Linked Open Data for the Italian PA: The CNR Experience

ALDO GANGEMI, ENRICO DAGA, ALBERTO SALVATI
GIANLUCA TROIANI, CLAUDIO BALDASSARRE

SUMMARY:


1. INTRODUCTION

An information system for organizations is traditionally thought as a mere technical tool for automation and management of administrative activities. In a scenario where semantic technologies are consistently proving that this idea is too restrictive, we want to reinforce the semantic web vision of aggregative information systems. We present the Semantic Scout, a software framework that offers semantic support to functionalities such as competence finding, social network discovery, etc.

The need for the Semantic Scout is motivated by the quest to provide a flexible decision making support within large organization, and in particular to support expert finding and project management. This is a common requirement within any organization with many stakeholders who are required to work in synergy, and to exploit internal resources, before looking for external competences. The hypothesis at the basis of this work is that the use of semantic technology, and in particular semantic search, automatic text categorization, linked data and ontologies, can make that requirement more easily achievable. In principle, the hypothesis is sensible for two reasons: firstly because semantic technology decouples knowledge from implemented systems, so that data can be consumed in ways closer to specific requirements or new scenarios; secondly, because semantic technology ex-

* The Authors belong to the Semantic Technology Laboratory, ISTC-CNR. This work has been supported by the CNR program Semantic IntraWeb, the Semantic Scouting project funded by the CNR Technology Transfer Office, as well as by the EU projects NeOn, funded within the 6th Framework Programme, and IKS, funded within the 7th FP.
explicitly represents the entities of an organization, which gather an own identity: such identity enables simple and effective data aggregation procedures, and nicely matches the way humans refer to relevant things in their environment. A conceptual level that is close to human knowledge management is additionally provided by explicit conceptual schemata for the data (ontologies)\(^1\).

In general, semantics improves the flexibility and adaptability of the systems, reducing the problems related to legacy and inconsistent data access, while augmenting the overall productivity. For example, the system described in our use case can be adapted to new requirements by simply changing the way the data are accessed, in a fully transparent and system-independent way.

Part of this work builds upon the results presented in two publications\(^2\), where the authors introduce an approach to migrate legacy data, in the domain of a large research institution, to a format that fosters interoperability and re-usability (RDF/OWL). Consistently with previous experiences\(^3\) we analyze the case of the Italian National Research Council (CNR)\(^4\), and capitalize the capability acquired to integrate information from different databases into an OWL knowledge base (KB). At the same time, we redefine the target goal\(^5\), expanding the request for tools that supports organizational research management both for internal needs, for opening organizational assets and data to the external world, and for assessing CNR research impact. By asset we mean humans, departments, research programs, scientific production (publications, patents), dissemination activities, etc. The objectives pursued by this work include:


Continua...