Link2Outcome: Coordinating Social Care and Healthcare using Semantic Web Technologies

Spyros Kotoulas, Vanessa Lopez, Pierpaolo Tommasi, Martin Stephenson, Marco Luca Sbodio, Wei Jia Shen, Gang Hu, Veli Bicer, Anastasios Kementsietsidis, M. Mustafa Rafique, Jason Ellis, Thomas Erickson, Kavitha Srinivas, Kevin McAuliffe, Guo Tong Xie, and Pol Mac Aonghusa

IBM Research

Abstract. We present a semantic infrastructure to augment existing enterprise applications with context coming from external systems. We showcase our infrastructure using a use-case around Care Coordination, based on a set of IBM solutions. An up-to-date version of this document, with additional media, can be found at `http://ibm.co/14nUY3i`.

1 Introduction

Healthcare and Social Care are unique domains in terms of cultural importance, economic magnitude and complexity. On a cultural level, the level of advancement of a society is often measured in terms of protection of the less able. In economic terms, for 2009, total expenditure on healthcare in the United States was 2.6 trillion USD or 17.4% of the GDP[1]. Total expenditure on social care was 2.98 trillion USD or 19.90% of the GDP[2]. In terms of US Federal government expenditure, social security, medicare and medicaid amount to 45% of total spending. In terms of complexity, organizations that are involved in providing social and medical care are numerous and span a very wide domain. For example, AHIP, the trade association of health insurers numbers some 1300 members[3]; the number of hospitals registered with the American Hospital Association is 5724[4] and the number of homeless shelters surpasses 4000[5]. In addition, medical information is vastly complex: Nuance reports that LinkBase®[6] contains more than 1 million concepts. Social care depends on information from a very broad domain, ranging from criminal records to housing.

Coordinating social care and health care has been identified both as a major pain point and a significant opportunity in modern health and social systems [1]. Several studies have shown that costs can be contained and outcomes improved

[1] http://dx.doi.org/10.1787/888932523215
[2] http://www.oecd.org/els/social/expenditure
[3] http://www.ahip.org
[4] http://www.aha.org/research/rc/stat-studies/fast-facts.shtml, retrieved 19/04/2013
[5] http://www.shelterlistings.org/
[6] http://www.nuance.com/for-healthcare/resources/clinical-language-understanding/ontology/index.htm

with a more holistic approach to care [2]. As a simple motivating example, consider an individual quartered in inappropriate housing while suffering from a relatively minor health issue, aggravated by the housing condition. As a result, the given individual frequently resorts to visiting emergency rooms, resulting in significant cost to the healthcare system and a less effective treatment. By itself, the housing situation does not warrant state intervention. Nevertheless, resolving it would dramatically improve the health situation, resulting in a better quality-of-life for the individual and lower costs for the health system.

Even in this simple example, the challenges presented are significant: *How do we access information in disparate systems, storing vastly heterogeneous information on various infrastructures? How do we cope with policy constraints disallowing replication or centralization of data? How do we abstract from the information and representation complexity?*

We propose a novel technical solution to augment applications with cross-domain context, in the domain of social care and health care based on business rules and contextual exploration. We claim that Semantic Technologies can uniquely address these problems because: (a) The distributed nature of RDF allows access to integrated information across silos. (b) Explicit and global semantics allow us to ground business rules across systems. (c) The distributed and incremental data integration paradigm advocated by linked data can help coping with the complexity of the data.

We present a demonstrator of a system that supports three key use-cases for this domain: (a) Displaying a view of the combined needs across several dimensions for a given person and people in their social context, based on a set of business rules. This allows a social/health worker to quickly assess the situation of an individual. From a knowledge management perspective, it requires grounding a set of business rules across several ontologies and instance data in several data sources. (b) Exploration of the context to surface information not directly covered by the business rules. Given the heterogeneity of the domain, the user will most likely need additional information around a given individual. Our demonstrator uses the business rules as a navigational aid to explore the semi-structured information. (c) Insight into the social context (e.g. family) of the individual, so as to get a more spherical view of related needs and competencies.

2 Scenario

In this section, we showcase the system through a business case for integrated care. The users of the system may have various roles within the health and social care administration, and they are not familiar with semantic technologies. We outline the features and functions the system provide in two use cases, as well as illustrating the central role played by semantic technologies to derive meaningful and useful information from heterogeneous and distributed data sources, published by different agencies or authorities.

In this use-case, we describe two main roles: the primary care provider (physician) and the care manager (or coordinator) [7]. In our particular case, a primary care provider can refer a person that potentially has vulnerabilities spanning beyond healthcare to a care manager and the latter can assess the needs of person and coordinate interventions across different dimensions (e.g. social, clinical).

