

Lightweight derivation and translation of XML vocabularies

- <http://dyomedea.com/papers/2004-extreme/>
- Extreme Markup Languages 2004
 - Eric van der Vlist, Dyomedea
(vdv@dyomedea.com)
 - Gérard Vidal, ENS Lyon
(Gerard.Vidal@ens-lyon.fr)
 - Sébastien Chaumat, ENS Lyon
(Sebastien.Chaumat@ens-lyon.fr)
 - Sébastien Pilloz, ENS Lyon
(Sebastien.Pilloz@ens-lyon.fr)

Context: ENS (Ecole Normale Supérieure) de Lyon

- The strategy
 - Open Standards for the Learning Content Management System (LCMS)
- The tactic
 - A customized vocabulary
 - Small subset
 - Translation of element and attribute names in French
 - As a single and simple step

The OpenOffice "schema"

- Element definitions
 - Name
 - DocBook name
 - Prototype name
 - Element content
 - Attributes
 - Miscellaneous

Compact RELAX NG

- How do you describe complex content models in a cell?
- That's a job for RELAX NG's compact syntax...

```
titre, métadonnéesDArticle?, (contenuBloc|partie)+
```

```
-----
```

```
attributsCommuns, alignement?, alignementVertical?
```

- Design decision: a named pattern per element and attribute

Other definitions

- Attribute definitions
 - Same principle
- Global definitions
 - Same principle

W3C XML Schema generation

- OpenOffice to RELAX NG (XSLT)
 - Straightforward
 - Beware of "@table:number-columns-repeated"
- RELAX NG to W3C XML Schema (trang)
 - Works out of the box
 - Substitution groups can be switched off
[tx:enableAbstractElements = "false"]

What kind of design pattern is that?

- Protocol adaptation
- Including a facade
- Implemented as a "derivation by redefinition"
- Could deserve a formal proof
- Other approaches: derivation by restriction, by annotation, ...

CSS stylesheet translation and augmentation

- Context aware translation
- PyParsing
 - A grammar based parsing library for Python
 - Hacked to perform punctual replacements
- Spreadsheet driven translation
 - Based on my "XML data driven classes" library
- CSS stylesheet augmentation

XSLT generation

- LivreDoc to DocBook
 - Gives access to the DocBook resources
 - Useful to test our derivation by redefinition
- DocBook to LivreDoc: more challenging

Conclusions

- A spreadsheet is a good XML design tool
- Restriction and translation can be done in one step
- Users still need to accept the design decisions of the base vocabulary