XVII[th] World Congress of the International Commission of Agricultural and Biosystems Engineering (CIGR)

Hosted by the Canadian Society for Bioengineering (CSBE/SCGAB)
Québec City, Canada June 13-17, 2010

REST-BASED WEB SERVICES FOR THE DISCOVERY AND DISTRIBUTION OF AGRICULTURAL PRODUCTION STANDARDS

R. NIKKILÄ[1], E. NASH[2], I. SEILONEN[1], K. KOSKINEN[1]

1 R. Nikkilä, Helsinki University of Technology, Department of Automation and Systems Technology, Information and Computer Systems in Automation P.O. Box 5500, FIN-02015 HUT, rnikkila@cc.hut.fi
[1] I. Seilonen, ilkka.seilonen@tkk.fi
[1] K. Koskinen, kari.o.koskinen@tkk.fi
2 E. NasH, Rostock University, Faculty of Agricultural and Environmental Sciences, Institute for Management of Rural Areas, Justus-von-Liebig-Weg 6, 18059 Rostock, Germany, edward.nash@uni-rostock.de

CSBE100584 – Presented at 8[th] World Congress on Computers in Agriculture (WCCA) Symposium

ABSTRACT Modern agricultural production is governed by standards which direct and restrict farming practices. The number and complexity of these standards is continuously increasing with considerable manual labour involved in the verification of farms' compliance to these standards. It is the vision of the FutureFarm project that in the future, parts of this verification could be done automatically given that the agricultural standards are represented as formal rules allowing computerised management and reasoning. To properly actualise this goal, an efficient mechanism and infrastructure for the discovery and distribution of these standards is required. Furthermore, since these standards are published by several companies and government bodies, it is unreasonable to assume extended co-operation or any centralised distribution medium. Our research studies a solution consisting of two REST-based web services operating in unison. The first of these services provides a catalogue functionality for the discovery of standards while the second service manages the actual distribution. REST-based web services were found to provide an efficient mechanism for both aspects of the research problem. Despite their inherent simplicity, REST interfaces were able to provide all required functionalities for both services, which was confirmed with a prototype client implementation. Besides and prior to the eventual and at least partially automated verification of compliance to agricultural production standards, the proposed infrastructure could be used to conveniently distribute these standards in natural language for all stakeholders of modern agriculture.

Keywords: REST, Web services, Production standards, Agricultural information systems

INTRODUCTION Agricultural production is increasingly governed by production standards which farms adhere to either voluntarily or obliged by law. These standards are commonly complex documents and validating the compliance of all individual farms requires a considerable amount of human labour. The FutureFarm project has focused on possible technical advances involving these production standards and agricultural

information systems. It is the vision of the FutureFarm project that the discovery and distribution of these standards can be automated and that with present and future information systems, compliance to these standards can be at least partly verified automatically based on the farm data stored in an information system. The feasibility of this goal is discussed in (Nash et al. 2009c) and further in (Vatsanidou et al. 2009).

This vision is not technically trivial to implement and requires numerous advances in agricultural technology and information systems. Having the agricultural standards presented formally as logical rules and made available for automatic selection and retrieval is merely one aspect of the technical problem. Additionally, farm data must be generated automatically in the machinery as part of the agricultural production process and this information must be combined and stored to an information system. All requirements combined, the vision of the FutureFarm project has similar technical requirements to precision, or information intensive, agriculture (Stafford 2000).

Agricultural production standards and their convenient availability would benefit most stakeholders of modern agriculture. Recent and relevant topics such as traceability and organic production are just examples that could utilise and benefit from the better deployment of standards.

The distinguishing feature in the use of standards in Agriculture, when compared to the use of standard elsewhere (e.g. industry), is their dynamic nature. Standards themselves tend to change often and farms can change their cultivar or customers which can cause spontaneous changes to the selection of standards the farmer has to consider.

