

Course Project

Assignment:

Implement a RSS Reader using the Micro XQuery engine

The reader should provide the following features:

- Retrieve RSS from one/many sources (e.g. NY Times, 20min, heise.de, ETH Life, ...)
- Understand multiple RSS formats (RSS 0.9, 1.0, 2.0, Atom)
- Render RSS “documents” into a suitable human-readable format, e.g. XHTML or plain text
- Sort/Group RSS “documents” according to date, author, ...
- Search for specific values of attributes/elements (e.g. author = “Duda”)
- Full text search over the complete contents
- Combine multiple feeds
- Feed Statistics
 - Most active authors
 - Most popular keywords
- Re-publish a combined RSS feed (to file, compliant to one of the formats above)
- Mark Read “news”, only display unread
- Subscription Management
- Offline Feed Cache

- Periodic polling
- Filter subscriptions (only show newly arrived with specific properties, expressed as XQuery)
- Filter “Spam”, e.g. annoying authors

Not all features need to be implemented, but the ones up to “offline feed cache” are. Features below are optional. The amount of Java code should be kept to an absolute minimum, as much as possible should be implemented in XQuery/XQuery Update and possibly XQueryP.

We provide a simplistic GUI Java program (Swing) with buttons/rendering area so that you can focus on implementing the RSS reader, which is available for download at http://www.dbis.ethz.ch/education/ws0708/xml_db_ws2007/project/RSSReaderGUI.zip

A basic solution for retrieving, storing, updating and displaying a single RSS feed is already provided.

Within the scope of this projects, use the MXQuery XQuery Engine. A current version is already included in the GUI, separate downloads are available via Sourceforge: <http://sourceforge.net/projects/mxquery/>

If there are any problems with the GUI or MXQuery, please contact Peter Fischer.

You can form teams with up to three students to work on this project. The project is due in the week between November 26th and November 30th. Please arrange an appointment with Cristian Duda (cristian.duda@inf.ethz.ch) or Tim Kraska (tim.kraska@inf.ethz.ch) to get a review & acceptance for the project.

An accepted project is the requirement to participate on the final exam of the course. There will be no grade on the project, just accept/not accept