The open discussion version of this paper is available at: Santos R, Huynh D, Anand S, Ray RV, Mayes S, Leibovici D. (2016) A geoprocessing modelling interoperable framework for AgriGIS using open data and open standards. PeerJ Preprints 4:e2136v2 https://doi.org/10.7287/peerj.preprints.2136v2

A geoprocessing modelling interoperable framework for AgriGIS using open data and open standards

Santos, Roberto[1,2], Huynh Chi, Dai[1,2], Anand, Suchith[2], Ray, Rumiana[3], Mayes, Sean[3], and Leibovici, Didier[2]

[1]**School of Geography, University of Nottingham, UK**
[2]**Nottingham Geospatial Institute, University of Nottingham, UK**
[3]**Plant and Crop Sciences, University of Nottingham, UK**

ABSTRACT

We retrace the construction of AgriGIS framework between multiple disciplines around a common objective of facilitating research on model simulations for sustainable food security. The geospatial media enabling multidisciplinary research in crop modelling but also supporting new types of hypothesis and analysis, is described with interoperability principles and seamless access and sharing for data, metadata and processing models. Designing the platform achieving this main objective generated a transdisciplinary vision of modelling and forecasting for sustainable agriculture.

Keywords: agriculture models, crop modelling, scientific workflow, geospatial information, genetic-trait, interoperability, GIS, OGC, web services

INTRODUCTION

Rapid developments in positioning, broad-band mobile communications, sensor platforms, sensor-web enablement, spatial search and pervasive computing fundamentally change the access to and use of location-based data for agriculture (Jackson et al., 2011). However, the necessary multi-disciplinary approach needed to transform raw data and information into useful intelligence and knowledge for scientists is still constrained by disciplinary and organisational silo's and legacy concepts. Geospatial interoperability and open source standards-based GIS and open data will help deliver holistic solutions in geospatial technologies in AgriGIS by enabling the ready integration of separate location relevant technologies and lowering costs. The expanding range of open source GIS tools and open data will greatly enhance the use of geospatial technologies in agriculture and facilitates the sharing of information across various stakeholders and collaborative work (Al-Azri et al., 2015; Nativi et al., 2013).

In agricultural modelling, when focusing on sustainable development and food security, the concepts of geographical space and time are fundamental components of the agronomical processes of interest (AgriGIS). To this respect what is required from the agronomical processes usually depends from their different spatial and time scales and resolution, e.g. micro scale (plot or fields and days or weeks), meso scale (landforms or landscapes and seasons or years) and macro scale (countries, regions or ecoregions and decades or periods). Their complementary focus are respectively to these scales on 1) production and the factors that might affect its yields; 2) sustainability and land use and 3) interactions with climate change (Loireau et al., 2005; Gaber et al., 2008).

We have been witnessing an expansion of georeferenced data available in the agriculture context this is still currently happening due to the growing use of mobile sensors, technologies and software applications as well as increasing easy access to public datasets of satellite remote sensing data. However, despite those agricultural models contexts of inherent spatial-dimension and increasing availability of georeferenced data, there is still a lack of transdisciplinary research to enable full potential of re-using those information within a shared environment allowing the knowledge to cross disciplines barriers beyond legacy concepts (Mirschel et al., 2014; Liu et al., 2013)

To meet these requirements and challenges for agricultural modelling taking in account the recent advances of georeferenced data in agriculture, we developed a framework integrating the domains of Environmental Science, Crop Science and Geospatial Information Science. This framework aims to provide concepts and tools to access available data, available geoprocessing models and allow to combine them using internal and external processing models to generate results based on 1) genetic and phenotypic and trait information; 2) agricultural and environmental information and 3)

Please cite this paper as: Roberto Santos, Dai Huynh, Suchith Anand, Rumiana V Ray, Sean Mayes, Didier Leibovici (2018) A geoprocessing modelling interoperable framework for AgriGIS using open data and open standards. In Marchesini I. & Pierleoni A. (Eds.) Proceedings of the 4th Open Source Geospatial Research and Education Symposium (OGRS2016). Perugia, 12-14 October 2016. https://doi.org/10.30437/ogrs2016_paper_07

geospatial architecture and data and models management. This first phase of this project has been recently finalised and we present here the design of the framework and first results. This initiative is an effort of team composed members from the University of Nottingham (Nottingham Geospatial Institute and Crop Science Department in United Kingdom and Malaysia), the Crops for the Future (Malaysia) and other partners in the United Kingdom and Australia.

