AKTive Food: Semantic Web based knowledge conduits for the Organic Food Industry

Christopher Brewster[1], Hugh Glaser[2] and Barny Haughton[3]

[1] Department of Computer Science,
University of Sheffield, Sheffield, S1 4DP, UK
C.Brewster@dcs.shef.ac.uk
[2] Dept. of Electronics & Computer Science,
University of Southampton, UK
hg@ecs.soton.ac.uk
[3] Quartier Vert, Bristol, UK

Abstract. We present a vision and a proposal for using Semantic Web technologies in the organic food industry. This is a very knowledge intensive industry at every step from the producer, to the caterer or restauranteur, through to the consumer. There is a crucial need for a concept of *environmental audit* which would allow the various stake holders to know the full environmental impact of their economic choices. This is a different and parallel form of knowledge to that of *price*. Semantic Web technologies can be used effectively for the calculation and transfer of this type of knowledge (together with other forms of multimedia data) which could contribute considerably to the commercial and educational impact of the organic food industry. We outline how this could be achieved as our essential objective is to show how advanced technologies could be used to both reduce ecological impact and increase public awareness.

1 Introduction

Much has been made in recent years of the knowledge economy, and the significance of knowledge for economic growth, competitive survival and corporate strategy[1]. In most such contexts, knowledge refers to intellectual property, competitor intelligence, and corporate internal knowledge management and transfer. However, in the wider knowledge economy there exist other types of knowledge which play a major role in economic activity. The great 20th century economist Friedrich Hayek was concerned with the analysis of, what he termed, distributed knowledge in rational economic systems. For him, one of the key issues in making economic decisions was how to utilise knowledge dispersed among many individuals [2]. His approach was to view *price* as a means for conveying a form of distributed knowledge i.e. it acts as a heuristic or summariser of the relevant knowledge which in fact is disseminated across many individuals. In a knowledge economy price is a very narrow vector for knowledge transmission, and there are many contexts in which price alone is no longer sufficient to guide economic decisions. For example, there is a significant segment of the financial markets which is concerned with the provision of "ethical" funds. Equally there is a growing consumer movement, especially in the UK, concerned with the environmental, ecological and ethical impact of everyday choices by consumers. This movement is dependent on specific types of knowledge concerning the products about to be purchased. The organic or ethical consumer cannot rely on Hayek's price alone to guide their economic behaviour.

It is in this context that we propose a project using the technologies provided by the Semantic Web to enable a relatively narrow domain (the organic food industry) to collaborate significantly by exchanging knowledge. Our purpose is to develop a scenario where advanced knowledge technologies are able to contribute both to facilitating restauranteur and consumer decisions, but also to reducing the ecological impact (the environmental footprint) of this particular industry. We propose to use Semantic Web technologies to allow the transmission of a form of knowledge which we shall term the *environmental audit*. This will be a representation of the environmental cost or impact of purchasing a particular product. The organic food industry and its consumers are both very aware of such factors and try to take them into account in their economic decisions. The environmental audit of a product is difficult to achieve and so we believe is a suitable subject for the application of Semantic Web technologies. One parallel to this concept of environmental audit is that of *ecological footprint* which calculates the total area of land equivalent that a nation or individual needs to support their lifestyle [3, 4]. This approach could provide an inspiration for the types of calculations our system would need. Semantic Web Services (SWSs) allow an architecture where different services are each contributing separately partial data for calculation of the overall environmental audit for a specific product.

This paper is organised as follows: Section 2 discusses the role of knowledge in the organic food industry, Section 3 presents our scenario describing the functionalities of the system. In Section 4 we discuss some of the relevant technologies which need to be used and this is followed by a conclusion in Section 5.

2 The Role of Knowledge in the Organic Food Industry

In many industrial sectors, analysts bemoan the fact that economic activity is driven entirely by price. One example of this is the personal computer industry where due to the standardisation of components, price is the determining factor as to whether a consumer chooses one product of equivalent specification over another. In contrast, there are industries where price is not alone in guiding decisions. One such is the organic food industry which has developed and grown as a result of an increasing knowledge on the part of consumers about how food is grown or raised and subsequently processed before arriving on the table. Although the movement dates back over 50 years, it is in the past decade that it has seen phenomenal growth. Between 1994 and 2004 sales of organic food grew in the UK from 100m to 1.12 billion and the number of organic farms grew to over 4000 in 2005. There are now over 600,000 ha of organic farmland in the UK[4].

Organic food is the quintessential knowledge dependent industry. Consumers want to know that the food they consume has been grown according to certain specific standards, under certain conditions and furthermore the organic food standards guarantee that growers, processors and importers are all registered and regularly inspected. Food certification mechanisms have arisen, originally organised by voluntary bodies, but since 1993 these have been codified by an EU directive. There are currently 10 certification organisations in the UK of which the most prominent is the Soil Association, with many more worldwide.

