Chapter 16
Ontology for Data Quality and Chronic Disease Management:
A Literature Review

Alireza Rahimi
University of New South Wales, Australia &
Isfahan University of Medical Sciences, Iran

Pradeep Kumar Ray
University of New South Wales, Australia

Jane Taggart
University of New South Wales, Australia

Siaw-Teng Liaw
University of New South Wales, Australia

Hairong Yu
University of New South Wales, Australia

ABSTRACT

Improved Data Quality (DQ) can improve the quality of decisions and lead to better policy in health organizations. Ontologies can support automated tools to assess DQ. This chapter examines ontology-based approaches to conceptualization and specification of DQ based on "fitness for purpose" within the health context. English language studies that addressed DQ, fitness for purpose, ontology-based approaches, and implementations were included. The authors screened 315 papers; excluded 36 duplicates, 182 on abstract review, and 46 on full-text review; leaving 52 papers. These were appraised with a realist "context-mechanism-impacts/outcomes" template. The authors found a lack of consensus frameworks or definitions for DQ and comprehensive ontological approaches to DQ or fitness for purpose. The majority of papers described the processes of the development of DQ tools. Some assessed the impact of implementing ontology-based specifications for DQ. There were few evaluative studies of the performance of DQ assessment tools developed; none compared ontological with non-ontological approaches.

INTRODUCTION

The growing use of electronic health records (EHRs) raises issues of semantic interoperability and the quality management/improvement of large datasets derived from multiple EHRs. Improved data quality in EHRs can improve the quality of decisions and lead to better policy that actually meet needs, strategies, evidence-based care and patient outcomes.

The acceptable level of data quality is not fixed in the system. Rather health professionals can provide it at different times and data users need to assess that quality contextually, based on the fitness for research, audit and quality assurance purposes (Devillers, Bedard, Jeansoulin, & Moulin, 2007). It is important to take a user view point of quality because it is the end user who evaluate whether or not data is fit for use. A focus is the quality of patient or disease registers derived from EHRs to support policy and practice. Patients registers need to have a level of completeness and the information contained, need a level of correctness and consistency to be useful for clinical, quality improvement and research purposes (S. Liaw, Taggart, Dennis, & Yeo, 2011).

DQ was conceptualised in terms of its "fitness for purpose/use" in a few papers (R. Wang, 1998; R. Y. Wang, D. M. Strong, & L. M. Guarascio, 1996). DQ can be described from two perspectives: (1) intrinsic quality of data elements and set of data elements (data set) and (2) how the set meets the user's needs i.e. fitness for purpose. The commonly approved definition of DQ has been epitomized in the International Standards Organisation definition: "the totality of features and characteristics of an entity that bears on its ability to satisfy stated and implied needs" (ISO 8402-1986, Quality Vocabulary). DQ also can be specified in terms of its "fitness for purpose/use" (R. Wang, 1998; R. Y. Wang, 1998; R. Y. Wang, et al., 1996).

Intrinsic DQ refers to the extent that data is free of defects as measured by specific DQ dimensions, including "accuracy, perfection, freshness and uniformity" (Redman, 2005) and "completeness, unambiguity, meaningless and correctness" (Choquet et al., 2010; Orme, Yao, & Etzkorn, 2007; Wand & Wang, 1996; Yao, Orme, & Etzkorn, 2005). The Canadian Institute for Health Information recommendations were the basis for an information quality framework comprising 69 quality criteria grouped into 24 quality characteristics, which was further grouped into 6 quality dimensions: accuracy, timeliness, comparability, usability, relevance and privacy & security (K. Kerr, A. Norris, & R. Stockdale, 2007). Research in DQ has tended to focus on the identification of generic quality characteristics such as accuracy, currency and completeness (Orme, et al., 2007; R. Wang, D. Strong, & L. Guarascio, 1996) or completeness, correctness, consistency and timeliness (S. Liaw, et al., 2011) as core dimensions of DQ that are relevant across application domains. However, a pervious review shows there is a lack of consensus conceptual framework and definition for DQ (S. T. Liaw et al., 2013).

Many studies regularly report a range of deficiencies in the collected information for professionals practice (Devillers, et al., 2007; B. K. Kahn, Strong, & Wang, 2002), clinical (Azaouagh & Stausberg, 2008; de Lusignan et al., 2010; Mitchell & Westerduin, 2008; Moro & Morsillo, 2004) and health promotion (Gillies, 2000b) purposes. Similar deficiencies exist with information in geographic (Devillers, et al., 2007; Ivanova, Morales, de By, Beshe, & Gebresilassie, 2013), hospital and general practice (ST Liaw et al., 2011) information systems, where the lack of coding rules meant that much of the data are often incomplete or in relatively inaccessible text format. The evidence is more encouraging for data for administrative purposes (Lain, Roberts,

Related Content

Artificial Intelligence and Its Integration with Regenerative Medicine Approach
Tungki Pratama Umar, Dessy Agustini, Abdelrahman M. Makram, Muhammad Muzzamil, Bella Stevanny,
Randa Elsheikh, N. L. Swathi, Tulika Garg, Aulia Ananditia Putriand Nityanand Jain (2023). *Integrating
Digital Health Strategies for Effective Administration (pp. 32-57).*
www.irma-international.org/chapter/artificial-intelligence-and-its-integration-with-regenerative-medicine-approach/323778

Good IT Requires Good Communication
Charles H. Andrusand Mark Gaynor (2013). *Cases on Healthcare Information Technology for Patient Care
Management (pp. 122-125).*
www.irma-international.org/chapter/good-requires-good-communication/73945

Obesity Epidemic Simulation Based on Behavioral Models and Intelligent Agents
Marina Massei, Alberto Tremori, Vera Novak, Simonluca Poggi, Christian Bartolucci, Angelo Ferrandoand
Alessandro Chiurco (2013). *International Journal of Privacy and Health Information Management (pp. 96-114).*
www.irma-international.org/article/obesity-epidemic-simulation-based-on-behavioral-models-and-intelligent-agents/102632

Wireless for Managing Health Care: The Wirhe Framework
Esko Alasaarela, Ravi Nemana, Steven DeMello, Nick S. Oliverand Masako Miyazaki (2010). *Health
Information Systems: Concepts, Methodologies, Tools, and Applications (pp. 790-810).*
www.irma-international.org/chapter/wireless-managing-health-care/49899

Supporting Health and Community Care with Multi-Agent Systems
Federico Bergenti, Monica Mordonini, Agostino Poggiand Michele Tomaiuolo (2017). *International Journal
of Privacy and Health Information Management (pp. 97-113).*
www.irma-international.org/article/supporting-health-and-community-care-with-multi-agent-systems/182881