THE EGRASP PLATFORM

The principle of this framework is the use and reuse of information and tools available, particularly the open data and open source tools.

Within the spatial infrastructure of the framework (see Figure 3), a PostGIS database called GeogermplasmDB, is responsible for archiving genetic information, mainly genotypic variation, associated with geolocation of the genotypes and utilises an extended version of the CropStoreDB based on OpenGIS (Castronova et al., 2013). The data is served using an OGC WFS.

Central to the eGRASP is the geoprocessing workflow composition which utilises the Business Process Model and Notation (BPMN2.0) to represent the combination of data sources and processing components (see Figure 2). This approach encourages the conceptual design of workflows representing large models (see Figure 1) which can be broken down in sub-workflows modelling and vice versa allowing to discover moels and sub-models to be combined. Designing components in such way permits the use and reuse of them for different locations or series of simulations of the workflow. For example, for a disease modelling, one re-use an external tool such as the Agricultural Production Systems sIMulator (APSIM) as a component (Holzworth et al., 2014).

The integration of Geospatial Information Science concepts in this framework allows for example the measure of spatial quality and spatial error propagation on the results and generates meta-data information for decision makers. The spatial error propagation is currently implemented using two approaches, 1) classical error propagation with multiple runs and 2) meta-propagation of uncertainty using WPS (Leibovici et al., 2009, 2013).

DISCUSSION

So far a prototype with the few components describe above have been tested but most of the component of the architecture designed in figure 3 are not yet operational. If the proof of concept has been successful, the eGRASP platform for AgriGIS is at its enfancy but other type of data sources such as sensor network and remote sensing imaging would fit seamlessly in this architecture.

The greatest body of knowledge often lies with the farmers who have grown the crops themselves. Crowdsourcing is a potential method to collect such knowledge (data) which can then be stored using semantics and ontologies (Huynh Chi et al., 2014). This type of data comes to enrich existing more authoritative data sources or replace them when none exist. The data can then be integrated within the workflow as shown on figure 1 in terms of enriching the geo-germplasm database. The ultimate goal for the eGRASP is to host open data and open geoprocessing facilities for the community of agriculture modellers in crop modelling and food security.

REFERENCES

Al-Azri, M., Leibovici, D., Karunaratne, A., and Ray, R. (2015). Simulating Eyespot Disease Development and Yield Loss Using APSIM for UK Wheat. *Procedia Environmental Sciences*, 29:256–257.

Castronova, A. M., Goodall, J. L., and Elag, M. M. (2013). Models as web services using the Open Geospatial Consortium (OGC) Web Processing Service (WPS) standard. *Environmental Modelling & Software*, 41:72–83.

Gaber, N., Laniak, G., and Linker, L. (2008). Integrated Modeling for Integrated Environmental Decision Making. Technical report, White Paper EPA 100/R-08/010, EPA (United States Environmental Protection Agency.

Holzworth, D. P., Huth, N. I., DeVoil, P. G., Zurcher, E. J., Herrmann, N. I., McLean, G., Chenu, K., van Oosterom, E. J., Snow, V., Murphy, C., Moore, A. D., Brown, H., Whish, J. P., Verrall, S., Fainges, J., Bell, L. W., Peake, A. S., Poulton, P. L., Hochman, Z., Thorburn, P. J., Gaydon, D. S., Dalgliesh, N. P., Rodriguez, D., Cox, H., Chapman, S., Doherty, A., Teixeira, E., Sharp, J., Cichota, R., Vogeler, I., Li, F. Y., Wang, E., Hammer, G. L., Robertson, M. J., Dimes, J. P., Whitbread, A. M., Hunt, J., van Rees, H., McClelland, T., Carberry, P. S., Hargreaves, J. N., MacLeod, N., McDonald, C., Harsdorf, J., Wedgwood, S., and Keating, B. A. (2014). APSIM – Evolution towards a new generation of agricultural systems simulation. *Environmental Modelling & Software*, 62:327–350.

