

## FOSS geospatial libraries in scientific workflow environments: experiences and directions

Graeme McFerren · Terence van Zyl · Anwar Vahed

G. McFerren. T. van Zyl. A. Vahed  
ICT4EO Research Group, Meraka Institute,  
CSIR, Pretoria, South Africa  
e-mail: gmcferrren@csir.co.za

**Abstract** In multiple research fields such as astronomy, bio-informatics, chem-informatics, geophysics and eco-informatics, scientists are increasingly turning to e-science and specifically scientific workflows as a way of improving, broadening, hastening and sharing their results. Enhanced collaboration, ad hoc access to tools, data and high-performance processing facilities are some of the gains to be made. Scientific workflows are concerned with, amongst others, supporting the repeatability and provenance of experiments. In context of three sets of research (wildfire research, flood modelling and the linking of disease outbreaks to multi-scale environmental conditions), we describe our efforts to provide geospatial capability for scientific workflow software environments to support researchers in exploring, integrating and visualising earth observation and geographic data in conjunction with other research data. We note that functionalities for data ingest (raster and vector), data transformation (reprojection and simplification), data export and spatial overlay operations commonly are required. We find a relative lack of support for geospatial data, services and these functions within several Free and Open Source Software (FOSS) scientific workflow packages. Furthermore, we highlight some software development and data encoding challenges faced when utilising various FOSS geospatial libraries within these scientific workflow environments. Finally, we offer suggestions for improving the integration of geospatial data as well as processing and analysis software tools into such environments.