OBJECTIVES The scope of this paper is restricted to the discovery and distribution of agricultural standards. The encoding of agricultural standards has also been addressed by the FutureFarm project and is discussed separately in (Nash et al. 2010). Also the actual evaluation of these standards with farm data is beyond the scope of this paper and a part of future work. This paper aims to identify the requirements of an infrastructure necessary for the discovery and distribution of standards and then propose a specification for an infrastructure fulfilling these requirements. This infrastructure is then evaluated based on both a prototype implementation of the infrastructure and a prototype implementation of a client software utilising the proposed infrastructure.

Research for this paper has the restricted focus of using the proposed infrastructure as part of the automated validation of compliance to standards but the equally important role of standards in the planning of agricultural work within information systems has also been identified. Regardless of any automated utilisation, the same infrastructure can be used to conveniently transfer the standards in natural language for farmers, thus helping farms in keeping up to date with any changes or updates to these standards.

The design objectives for the infrastructure are, besides fulfilling the requirements of the stakeholders and providing the required functionality, also keeping the specification, implementation and use of the infrastructure as simple as possible while providing an extensible system with a possibly long lifespan. The implementation of the infrastructure in the form of Web services is seen as the only practical and modern approach in order to achieve interoperability and any level of common acceptance.

RELATED RESEARCH Little to no direct research has been done on the automated utilisation of agricultural standards regardless of the fact that these standards are partly utilised in existing information systems, albeit hard-coded in the application logic.

Logical and formal rules have been studied extensively in the field of theoretical computer science. Through this research, it is possible to formulate rule systems that are expressive while retaining computability and the possibility of efficient automated management.

Service discovery, being an essential element of several systems, such as Semantic Web, has also been under recent study, with various proposed protocols and systems (Bettstetter and Renner 2000). However, many of these protocols and systems are complicated and usually intended to provide much more general solutions than what is required for the discovery of agricultural standards.

Business rules (Ross 2003) are partly analogous to the research problem presented here as they act as constraints on business structure in the manner which agricultural production standards govern agricultural production. Hence, research on business rules, their representation and evaluation, is relevant to research on agricultural production standards.

Agricultural information systems Agricultural production standards are used in present information systems since these systems must incorporate planning and reporting features, many of which are regulated by standards. However, there is no indication that any present information system would utilise standards in any dynamic manner rather than hard coding parts of the standards in the application logic. Neither is there any previous research nor work on automatic compliance checking for agricultural standards. The reason for which, is likely to be a combination of immature information systems and the lack of agricultural production standards in a format suitable for automated handling. The general requirements of the technical framework are further discussed in (Nash et al. 2009b)

Besides the low automated utilisation of production standards, there would appear to be no existing infrastructure for distributing these standards besides having them published on the World Wide Web in, what is essentially, a natural language text format. Due to the increasing number and complexity of the production standards, a more mature form of distribution is soon required for their present use in agriculture and especially for their computerised handling.

REQUIREMENTS OF SERVICES

General requirements As a fundamental requirement, the infrastructure must provide the means for an efficient discovery and distribution of agricultural production standards. Additionally, it should be possible to do filtering and searches on each step of the process based on various criteria, many of which are unique to the field of agriculture. Technically, the most complicated requirement is the support for spatial searches and filtering, because the standards of agriculture often affect a specific country, region or other geographic area. Furthermore, it should be possible to obtain and search standards based on the agricultural operation they affect, this operation can be anything from management to a specific field operation, such as fertilisation.

The interfaces of the infrastructure should be sufficient for all relevant operations involving both discovery and distribution as well as the routine adding, removing and modifying of information on the services. Furthermore, simple interface should be preferred and in case of multiple interfaces, the interfaces should be as similar as reasonably possible.

A certain degree of security is necessary throughout the infrastructure, in practice this can be achieved with authentication of services with certificates and the use or at least the possibility of using secure connections in communication. Authentication of users is required only for operations that alter the information stored on the services, thus retrieval of information is possible without an existing account or access privileges to any service.

Stakeholders Any infrastructure for the distribution of agricultural standards involves several stakeholders, though the list is somewhat shorter than that of a general information system for agriculture (Nikkilä et al. 2010), many distinct stakeholders exists. The obvious stakeholders are farms and farmers as well as any producers of agricultural production standards, which include both government bodies and companies. Also companies that produce agricultural information systems are stakeholders as they are required to utilise the services and standards in their software and systems. Few stakeholders have specific requirements for the system, for all stakeholders the availability and function of the system is their primary concern.