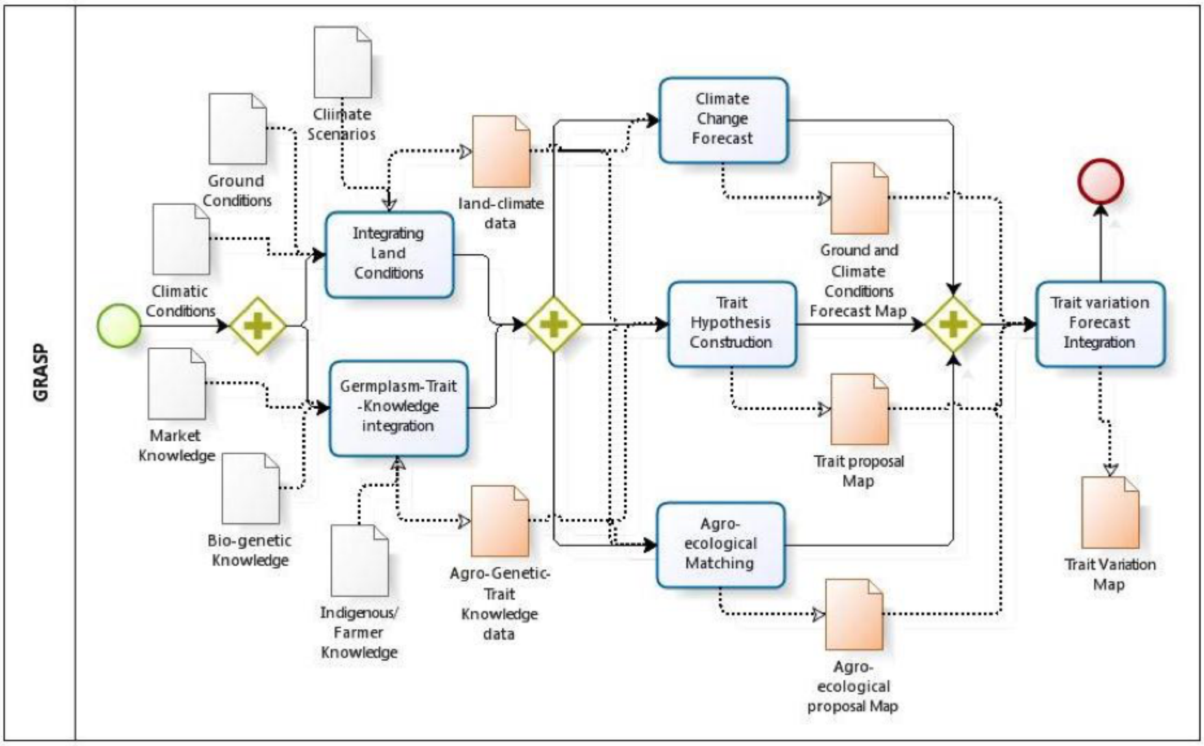

Figure 1. Generic workflow of the framework with the use and reuse of available information.

Huynh Chi, D., Vu, T., Jackson, M., Leibovici, D., and Mayes, S. (2014). Towards a User Centric Approach for Assessing Quality of Crowd-Sourced Data: an application in Agriculture. In *Workshop "Role of Volunteered Geographic Information in Advancing Science: Effective Utilization"/GIScience2014*, Vienna, Austria.

Jackson, M., Hobona, G., Bernard, L., Brauner, J., and Higgins, C. (2011). The role and development of a persistent interoperability test bed for geosciences research. pages 69–75.

Leibovici, D., Hobona, G., Stock, K., and Jackson, M. (2009). Qualifying geospatial workflow models for adaptive controlled validity and accuracy. In *2009 17th International Conference on Geoinformatics*, pages 1–5. IEEE.

Leibovici, D. G., Pourabdollah, A., and Jackson, M. J. (2013). Which spatial data quality can be meta-propagated? *Journal of Spatial Science*, 58(1):3–14.

