T. Wettstein. Computer-based Performance Measurement in SME's: Is there any option. In *Proceedings of the International Conference on Systems Thinking in Management*, pages 8–10, 2000.
[26] S. T. March and A. R. Hevner. Integrated decision support systems: A data warehousing perspective. *Decision Support Systems*, 43(3):1031–1043, 2007.
[27] D. W. McCoy. Business activity monitoring: Calm before the storm. Technical Report LE-15-9727, Gartner, April 2002.
[28] E. Motta. *Reusable Components for Knowledge Modelling. Case Studies in Parametric Design Problem Solving*, volume 53 of *Frontiers in Artificial Intelligence and Applications*. IOS Press, 1999.
[29] M. Mühlen. *Workflow-based Process Controlling. Foundation, Design, and Implementation of Workflow-driven Process Information Systems.*, volume 6 of *Advances in Information Systems and Management Science*. Logos, Berlin, 2004.
[30] J. Nitzsche, T. van Lessen, D. Karastoyanova, and F. Leymann. BPEL for Semantic Web Services (BPEL4SWS). *On the Move to Meaningful Internet Systems 2007: OTM 2007 Workshops*, pages 179–188, 2007.
[31] B. Norton, C. Pedrinaci, J. Domingue, and M. Zaremba. Semantic Execution Environments for Semantics-Enabled SOA. *it - Methods and Applications of Informatics and Information Technology*, Special Issue in Service-Oriented Architectures:118–121, 2008.
[32] Oracle Corporation. Oracle Business Activity Monitoring (BAM). Available at http://www.oracle.com/appserver/business-activity-monitoring.html, 2008.
[33] C. Pedrinaci, J. Domingue, and A. K. Alves de Medeiros. A Core Ontology for Business Process Analysis. In *5th European Semantic Web Conference*, 2008.
[34] H. Roth, J. Schiefer, and A. Schatten. Probing and Monitoring of WSBPEL Processes with Web Services. In *CEC-EEE '06: Proceedings of the The 8th IEEE International Conference on E-Commerce Technology and The 3rd IEEE International Conference on Enterprise Computing, E-Commerce, and E-Services*, page 30, Washington, DC, USA, 2006. IEEE Computer Society.
[35] M. Sayal, F. Casati, U. Dayal, and M.-C. Shan. Business process cockpit. In *VLDB '02: Proceedings of the 28th international conference on Very Large Data Bases*, pages 880–883. VLDB Endowment, 2002.
[36] K. Schlegel, B. Hostmann, and A. Bitterer. Magic Quadrant for Business Intelligence Platforms, 1Q07. Gartner RAS Core Research Note G00145507, Gartner, January 2007.
[37] G. Schreiber, H. Akkermans, A. Anjewierden, R. de Hoog, N. Shadbolt, W. V. de Velde, and B. Wielinga. *Knowledge Engineering and Management: The CommonKADS Methodology*. MIT Press, 1999.
[38] R. Studer, R. Benjamins, and D. Fensel. Knowledge Engineering: Principles and Methods. *Data Knowledge Engineering*, 25(1-2):161–197, 1998.
[39] Supply-Chain Council. Supply-Chain Operations Reference-model. http://www.supply-chain.org, 2008.
[40] M. Uschold, M. King, S. Moralee, and Y. Zorgios. The Enterprise Ontology. *Knowledge Engineering Review*, 13(1):31–89, 1998.
[41] W. M. P. van der Aalst, A. H. M. ter Hofstede, and M. Weske. Business Process Management: A Survey. In W. M. P. van der Aalst, A. H. M. ter Hofstede, and M. Weske, editors, *Business Process Management*, volume 2678 of *LNCS*, pages 1–12. Springer, 2003.
[42] B. van Dongen, A. K. Alves de Medeiros, H. Verbeek, A. Weijters, and W. van der Aalst. The ProM Framework: A New Era in Process Mining Tool Support. In *Applications and Theory of Petri Nets 2005. 26th International Conference, ICATPN 2005*, pages 444–454, Miami, USA, June 2005. Springer-Verlag.
[43] T. van Lessen, J. Nitzsche, M. Dimitrov, D. Karastoyanova, M. Konstantinov, and L. Cekov. An Execution engine for BPEL4SWS. In *2nd Workshop on Business Oriented Aspects concerning Semantics and Methodologies in Service-oriented Computing (SeMSoc) in conjunction with ICSOC*, Sept. 2007. Vienna, Austria.
[44] U. Wahli, V. Avula, H. Macleod, M. Saeed, and A. Vinther. *Business Process Management: Modeling through Monitoring Using WebSphere V6.0.2 Products*. Number SG24714801. IBM RedBooks, 2007.
[45] H. J. Watson and B. H. Wixom. The Current State of Business Intelligence. *Computer*, 40(9):96–99, 2007.