Computers and Electronics in Agriculture 140 (2017) 433–442

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Original papers

A generic ontological network for Agri-food experiment integration – Application to viticulture and winemaking

Aunur Rofiq Muljarto [a,d], Jean-Michel Salmon [b], Brigitte Charnomordic [a], Patrice Buche [c,e,*], Anne Tireau [a], Pascal Neveu [a]

[a] INRA, MISTEA Joint Research Unit, UMR729, F-34060 Montpellier, France
[b] INRA, Unité Expérimentale de Pech Rouge, UE0999, F-11430 Gruissan, France
[c] INRA, IATE Joint Research Unit, UMR1208, F-34060 Montpellier, France
[d] Dept. of Agroindustrial Technology, Brawijaya University, Malang 65145, Indonesia
[e] INRIA GraphiK, LIRMM, UMR5506, F-34095 Montpellier, France

ARTICLE INFO

Article history:
Received 20 September 2016
Received in revised form 14 February 2017
Accepted 21 June 2017
Available online 29 June 2017

Keywords:
Ontology network
Data integration
Viticulture
Winemaking

ABSTRACT

This paper presents an ontological approach of scientific experimental data integration across complementary sub-domains, i.e., agricultural production and food processing, with an application to viticulture and winemaking. The two main steps in this approach are (i) to integrate preexisting ontologies to create a so-called ontology network and (ii) to populate the ontology network with experimental data from various sources. The Agri-Food Experiment Ontology (AFEO), a new ontology network, was developed, based on two ontological resources, i.e., AEO (Ontology for Agricultural Experiments) and OFPE (Ontology for Food Processing Experiments). It contains 136 concepts which cover various viticulture practices, as well as winemaking products and operations. AFEO was used to guide the data integration of two different data sources, i.e., viticulture experimental data stored in a relational database, and winemaking experimental data stored in Microsoft Excel files. Two applications illustrate the approach. The first one is on wine traceability and the second one is related to the influence of irrigation practices and winemaking methods on GSH concentration in wine. These examples show that data integration guided by an ontology network can provide researchers with the information necessary to address extended research questions.

1. Introduction

Research in Agri-food and related fields dealing with sustainability has undergone important changes in the past years and tends to be more integrative, collaborative, and interdisciplinary (Perrot et al., 2016). In these lines of research, the Agri-food domain is considered as an interconnected system with various entities and complex relationships among them (Wolfert et al., 2010). More and more numerous data sources cover the whole food chain and can be combined to address new questions. For example, to test a hypothesis about the effects of different viticulture treatments on wine quality, researchers need to access and analyze various data sources at different scales, from the field to the bottle.

Data integration is not so easy and researchers are confronted with some obstacles. Data are commonly stored in scattered places and their formats, naming, storage and query or retrieval mechanisms are very diverse. The heterogeneity of scientific data may come from many factors, such as (i) They are collected separately based on independent research projects; (ii) the data structure is frequently selected according to the collection method (e.g., to make data easier to record) or the format varies in function of future analysis, instead of using standard data representation (e.g., relational database schema); (iii) the terms and concepts used to label data are not standardized, neither within nor across scientific disciplines and research groups (Bowers, 2012). The difficulties in organizing data and knowledge in a unified way do not only limit research productivity but also reduce data traceability (Gardner, 2005).

Research experiments are commonly divided into sub-domains, such as agricultural production, post-harvest, and food transformation. Even though the explicit relation between them is easy to explicit and understand, each of them has different objectives, scopes, and circumstances. For instance, agricultural experiments are usually conducted in the fields, where environmental factors

* Corresponding author at: INRA, IATE Joint Research Unit, UMR1208, F-34060 Montpellier, France.

E-mail address: patrice.buche@inra.fr (P. Buche).

http://dx.doi.org/10.1016/j.compag.2017.06.020
0168-1699/

are difficult to control, while food processing experiments are generally carried out in laboratories with controlled environmental conditions. From a practical point of view, they require different methods for collecting and organizing observational data that yield differences in data format, structure and storage. The heterogeneity also occurs due to the vast scope of Agri-food sub-domains. Each academic discipline uses its own way to express knowledge, terms, concepts and semantic relations, which results in difficulties to share the observational data.

Studies in the last two decades have shown that ontologies represent a flexible way to link the information contained in heterogeneous data sources within or across domains (Gardner, 2005; Seedah et al., 2015). Ontologies also provide standard concepts for data integration, opening the possibility to draw more comprehensive conclusions and to view data from different perspectives. Ontologies also allow certain types of automated reasoning to be performed. These features will help to develop advanced Information Systems able to manage heterogeneous data sources and to design platforms for more collaborative scientific data analysis.