As an illustrative example, a patient named Bob visits the physician and is diagnosed with pre-CHF (Congestive Heart Failure) condition and obesity. He has an unhealthy diet and sedentary life style. The physician can access the different medical information from various systems to obtain a complete clinical view. The primary care provider can see vitals (e.g., weight, pressure), symptoms (such as cough, wheezing and shortness of bread), medications, allergies, lab tests, care plans, and family history of CHF as well as analytics results (risk stratification). His care plan requires implementing lifestyle changes, such as healthy heart diet and exercise.

Bob's clinical issues are complicated with social issues - unemployment - and signs of depression. Thus, Bob is referred to integrated care by the physician to help him achieve necessary lifestyle changes and receive proper counseling and services (e.g., social benefits or referral to a nutritional programme). Bob has been registered on the system and he has given his consent to accept services through the health home[8]. The assigned care manager can review Bobs medical, social and behavioral situation, across both health and social care systems. Individual risks variables that influence care include: social needs (e.g., does the person own a house, are they homeless?), personal details (age, gender, ethnicity), health history, entire family situation and even places or communities where she belongs (deprivation and morbidity indexes). Based on the information above, the care manager can make informed decisions concerning treatments plans, involving additional services (e.g. social service, income support) or including family members.

2.1 Assessing the needs and vulnerabilities of an individual through business rules

Evidence-based models and vulnerability indexes are used to help care coordinators in understanding individual cases. Although there are many ways to models vulnerability indexes, in our work, we have chosen a model similar to the ones typically found in Key Performance Indicators (KPIs), designed by domain experts to ground business rules to data. For care coordination there are six main dimensions of interest i.e., health, safety, income, shelter, food and education.

A hierarchical tree of factors contributes to each vulnerability. Considering the index for food vulnerabilities, the factors contributing to food insecurity span through several dimensions from both clinical and social sources. For instance,

[7] A care manager is responsible for coordinating various aspects of care for a given person, with the goal of providing better outcomes across health, social and other needs.

[8] A health home is an organization to coordinate care

Fig. 1: Vulnerabilities of Bob

low income, family size, physical or behavioral health problems, in particular, obesity, disabilities or chronic diseases that affect diet (e.g., diabetes and CHF); substance abuse, which can affect eating habits and money available for food; poor housing, which affects access/preparation/store of food; lower levels of education; social services received, in particular food stamps or school lunches; safety issues, such as domestic violence or youth in foster care; or other barriers related to the population, such as single parent households and access to food in impoverished urban and rural areas.

External open data sources are used to generalize specialized terms (e.g., smoking and alcohol abuse are of type substance misuse and mental health problems, pica disease is an eating disorder, etc.), providing a common vocabulary and meaningful relationships to other relevant entities. A social care ontology has been created based on the UK social care online taxonomy [9] - a large open social care taxonomy of more than thousand social care topics created to provide a controlled vocabulary to index and annotation social care content topics. Terms in this taxonomy are used as common anchors across sources.

Figure 1 shows the vulnerability index for Bob on the left. At a glance, the care manager can see that Bob is facing problems in multiple dimensions, with Health and Food being the most important ones. The care manager can also see that the food problems are related to clinical problems.

2.2 Contextual exploration

The contextual exploration allows to surface information not directly covered by the business rules and to uncover relevant paths and connections across factors. Additional semi-structured sources can be used to uncover connections between factors and records. For all factors that are also described as entities in DBpedia, we have added links (owl:sameAs) to the corresponding DBpedia term. Furthermore, for each topic linked to a DBpedia term, we extract the social care topics

[9] http://www.scie-socialcareonline.org.uk/

that are also mentioned in the descriptions of the corresponding Wikipedia page, so as to add links to commonly co-related factors. These correlations, although do not necessarily imply causality, can be used to uncover and present relevant content for the user to explore.

Figure 1 shows the exploration pane on the right. The care manager can drill into the specific vulnerabilities of Bob. Hovering over a node provides a description of the node and provenance information while clicking on a node updates the navigation pane on the right. For example, the care manager can see that Bob smokes, and that smoking is a type of substance abuse. More importantly, the care manager can see that smoking is correlated to CHF, so it is particularly important to try to convince Bob to implement a life-style change (i.e. quit smoking).

2.3 Social Context

In order to make an informed decision, it is often required that the care manager has visibility at the situation of the social circle of the person. The care workers can find more information about the household and social connections by looking at the genogram [3] visualization, that provides visibility into family relationships.

For Bob, the genogram will show that he is divorced and living with his two kids. It can also show the extended family (parents, brothers) and family medical history (Bob mother has also had CHF and there is a history of depression in the extended family).