Requirements for discovery Discovery of the standards should be possible with minimal amount of initial information, i.e. the address of a single service. Starting with this initial information, it should be possible to locate all necessary services and standards. Since the stakeholders of the infrastructure include several nations, organisations and companies, the discovery should not except nor rely on extensive co-operation between these stakeholders. However, discovery does require at least some services to know of each other but this knowledge can be limited to a few items of metadata and the address of the other service.

During discovery, spatial criteria should be available for filtering and searching to limit the scope of the search to the relevant area of interest for standards.

Requirements for distribution For the distribution, it is assumed that all agricultural standards can be represented in one or more unified formats. One such format is the XML schema for the encoding of agricultural standards proposed in the FutureFarm project deliverable 4.1 (Nash et al. 2009a). The format in which the standards are stored on the services is left unspecified but regardless of that format, the service should still provide the standard in a common format.

One specific requirement of distribution is that it should not rely on external resources or other publishers of standards so that any agricultural standard can be published independently with the discovery of the service being the only concern. Besides obtaining the full encoded standard, it should be possible to obtain only the header of a standard, individual rules or only the rules that pertain to a certain agricultural operation.

DESCRIPTION OF THE SERVICES To fulfill the requirements identified in the previous section, an infrastructure consisting of two services is proposed. Both of these services are based on REST (Representational state transfer) (Fielding 2000) which is a style of software architecture well suited for simple online services. REST is often implemented using the family of protocols for HTTP that provides a well-supported base of functionality.

The services, known hereafter as the catalogue service and the rule service, fulfill the requirements of discovery and distribution correspondingly. Both have REST-based interfaces that provide access to their resources and operations. All communication occurs in XML using schemata specific to the services. However, certain operations, such as obtaining a list of catalogue services requires no actual input from the client using the service, thus a simple GET request to the desired resource is sufficient without any parameters.

The structure of both services is described graphically in a tree-like structure, which shows the resource hierarchy, object identifications and valid operations for each resource.

Use case Following is a generic use case for the whole infrastructure, certain operations, such as selecting which services to use are intentionally left abstract. Additionally, obtaining a list of services or standards can contain filtering or searching which are intentionally omitted from the use case.

1. A catalogue service is contacted and a list of catalogue services is obtained
2. Repeat step 1. for each new catalogue service that was discovered
3. Select some of the catalogue services to query
4. Query the selected catalogue services and construct a list of rule services
5. Select some of the rule services to query
6. Query the selected rule services and construct a list of standards
7. Select some of the standards
8. Query the rule services and obtain rules for the standards
9. Use the standards or rules

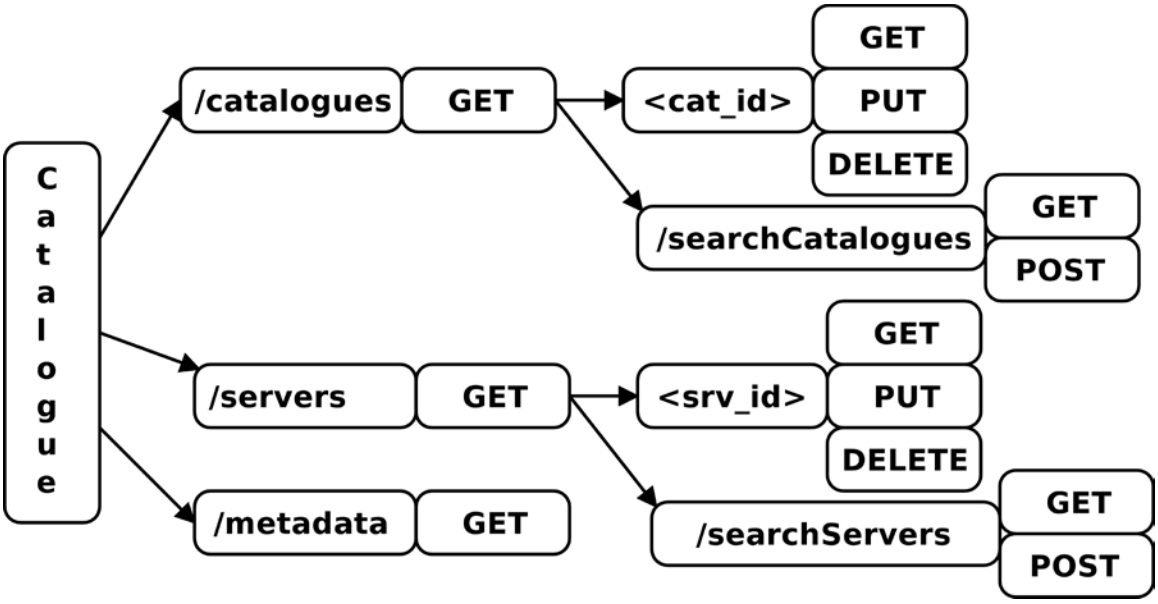
Catalogue service The catalogue service is described in detail and defined as part of the FutureFarm project deliverable 4.2 (Nikkilä and Nash 2009) The catalogue service provides essentially a listing of other catalogues and rule services and does not contain any standards or rules. Metadata is stored to provide searching and filtering for both listings. The functionality is provided by two REST-resources, one for other catalogue services and one for rule services.

