**PODD - An Ontology-Driven Data Management System for Plant Phenomics Research**

**Gavin Kennedy1,2, Yuan-Fang Li1, Faith Davies1,Jane Hunter1**

1. The Australian Plant Phenomics Facility, Canberra, Australia

Gavin.Kennedy@csiro.au

1. The University of Queensland, Brisbane, Australia

{ g.kennedy1,uqyli4, f.davies, j.hunter1}@uq.edu.au

Data management has become a critical challenge faced by a wide array of scientific disciplines in which the provision of sound data management is pivotal to the success and impact of research projects. The huge and rapidly growing amounts of data to be managed and the fact that the models of data evolve over time contribute to making data management an increasingly complex undertaking that warrants a rethinking of its design. A number of intrinsic characteristics of Semantic Web ontology languages OWL and RDF Schema, such as semantic rigor and the extensible nature, make them an ideal conceptual platform on which effective data management systems can be developed. We have designed the Phenomics Ontology Driven Data repository (PODD), a domain independent, ontology-centric architecture for data management systems that is open and extensible. In this architecture, the behaviors of domain concepts and objects are captured entirely by ontological entities, around which all data management tasks are carried out. Moreover, the open and semantic nature of ontology languages also makes such systems amenable to greater data reuse and interoperability. An ideal domain for applying these principles is phenomics, the systematic study of phenotypes of organisms. Plant phenomics in particular generates high volumes of heterogeneous data and makes use of emerging imaging and measurement technologies and processes, thus making it an ideal domain for the ontology approach to data management. In this context, we describe the development of the PODD data management system for phenomics research, as a step towards validating the practicality of the ontology-centric architecture.