



Chapter 90

The Big Data Research Ecosystem: An Analytical Literature Study

Moses John Strydom

 <https://orcid.org/0000-0002-8865-7474>
University of South Africa, South Africa

Sheryl Buckley

 <https://orcid.org/0000-0002-2393-4741>
University of South Africa, South Africa

ABSTRACT

Big data is the emerging field where innovative technology offers new ways to extract value from an unequivocal plethora of available information. By its fundamental characteristic, the big data ecosystem is highly conjectural and is susceptible to continuous and rapid evolution in line with developments in technology and opportunities, a situation that predisposes the field to research in very brief time spans. Against this background, both academics and practitioners oddly have a limited understanding of how organizations translate potential into actual social and economic value. This chapter conducts an in-depth systematic review of existing penchants in the rapidly developing field of big data research and, thereafter, systematically reviewed these studies to identify some of their weaknesses and challenges. The authors argue that, in practice, most of big data surveys do not focus on technologies, and instead present algorithms and approaches employed to process big data.

INTRODUCTION

The International Data Corporation (Reinsel et al., 2017) forecasts that by 2025 the global data sphere or ecosystem will surge to 163 Zettabytes, id est, a trillion Gigabytes. This signifies that the new data lake was foreseen to be ten times the 16 Zettabytes worth of data that its devices, storage systems and data centers generated in 2016, singly. As expected, research in the domain has likewise literally exploded.

DOI: 10.4018/978-1-7998-7705-9.ch090

The direct consequence of these expectations is that there has been a parallel paradigm shift in research activity: a literatim explosion!

Some changes seem to be running at the speed of light. In this morass of hard trends and cosmetic swirls - an unquestionable tsunami - how do we distinguish the trend from the trendy? How do we know which changes are truly disruptive, and which are merely potentially transformative? Which technologies are going to have the biggest impact on decision making? And, of the many initiatives put forward, which will be pivotal to the researcher's and the practitioner's journey?

Predictably, big data governance has thus become increasingly critical, indeed decisive, as more organizations learn how to leverage their data and exploit it to make better decisions, optimize operations, create new products and services, and improve profitability.

Beyond their sheer magnitude, these data-sets and affiliated applicatory considerations pose significant challenges for method and software development.

Hitherto, the bulk of big data reports have focused on discoursing opportunities, applications, challenges and issues (Wang et al., 2016). Others favored to survey and study the algorithms and techniques used in such contexts (Raguseo, 2018).

Only a limited number of authors treat big data technologies with respect to their ecosystem by focusing on the phases of the data life-cycle model (Cavanillas et al., 2015).

Additionally, there is strong need to systematically review these studies in order to render them accessible to researchers and practitioners.

By clearly defining the opportunity in big data, by examining the big data value chain, and by undertaking a comprehensive inspection into industry sector applications, this chapter charts a way forward to new value creation and new opportunities from big data. Decision makers, policy advisers, researchers, and practitioners on all levels can benefit from this.

The novelty of this study is that it explores the current trends in the field of big data research and most relevant research areas, during the past triennial, that is, 2015 to 2017.

In doing so, the authors introduced the big data value chain, opportunities and applications. They presented the main challenges encountered when facing the complexity of imbalanced data-sets and difficulties contingent on the V's of big data.

We are reminded that big data is not a single technology or initiative. Rather, it is a contiguous derivative of all our interactions with rapidly evolving technologies. This latter reality presages that big data is not a stable platform but is as dynamic as is the technology that defines it.

The tendencies, based on a systematic literature review methodology were identified by an extensive audit of peer-reviewed scholarly journals where the objective was to congregate, elaborate and synthesize an exhaustive systematic literature review of current issues related to the area of big data research trends. In the same vain, topics and proclivities in the areas of exascale computing (exascale ecosystem?) and social data analysis were reported.

Primarily, with the objective of analyzing current research, content analysis was employed.

Building a representative dictionary of the subjects being analyzed was a key to the proposed procedure. Likewise, social network analysis tools were employed to interpret the interrelationship between the big data taxonomy of terms.

The results were interpreted using descriptive analysis (frequencies) and social network analysis

This systematic literature review is foreseen to contribute to the scientific knowledge on computer science and information system management by (i)studying in detail the current issues related to big

29 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/the-big-data-research-ecosystem/270684

Related Content

Declarative Planning and Knowledge Representation in an Action Language

Thomas Eiter, Wolfgang Faber and Gerald Pfeifer (2008). *Intelligent Information Technologies: Concepts, Methodologies, Tools, and Applications* (pp. 192-221).

www.irma-international.org/chapter/declarative-planning-knowledge-representation-action/24279

Beyond the Hype: The Realities of AI in Clinical Practice

Adan Khan, Venus Oseyi Omokhoa and Safa Inaya Afzal (2024). *Clinical Practice and Unmet Challenges in AI-Enhanced Healthcare Systems* (pp. 1-22).

www.irma-international.org/chapter/beyond-the-hype/352910

The WASP Framework: Bridging the Gap Between the Web of Systems, the Web of Services, and the Web of Semantics with Agent Technology

Thomas Biskup, Jorge Marx Gomez and Claus Rautenstrauch (2005). *International Journal of Intelligent Information Technologies* (pp. 68-82).

www.irma-international.org/article/wasp-framework-bridging-gap-between/2385

AI in Health and Safety Management for Real Estate 4.0

Soumi Majumder (2022). *International Journal of Ambient Computing and Intelligence* (pp. 1-18).

www.irma-international.org/article/ai-in-health-and-safety-management-for-real-estate-40/311061

Generative AI in Predicting and Preventing Foodborne Illnesses

Shweta Ashish Koparde, P. Selvakumar, Ravi Prakash Jha, K. Deepthi Reddy, Jyothsna Devi Kuchipudi, T. C. Manjunath and Mohit Sharma (2026). *Generative AI in Food Systems: Predictive Demand, Smart Supply Chains, and Sustainable Service Futures* (pp. 205-232).

www.irma-international.org/chapter/generative-ai-in-predicting-and-preventing-foodborne-illnesses/383884