For both of these resources, the same set of operations is provided.

- Listing all stored entries
- Fetching a stored entry
- Adding and removing entries, both of which require authentication of the user

- Searching entries trough a search resource

Additionally, there exists a metadata resource that returns the information of the current



catalogue service as it would appear in the listing of other catalogue services.

Figure 1. The catalogue service resources and operations

Rule service

Like the catalogue service, the rule service is specified as part of the FutureFarm project deliverable 4.2 (Nikkilä and Nash 2009). However, while similar in its interface to the catalogue service, the rule service contains no listings or references to any external service. This is with the intention that standards publishers can run their own rule services independently of any other organisations or companies. The only requirement is that one or more catalogue services know of this rule service to facilitate the discovery and access by client software.

The core functionality of the rule service is to provide agricultural standards, parts of the standard or individual rules from a standard. The interface to this functionality is similar to that of the catalogue service despite the considerable differences of the elements provided by the resource. The rule service provides one resource for which the following operations are supported.

- Listing all stored standards
- Fetching a standard
- Adding and removing standards, both of which require authentication of the user
- Searching standards based on some criteria
- Fetching all rules of a standard pertaining to a certain agricultural operation

- Fetching an individual rule from a standard

Identically to the catalogue service, the rule service contains a metadata resource that returns the information on the current rule service as it would appear on catalogue services. Some resources of the rule service return elements directly from the schema encoding the agricultural standards.

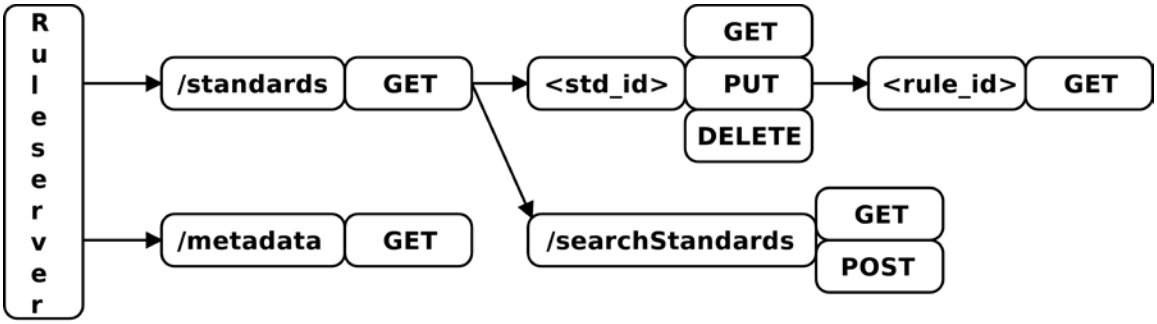


Figure 2: The rule service resources and operations

EVALUATION OF SERVICES For the evaluation of the services and due to the practical need of having instances of both services available for further work, both services were implemented according to the extent their specifications.

