e-Health – For Continuity of Care
C. Lovis et al. (Eds.)
© 2014 European Federation for Medical Informatics and IOS Press.
This article is published online with Open Access by IOS Press and distributed under the terms of the Creative Commons Attribution Non-Commercial License.
doi:10.3233/978-1-61499-432-9-692

Enabling Person-Centric Care using Linked Data Technologies

Spyros KOTOULAS [a,1], Walter SEDLAZEK [b], Vanessa LOPEZ [a], Marco SBODIO [a], Martin STEPHENSON [a], Pierpaolo TOMMASI [a] and Pol MAC AONGHUSA [a]

[a] *IBM Research*
[b] *IBM Cúram*

Abstract. Patient-Centric Care requires comprehensive visibility into the strengths and vulnerabilities of individuals and populations. The systems involved in Patient-Centric Care are numerous and heterogeneous, span medical, behavioral and social domains and must be coordinated across government and NGO stakeholders in Health Care, Social Care and more. We present a system, based on Linked Data technologies, taking first steps in making this cross-domain information accessible and fit-for-use, using minimal structure and open vocabularies. We evaluate our system through user studies.

Keywords. Patient Centered Care, Smarter Care, Information Sharing, Integration

Introduction

Organizations seeking to improve health outcomes and lower costs are facing unique challenges such as aligning care delivery to population needs, creating and managing holistic, individualized care plans and care coordination to produce positive and sustainable outcomes at reduced cost. In [1], it is reported that 5% of individuals face complex issues spanning multiple domains and accounting for 50% of the cost. Identifying these individuals early on is key to reducing costs. The impact of social determinants for health dictates that multi-domain information is needed for holistic and individualized care delivery [2]. Furthermore, coordination across care agencies and stakeholders requires an integrated view of the individual, their vulnerabilities and their environment [3].

The common denominator is the need for *fit-for-use* information spanning *multiple domains*. In general, needs span six core areas: health, food, shelter, safety, education and income. Potential information sources are diverse and numerous (e.g. the American Hospital Association numbers 5724[2] members and the number of homeless shelters surpasses 4000[3]). The complexity of Health Care data is vast and Social Care systems have a very broad scope.

Relevant use-cases are abundant: In a New York hospital, a survey has shown that 9.2 minutes out of a 15-minute doctor's visit were spent on social needs, crowding out clinical care [4] and illustrating that "social context" of individuals is critical in improv-

[1]Corresponding Author: Spyros Kotoulas, IBM Technology Centre, Dublin 15, Ireland; E-mail: spyros.kotoulas@ie.ibm.com.
[2]http://www.aha.org/research/rc/stat-studies/fast-facts.shtml, retrieved 19/04/2013
[3]http://www.shelterlistings.org/

ing care (for example, consider asthma triggered by sub-standard housing, depression or chemical dependency affecting medication adherence, lack of food impacting diabetes). A single social worker may be responsible for thousands of people [4]. Providing timely, relevant, multi-dimensional and fit-for-purpose information about vulnerability to all care workers is critical.

1. Methods

Typical Health Care data integration approaches use an all-or-nothing model, i.e. data is either part of the (mediated) model or not accessible at all [5]. We argue that this is unrealistic in the Care Coordination domain and propose a Linked Data-based approach, in which a *minimal model is used to provide a high-level navigation structure* and users are able to *further explore* the entire data space, adapted to user expertise. We concede that a shared model capturing all information that needs to be shared is unrealistic. We develop an approach relying on *readily available vocabularies or ontologies* and a *navigational model* for minimal integration. We use *exploration* to access the rest of the information, based on this model. In addition, information is made *fit-for-use* across specialist domains by exploiting domain ontologies, published as open data.

1.1. Vocabularies and Ontologies as drivers for integration

Given that we do not develop a full model for the data to be exchanged, or at least a single centralized model, how do we effectively retrieve and process relevant data? In our approach, we rely on a set of *reference ontologies*, acting as integration points across systems and providing context for users, although we make little use of their formal semantics. In general, these ontologies should have good coverage of the domains (e.g. health, social services, homelessness), but we do not expect them to be fully integrated or to cover the full spectrum of the information exchanged. In addition, re-using existing ontologies reduces both the development and integration cost. Some examples of the ontologies/taxonomies we have used are WSG84[4] and the Time Ontology[5] for spatiotemporal representation, the human disease ontology[6], FOAF and VCARD for basic personal information, the family ontology [6], the social care subject taxonomy[7] as well as ontologies derived from HL7.

