FAQIH: Framework for Agent-based Query-enabled Integrated Information for Health and Nutrition

Yuri A Tijerino
Kwansei Gakuin University
Kobe-Sanda, Japan
yuri@tijerino.net

Tarek Helmy El-Bassuny
King Fahd University of Petroleum and Minerals
Dhahran, Saudi Arabia
helmy@kfupm.edu.sa

Jeffrey M. Bradshaw
Florida Institute for Human and Machine Cognition
Pensacola, FL, USA
jbradshaw@ihmc.us

Abstract

As the processes of urbanization and globalization continue to gain momentum, issues relating to interdependencies among water, food, nutrition and public health policy have increased in complexity and scope. Thus, it is no longer sufficient for subject-matter experts to develop models for each of their underlying domains of expertise (e.g., nutrition, healthcare, and land and water management) independently. Both expert stakeholders and ordinary citizens would benefit from reliable and continually up-to-date models on which to base their decisions. In this preliminary overview paper, the authors propose to develop a Framework for Agent-based Query-Enabled Integrated Information for Health and Nutrition (FAQIH) to support healthcare professionals, policy makers, and citizens to obtain the knowledge they seek.

1. Introduction

The FAQIH framework supports decision-making by integrating structured and semi-structured, distributed information sources using ontologies on a variety of ubiquitous delivery platforms, such as PCs (home and office), mobile and wearable devices (laptops, mobile phones, PDAs), and location-based information kiosks [7]. Our initial development focuses on health management for individuals by providing active and passive advice on water, nutrition requirements and exercise choices, and will later extend to interdisciplinary problems faced by diverse communities of expert professionals. Advice will be supported through a networked community of software agents that collaboratively attempt to meet individual goals while meeting policy constraints defined by individual citizens, healthcare providers, public health policy makers, and other stakeholders.

2. Information fusion

Semantic web technologies are increasingly becoming the choice when it comes to integrating distributed information sources that can be exploited through software agents. In particularly, ontologies provide formal semantics, which allow agents to perform reasoning tasks based on the distributed and disparate information sources on the Web. However, as proposed in this framework, the use of ontologies need not be constrained to resources on the traditional Web, but instead they can be weaved to include other devices such as mobile phones, home appliances, RFID tags, and information systems, such as inventory systems, databases and any kind of information appliance or sensors available to the general public and decision makers. As proposed in this framework, ontologies can provide the means for describing and reasoning about sensor data, objects, relations and general domain theories.

3. Ontological Foundation

In computer science and ontology refers to explicit specification of a con-ceptualization [5] Commonly, this involves a vocabulary for describing the concepts that exist in a field of knowledge, the relationships that exist between them, and the constraints imposed on their meanings. One of the key foundations of this framework consists of allowing agents to reason on a variety of ontologies, starting from personal in-formation ontologies to domain specific ontologies.

3.1. Personal Information Ontologies

We place particular emphasis on the development of personal information ontologies to allow the efficient capture, storage, retrieval and discovery of situation-relevant information from data such as email, contact information, scheduling, to-do-lists, and other multimedia information available in personal mobile devices.

We propose policy-based agents capable of semantic understanding outside the context of the traditional Web. This need is addressed by the FAQIH framework through personal information ontologies capable of capturing the idiosyncrasies of the way users employ their personal devices to communicate, search, store and retrieve information.

In order to semantically understand personal in-formation, the framework will employ three aspects of ontology engineering at the personal level, namely, personal ontology creation, ontological annotation of personal information, and ontological mapping of situation-relevant information queries.

3.1.1 Ontology Generation: For personal ontology creation the framework will use an adaptation of ongoing work on ontology generation from structured data using information extraction ontologies [8]. First, the framework calls for the engineering of a common personal device ontology that describes the common aspects the device and its general use. A common ontology is engineered for information stored in personal mobile devices, such as smart phones, using a combination of OWL DL for general text-based information such as email, Web pages, personal notes and other text documents; MPEG-7 for multimedia files such as music, video clips and photos; and FOAF for information found in address books, browsing histories, call histories, to-do-lists, agendas and the such. These three formats will be integrated into OWL DL to form a single merged ontology. Once this common personal device ontology is created the framework will employ the information extraction ontology generation approach [9] to create a personalized version of the ontology for each user by mining other structural aspects present in the information stored in the device. This engineering process results in a coherent personal information extraction device ontology that captures the semantic idiosyncrasies of the information stored in the personal mobile device. Figure 1, illustrates an example of how the device ontology can be mapped across individuals.

