Workflow Modeling Technologies

Maria N. Koukovini
National Technical University of Athens, Greece

Eugenia I. Papagiannakopoulou
National Technical University of Athens, Greece

Georgios V. Lioudakis
National Technical University of Athens, Greece

Nikolaos L. Dellas
SingularLogic S.A., Greece

Dimitra I. Kaklamani
National Technical University of Athens, Greece

Iakovos S. Venieris
National Technical University of Athens, Greece

INTRODUCTION

Workflows, that is, well-defined sequences of tasks coordinated in order to achieve a variety of business, scientific and engineering goals, have emerged as a prominent technology in current distributed and dynamic environments, fuelled to a large extent by the development of Service Oriented Architectures (SOA) and their loose-coupling nature. Emanating from the first office automation systems, workflows originally had a purely business orientation and have in the meantime evolved to what is being referred to today as business workflow or, more broadly speaking, Business Process Management (BPM) technology. Indeed, the Workflow Management Coalition (WfMC) defined a workflow as:

The "computerized facilitation or automation of a business process, in whole or part" during which "documents, information or tasks are passed between participants according to a defined set of rules to achieve, or contribute to, an overall business goal" (WfMC, 1995).

However, it later became apparent that the workflow paradigm could also benefit the sciences domain and their complex and data-intensive operations. This has led to the emergence of a new family of workflows referred to as scientific; in general terms, a scientific workflow is:

A formal description of a process for accomplishing a scientific objective, usually expressed in terms of tasks and their dependencies (Ludäscher et al., 2009).

Business and scientific workflows present similarities but also differences, mainly stemming from the purposes they serve and their historical context. In any case, a fundamental common characteristic is that all workflows rely on models, providing their conceptual representation, or, in other words, the "blueprint" from which the eventually executed workflow instances are derived.

In this context, the goal of this article is to comprehensively present the most influential technologies in the area of both business and scientific workflow modeling. The next section provides a brief overview

DOI: 10.4018/978-1-4666-5888-2.ch528

on their evolution and associated key concepts, followed by three sections constituting an overview of the currently most prominent approaches. Before concluding, the article highlights main future research directions.

BACKGROUND

Most researchers agree that workflow management and modeling technologies have their roots in the office automation systems that emerged in the 1970s (Ellis & Nutt, 1980). At those times, variants of Petri Nets (Petri, 1962) have been used in order to model related procedures. However, it took two more decades before they came to the spotlight.

In 1993, the Workflow Management Coalition (WfMC) was founded, and two years later the Reference Model (WfMC, 1995) was published, describing the major components and functions involved in a workflow's lifecycle. It was followed, three years later, by the first specification of Workflow Process Definition Language (WPDL), which evolved to the contemporary XML Process Definition Language (XPDL) (WfMC, 2012). In the meantime, various languages emerged, often focusing on different aspects. Among them, the Business Process Model and Notation (BPMN) (OMG, 2011) is the de facto modeling standard, whereas Yet Another Workflow Language (YAWL) (van der Aalst & ter Hofstede, 2005) is a particularly noteworthy approach coming from academia.

At this point, the use of the term workflow should be clarified, especially against the term business process. According to the WfMC, a business process is related to any kind of activity, manual or automated, that realizes a business objective. A workflow, on the other hand, is an (partial) automation of a business process. Following this distinction, and as explained in (van der Aalst, ter Hofstede & Weske, 2003), Workflow Management (WFM) focuses on creation and enactment of operational processes, whereas the more recent term of Business Process Management (BPM) constitutes a superset of the traditional WFM approach, extending it by support for other important aspects, as are Business Process Analysis (BPA), but also new ways to support operational processes.

While business workflows are clearly process-oriented, scientific workflows emphasize on data; in fact, scientific research is driven by analysis of large data sets and the application of computational methods. Therefore, scientific workflows have been motivated by the need to automatize the execution of computation-intensive operations over enormous data. In this context, while their influence by business workflows is not negligible, their roots are primarily found in database research (e.g., Medeiros, Vossen & Weske, 1995) and problem-solving environments, whereas the technology underpinning their uptake has been the computational grid (Yu, & Buyya, 2005).

A significant milestone in the discipline of workflow management has been the establishment of the Workflow Patterns Initiative in 1999, aiming at the identification of the generic, recurrent constructs (patterns) that are often met in workflows and business processes, and their description in a language-independent manner. The seminal outcome of the initiative has been the codification of the most important control-flow patterns (van der Aalst et al., 2003), describing fundamental modeling features for managing the flow of control among tasks, such as branching, synchronization, concurrency and termination. This initial set of 20 control-flow patterns has been later revised (Russell et al., 2006) in order to include more sophisticated constructs. Further, data patterns (Russell et al., 2005b) deal with the definition and management of data in terms of visibility, interaction, transfer and data-based routing, whereas resource patterns (Russell et al., 2005a) are concerned with the distribution of work to the available resources. Despite some criticism they have been subject to (Börger, 2012), the current set of 126 patterns comprises a widely used and well-accepted evaluation framework for workflow systems. Essentially, the different categories of workflow patterns reflect the so-called workflow perspectives (Jablonski & Bussler, 1996), the most important being the control perspective, reflecting tasks ordering and control dependencies among tasks, the data perspective, dealing with data tasks produce and require, and the resource perspective, concerned with the allocation of tasks to resources.

In light of the above, forthcoming sections elaborate in particular on BPMN and YAWL, that, as aforementioned, constitute reference business workflow modeling approaches, while scientific workflows are separately discussed, mainly on the basis of the Kepler system (Ludäscher et al., 2006).

Related Content