Chapter VIII

Querying the Web Reconsidered: Design Principles for Versatile Web Query Languages

François Bry, University of Munich, Germany

Christoph Koch, Vienna University of Technology, Austria

Tim Furche, University of Munich, Germany

Sebastian Schaffert, University of Munich, Germany

Liviu Badea,
Nat. Inst. for Research & Development Informatics, Bucharest, Romania

Sacha Berger, University of Munich, Germany

Abstract

A decade of experience with research proposals as well as standardized query languages for the conventional Web and the recent emergence of query languages for the Semantic Web call for a reconsideration of design principles for Web and Semantic Web query languages. This chapter first argues that a new generation of versatile Web query languages is needed for solving the challenges posed by the changing Web: We call versatile those query languages able to cope with both Web and Semantic Web data expressed in any (Web or Semantic Web) markup language. This chapter further suggests that well-known referential transparency and novel answer-closedness are essential features of versatile query languages. Indeed, they
allow queries to be considered like forms and answers like form-fillings in the spirit of the query-by-example paradigm. This chapter finally suggests that the decentralized and heterogeneous nature of the Web requires incomplete data specifications (or incomplete queries) and incomplete data selections (or incomplete answers); the form-like query can be specified without precise knowledge of the queried data, and answers can be restricted to contain only an excerpt of the queried data.

Introduction

After a decade of experience with research proposals as well as standardized query languages for the conventional Web, and following the recent emergence of query languages for the Semantic Web a reconsideration of design principles for Web and Semantic Web query languages is called for.

The Semantic Web is an endeavor widely publicized in 2001 by an influential but also controversial article from Tim Berners-Lee, James Hendler, and Ora Lassila (Berners-Lee et al., 2001). The Semantic Web vision is that of the current Web which consists of (X)HTML and documents in other XML formats extended by metadata specifying the meaning of these documents in forms usable by both humans and computers.

One might see the Semantic Web metadata added to today’s Web documents as semantic indices similar to encyclopedias. A considerable advantage over paper-printed encyclopedias is that the relationships expressed by Semantic Web metadata can be followed by computers, very much like hyperlinks, and be used for drawing conclusions using automated reasoning methods:

*For the Semantic Web to function, computers must have access to structured collections of information and sets of inference rules that they can use to conduct automated reasoning.* (Berners-Lee et al., 2001)

A number of formalisms have been proposed in recent years for representing Semantic Web metadata (e.g., RDF [Klyne et al., 2004], Topic Maps [ISO, 1999], and OWL [Bechhofer et al., 2004]). Whereas RDF and Topic Maps provide merely a syntax for representing assertions on relationships like “a text T is authored by person P,” schema or ontology languages such as RDFS (Brickley et al., 2004) and OWL allow one to state properties of the terms used in such assertions (e.g., that no person can be a text). Building upon descriptions of resources and their schemas, as detailed in the architectural road map for the Semantic Web (Berners-Lee, 1998), rules expressed in SWRL (Horrocks et al., 2004) or RuleML (Boley et al., 2002),
Social Networks Applied to E-Gov: An Architecture for Semantic Services
www.igi-global.com/chapter/social-networks-applied-gov/35758?camid=4v1a