Liu, J., Folberth, C., Yang, H., Röckström, J., Abbaspour, K., and Zehnder, A. J. B. (2013). A Global and Spatially Explicit Assessment of Climate Change Impacts on Crop Production and Consumptive Water Use. *PLoS ONE*, 8(2):e57750.

Loireau, M., Leibovici, D., Desconnets, J.-C., and D'Herbés, J.-M. (2005). LEIS, a Tool for Diagnosis and Prevision of Antropogene Pressure on Natural Vegetation. In Roder, A. and Hill, J., editors, *Proceedings of 1st International Conference on Remote Sensing and Geoinformation Processing in the Assessment and Monitoring of Land Degradation and Desertification*, Trier, Germany.

Mirschel, W., Wieland, R., Wenkel, K.-O., Nendel, C., and Guddat, C. (2014). YIELDSTAT – A spatial yield model for agricultural crops. *European Journal of Agronomy*, 52:33–46.

Nativi, S., Mazzetti, P., and Geller, G. N. (2013). Environmental model access and interoperability: The GEO Model Web initiative. *Environmental Modelling & Software*, 39:214–228.

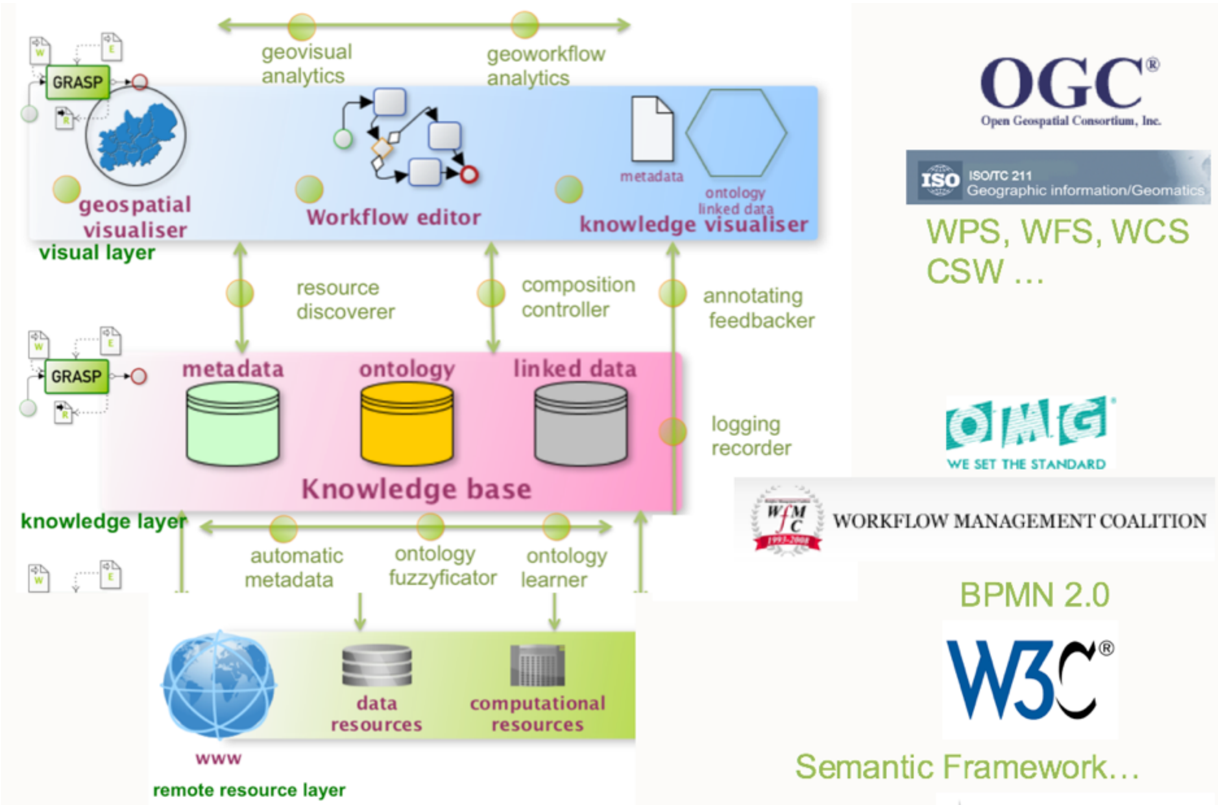

Figure 2. Functional view of the eGRASP framework.