The ease of implementation can be used as an indicator of the general suitability of REST as the interface of choice for the infrastructure. Furthermore, client software and its implementation can be used to evaluate the infrastructure and its interfaces in general.

An example of the request XML content for searching standards. This XML is the content of a HTTP POST request to the resource /standards/search_standards.

```
<?xml version="1.0" encoding="UTF-8"?>

<rs:ff_rules_request version="0.0.2" xmlns:rs="http://www.futurefarm.eu/ruleserver">

  <rs:search_standards>

    <ff:validRegion>Deutschland</ff:validRegion>

    <ff:classification>mandatory legal requirement</ff:classification>

  </rs:search_standards>

</rs:ff_rules_request>
```

Implementation Both services were implemented using the Ramaze modular Web framework (http://ramaze.net) together with various libraries in the Ruby programming language. The rationale for using this particular platform for the implementation was simplicity and the good availability of supporting libraries. While the services are independent, they share some common parts in their implementation due to their general similarity and in particular, the similarity of their interfaces.

While not intended for the implementation of REST services per se, Ramaze proved to be a very suitable platform for the task. Using their specifications (Nikkilä and Nash 2009), the implementation of the services was straight-forward with service interfaces implemented as Ramaze controllers. A PostgreSQL relational database was used for both services, though particularly the rule service would most likely have benefitted from the use of a native XML database.

Functionality The specified functionality for both services was achieved with a surprisingly low amount of work with the greatest technical difficulties involving the handling and management of XML documents rather than the REST interface or service communications. In their current state, both of the services constitute of roughly 500 lines of Ruby program code.

Features and functionality not included in their specification were not implemented for either service. In practice this means that some functionality, such as adding new authentication entries to services, is achieved through means other than the specified REST interface. Likewise, error reporting could be more verbose and error management in general could be implemented in a more flexible manner, since at present, the services identify error conditions but make little to no attempt to recover from the error rather than just reporting it to the client software.

Case: simple client software To test the implementation of both services and to provide a prototype implementation for the next phase of the project, a simple client software was implemented. The client is a Web application based on the same Ramaze web framework as the catalogue and rule services but with the difference of being an actual web application instead of a REST-based service.

Communication with the services was achieved with a very small amount of work and the functionality described in the previously presented use case was achieved. This prototype was further demonstrated at the Agritechnica exhibition 2009 in Hannover, Germany, utilising several instances of the implemented catalogue and rule services. All services operated as expected and their interfaces proved sufficient for all operations required by the prototype.

CONCLUSIONS An infrastructure for the discovery and distribution of agricultural production standards can, with a good degree of technical elegance, be achieved using REST-based web services. Based on the requirements for the infrastructure, the separation to catalogue and rule services is reasonable and does not increase the complexity of the infrastructure. Futhermore, it was possible to implement both of these services with little work even with a technical infrastructure not primarely intended for the implementation of REST-based services. REST has also succesfully been used for agricultural services in related research (Martini et al., 2009).

DISCUSSION It is reasonable to assume that the role of agricultural production standards can only increase in the present information systems and the systematic use of these standards in the planning features of these information systems is fore-seeable in the near future. Eventually, when the availability of farm data is better and the information systems of agriculture have matured, compliance to production standards will be automatically evaluated to at least some degree. Hence, having some common representation for these standards and preferably a common route of distribution, can be beneficial for all stakeholders of modern agricultural production.

Similar functionality could be provided with other technical solutions, such as SOAP and related technologies. However, the simplicity of specifying and utilising the REST-based solution suggest that it is a preferable technology in fulfilling the requirements of the infrastructure over the more complicated alternatives. REST has also been identified as a suitable platform for smaller web services in research (Pautasso et al. 2008). Instead of a dedicated discovery and distribution framework, it would be possible to have the agricultural standards available as ordinary XML files over the World Wide Web but this solution would most likely be extremely shortsighted as it would hinder the automated management of the standards in the future.