In coordinated care, it is sometimes beneficial for other family members to be involved. To this end the care manager needs visibility on the situation of family members. The Social Context tab shows related information concerning the vulnerabilities, relationships and home locations of family members.

The care manager can look at the family situation of Bob, using the social context analytics to find someone in the family with no vulnerability issues and supportive to help Bob with all lifestyle changes (Fig. 2). Furthermore, using the map interface, the care worker can also choose the family members living geographically close. She invites Bob's uncle to be part of the care plan team, since he is located close to Bob, his vulnerability scores are low and, in the past, Bob has mentioned that he has a positive relationship with him.

3 Design, Architecture and Deployment

Due to the heterogeneity and complexity of the domain (there is no single schema) we adopted a set of reference ontologies for defining mappings from the different data sources to virtual RDFs. The current different ontologies model the following: (1) family relationships (including temporal relations to model marital status); (2) personal and contact details, reusing widely used schemas such as FOAF and VCARD; (3) social care evidences and case records; (4) a social care taxonomy (representing a hierarchical arrangement of social care topics);

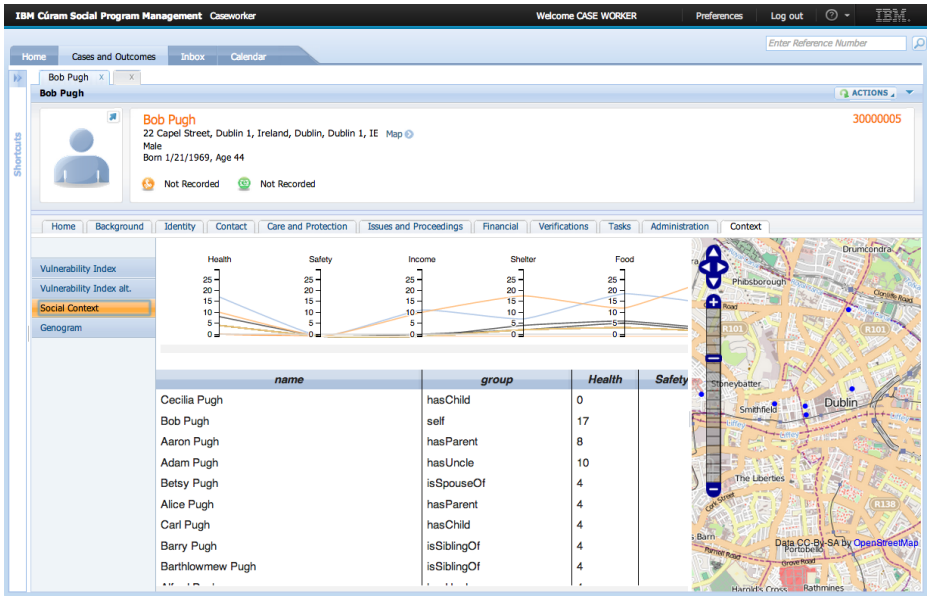

Fig. 2: Social Context of Bob

(5) clinical schemas. The semantic layer hides the distribution and heterogeneity of underlying sources, allowing applications to formulate queries guided by the reference schemas. Data is federated on-demand, only exposing as RDF the data that is needed for care coordination. Data owners retain control of their data respective data (a business requirement in this particular case) and share information by translating distributed SPARQL queries to their proprietary representation at runtime. Although access control is a contested subject and an active research field, due to the highly sensitive nature of the shared information, in this work, we are working on the prevalent model of blanket consent of participants in health home programmes (i.e. consent to share all information regarding an individual across all organizations).

An enterprise architecture supporting our approach is shown in Fig. 3. Due to space restrictions, we describe only the components necessary to understand the basic operation of the system. Web-facing services use a set of REST services, implemented on a custom application running on IBM WebSphere Application Server. The main components for these services are the *Node registry*, which tracks nodes in the *Federated Query Engine*, the *View definitions*, that are used to project information out of the graph model for use by analytics widgets and UI elements. *Data Sources* are exposed as virtual RDF, using SeDA, an IBM technology to execute R2RML mappings. The virtual RDF Data Sources, the Metadata Repository and the Ancillary Indexes are accessed through the *Federated Query Engine*, providing transparent access to the distributed information. All core components in this architecture can be clustered, for high availability and performance.