One of the reasons that knowledge is so important in this industry is because so much depends on accurate trustworthy knowledge. The consumer chooses to buy organic food specifically because they have informed themselves of the benefits in environmental, ethical and health terms. It is this knowledge that allows the consumer choice and it is the relevant communication of that knowledge that allows the producer to fulfill that need, and allows the restauranteur to provide for it.

3 A Knowledge Enabled Organic Food Scenario

In this section we present our scenario, our vision of how a number of actors involved in the production, preparation and consumption of organic food could interact and the role technology could play. A pivotal role is played in our conception by the restaurant. There are a number of reasons for this. One of the co-authors of this paper (Haughton) is a restauranteur himself and his vision of a newly planned complex informs the ideas presented here. In particular, a restaurant/cookery school complex, suitably designed and which attempts to integrate Semantic Web technologies can fulfill a number of important objectives:

– Facilitate the education of consumers about the origin, use and characteristics of food (both raw and cooked).
– Provide a means for tracking the environmental origin and impact of all resources used in the business so as to minimise the ecological impact of the business' activities.
– Provide a resource for a cookery school to research food and catering related topics, and provide a complement to the skill-based training the school will undertake.
– Enable food suppliers to provide information to the end consumers of their products, educationally relevant material to the cookery school, and commercially relevant information to the restaurant(s) involved in the project.
– Enable the restaurant's chefs and management to design menus and seamlessly obtain availability and cost data (both financial and ecological i.e. environmental audit) on the one hand, and present their choices to customers with the related information.

[4] source: The Soil Association and [5]

- Close the circle by providing feedback to suppliers and others in the supply chain about the downstream use of their products, thus enabling them to modify their production and distribution to reduce environmental impact, for example by modifying packaging.

A large two level complex by the waterfront of a major UK city will contain two restaurants, a bakery, and a cookery school. The enterprise will be a not for profit "community interest company" whose brief ranges from providing good food at a range of prices, educating a new generation of food lovers and chefs, educating the local community of the possibilities of ecologically aware food consumption, and providing opportunities for food and catering suppliers to both sell their goods and promote themselves. Clearly, using SW technologies multiple restaurants and catering schools could be linked in as and when appropriate. A number of different users will have different perspectives on the same underlying information and knowledge contained in the system:

The Customer. The customer comes to one of the restaurants and is presented with a menu which details not only the dishes available but also the ingredients and provenance. Discretely placed, there are touch interface computers giving access to further details, which the customer can navigate in a number of ways. Today's menu can be accessed and from that the ingredients, where they come from, details about the suppliers, their locations and the environmental audit associated with the menu items. Alternatively they can look at a map and see where the different suppliers are located, what they supply including images and other multimedia components. The overall intent is to educate customers with a light hand, and, if anything, should arouse their curiosity without being imposing.

The Chef/ Manager. Using a suitably designed interface, the chef can consider what ingredients are available both in stock and by order. They can design a menu, which on the one hand will automatically generate both a printed and on screen version for the customers, and on the other hand allow the chef to make the orders necessary for that day. The orders can be transmitted electronically to the suppliers who are part of the network, otherwise printed out to be fulfilled on the telephone etc... Depending on the degree of electronic tracking available greater or lesser degrees of stock control can be applied. The data provided to the chef will allow for culinary, financial and ecological calculations to be made, e.g. enabling them to choose the closest supplier for a given product or quality.

The Cookery School Student. Between active skills training, the student will be able to use the system to educate themselves on a range of issues. They will be able to use the system to learn about the various foods used, their seasons, producers, the ecological impact of their use, who are local suppliers, etc. Small research projects can be undertaken either solely with the data in the system, or going outside collecting relevant information and putting this into the system. The introduction of this kind of use of computers in this context will be highly innovative.

The Supplier. A core intent of this project is to gradually include a large range of suppliers in a Semantic Web based e-business infrastructure. Currently many organic food producers do not have web sites let alone Semantically enabled web sites. Thus for those which do have web sites, information can be scraped initially. Eventually the supplier will be able to use a Semantic Web enabled web page construction environment which will allow them to: a) upload general information about themselves and their products (available to the public), and b) upload data about the actual availability of products and price (for professional purposes). The system will automatically generate a semantically enable public web site and also publish professional data via a Semantic Web Service.

The system should be accessible both to schools (for educational purposes) on the one hand and multiple restaurants and other interested parties (certification inspectors, DEFRA) on the other hand.