A critical requirement for any automatic verification of compliance to a standard is having the standard encoded as a set of rules. While the standards themselves are written in natural language, the wording of the standard has a dramatic effect on any attempt to encode the standard as rules. For a clearly written and logical standard, encoding can be very straightforward but for a standard, for which a several hundred page long addendum is retroactively released in an attempt to shed light on various inconsistencies and vague expressions, encoding is likely to be difficult. Hopefully, increased automated use of standards will encourage publishers to using a more clear, precise and logical language in their standards.

FURTHER RESEARCH The encoding of standards is still ongoing research with one possible representation provided by the FutureFarm project in the form of an XML schema for both the standards and rules. An interesting question is the extent to which a standard can be encoded as rules and with what expressions while still maintaining the possibility of convenient automated management. The solution presented here takes the traditional and safe route of relying on an encoding that is based on a subset of first order logic to guarantee computationality. With advances in the evaluation and encoding of the rules, additional logical elements could be included to allow even more extensive encoding of standards. However, this is research more commonly associated to theoretical computer science.

The proposed infrastructure or a similar solution could be utilised in fields other than agriculture to provide the discovery and distribution of standards and rules for some processes. However, few fields use standards in the same dynamic manner as agriculture and would thus benefit less from not having the production standards hard-coded in application logic.

Acknowledgements The research done for this paper was for the FutureFarm project which was funded as part of the Seventh Research Framework programme by the European Union.

REFERENCES

Bettstetter, C., Renner, C. 2000. A comparison of service discovery protocols and implementation of the service location protocol. EUNICE Open European Summer School, Twente, The Netherlands.

Fielding, R.T. 2000. Architectural Styles and the Design of Network-based Software Architectures, Doctoral dissertation, University of California, Irvine.

Martini, D., Schmitz, M., Frisch, J., Kunisch, M. 2009. A Service Architecture for Facilitated Metadata Annotation and Ressource Linkage Using agroXML and ReSTful Web Services. Communications in Computer and Information Science, volume 46:

Nash, E. et al. 2009. Machine Readable Encoding for Definitions of Agricultural Crop Production and Farm Management Standards. FutureFarm Deliverable 4.1.1. Available at http://www.futurefarm.eu/node/183. Accessed 15/02/2010.

Nash, E., Nikkilä, R., Pesonen, L., Sørensen, C.G. 2009. Technology requirements for a standard information infrastructure to assist compliance with crop production standards. In: Edited by E.J. van Henten, D. Goense and C. Lokhorst. Precision agriculture "09. Wageningen Academic Publishers: 935-942.

Nash, E., Vatsanidou, A. and S. Fountas. 2009. Can compliance to crop production standards be automatically assessed? In: van Henten, E.J., Goense, D. and C. Lokhorst, Precision agriculture ’09, Proceedings of the 7th European Conference on Precision Agriculture, Wageningen, the Netherlands, 6-8 July 2009. Wageningen Academic Publishers: 899-906.

Nash, E., et al. 2010. Formal representation of agricultural production standards. 17th CIGR World Congress 2010, Canada.

Nikkilä, R., Nash, E. 2009. Definition of an interface for a knowledge repository. FutureFarm Deliverable 4.2. Available at http://www.futurefarm.eu/node/196. Accessed 15/02/2010.

Nikkilä, R., Seilonen, I., Koskinen, K. 2010. Software architecture for farm management information systems in precision agriculture. Computers and Electronics in Agriculture Volume 70, Issue 2, March 2010: 328-336.

Pautasso, C., Zimmermann, O.L.F. 2008. RESTful Web Services vs. Big Web Services: Making the Right Architectural Decision. 17th International World Wide Web Conference (WWW2008) (Beijing, China)

257-262.

Ross, R.G. 2003. Principles of the Business Rule Approach. Addison-Wesley Longman Publishing Co., Inc.

Stafford, J.V. 2000. Implementing Precision Agriculture in the 21st Century. Journal of Agricultural Engineering Research, volume 76, issue 3: 267-275.

Vatsanidou, A. et al., 2009. Compliance to standards specifications. FutureFarm Project Deliverable 2.1.1. Available at http://www.futurefarm.eu/node/159. Accessed 15 February 2010.