Evolution of Linked Data Application Domains From 2009 to 2015: A Systematic Literature Review

Alia O. Bahanshal, King Saud University, Riyadh, Saudi Arabia & King Abdulaziz City for Science and Technology, Riyadh, Saudi Arabia
Hend S. Al-Khalifa, King Saud University, Riyadh, Saudi Arabia

ABSTRACT

Nowadays, research in Linked Data has significantly advanced and it now entails enormous area of applications like research publications and data sets. Its flexibility and effectiveness in handling and linking data from numerous sources has made Linked Data more popular. The aim of this article is to systematically present literature review of Linked Data and its development since 2009. Moreover, cumulative experiences and lessons learned from recent years will be highlighted. Findings showed that Linked Data has grown in the past five years in terms of number of datasets, research publications and domain-specific applications.

KEYWORDS

Linked Data, Linked Data Applications, Open Data, Web of Data

1. INTRODUCTION

Recent years have witnessed a sharp increase in research of Linked Data in various domains and areas of applications with many standards, datasets and publishing tools. The reasons for the technology’s popularity are rooted in its openness, capability of managing and linking heterogeneous data from several sources as well as flexibility, which allows its application to different domains (Heath & Bizer, 2011).

Despite the number of relevant publications on Linked Data usages, the demand for more up-to-date and structured reviews on current developments in the field remains. Examining the most developed domains, data sets and applications will provide academics, researchers, students, companies, developers and contributors with the competence to get involved in the field.

From its inception, many efforts were made to track the development of Linked Data. In 2009, Bizer et al. (2009) presented the accomplishments that were made and explored the engagements in different applications areas. The publishing tools of that time were also reviewed. In (Bizer, 2009), Bizer recognized the potential of the usage of Linked Data in specific domains, such as Publications, Media, Life Sciences, Geography and emphasized the amount of Web 2.0 User-Generated Content that is presented as Linked Data.

DOI: 10.4018/IJTD.2018040101
In 2009, the W3C RDB2RDF Incubator Group (Sahoo et al., 2009) conducted a survey about existing approaches of mapping between Relational Databases (RDB) and the Resource Description Framework (RDF) (Beckett, 2004). In 2011, Heath et al. (2011) made efforts covering most of Linked Data related topics, technologies, and the different types of applications.

However, in the years since then, things have evolved enormously in the field of Linked Data with more research penetrating many domains and more publishing tools used in the respective domains. Furthermore, the previous efforts to track Linked Data growth provided examples of its usage, but a complete list of domain-specific applications was not found. Also, fully categorized Linked Data domain-specific applications are still needed to track the change in each domain independently, which would assist contributors and researchers to easily identify the current efforts and future trends in the specific domains.

Bizar et al. (2009) explained the applications that exploited Linked Data in 2009 and publishing approaches of that time. In this research, we analyzed and categorized academic journals articles related to Linked Data covering the time span from 2009 to 2015 for the purpose of tracking the development of Linked Data since 2009. Linked Data domains of application were categorized and explained, and each domain published articles were provided. Also, evidence for the growth of Linked Data was reported. This systematic review were conducted to seek answers to the following research questions:

1. How to track the development of Linked Data during the time span from 2009 to 2015?
2. What are the new Linked Data domains of application?

Therefore, the main contribution of this study lies in the following points:

1. Tracking the growth of Linked Data research and data sets;
2. Categorizing the existing Linked Data research into application domains, for the purpose of offering understanding and prospects to analysts and researchers with regards to each domain;
3. Providing a list of 740 research paper in Linked Data research.

The rest of this paper is organized as follows: Section 2 presents Linked Data background and its growth since its inception. Section 3 explains the research methods applied in this study. Section 4 reports the literature review findings. Section 5 summarizes the obtained results and provides future forecast.

2. BACKGROUND

In 2006, Linked Data (aka Web of Data) was introduced by Tim Berner Lee as a method to publish interlinked data forming a single global space of data from various sources (Heath & Bizer, 2011). Linked Data format is understood by machines, and thus raw data can be retrieved. The term Linked Data refers to “the best practices for connecting and publishing structured data on the Web” (Bizer et al., 2009). Linked Data, same as web of documents, connects different online resources. However, Linked Data interlinks both data and documents in a predefined standard format. To make your data linked, four rules were defined by Tim Berner Lee (Bizer et al., 2009), which are: “Use URIs as names for things”, “Use HTTP URIs, to look up things”, “Provide useful information in RDF”, and “Include links to other URIs”.

Resource Description Format (RDF) is the standard model for the expression of data and relations in Linked Data. RDF data model consists of subject - predicate - object triples (Beckett, 2004). Subject and object, could be URIs referencing resources. The object could be a string literal, while the predicate represents the relation between subject and object.
Jesus Cano and Roberto Hernández (2017). Securing Government Information and Data in Developing Countries (pp. 1-26).
www.igi-global.com/chapter/managing-software-architecture-in-domains-of-security-critical-systems/178657?camid=4v1a

Interactive Self-Assessment Questions within a Virtual Environment
Chris Evans and Luis Palacios (2013). Adoption of Virtual Technologies for Business, Educational, and Governmental Advancements (pp. 218-227).
www.igi-global.com/chapter/interactive-self-assessment-questions-within/72411?camid=4v1a