Journal of Biomedical Informatics xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

Ontological modeling of electronic health information exchange

J. McMurray [a,c,*], L. Zhu [a], I. McKillop [a,b], H. Chen [a,b]

[a] School of Public Health and Health Systems, University of Waterloo, Canada
[b] David R Cheriton School of Computer Science, University of Waterloo, Canada
[c] School of Business & Economics/Health Studies, Wilfrid Laurier University, Canada

ARTICLE INFO

Article history:
Received 4 January 2014
Revised 19 April 2015
Accepted 28 May 2015
Available online xxxx

Keywords:
Ontological model
Data quality
Health information exchange
Interoperability
Health system performance

ABSTRACT

Introduction: Investments of resources to purposively improve the movement of information between health system providers are currently made with imperfect information. No inventories of system-level electronic health information flows currently exist, nor do measures of inter-organizational electronic information exchange.
Methods: Using Protégé 4, an open-source OWL Web ontology language editor and knowledge-based framework, we formalized a model that decomposes inter-organizational electronic health information flow into derivative concepts such as diversity, breadth, volume, structure, standardization and connectivity.
Results: The ontology was populated with data from a regional health system and the flows were measured. Individual instance's properties were inferred from their class associations as determined by their data and object property rules. It was also possible to visualize interoperability activity for regional analysis and planning purposes. A property called Impact was created from the total number of patients or clients that a health entity in the region served in a year, and the total number of health service providers or organizations with whom it exchanged information in support of clinical decision-making, diagnosis or treatment. Identifying providers with a high Impact but low Interoperability score could assist planners and policy-makers to optimize technology investments intended to electronically share patient information across the continuum of care. Finally, we demonstrated how linked ontologies were used to identify logical inconsistencies in self-reported data for the study.

1. Introduction

While the use of local data management solutions (such as EHRs, PACs, and pharmacy systems) are now very prevalent in hospitals and other care provision settings [1], the investment in technologies to enable inter-provider exchange of information remains low in the healthcare domain [2–5]. The fact that electronic information often cannot flow freely between providers represents an obstacle to the vision of a fully integrated health system. The consequences are felt, sometimes acutely. Providers make clinical decisions based on partial information, service can be delayed [6,7], and information is sometimes missing or received too late to be useful [8,9].

While there is almost universal acceptance of the value of integrated, electronic health information systems, investments to improve the movement of electronic health information between providers are often made with imperfect information. There do not appear to be any inventories of regional digital information flow nor any broadly developed measures of inter-organizational health information exchange [10]. Equally concerning is our limited understanding of the scope of, or capacity for, information exchange between the sophisticated systems now common in hospitals and medical clinics, and the massive amounts of siloed data collected by other players in the system who are part of the circle of care, such as pharmacies, home care providers, and allied health care providers.

There is a clear need to be able to systematically describe and then assess progress toward interoperability and information exchange among healthcare providers. Without this ability, policy makers, developers and users remain blind as to where investment and effort is needed. This paper introduces a promising first step to help address this challenge by demonstrating how an ontological approach to describing information flows between disparate healthcare systems can be unambiguously described and later measured.

* Corresponding author at: Business Technology Management/Health Studies, Wilfrid Laurier University, 73 George St, Brantford, Ontario N3T3Y3, Canada. Tel.: +1 519 242 7477; fax: +1 519 751 7589.
E-mail address: jmcmurray@wlu.ca (J. McMurray).

http://dx.doi.org/10.1016/j.jbi.2015.05.020
1532-0464/

2. The problem

An established body of literature exists that describes and tests interoperability frameworks and models, a good portion of which comes from the military and business domains [11–16]. Unfortunately, to date there is no framework or model to describe system-level information flows between organizations in the healthcare sector. Moreover, the collection and analysis of large amounts of data from heterogeneous sources presents numerous challenges which are compounded when the data set is large and complex, or needs to be systematically analyzed. In these cases it is common to make data machine-readable (searchable and able to be understood by computers) through the use of standardized, unambiguous terminology for concepts in the knowledge base (say by defining "name" to mean surname), along with rules (axioms) that impose constraints on the data (such as "name" cannot contain integers) [17–20]. In practice, the localized implementation of terminology and rules (such as "name" in one setting may or may not have the same meaning as "PatName" in another setting) still limits the vision of an unambiguous, broad exchange of health information. Semantic interoperability between clinical information systems, where exchanged information is explicitly understood by both sender and receiver, must occur if we wish to achieve integration through "seamless data exchange" [21].