4 Relevant Technologies

A large number of technologies can come into play in this scenario. There already exist a number of e-commerce web sites (such as Amazon.com) which show signs of developing Semantic Web type technologies to facilitate and augment e-commerce. Here we outline a number of technologies and the roles they might play.

4.1 Ontologies

As we have argued, the organic food industry and the associated social movement is a knowledge driven industry. Consequently appropriate knowledge representations in the form of ontologies lie at the core. One of the advantages of this domain is that it is reasonably clear cut (there are a relatively finite number of agricultural products). There already exist considerable efforts to develop agricultural and food-related ontologies, in particular the conversion of the AGRIVOC thesaurus to a fully fledged ontology [6]. There also exist some environmental and ecological ontologies although these tend to be focused on the earth science [7] or bioscience [8] end of environmental studies. There also exist extensive ontologies for price and products.

Clearly a specific integration of existing ontologies and tailoring to the specific needs of the domain is required. Apart from existing work, the domain has certain advantages in that it is relatively stable (i.e. new food product types are relatively rare) although allowance must be made for new plant varieties and the extensive creativity of chefs in creating recipes.

The nature and complexity of the different participants in this proposed system raises the issue of whether one or more ontologies are needed. It is expected that both common and different ontologies will be needed by the customer of a restaurant, by the chef, and yet another by the food producer/supplier. This will be an issue to be decided in practice for reasons of efficiency of reasoning and inference. However, clearly there must be closely regulated mappings between the different ontologies to allow inter-operability and effective use of Semantic Web Services throughout.

4.2 Semantic Web Services

A key component of our vision is the use of Semantic Web Services [9, 10]. There are two major roles: First to identify and collate organic food products so as to allow a continuously updated catalogue of what is available, at what price, distance and environmental audit. This is based on the idea that each food producer will be able to easily publish their Semantic Web data on appropriately designed web sites. From a Semantic Web Service perspective, the fact that this industrial sector has registries of approved suppliers means that "discovery" process is, in part, significantly simplified. Secondly, the calculation of environmental audit requires several SWSs computing such aspects as distance, means of transportation, conditions of production etc. In spite of such efforts as [3] and even the ISO 14000 (`http://www.iso14000.com/`), the calculation of environmental or ecological impact in terms which go beyond pricing is still complex and in need of research. Thus we foresee different SWSs specialising in different calculations relevant to the overall environmental audit: one may calculate the environmental impact of transportation dependent on distance, means, time, etc.; another may calculate the impact of packaging used; yet another the production methods and the effect of food processing, etc.

4.3 Semantically Enabled Web Sites

From the food producer's side, the most important aspect is to be able to create and maintain a Semantic Web enabled web site, easily and with a minimum of manual intervention. We believe that such technologies are approaching maturity. For example the the KMI Semantic Web site[5] is generated and kept up to date entirely automatically. Data is integrated from heterogeneous sources including existing html pages and databases, and uses information extraction and data integration techniques. In a similar manner, the food producers' semantically enabled web sites can be auto-generated and kept up to date. In the initial phases of such a project technologies like Armadillo can be used to collate the necessary data where web sites are not semantically enabled yet [11].

4.4 Knowledge Visualisation and Presentation

One of the major objectives of this project is make it possible to easily educate and inform the wider public (customers and other interested parties) about the food available, its origins, productions methods etc. Equally the cookery school component of the intended complex will need to educate the students not just in the practical side of cookery but in the theoretical aspects. Computers allow the

[5] `http://semanticweb.kmi.open.ac.uk:8080/ksw/`

integration of multiple media which, using SWSs, can be linked to the source data in food producer web sites.

One of the key issues in the organic food movement is the need to obtain locally produced food, because of the growing awareness of the cost of food transportation and the tendency of major supermarkets to source all types of food (including organic) from abroad. Both the customer and the student in the cookery school need to visualise the distance covered by the food used. The CS AKTive Space project [12] has shown how SW technologies can be integrated into a high impact visual environment which could be adapted for the purposes of this project. This was designed to present geographically related information concerning the computer science research community in the UK, but could equally be adapted to the presentation of information concerning local suppliers, mapping local food producers and allowing the linking of items in the menu with the production sources.

4.5 Trust Technologies

Trust in Semantic Web technologies is a major issue [13] and of special significance in this domain. In one sense it is substantially less complex, in that the registration system ensures that producers are guaranteed to be providing "organic" produce and thus a key component of the trust challenge is obviated. Suitable representations would need to be developed to establish what it is that is trusted by the existence of certification, and what is not necessarily trusted. We envisage the applicability of most e-business trust models

5 Final Thoughts

We have presented in this paper a vision for the integration of a large body of heterogenous knowledge (textual, numerical, multimedia) but for a relatively finite domain. We believe that this type of business to business to consumer integration has not been proposed elsewhere and is only feasible by the appropriate application of Semantic Web technologies. This is a major opportunity to apply advanced technologies to achieve two important aims viz. to minimise the ecological impact of a specific industry and concurrently to educate the end user and trainee in that industry. There are challenges to be overcome. These include persuading the producers that involvement in such a project is beneficial to them, finding the most appropriate means to communicate the knowledge to restaurant clients, and envisaging ways to use the knowledge collected and available for the best possible educational effect. These are challenges which are tractable in the current technological context to a large extent because of the sophistication of the various potential users and the significance of knowledge in this industry.