The contribution of this work is to provide a methodology to prepare and to integrate data sources prior to further analysis, in order to answer complex questions that require access to various Agri-food scientific data sources. The approach is illustrated by two applications in viticulture and winemaking using heterogeneous experimental data sources. To achieve this purpose, we developed the Agri-Food Experiment Ontology (AFEO), a new ontology network resource, based on two ontology resources, i.e., AEO[1] (Ontology for Agricultural Experiments), which is also an original contribution of this paper, and OFPE[2] (Ontology for Food Processing Experiments). An ontology network (Suárez-Figueroa et al., 2012) is a new ontology engineering concept, which allows ontology re-use and avoids custom-building new ontologies from scratch. The AEO and OFPE have been developed separately in research laboratories to provide generic knowledge representations of agricultural production and food transformation processes. By following the NeOn (Suárez-Figueroa et al., 2012) methodology, we integrated these two ontologies into an ontology network, in order to facilitate data integration all through the food chain.

Although in this paper the proposed ontology is specialized and tested for viticulture and winemaking experiments, the core elements of AFEO are fairly generic and can be adapted to other food products. Furthermore, the ontological definitions (concepts and relations) can be useful to impose and preserve a logical structure for new types of scientific data, that may appear due to new sensors, protocols and analyses.

This paper is structured as follows: Section 2 gives a brief survey on recent work related to scientific data integration and Agri-food ontologies; Section 3 describes the global Agri-food experiment ontology network design process, introduces the two ontology resources used to build the network, their integration and specialization to viticulture and winemaking experiments; Section 4 presents how the new proposed ontology network is instantiated to integrate two different data sources; Section 5 is dedicated to two potential uses in the domain of viticulture and winemaking; Finally, in Section 6, conclusions are drawn and further work outlined.

2. State of the art

Driven by the need to communicate and share knowledge, research work on ontologies has increased recently not only in computer science but also in various fields, including Agri-food related ones. This section presents the state of the art, divided into two subsections: (i) scientific data integration with respect to an ontological approach and (ii) current work on Agri-food ontologies.

2.1. Scientific data organization and integration

Over the recent years, researchers have faced significant problems to manage scientific data due to their increasing volume and complexities (Lapatas et al., 2015; Luyen et al., 2016). It causes them to spend quite a bit of time organizing and accessing scientific data rather than directly focusing on their analysis (Madin et al., 2007). Scientific data are generally measurements directly linked to real-world phenomena (Shawn Bowers and Madin, 2008). Within the Agri-food domain, cross-disciplinary scientific data are required to explore complex and temporal aspects of food quality and the impact of practices and operations.

The need for a more adaptable mechanism to organize scientific data has been addressed in the literature, both in non-ontological based approaches such as, Lab Key Server (Nelson et al., 2011) and Sci Port (Wang et al., 2007); and in ontological based approaches such as the ones proposed in Luduscher et al. (2006), Madin et al. (2007), Leonelli (2013), Bowers (2012), Li et al. (2013) and Fox et al. (2009). Some of these approaches are targeted to specific scientific domains while the others are developed to be more generic and extensible. The non-ontological approach mostly relies on data models, such as database or XML schema, where attributes and relationships of domain concepts are captured in standardized structures. The ontological approach has some additional advantages in terms of data interoperability and knowledge reasoning (Daraio et al., 2016; Imran and Young, 2016; Arch-int and Arch-int, 2013; Lousteau-Cazalet et al., 2016).

To the best of our knowledge, there is no ontological approach to represent scientific experiments as well as observational data which fulfill all our needs. An ontology of scientific experiments, the EXPO (Soldatova and King, 2006), has been proposed to formalize the generic concepts of experimental design, methodology and result representation. However, this ontology does not provide a clear explanation about scientific data representation. Neither does this ontology describe how to manage a set of experiments in which several interrelated experiments have to be conducted in a given order. The Extensible Observation Ontology (OBOE) has been developed to serve as a formal and generic conceptual framework for describing the semantics of observational data sets (i.e., data sets consisting of observations and measurements) (Madin et al., 2007). The basic concepts of the observational model consists of five classes and six properties (Bowers, 2012), and it can be applied to various types of observations. Nevertheless, though well suited to representing scientific data which are generated from measurement by sensors, it does not provide other types of observational data such as expert judgments or results from a calculation procedure.

2.2. Agri-food ontologies

A line of research focused on building ontologies to conceptually model agricultural practices related to crop productions, such as, hilly citrus production ontology (Wang et al., 2015), precision agriculture ontology (Song et al., 2012), crop-pest ontology (Beck et al., 2005) and potato ontology (Haverkort et al., 2007). Other works centered on food taxonomy have also been carried out, such as FOODS (Food-Oriented Ontology-Driven System) (Snae and Bruckner, 2008) and wine classification (Graça et al., 2005).

The definition of valid concepts and a complete taxonomy for agricultural practices and food processing are both important to model Agri-food experiments. Most of the above ontologies refer to the terms or concepts listed in AGROVOC. AGROVOC is a

[1] http://agroportal.lirmm.fr/ontologies/AEO.
[2] http://agroportal.lirmm.fr/ontologies/OFPE.