From a methodological perspective, the nature of distributed data can make it equally challenging to assemble system-level performance information to objectively assess the impact of policies or investments that enable networked health information systems. The tools used for the collection of such data often involve self-reporting and questions that allow a subjective response, particularly where there are no independent sources of the data outside the reporting organization. Cognitive biases may impact the quality of the data as a result [22]. Clearly, methods for automatic cross-checking of self-reported data to ensure reliability are also required [23].

2.1. An ontological solution

A promising solution may lie in the use of an ontological approach. This methodology allows for the conceptualization of information flows between providers. Furthermore, flows that can be modeled can also be measured – thus helping to solve the associated performance measurement conundrum [10,24,25].

Ontologies in information science have their origins in artificial intelligence where the lack of a common understanding of a domain of interest presented a significant barrier to building and sharing accumulated knowledge bases, and thus to building interoperable systems [26]. Ontologies are an abstraction of reality; they describe the concepts associated with a domain or artefact, and through a systematic descriptive process highlight the interrelationships and constraints between those concepts [27]. The use of ontologies in biomedical informatics is well established [28,29], however in the performance measurement field the approach is nascent; thus sharing and collaborating to construct common ontologies is an important component of developing this domain [30].

In both their syntactical and semantic forms, ontologies allow for formal definition and computation of the relationships between objects being described. By employing an inference engine, or "reasoner", the asserted relationships between objects can be computed. Any violation of axioms and logic constraints will be identified [27]. The modular approach inherent in ontologies (achieved through the use of a common vocabulary of representational terms and explicit declarative knowledge) means an ontology is extendable for use with other related ontologies. Furthermore, ontological models are logically computable meaning it is possible to infer new information based on the underlying model [31].

While ontologies are common in biomedicine and are increasingly used in designing controlled vocabularies [32] and reference terminologies [33], there is little research into the use of ontologies applied to health information systems themselves [34], or in support of healthcare performance measurement [35]. Without shared or standardized lexical and semantic models, reliable and accurate comparison of performance across sectors and geographic borders is unachievable.

Here we explore the utility of an ontological model designed both to conceptualize and measure electronic health information exchange between health entities in a large regional healthcare system in Ontario, Canada, and to test the reliability of the collected data.

3. Methodological approach to develop the ontology

In this paper we describe the conceptual framework of health system information exchange and its related ontology. We confirm the veracity of the framework using Protégé 4, an open-source OWL Web ontology language editor and knowledge-base framework (Stanford Center for Biomedical Informatics Research, Version 4.1.0 Beta, 2011, from http://protege.stanford.edu). HermiT Reasoner, a semantic rules engine (Information Systems Group, Department of Computer Science, University of Oxford, Version 1.3.6, 2011, from http://hermit-reasoner.com), was used for classification and rule validation.

The process of designing the ontology was iterative. First, a review of the literature and an informal conceptualization of a measurement model were used to derive a formula to calculate an individual organization's interoperability score (the electronic health information exchange indicator or eHIE – the explanation of which is outside the scope of this paper). We populated the ontology with instances from a health region to validate that the ontology is a fair representation of the real world [10]. All data collection received ethics clearance from the authors' university Research Ethics Board in accordance with protocols for research with human participants, including informed consent.

The ontology (referred to here as Health Exchange Interoperability Ontology or HEIO) provides a permanent artefact of the specification of each of the concepts in the measurement model, and allows for a more explicit visualization of the model and instances. Access to the OWL version of the HEIO ontology is available at BioPortal: HEIO v1.15 http://bioportal.bioontology.org/ontologies/HEIO. An adaptation of CamelCase is used throughout, whereby all class and compound class names start with a capital letter for each word with no spaces between them. The naming convention for properties is similar except that the first letter of the first word is not capitalized [36]. In some cases, class names used in HEIO are modeled after the class structure from the original source i.e. the type of information classes reflect Canada Health Infoway's Blueprint [37], or were given class names that reflected common and recognizable terminology i.e. application classes.

3.1. An informal model of electronic health information exchange

An informal conceptual model of electronic health information exchange between a system of regional healthcare providers was developed based on an iterative process of brainstorming, reference to literature, and consensus development. A representation, using IHMC ® CMap Tools (v5.04.02, Florida institute for Human & Machine Cognition available from http://www.ihmc.us/cmap-tools.php) is provided in Fig. 1.