

A Proposal for OSGEO Educational Authoring Framework using DocBook and Subversion

Reactions/Feedback Appreciated!

By

Alexander Stepanov, Charlie Schweik, Maili Page – University of Massachusetts, Amherst
Ned Horning – American Museum of Natural History

1. INTRODUCTION

In order to move forward in a more organized manner, the OSGeo Education group needs two components in place: a standard method for creating educational content, and a web-based versioning system to manage this content. After conducting substantial research and exploration on alternatives, we are proposing that OSGeo education use:

- (1) **XMLMind** (<http://xmlmind.com/xmleditor/persoedition.html> or another XML editor) for creation and editing educational content;
- (2) the **Apache DocBook Framework** (<http://velocity.apache.org/docbook/>) for rendering documents in PDF and HTML formats; and
- (3) **Subversion** for content version control and to promote searching capabilities (OSGeo already has this in place on their server).

This document provides a short description of our “DocBook” proposal. We have successfully implemented this process and are currently writing tutorials on this process. But we want to solicit feedback from the OSGeo edu community before we go much further. We should note that what we are proposing is a *standard* (not an absolute requirement) that we hope other OSGeo educational content providers would follow so we can build a content repository.

2. THE BENEFITS OF DOCBOOK

DocBook is a markup language for technical documentation that uses XML as its foundation. The benefits of using the DocBook format are:

- 1- It enables users to create document content in a neutral form that captures the logical structure of the content; this content can then be published in a variety of formats, including HTML and PDF without requiring users to make any changes to the source.
- 2- Documents (article or book) can be created relatively easily.
- 3- It provides a standard method for OSGeo authors to collaborate on modules/materials, whether they be educational modules or complete books (like the Free GIS book effort).
- 4- Several authors can work on different sections of the same document at the same time and generate a final output where the formatting is the same throughout the document; Subversion, the version control system, tracks changes and merges them into the document.
- 5- Revision history and authorship are maintained in the document through our selection of XML tags. In particular, there are metadata tags in DocBook document itself. In addition, Subversion tracks document history and introduced modifications.

- 6- It stores content in an XML format thereby providing flexibility. DocBook standard/tags are well documented.
- 7- By having content conform in this standard, it will enable OSGeo users to pull out and combine content of interest (perhaps from a number of OSGeo contributors) and have them generally follow a similar structure and look.
- 8- The authoring software runs on all platforms (Linux, Windows, Mac)
- 9- It will allow us to develop an indexing/search system (we think using Subversion) to improve the searching of modules, and to avoid duplication
- 10- DocBook appears to be becoming a defacto standard; it is used by other open source communities. For example, the Ubuntu project just provided a new set of tutorials using DocBook and it is being used in the Eclipse project. A list of projects, relying on DocBook, may be found at <http://wiki.docbook.org/topic/WhoUsesDocBook>

3. THE PROPOSED GENERAL PROCESS: DOCBOOK AND SUBVERSION

Figure 1 below summarizes the proposal. Starting at the bottom of the graphic (STEP 1, FIG 1), a variety of OSGeo authors globally would be writing their educational content using an XML template we will provide and either “XMLMind” (<http://www.xmlmind.com/xmleditor/persoedition.html>) or other XML editor (Many popular open source editors, for example Vim and Emacs, have support for XML file editing. There is some potential to use Open Office Writer, though it would take some work for us to figure this out and we are not sure it is worth it). When complete, the author would “check in” the completed document (these are XML modules) into the OSGeo Educational Material Subversion system that resides on the OSGeo server (STEP 2, FIG 1). Periodically the Subversion system could run a script to convert the modules to standard OSGeo educational output (STEP 3, FIG 1). Pdf and HTML format is likely. Like in open source software management, documents will have designated “owners” or “maintainers” who have check out and “commit” authority. If someone wants to create a new derivative of a document they will need to contact the owner of that educational material to get access to the XML files and the authors will have to point them to the tutorials we are creating on document authoring.

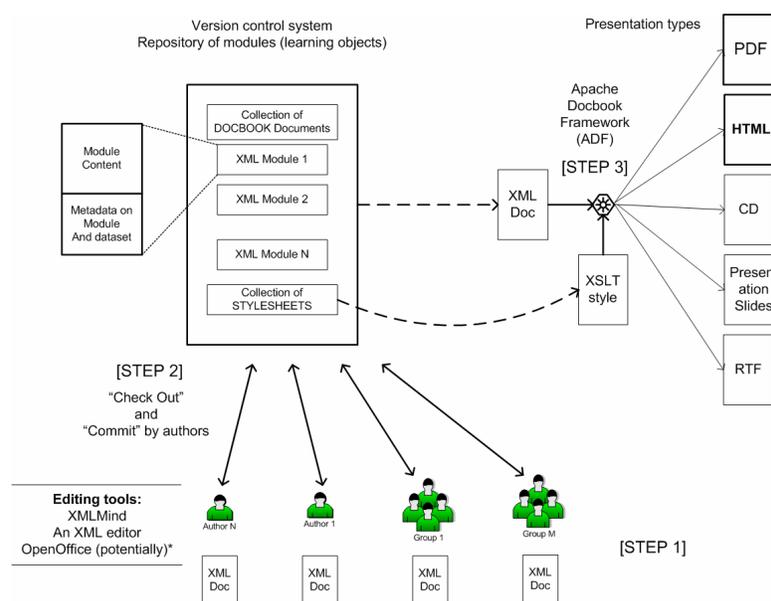


Figure 1. Proposed OSGeo Educational Document Authoring and Version Control Framework

4. PROPOSED PROCEDURE FOR ALL OSGEO EDU AUTHORS TO FOLLOW

The steps in Figure 1 are ultimately where we'd like to go. But Step 3 in Figure 1 will require setting up the Subversion server to regularly convert the XML Docbook files to PDF and HTML. We need to talk to Tyler about whether this is feasible. In the meantime, OSGeo education authors should first download and install XMLMind (<http://xmlmind.com/xmlmind/persoedition.html>) or another XML editor and the Apache DocBook Framework (<http://velocity.apache.org/docbook>) on your computer.

Once the above is completed, ***the steps to create OSGeo educational content are as follows:***

1. Write your article or book using the XML (DocBook format) template we provide using your XML editor (e.g., XMLMind).
2. Process the DocBook XML file (we recommend ANT included in Apache Framework) to transform your article or book into a pdf and HTML page.
3. When ready, upload both the XML source file of the module and the pdf/html files to the OSGeo Edu's Subversion system. If we can implement STEP 3 (FIG 1) you will only need to upload the XML file.

5. AN EXAMPLE OF THE OUTPUT

An pdf example of one of our tutorials that is output from this process using our XML template can be found at http://people.umass.edu/cschweik/pdfs/db_foss_m1.pdf. An example of both the html and pdf output are in a .zip file that can be downloaded at: <http://people.umass.edu/cschweik/pdfs/dbmodule.zip>.

6. CONCLUSIONS

IMPORTANT! WE WANT TO MOVE FORWARD WITH THIS PROPOSED PROCESS. BUT BEFORE PROCEEDING TOO MUCH FARTHER, WE WANTED TO GET THE PERMISSION OF OUR OSGEO EDU COLLEAGUES.

ANY OBJECTIONS TO US MOVING FORWARD?

6. SOME USEFUL URLS:

<http://velocity.apache.org/docbook> - Download DocBook software from here

<http://www.e-ucm.es/drafts/7.pdf> - A paper "Using DocBook and XML Technologies to Create Adaptive Learning Content in Technical Domains"

<http://en.wikipedia.org/wiki/DocBook> - Wikipedia overview of DocBook with more links