The Workflows of Ontology Authoring: Controlled vs Naturalistic Settings

Markel Vigo, Nicolas Matentzoglu, Caroline Jay, Robert Stevens
{markel.vigo | matentzn | caroline.jay | robert.stevens}@manchester.ac.uk
School of Computer Science
University of Manchester

In the last editions of the UKON meeting we have reported the results of how individuals go about authoring ontologies: in an interview study in 2014, we learned about the problems encountered by ontologists and the strategies they employed to circumvent them. In 2015, the log files generated by Protégé4US –an instrumented version of Protégé for User Studies– and eye-tracking data allowed us to identify ontology authoring workflows in a controlled environment.

We report the results of a more naturalistic, ecologically-valid study with Protégé4US in which the interaction of ontology authors' with Protege was collected for a few days. In order to evaluate the external validity of previous studies no instructions were given this time: ontologists carried out their own authoring tasks on their own ontologies at their own pace.

We find that the workflows exhibited in the lab do also occur into the wild: adding and editing restrictions to classes is again the most common activity, exploration of the asserted class hierarchy is often a precursor to entity addition and modification, and the inferred hierarchy is explored after running the reasoner.

We also identify new behaviours we did not observe before: the environment is not only saved before running the reasoner, but also after heavy editing and before loading another OWL file. The search functionality is frequently used and annotations are added as soon as individuals are created. Interestingly, 'undoing' is a precursor of heavy editing.