1.2. Minimal model-driven presentation

Although reference models are enough to provide basic linking of information, they make information consumption challenging, since they do not capture the relative importance of various pieces of information or the user task, as we will also show in our evaluation. On the other hand, a presentation layer, such as a collection of views that can be rendered on the screen, would nullify any advantage of not having a rigid, central model and would essentially reduce the addressable information space to what is covered by those views. We propose using a business rule-like structure to unify models and help users quickly assess the severity and navigate through vulnerabilities of individuals. Such a structure can be seen in Figure 1. At the top level, we have a logical model, consisting, in this case, of a hierarchy of weighted factors contributing to an individual's vulnerability. Each node on top is connected by means of SPARQL queries to the Semantic Layer

[4]http://www.w3.org/2003/01/geo/
[5]http://www.w3.org/TR/owl-time/
[6]http://disease-ontology.org/
[7]http://www.scie.org.uk/publications/misc/taxonomy.asp

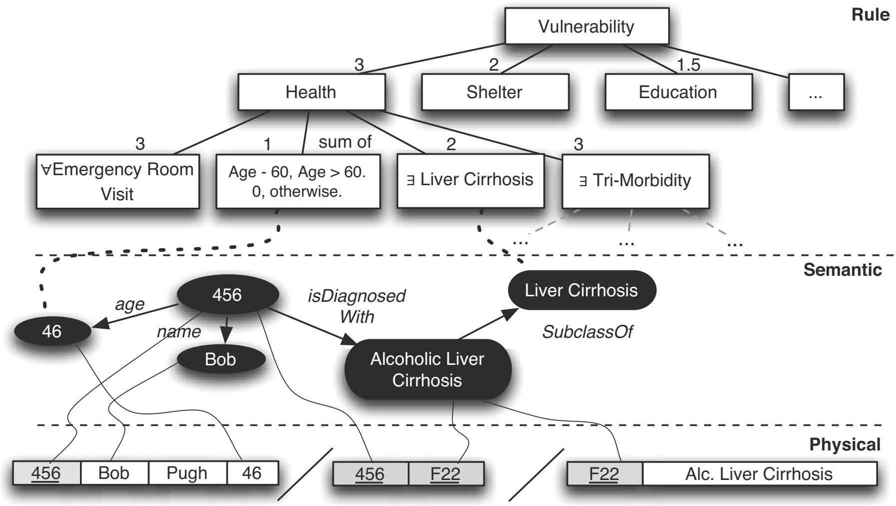

Figure 1. Example model for vulnerability

in the middle. The Semantic Layer abstracts from the representation of the information in the source systems (systems for Social Care, EHRs etc). We elaborate using Figure 2:

Conciseness. The structure shown on Figure 2(α) shows a concise navigational structure of the vulnerabilities of an example individual. The circles represent the relative importance of the vulnerabilities in a set of dimensions that are deemed important in the domain of the user (in this case, a multi-disciplinary team worker). Clicking on a circle expands it so as to display its contributing factors. In this way, the user gets an immediate impression of the key weaknesses of an individual (in this case, it would be problems regarding Health and Food) and retains the capability to explore further (in this case, the user has chosen to expand on the main factors affecting income).

Exploration. Our reference ontologies and navigation structure do not cover the entire information space, since we impose no restriction to the models used by each organization. To get more fine-grain information, an exploration pane (Figure 2(β)) is used to navigate the entire space, based on the Linked Data structure and potential ontology overlap. In the example, the user can see that the individual is receiving child benefit amounting to 170 euros weekly. In addition, it displays relevant information from Linked Open Data, such as complicating factors for diseases (not shown in example).

Fitness-for-use. Specialist knowledge inhibits information consumption across domains. For example, although *Pica Disease* may be the most appropriate term to describe a condition to a physician, it has little fitness-for-use for a social care worker. For the latter, the super-class *Eating Disorder* is more useful. In our system, based on the clicks of each user group when exploring data, we adapt the terminology presented by default (not shown in example).