3**.1.2. Annotation**: For annotation, the framework will use the personalized version of the information extraction ontology to annotate the information in the device using

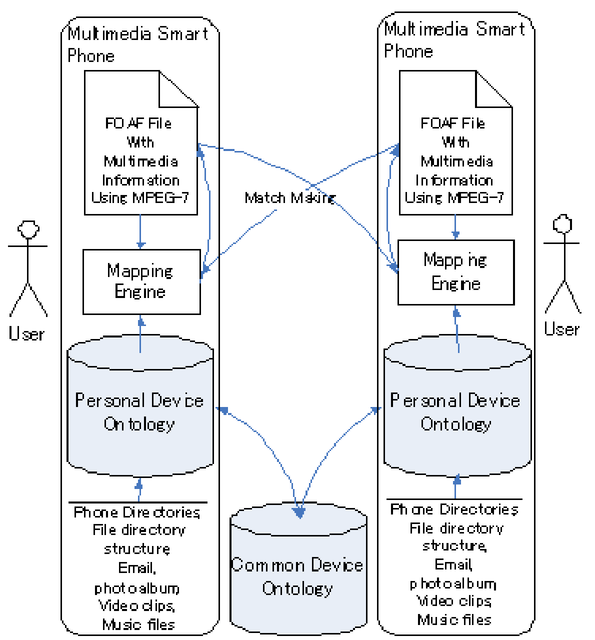

Figure 1. Mapping of personal mobile device ontologies.

data frames [3 & 4] for text-based information and data frames for multimedia information in the device. Media frames are an extension of data frames that incorporate the multimedia description capabilities found in MPEG-7. The framework also employs the personal information extraction ontology for the actual creation of a simple FOAF profile from the address book, agenda, call history, to-do-list, and email headers. Since the FOAF profile is intended for public sharing, the framework interactively allows the user to modify it, add and edit privacy and security constraints in an interactive manner.

3.1.3. Query Understanding: Since at this stage the framework uses OWL DL for ontology representation, it can support simple queries that do not require reasoning. However, policy-based agents are also capable of reasoning on annotated information in the devices. Queries are parsed and mapped to the personalized ontology using a approach based on information extraction ontologies [1] that does not require natural language processing for semantic understanding and mapping of the queries.

3.2. Related Ontological Efforts

In our goal to provide a viable framework to facilitate decision-making across the nutritional, biotechnology and healthcare domains through ontologies, we are not alone. There are numerous government, and non-government organizations already working to provide useful modeling tools and plat-forms. In the nutrition domain there exist many well-established and authoritative controlled vocabularies, such as the AGROVOC Multilingual Thesaurus[1] and the AGROVOC Concept Server[2] which is an ontology-based framework that provides vocabulary, guidelines and standards to facilitate the integration of agriculture data from different sources scattered throughout the world. Other food-related ontologies include LanguaL [6], an international framework for food description originally developed by the Center for Food Safety and Applied Nutrition (CFSAN) under the United States Food and Drug Administration (FDA), and the FAO-related INFOODS Guidelines for Describing Foods [9] is a systematic approach to describing foods, designed to facilitate international exchange of food composition data.

Besides the above ontologies, we plan to explore whether the MyPyramid dietary guidelines developed by the US Department of Agriculture (USDA)[3] could be used to create a category of INFOODS-based food descriptions. This category, in turn, could be turned into a full-blown food ontology using tools from our framework. Moreover, it is easy to imagine how specific guidelines

[1] http://www.fao.org/aims/ag_intro.htm
[2] http://www.fao.org/aims/aos.jsp
[3] http://www.mypyramid.gov/

from the FDA's Center for Food Safety and Applied Nutrition (CFSAN), which dictates labeling regulations for foods, could be used to generate an initial set of agent policies to address the safety concerns of individuals.