6 Acknowledgments

This work is supported under the Advanced Knowledge Technologies (AKT) Interdisciplinary Research Collaboration (IRC), which is sponsored by the UK Engineering and Physical Sciences Research Council under grant number GR/N15764/01. The AKT IRC comprises the Universities of Aberdeen, Edinburgh, Sheffield, Southampton and the Open University. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing official policies or endorsements, either express or implied, of the EPSRC or any other member of the AKT IRC. It has also been supported by the ABRAXAS project (www.nlp.dcs.shef.ac.uk/abraxas) also funded by the UK Engineering and Physical Sciences Research Council under grant number GR/T22902/01.

References

1. Frederick, H., Beattie, D., McIlroy, D.J.: The knowledge economy. Technical report, Information Technology Advisory Group (1999)
2. Hayek, F.A.: The use of knowledge in society. American Economic Review **35** (1945) 519–530 Reprinted in F.A. Hayek (ed.), Individualism and Economic Order. London: Routledge and Kegan Paul.
3. Wackernagel, M., Schulz, N.B., Deumling, D., Linares, A.C., Jenkins, M., Kapos, V., Monfreda, C., Loh, J., Myers, N., Norgaard, R., Randers, J.: Tracking the ecological overshoot of the human economy. PNAS **99** (2002) 9266–9271
4. Wackernagel, M., Rees, W.E.: Our Ecological Footprint: Reducing Human Impact on the Earth. New Society Publishers, Gabriola Island, BC (1996)
5. DEFRA: Organic statistics united kingdom. Technical report, Department for Environment, Food and Rural Affairs (2005)
6. Fisseha, F., Liang, A., Keizer, J.: Reengineering agrovoc to ontologies - step towards better semantic structure. In: Building a Menaingful Web: From Traditional Knowledge Organization Systems to New Semantic Tools - The 6th NKOS Workshop at ACM-IEEE Joint Conference on Digital Libraries (JCDL) May 31, 2003 Houston, TX. (2003)
7. Raskin, R.G., Pan, M.J.: Ontology development in the semantic web for earth and environmental terminology (sweet). In: Geological Society of America Annual Meeting, Seattle, November 2-5,. (2003)
8. Keet, C.M.: Factors affecting ontology development in ecology. In Ludaescher, B., Raschid, L., eds.: Data Integration in the Life Sciences (DILS2005). Lecture Notes in Bioinformatics, San Diego, USA, Springer (2005) 46–62
9. Cabral, L., Domingue, J., Motta, E., Payne, T.R., Hakimpour, F.: Approaches to semantic web services: an overview and comparisons. [14] 225–239
10. Paolucci, M., Sycara, K.: Autonomous semantic web services. IEEE Internet Computing (2003) 34–41
11. Ciravegna, F., Chapman, S., Dingli, A., Wilks, Y.: Learning to harvest information for the semantic web. [14] 312–326
12. Glaser, H., Alani, H., Carr, L., Chapman, S., Ciravegna, F., Dingli, A., Gibbins, N., Harris, S., Schraefel, M.M.C., Shadbolt, N.: Cs aktivespace: Building a semantic web application. In Bussler, C., Davies, J., Fensel, D., Studer, R., eds.: ESWS. Volume 3053 of Lecture Notes in Computer Science., Springer (2004) 417–432
13. O'Hara, K., Alani, H., Kalfoglou, Y., Shadbolt, N.: Trust strategies for the semantic web. In Golbeck, J., Bonatti, P.A., Nejdl, W., Olmedilla, D., Winslett, M., eds.: ISWC Workshop on Trust, Security, and Reputation on the Semantic Web. Volume 127 of CEUR Workshop Proceedings., CEUR-WS.org (2004)
14. Bussler, C., Davies, J., Fensel, D., Studer, R., eds.: The Semantic Web: Research and Applications, First European Semantic Web Symposium, ESWS 2004, Heraklion, Crete, Greece, May 10-12, 2004, Proceedings. In Bussler, C., Davies, J., Fensel, D., Studer, R., eds.: ESWS. Volume 3053 of Lecture Notes in Computer Science., Springer (2004)