PopHR: An Integrated Semantic Framework for Population Health Surveillance

Arash Shaban-Nejad[*], Christian Jauvin, Maxime Lavigne, Masoumeh T. Izadi, Luke Mondor, Anya Okhmatovskaia, and David L. Buckeridge

McGill Clinical & Health Informatics, Department of Epidemiology and Biostatistics, McGill University, Montreal, Quebec, H3A 1A3 Canada

ABSTRACT

This is a progress report on our effort to design and development of a semantic backbone using an integrated ontological framework to support calculation, interpretation and querying population health indicators within the scope of the PopHR project.

1 OBJECTIVES AND SYSTEM DESCRIPTION

The PopHR project (Buckeridge et al., 2012) aims to deliver an informatics platform that integrates multiple clinical and administrative data sources to provide a coherent view of the health of populations in the province of Quebec, Canada (Figure 1). Users of the PopHR can develop detailed portraits of the health status and healthcare utilization patterns for a population, monitor various health indicators to detect temporal and spatial variations in disease activity, and evaluate the effectiveness of interventions or other events on population health. The platform provides representative information in near-real time with high geographical resolution, thereby assisting public health professionals, clinicians and the public in diagnostic and therapeutic decision-making. The design of the PopHR system is centered around a population health indicator ontology (PHIO) (Shaban-Nejad et al., 2013). Also we have developed the Geo-PopHR ontology, to assist establishing the relationships between administrative data to geospatial data, with emphasize on mapping variations in disease rates.

Two different approaches are used to calculate health indicators in the PopHR. Indicators related to disease status and health care utilization are generally calculated using administrative and clinical data from a representative cohort. Indicators related to determinants of health are generally calculated using area level data from the census, surveys, and other sources.

In order to facilitate semantic question answering about health indicators, determinants, disease outcomes and interventions we are developing a natural language (query) interface (NLI). To derive a semantic tree for each of the queries, interface, at the current stage of our experiment, we are using a formal syntax and grammar through the following steps: (i) Production of a Lex file to generate a custom lexer;

(ii) Implementation of a non-deterministic finite automaton for grammar parsing.
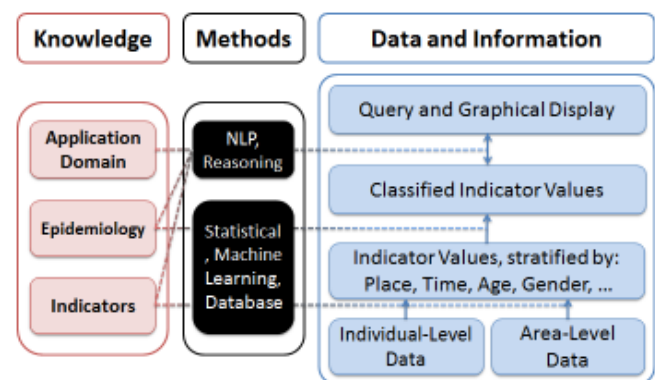
Fig. 1. Schematic architecture for the Population Health Record (PopHR), illustrating how knowledge, methods, data, and information are maintained separately, but related within the system. Knowledge is encoded using an ontology and the methods interact with the ontology to obtain the knowledge needed to process data and facilitate the interpretation of information. In the public health domain, where the PopHR is currently implemented, relevant knowledge relates to the determinants and outcomes of diseases.

REFERENCES

Buckeridge DL, Izadi MT, Shaban-Nejad A, Mondor L, Jauvin C, Dubé L, Jang Y, and Tamblyn R. An infrastruc-ture for real-time population health assessment and moni-toring. IBM Journal of Research and Development, 2012, 56(5): 2.

Shaban-Nejad A, Okhmatovskaia A, Izadi, MT, Naderi N, Mondor L, Jauvin C, Buckeridge DL. PHIO: A Knowledge base for Interpretation and Calculation of Public Health Indicators. In proc. of the 14th World Congress on Medical and Health Informatics (MedInfo'13), Copenhagen, Denmark, 20-23 Aug 2013.

[*] To whom correspondence should be addressed:
  arash.shaban-nejad@mcgill.cs