

# RDF Containers – A Framework for the Integration of Distributed and Heterogeneous Applications

Richard Mordinyi<sup>1,2</sup>, Thomas Moser<sup>2</sup>, Martin Murth<sup>1</sup>, Eva Kühn<sup>1</sup>, and Stefan Biffl<sup>2</sup>

<sup>1</sup> Space-Based Computing Group, Vienna University of Technology, Austria

<sup>2</sup> Christian Doppler Laboratory “Software Engineering Integration for Flexible Automation Systems”, Vienna University of Technology, Austria  
`{firstname.lastname}@tuwien.ac.at`

**Abstract.** Current trends like globalization and virtual enterprises result in an increasing need for on-the-fly integration of distributed applications grown over the past decades and originally not intended for integration and cooperation. In order to enable the collaboration of such applications, aspects of distribution, such as heterogeneous data sources, heterogeneous network technologies and coordination requirements, have to be addressed in the integration process. In this position paper, we introduce a framework, the so-called RDF Containers, for technical and semantic integration of distributed and heterogeneous data sources considering integration requirements. We discuss the benefits and limitations of the proposed framework based on a real-world use case from the e-health domain. The major benefit is that both semantic and technical integration are supported by a single framework, while complexity aspects related to the integration process do not affect the integrated applications.

**Keywords:** semantic integration, space-based computing.

## 1 Introduction

Today’s communication network technologies offer enough bandwidth and interconnection facilities for the development of applications based on combining well-established distributed applications, resulting in the increasing need for efficient and effective system integration methodologies [1]. Core questions are on the one hand side how to satisfy communication and coordination requirements between technologically different application platforms probably belonging to different autonomously acting enterprises. On the other hand side how to integrate data models across platform and domain boundaries, and finally how software engineers can be supported in designing and implementing distributed applications.

Hence, typical system integration challenges result from both technical heterogeneities, e.g., tools from different sources may use a range of technologies that become expensive and error-prone to integrate in traditional point-to-point ways; as well as from semantic heterogeneities, e.g., project participants may use different terms for common concepts in the application domain [2].

Modern technical integration approaches, such as the Enterprise Service Bus (ESB) concept [1], rely on message-based infrastructures and are capable of abstracting

complexity issues of distributed systems from the application. However, coordination logic (e.g., message ordering or coordination patterns) still remains in the application, hindering the application designers to focus entirely on the application itself [3]. Current alternative solutions for semantic integration like standards for data models [4], data-driven tool integration [5], or complete data transformation [6] work in principle, but pose their own challenges, such as inefficient and complex data access and query definitions, solutions which are not robust enough, or take considerable effort to develop and modify.

In this position paper, we introduce the RDF Containers approach as a platform for technical and semantic integration of distributed and heterogeneous data sources. RDF Containers combine the space based computing paradigm (SBC) for solving technical integration and coordination challenges [7], and the Engineering Knowledge Base (EKB) framework for supporting semantic integration [8]. SBC is a coordination middleware based on concepts of virtual shared memory that explicitly distinguishes between computation and coordination logic. This allows shifting coordination complexities into the middleware layer, thus minimizing the implementation effort for the application developers needed to control these coordination complexities and therefore allows focusing on application development entirely. Additionally, SBC provides mechanisms to abstract issues of technical heterogeneity in distributed systems, and therefore facilitates technical integration [7]. The EKB framework uses ontologies for explicitly modeling common and local concepts as well as mappings between these concepts, thus enabling semantic integration in multi-organizational scenarios. Standards are hard to apply for projects with experts from different organizations, who have independently invested efforts into the development of different kinds of local data standards or notations. The EKB framework allows these experts to use their established and well-known local tools and data models, while additionally providing access to data originating from tools of other organizations within their local tools using local data standards or notations [8, 9].

We discuss the benefits and limitations of RDF Containers based on a real-world use case from the e-Health domain, in which data regarding accident victims is dynamically consolidated from various sources (e.g., hospitals, doctors) and used for the coordination of their treatments. Major benefit of RDF Containers is that both semantic and technical integration are supported by a single framework, while complexity aspects related to the integration and coordination of processes do not affect the integrated applications.

The remainder of this paper is structured as follows: Section 2 summarizes related work on technical and semantic integration. Section 3 identifies the research issues, while section 4 introduces the use-case. Section 5 presents the proposed approach, section 6 discusses the findings, and finally section 7 concludes the paper and presents further work.

## 2 Related Work

This section summarizes related work on technical and semantic integration.