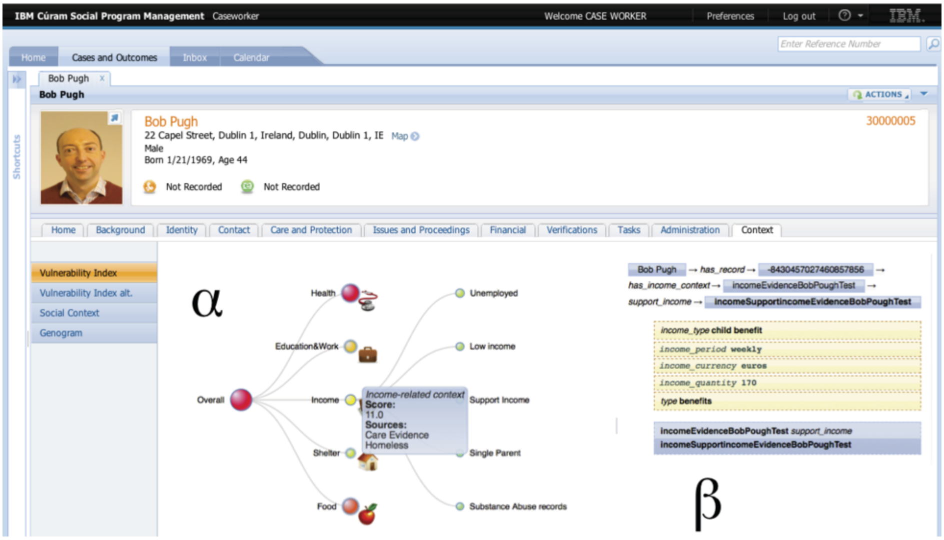

Figure 2. User interface for navigation and exploration

1.3. *Realization and Deployment*

We developed a prototype implementation on top of enterprise IBM systems for Social Care and Health Care. The IBM Cúram Social Program Management Platform[8] is a business and technology solution that delivers prebuilt social program components, supporting users from multiple disciplines. IBM software Patient Care and Insights[9] integrates and analyzes patient information providing information such as care plans, risk assessment and clinical summaries. We have integrated information from the above systems, along with Open Data sources and other data (e.g. information typically managed by a homeless shelter), for a total of 3 relational databases (some enterprise systems have multiple databases), 1 non-relational database, 1 RDF data source and 12 ontologies.

2. Results

We have performed a user study to evaluate the time required to get a business result. We use a standard Web-based exploration tool for RDF[10] as a baseline, also addressing the entire information space (i.e. using the same input). Table 1 shows a set of questions, taken from an actual intake for coordinated care, for general vulnerability and risk of homelessness. Also shown is the average time to retrieve an answer using our system and the average time to retrieve an answer using the baseline Web-based RDF exploration tool, using a randomized question order. Results for one-way ANOVA ($F(1,104)=8.366, p=0.005$) for 10 respondents indicate a statistically significant difference. The baseline system took 54.1% more time to retrieve an answer than our system, on average. The number of clicks was comparable while, for the baseline, on 3 occasions the user gave an incorrect answer and on 3 more they gave up without giving an answer. We note that our models and system have in no way been adapted for the particular tasks.

[8] `http://www-03.ibm.com/software/products/en/social-programs/`
[9] `http://www-01.ibm.com/software/ecm/patient-care/`
[10] http://wifo5-03.informatik.uni-mannheim.de/pubby/. Deployed on the same platform as our approach.

ID	Domain	Question	Baseline	Ours
A1	Employment	When did you last work?	51 *sec*	**43** *sec*
A2	Education	What is the highest level of education completed?	85 *sec*	**32** *sec*
A3	Income	What type of services is your family receiving?	76 *sec*	**40** *sec*
A4	Housing	What is your current housing situation?	**37** *sec*	51 *sec*
A5	Substance abuse	Have you ever had a problem with alcohol or drugs?	**42** *sec*	51 *sec*
A6	Physical health	Do you have any health or medical conditions?	157 *sec*	**42** *sec*
B1	Housing	What is your primary reason for homelessness?	41 *sec*	**15** *sec*
B2	Housing	How long have you been living on the street?	**32** *sec*	46 *sec*
B3	Child Welfare	Do you currently have children in foster care?	70 *sec*	**62** *sec*
B4	Mental health	Do you have emotional, personal problems or stress?	30 *sec*	**19** *sec*

Table 1. Intake questions. A1 to A6 refer to vulnerability, B1 to B4 refer to homelessness.

We show in isolation that adapting terminology yields improved results for the user. We start from a condition in the Human Disease Ontology, and ask the user whether they are familiar with the term. If not, we select the more generic condition (i.e. the parent in the hierarchy of diseases) and repeat. For 10 non-specialist users (with a computer science background), and 21 distinct terms, there was user agreement on what the understandable term is in 120 out of 210 observations. As a consequence, adapting the original term to the most identifiable term means that 57% of the time, users would immediately identify the term. When we do not, the corresponding percentage was 21%.

3. Discussion

We have shown some first steps towards information sharing for Person-Centric Care across highly diverse and challenging domains, using Linked Data. Our results show that we can address the same information space as a generic graph exploration approach while reducing the time to reach a business result. To put our results into context, consider the study reporting that 9.2 minutes were spent on social needs in each doctor's visit [4].

There are many remaining challenges in the area. Although current practice in Care Coordination follows a blanket consent model (i.e. persons need to agree to share all their data across care agencies), privacy and consent management is a challenging and exciting direction. In addition, better exploration interfaces that hide the complexity of the RDF graph, yet remain generic, are needed. User context (e.g. location) can be taken into consideration and the context captured by our system can be used as input for analytics.

References

[1] Conwell LJ, Cohen JW. Characteristics of persons with high medical expenditures in the US civilian non-institutionalized population, 2002. Medical Expenditure Panel Survey, Agency for Healthcare Research and Quality; 2005.

[2] Marmot M, Wilkinson R. Social determinants of health. Oxford University Press; 2009.

[3] Kodner DL, Spreeuwenberg C. Integrated care: meaning, logic, applications, and implications–a discussion paper. International journal of integrated care. 2002;2.

[4] Onie R, Farmer P, Behforouz H. Realigning health with care. Stanford Social Innovation Review. 2012;10:28–35.

[5] Nodine M, Hee A, Ngu H, Bohrer W. Semantic Brokering over Dynamic Heterogeneous Data Sources in InfoSleuth. In: ICDE; 1999. .

[6] Stevens R, Stevens M. A Family History Knowledge Base Using OWL 2. In: OWLED; 2008. .