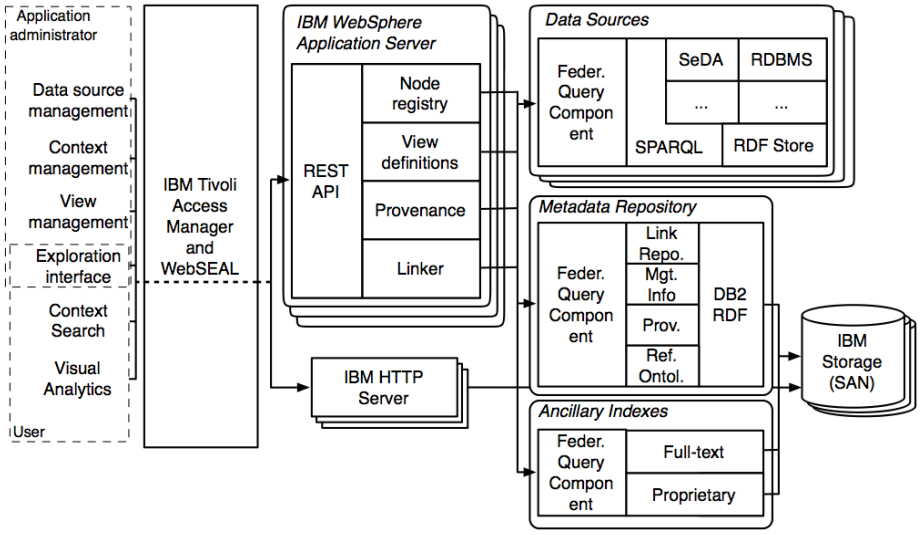

Fig. 3: System architecture

We will present a proof of concept based on the above architecture, integrating a set of IBM solutions for clinical and social program information: **IBM software Patient Care and Insights** provides data driven population analysis to support patient centered care processes. It integrates and analyzes the full breadth of patient information sourced from multiple systems and different care providers. It stores three categories of data: extracted patient medical history called clinical summary; medical data analytics results from an analytics component called care insights and personalized electronic care plans. **IBM Cúram** is a business and technology solution to help social program organizations provide optimal outcomes for citizens, satisfy increasing demand, and lower costs for organizations. We will demonstrate a user interface that integrates clinical information and a user interface integrated with **IBM Cúram** to augment its care coordination capabilities with context coming from other enterprise applications.

References

1. Rigby, M., Hill, P., Koch, S., Keeling, D.: Social care informatics as an essential part of holistic health care: A call for action. I. J. Medical Informatics **80**(8) (2011) 544–554
2. Peikes, D., Chen, A., Schore, J., Brown, R.: Effects of care coordination on hospitalization, quality of care, and health care expenditures among medicare beneficiaries. JAMA: the journal of the American Medical Association **301**(6) (2009) 603–618
3. Jolly, W., Froom, J., Rosen, M., et al.: The genogram. The Journal of family practice **10**(2) (1980) 251

Criterion	Rating	Explanation
End-User Application	H	Web-based enterprise application. Main audience is specialists, although the citizen portal is available to the general public.
Diverse ownership or control of sources	H	We use data for Social Care, Health Care and Linked Data. Data is controlled and located in the IT infrastructures of multiple organisations.
Heterogenous sources	H	Our system handles RDF data as well as multiple, complex relational databases.
Real-world data	M	Our system is deployed on top of real, complex IBM solutions (IBM Curam and Smarter Healthcare). The real data is extremely sensitive, thus all data presented is synthetic.
Use of Semantic Web Technologies	H	RDF is used as the interchange format and to provide grounding for terms.
Data processed to derive useful information	H	We show the business value of our system through a motivating use-case.
Suitability of Semantic information processing	H	Semantic technologies are used for flexible data representation and incremental integration and modeling. Traditional data management tools (e.g. RDBMS) would be very cumbersome to use, since we are lacking a global schema.
Attractive and functional Web interface	H	We make extensive use HTML5 and visualization tools to produce an attractive interface. The user interface is highly customizable.
Scalable application	H	The enterprise software components in our application can be clustered for scalability. More distributed systems can be added transparently. Ontologies can be changed.
System evaluation and validation of results	M	We validate our approach through a business use-cases. We are currently working on obtaining quantitative results.
Novelty in applying semantic technology	H	We apply semantic technologies to a new and highly challenging domain. The combination of complexity in Healthcare and diversity is Social care is unique.
Functionality goes beyond information retrieval	H	Our system is used to integrate and provide insights from several sources.
Commercial potential	H	We show the business value of our system. Our demo is based on an actual IBM solution.
Contextual information for ratings or rankings	H	Information for vulnerabilities is acquired from several sources and aggregated to provide meaningful results.
Multimedia documents	N	-
Use of dynamic data	H	The system pulls data live from multiple enterprise applications.
Accurate results	H	Since the data sources are enterprise systems, data accuracy is generally high.
Support for multiple languages and accessibility on a range of devices	M	IBM Curam is currently available in 8 languages. Current application is not mobile-enabled.