Chapter 12
Towards Semantic Data Integration in Resource-Limited Settings for Decision Support on Gait-Related Diseases

Olawande Daramola
Cape Peninsula University of Technology, South Africa

Thomas Moser
https://orcid.org/0000-0002-9220-649X
St. Pölten University of Applied Sciences, Austria

ABSTRACT

Resource-limited settings (RLS) are characterised by lack of access to adequate resources such as ICT infrastructure, qualified medical personnel, healthcare facilities, and affordable healthcare for common people. The potential for the application of AI and clinical decision support systems in RLS are limited due to these challenges. Towards the improvement of the status quo, this chapter presents the conceptual design of a framework for the semantic integration of health data from multiple sources to facilitate decision support for the diagnosis and treatment of gait-related diseases in RLS. The authors describe how the framework can leverage ontologies and knowledge graphs for semantic data integration to achieve this. The plausibility of the proposed framework and the general imperatives for its practical realisation are also presented.

INTRODUCTION

Accurate and prompt treatment decision-making by medical practitioners in dealing with patients is critical to ensuring the health and well-being of patients. It is particularly important when this is done in resource-limited settings (RLS) where several challenges impede good healthcare conditions. Some of the challenges of RLS, which any viable healthcare solution must overcome include the shortage of qualified

DOI: 10.4018/978-1-7998-6697-8.ch012

personnel, lack of good infrastructure, lack of ready access to technology, and high cost of healthcare for common people (Zargaran et al. 2014; Fritz et al., 2015, Daramola & Moser, 2019; Siow et al., 2020). Another challenge is the scarcity of quality data to support Artificial Intelligence (AI) operations. Even when such data exist, they are often in disparate health information systems, which translate to inconsistencies of representations in the schema, semantics, terminologies, data types, and data formats that are used (Dhayne et al., 2018). Also, health data exist in both structured and unstructured forms all of which should be harnessed for effective data-driven decision-making. Thus, the semantic integration of health data sources is necessary for effective data-driven decision support in healthcare. Semantic data integration (SDI) enables heterogeneous health information systems, and health data to be harnessed for meaningful communication and exchange of data while the context of the individual data and systems are preserved (Cheatham & Pesquita, 2017). SDI will ensure that the exchange of data and interpretation of data are consistent irrespective of differences in data labels, data schema, and terminologies that are used by different databases or data (Cheatham & Pesquita, 2017; Vidal et al., 2019; Asfand-E-Yar & Ali, 2020). This will ensure access to all relevant data for accurate data-driven decision-making in healthcare (Shi et al., 2017; Balakrishna et al., 2020).

The advent of new digital technologies such as cloud computing, mobile computing, wearable sensors, Internet of Things (IoT), and linked open data has created new opportunities for semantic integration of heterogeneous data that can facilitate intelligent decision-making. This could aid important clinical processes in the aspects of disease diagnosis, disease treatment, and patient's rehabilitation even if there is a physical distance between the patient and the medical practitioner (Balakrishna et al., 2020). One of the state-of-the-art semantic technologies being used for semantic integration is a knowledge graph. Sundry definitions of a knowledge graph exist and it is difficult to find a consensus definition. However, based on some of the more quoted definitions, we can claim that a knowledge graph describes real-world entities and the relationships that exist between them by using a graphical model of representation. It has a schema that defines classes and entity relations; it captures potential interrelationships between arbitrary entities; and spans several topical domains (Paulheim, 2017). A knowledge graph stores interlinked descriptions of real-world concepts, objects, event, entities, things, situations, and their semantic definitions, which provides a basis to reason on them and derive new knowledge (Ehrlinger & Wöß, 2016; Hogan et al., 2020). It is essentially an ontology plus data instances that are organized based on a graph-based data structure, such that reasoning can be applied on it to generate new knowledge. Well-known organizations like Google, Facebook, Amazon, Microsoft are known to have their AI search operations powered by knowledge graphs (Hogan et al., 2020). The application of AI in healthcare can benefit from semantic data integration through the use of ontologies and knowledge graphs because it will provide a solid basis for data-driven support for clinical healthcare operations. One aspect that is of interest is clinical gait analysis, which is being increasingly used for diagnosis and treatment of neurological and cardiovascular diseases, and chronic diseases in general (Whittle, 2014; di Biase et al., 2020; Dugan et al., 2020).

Gait analysis (GA), which is also known as locomotive analysis entails collecting quantitative data on the pattern of physical movements of a human to understand the etiology of gait defects and the formulation of an appropriate treatment plan (Whittle, 2014; Buongiorno et. al, 2019). GA has been used for disease diagnosis for patients with walking impairment, and the treatment of many neurological disorders and cardiovascular diseases (Tang and Su, 2013; Camps et al., 2018; Buongiorno et al, 2019). Currently, it is possible to obtain real-time information on a patient's gait pattern while the patient is walking both within or outside the hospital environment. Critical gait parameters such as gait speed,