

ES-DOC CIM 2.0 & CMIP6 Report

Date: 22-07-2015
Author: Mark A. Greenslade (MG)

Notebook

Bryan Lawrence (BL) has been developing a python based desktop tool called the **esdoc-notebook** (currently [hosted on BitBucket](#)). The tool allows a user to create, open, edit & associate “documents”. All documents are stored to the local file system in the python pickle format.

The notebook outputs CIM v2 compliant documents. The notebook defines the [CIM v2 schema](#) using a similar syntax to the [CIM 1 schema](#) definitions found in the esdoc-mp (meta-programming) framework. The notebook also defines [schema adaptations for CMIP6](#). BL has also attempted to represent constraints such as controlled vocabularies using a similar pythonic syntax.

Notebook Shell

In order to streamline notebook installation & usage MG developed the [esdoc-nb-shell](#). The shell pulls together the notebook, esdoc-mp & pyesdoc libraries thereby supporting the eco-system integration process. Currently the shell supports the following commands:

```
esdoc-notebook-install  
esdoc-notebook-generate  
esdoc-notebook-open  
esdoc-notebook-uninstall  
esdoc-notebook-update  
esdoc-notebook-validate
```

NOTE - The notebook installation relies upon the user’s system python as the GUI library pygtk cannot be compiled against a virtual environment. Thus before installing the notebook the user will need to ensure that the gtk+ libraries are installed.

CIM v2 Schema

As mentioned above, the CIM v2 schema syntax defined within the notebook is similar but not identical to the CIM v1 schema syntax. As a result MG was obliged to develop a schema adaptor so that the CIM v2 schema can be recognised by the esdoc-mp framework. This adaptor was an essential step to extending pyesdoc with the CIM v2 types within pyesdoc (the esdoc-mp code generator was used to generate the CIM v2 types from the CIM v2 schema). Now that the CIM v2 schema can be recognised by esdoc-mp it can be ported to esdoc-mp directly.

CMIP6 Vocabularies / Constraints

As mentioned above BL has attempted to represent constraints such as controlled vocabularies using a similar pythonic syntax. BL has developed an initial version of a [sea-ice vocabulary](#). BL is approaching the sea-ice community in order to develop this vocabulary further (using the same pythonic approach). It should be noted that both the esdoc-mp and pyesdoc libraries will need to be extended to support this constraint language, the work is considered to be in the order of 2 weeks from start to finish.

The notebook is initially to be used by Charlotte Pascoe and Sebastien Denvil in order to define the CMIP6 experiment definitions. The notebook is extensible as it is based upon meta-programming techniques and thus the notebook can (and does) support other document types.

Issues / Tasks

1. Notebook maintenance is an issue. It will take someone with very good python skills to be able to take it over. At present only BL really understands it.
2. The CIM v2 schema definition should be moved into the esdoc-mp package. The notebook will take a dependency hit upon esdoc-mp. This is a low complexity task.
3. The notebook should be updated so that instead of outputting documents in python pickle format the documents are written to the file system in xml/json formats using pyesdoc. This will make the documents more resilient to schema changes. This is a medium complexity task.
4. The notebook-shell should be extended to support publishing to the remote ES-DOC web service, this will mean that the documents will become available to other tools such as the ES-DOC viewer and/or comparator. A new command can be developed to support this - i.e. esdoc-notebook-publish. This is a medium complexity task.
5. The notebook should be moved from BitBucket into the ES-DOC GitHub space. This is easy to do but the tickets will probably need to be kept in BitBucket in the short term. This is a low complexity task.
6. Extending esdoc-mp & pyesdoc to support the constraint language will take some effort. There are some low hanging fruit to attack first and then afterwards there is the prototypical sea-ice vocabulary. It might be of interest to use the prototype pyessv library to provide CV support. This is a medium complexity task.

Links

ES-DOC Notebook: <https://bitbucket.org/bnlawrence/esdoc-nb>

CIM v1 Schema: <http://bit.ly/1GOuCpI>

CIM v2 Schema: <http://bit.ly/1Kqq6To>

CIM v2 Schema Adaptor: <http://bitly.com/1Kqq6Tp> & <http://bit.ly/1fupQle>

CIM v2 Types (generated using esdoc-mp): <http://bit.ly/1OFn9zU>

CMIP6 Sea Ice Vocabularies: <http://bit.ly/1luW3ML>

pyessv library: <https://github.com/Prodiguer/pyessv>