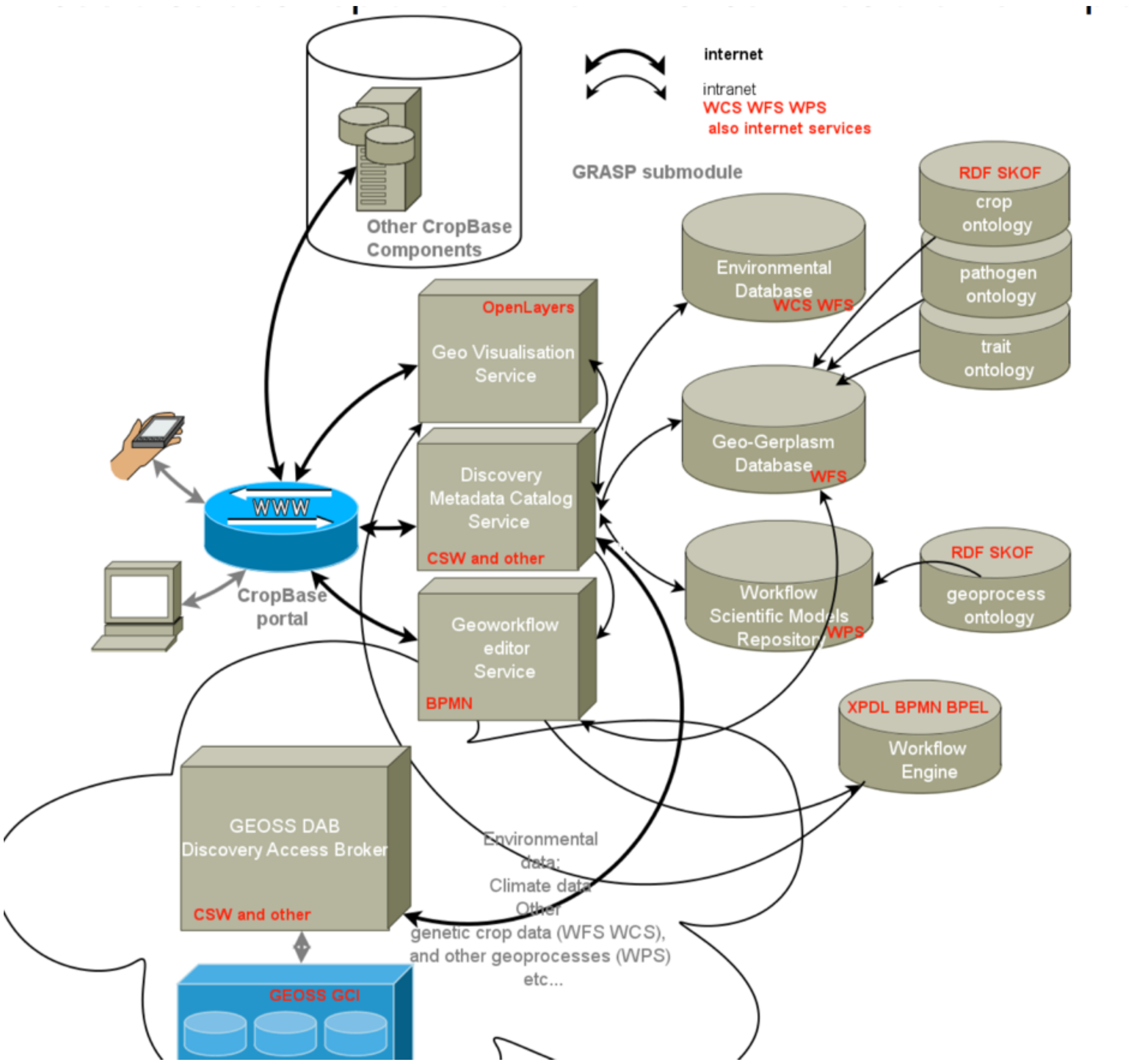

Figure 3. The eGRASP framework platform.