4. Policy-Governed Software Agents

Agents should be able to process and semantically interpret the contents of the personal device and Web pages, and exchange the semantic policies/ontologies freely with each other. Our ontologies can enhance the functionality of the devices and Web in many ways. They can be used in a simple fashion to improve the accuracy of searching the contents of the devices and the web. For instance a Web search agent can look for only those pages that refer to a precise concept instead of all the ones using ambiguous keywords, while a personal device search agent can search for information in the device based on the device ontology. More advanced applications will use ontologies to relate the information on a page or device to the associated knowledge structures and inference rules. This makes it much easier to develop agents that can tackle complicated questions whose answers do not reside on a single Web page. Suppose I wish to find Dr. Bradshaw who I met at the IAT conference last year. I don't remember his first name, but I remember that he works for the west Florida University and his main interest is the autonomous agents. An intelligent Web search agent can sift through the device for people whose name is "Bradshaw" and find the ones that mention working for the west Florida University and follows links to the Web pages of their publications to track down if any are in IAT conference.

The increased capabilities afforded by software agents are both a boon and a danger. By their ability to operate independently, agents can perform tasks that would be impractical or impossible earlier. On the other hand, this autonomy has the potential of compromising information privacy or effecting severe damage to operations (e.g., through buggy or malicious code). Therefore, FAQIH framework will implement the KAoS policy-based agent services to address security and privacy concerns. The KAoS services framework relies on a collection of user-extensible ontologies designed to assure that agents and other software components will always operate within the bounds of explicit policies and will be continually responsive to human control [2]. To adapt to changing situations, such as disease outbreaks, the policy framework supports dynamic runtime policy changes, and not merely predetermined static policies.

5. Sample Problem Scenario

The ups and downs in energy and nutrition requirements during various phases of pregnancy and nursing can place significant stress on a mother. In our scenario, Kamila is a young 20-year-old mother expecting her first child. As is her extended family's tradition, she is receiving much needed advice from her elder sisters, her mother, and other female relatives. However, she is in the local market doing her own shopping for the first time. She connects to the FAQIH framework using her mobile phone, and checks her nutrition requirements for the day, in order to buy the correct ingredients for supper. Today Kamila needs to cook for her visiting parents-in-law and husband, but needs to make sure that she can create a menu that fits her nutrition requirements, while still impressing her in-laws and satisfying her husband. For an appetizer, she decides to start with *betingan makdous* (stuffed eggplant pickle in olive oil). Some *shourabat el queema* (meatball soup) also sounds good. For the main entrée, she decides on al *kharoof bel roaz* (lamb with rice). She inputs the menu into her mobile phone FAQIH interface, and the system retrieves all the nutritional information and warnings related to that menu from the system. In addition to retrieving all the nutritional information about her menu, her personal agent compares it with the specific nutritional plan answering her needs. In this case, the agent discovers that her menu contains too much fat and not enough calcium, so it suggests replacing al *kharoof bel roaz* with *sayyadieh* (fish with rice) and suggests *nooranoush* (pomegranates and yogurt) for dessert. However, FAQIH also warns her not to buy yogurt of a specific brand, because it seems to be causing an outbreak of diarrhea in a close-by locality. In order to perform this comparison, make suggestions and provide a warning, the agent needs to perform collaborative reasoning with the concepts and relationships defined in multiple ontologies, including the nutritional ontology, a recipe ontology, and a general-purpose food ontology.

The nutritional ontology, recipe ontology, and general-purpose food ontology will be developed by exploiting existing thesauri described above and populated with instance data using our unique ontology generation and ontology-based annotation approaches. The resulting ontologies are also complemented through a network of policies defined by the individual to meet her own preferences (including religious and cultural concerns), by her nutritionist or healthcare provider to meet her nutritional needs, and by public health decision-makers to meet requirements imposed by local regulations and specific health concerns.

Although Kamila interacts with FAQIH through her mobile phone while on the go, she can also interact through her home PC, and even through her smart refrigerator, which keeps track of its contents, the nutritional value of each ingredient in the refrigerator, how nutritional values vary over the shelf time, and the expiration dates of all ingredients. Other actors, e.g., nutritionists and healthcare providers, can interact with FAQIH agents through their own interfaces. Such professionals will also be able to search the literature for specific information through their personal semantic search agents. The information residing on various kinds of pages will be annotated with our automated and semi-automated annotation techniques. These agents will be capable of answering free-form queries, benefiting from our ontology-based query disambiguation techniques [3].

6. Architecture

The architecture for the framework is depicted in Figure 2. The architecture consists of:

An interface layer, which can be used to inter-face end users via fixed or mobile platforms. Us-ers can either be information consumers and in-formation producers. Both can be end-users, software artifacts or experts who consume or produce information or knowledge.

A data integration layer, which consists of mul-tiple data integration components for various media formats. This layer also helps to manage the information sources, such as Web, databases, and other distributed or repositories.

