

Chapter 5

Linked Data: A Manner to Realize the Web of Data

Leila Zemmouchi-Ghomari

Ecole Nationale Supérieure de Technologie (ENST), Algeria

ABSTRACT

The data on the web is heterogeneous and distributed, which makes its integration a sine qua non-condition for its effective exploitation within the context of the semantic web or the so-called web of data. A promising solution for web data integration is the linked data initiative, which is based on four principles that aim to standardize the publication of structured data on the web. The objective of this chapter is to provide an overview of the essential aspects of this fairly recent and exciting field, including the model of linked data: resource description framework (RDF), its query language: simple protocol, and the RDF query language (SPARQL), the available means of publication and consumption of linked data, and the existing applications and the issues not yet addressed in research.

INTRODUCTION

Web data is often stored in isolated silos; these isolated silos are unable to exchange their content with other web systems. Incompatibility between data formats leads to low level of communication and data sharing (Herman, 2010). The contextual interpretation of data sets can be expensive, difficult, and time-consuming; in addition, it requires the intervention of developers. Furthermore, data is integrated into web pages and is only available for human consumption; thus, it is essential to adopt a framework based on coherent (homogeneous) principles that allow automatic processing, information sharing, cooperation, and collaboration, in lieu of the current framework. The primary objective of the Semantic Web is to enable machines to understand and process data contained in web pages and documents; it aims to create a structured Web of Data to complement the existing Web of Documents.

According to many researchers, including Bizer and Heath (2009), Wood et al. (2014) and Hogan (2014), The Linked Data Initiative¹ is a promising approach, which includes a set of standards and tools that enable publishing and interlinking structured data on the web. The Linked Data technologies render the conceptual models on which these data are based to be explicit and transparent. In particular, these

DOI: 10.4018/978-1-5225-6367-9.ch005

Linked Data technologies support data integration in dynamic and distributed environments, such as large enterprises, intergovernmental organizations, and the World Wide Web (Mendez & Greenberg, 2012).

In fact, if the Resource Description Framework (RDF) is adopted and shared by data providers and consumers, automated data exchange would be considerably easy with common tools for accessing data, such as the Simple Protocol and the RDF Query Language (SPARQL). Linked Data enables development of new web pages using data extracted from multiple other web pages, i.e., Linked Data allows for uncoordinated cooperation (Wood et al., 2014).

However, the publication of Linked Data does not guarantee their publication, because publication requires the provenance, quality, credit, attribution, and implementation of methods to ensure reproducibility for the validation of results (Bechhofer et al., 2013; Schmachtenberg et al., 2014); thus, several challenges still need to be addressed by researchers in this field.

Exploration, description, and explanation are the main purposes of survey research (Pinsonneault & Kraemer, 1993). This work involves exploration, which is a method that aims to make readers more familiar with a topic and explain the related basic concepts; thus, the aim of this chapter is to highlight the efforts of the Semantic Web community towards the realization of the Web of Data using Linked Data principles for publishing and consuming data on the web. The best practices and recommendations of the World Wide Web Consortium (W3C) are explored; in particular, web data modeling and querying, interlinking techniques, dataset reuse, and challenges for future research, are discussed.

The remainder of this chapter is organized as follows. Section 2 describes the features of the Web of Data and compares them with those of the Web of Documents; in addition, several approaches for Web of Data are presented. Section 3 defines the Linked Data principles. Section 4 describes the Linked Data model: RDF. Section 5 presents SPARQL, the Linked Data query language, and its capabilities. The method of consumption of Linked Data datasets is explained in Sections 6. Then, Section 7 discusses the existing approaches and tools for the publication of Linked Data. Section 8 presents some of the most common applications of Linked Data technologies. The issues with Linked Data along with their probable causes and possible solutions are summarized in Section 9. Finally, Section 10 concludes the chapter by highlighting the role of Linked Data in realizing the full potential of the web.

WEB OF DATA DESCRIPTION AND APPROACHES

Bizer and Heath (2011) define the Web of Data as another layer that is linked with the classic document Web, and it has the following features:

- **Generic:** Can contain any type of data.
- **Interlinked:** Entities are connected by links that enable the visualization and the exploration of a unique giant global graph.
- **Open:** Data sources are dynamic and can be exploited in the framework of applications, in addition, anyone can publish data to the Web.

A comparison between the Web of Data and the Web of Documents is presented in Table 1 (Zem-mouchi-Ghomari, 2015).

25 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/linked-data/208794

Related Content

Knitting Patterns: Managing Design Complexity with Computation

Dermott John James McMeeland Robert Amor (2021). *Research Anthology on Recent Trends, Tools, and Implications of Computer Programming* (pp. 2055-2069).

www.irma-international.org/chapter/knitting-patterns/261116

Delay Faults Testing

Marcel Baláž, Roland Dobaia and Elena Gramatová (2011). *Design and Test Technology for Dependable Systems-on-Chip* (pp. 377-394).

www.irma-international.org/chapter/delay-faults-testing/51410

Analyzing Coordinate Relations in Handwriting Activity: Tacit Skill and Individuality

Yusuke Manabe and Kenji Sugawara (2011). *Kansei Engineering and Soft Computing: Theory and Practice* (pp. 287-302).

www.irma-international.org/chapter/analyzing-coordinate-relations-handwriting-activity/46404

SOI Technology in Designing Low-Power VLSI Circuits

Srividya P. (2023). *Energy Systems Design for Low-Power Computing* (pp. 17-28).

www.irma-international.org/chapter/soi-technology-in-designing-low-power-vlsi-circuits/319987

Cyber Space Security Assessment Case Study

Hanaa. M. Said, Rania El Gohary, Mohamed Hamdy and Abdelbadeeh M. Salem (2018). *Cyber Security and Threats: Concepts, Methodologies, Tools, and Applications* (pp. 1060-1092).

www.irma-international.org/chapter/cyber-space-security-assessment-case-study/203548