Improvements to Fedora Content Model Architecture

Blekinge, Asger Askov

With the release of Fedora Commons 3.0, the Content Model Architecture (CMA) was added to Fedora. It was not meant as an end-all solution, but rather as a starting point for building more advanced content models. The Fedora team expected the user community to figure out the best ways to use and amend this design. Now, the CMA have been around for a while, and certain improvements have, by agreement of the Fedora Committer Team, matured enough to be brought back into the core Fedora system, probably with the coming Fedora 3.4 release. This proposal aims to present these improvements to the Fedora community. Most of this come from the Enhanced Content Models, developed by the State and University Library, Denmark. The improvements include the following; Datastream contents can be specified with XML Schema, and more schema languages are to be supported. Ontologies are specified in a subset of OWL Lite, and allows the user to put constraints on the relations from the data object, or relations from the datastreams themselves. In addition, Fedora will have a new API method for validating data objects against their content models. It will be possible to mark a datastream as optional, so that the validator will not report an error, if the datastream is not present, but still validate the datastream if it is present. Lastly, it will be possible to specify further descriptive information about the datastreams of data objects in the content models.