An ontology services layer, which allows creation and maintenance of ontologies and keeps tracks of modification and provemance. It also helps in the integration and mapping of various ontologies for more efficient sharing and resuse.

A Semantic layer, which provides tools and me-chanisms to annotate data in the data layer, with ontologies in the ontology layer. It also provides multiple mechanisms to perform automated an-notations, querying and logic reasoning.

A policy layer, which allows the creation and maintenance of agent interaction policies. It also provides tools to allow modification and tracking of policy provenance.

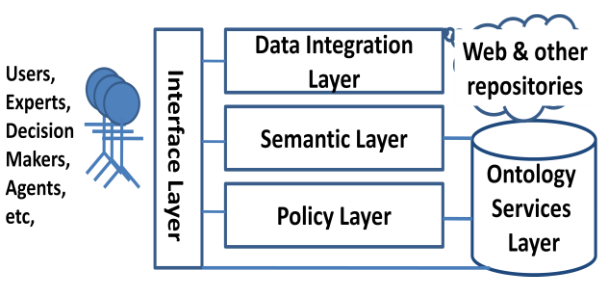

Figure 2. FAQIH Architecture.

7. Concluding Remarks

Although we are not alone in trying to combine knowledge from these fields for problem solving, our ontology generation framework and policy-based agent approach enables us to combine knowledge from various domains dynamically and seamlessly to provide answers to specific questions, both in a proactive and reactive manner through personal devices as well as standard web interfaces. With the proposed framework, the policy and decision makers are provided with a one-stop solution that will empower them to act upon the knowledge in a timelier manner whenever and wherever required.

References

[1] Al-Muhammed, M.J., D.W. Embley, S.W. Liddle, and Y.A. Ti-jerino. 2007. Bringing Web principles to services: Ontol-ogy-based Web Services. Presented at the *Fourth International Workshop on Semantic Web for Services and Processes*, Salt Lake City, UT, 9 July, 2007.

[2] Bradshaw, J.M., P. Beautement, M.R. Breedy, L. Bunch, S.V. Drakunov, P.J. Feltovich, R.R. Hoffman, R. Jeffers, M. Johnson, S. Kulkarni, J. Lott, A. Raj, N. Suri, and A. Us-zok. 2004. Making agents acceptable to people. In *Intelligent Technologies for Information Analysis: Advances in Agents, Data Mining, and Statistical Learning*, edited by N. Zhong and J. Liu, 361-400. Berlin: Springer Verlag, 2004.

[3] Ding, Y., D.W. Embley, and S.W. Liddle. 2007. Enriching OWL with instance recognition semantics for automatic semantic annotation. Presented at the *26th International Conference on Conceptual Modeling (ER 2007)* ONISW Workshop, Auckland, NZ, November, 2007.

[4] Embley, D.W., E.M. Campbell, Y.S. Jiang, S.W. Liddle, D.W. Lonsdale, Y.K. Ng, and R.D. Smith. 1999. Conceptual-model-based data extraction from multiple-record Web documents. *Data and Knowledge Engineering* (November 1999).

[5] Gruber T. R. 1993. Toward principles for the design of ontolo-gies used for knowledge sharing. *Presented at the Padua workshop on Formal Ontology*, March 1993, later published in *International Journal of Human-Computer Studies*, Vol. 43, Issues 4-5, November 1995, pp. 907-928.

[6] Hendricks, T.C. 1992. LanguaL: An automated method for describing, capturing, and retrieving data about food. World Review on Nutrition and Diet 68 (1992): 94-103.

[7] Tijerino, Y.A. 2005a. Next Generation Multimedia Devices and the Semantic Web. In *Poster and Demo Proceedings of 4th International Semantic Web Conference* (ISWC2005). PID-89. Galway, Ireland, 2005.

[8] Tijerino, Y.A., D.W. Embley, D.W. Lonsdale, Y. Ding, and G. Nagy. 2005b. Towards ontology generation from tables. *World Wide Web Journal* 8, no. 3 (September 2005): 261-285.

[9] Truswell, A.S., D.J. Bateson, K.C. Mdafiglio, J.A.T. Pennington, W.M. Rand, and J.C. Kleinsin. 1991. INFOODS guidelines for describing foods: A systematic approach to describing foods to facilitate international exchange of food composition data. *Journal of Food Composition and Analysis* 4, no. 1 (1991): 18-38.