Subject classification

Metadata delivered via OAI-PMH contain a broad range of subject headings and classification information. The used classification and subject heading systems and the presentation formats vary broadly. In most cases this information appears in simple dc format in the subject element. Classification information is often used for groupng a repository into items under discipline orientated aspects. Therefore such information appears frequently in the OAI setSpec element. EPrints repositories (LoC classification) and DINI-certificated repositories (DDC) are examples for this approach.
Most frequent used classification schemes in OAI context are

- Library of Congress Classification http://www.loc.gov/catdir/cpso/lcco/
- Dewey Decimal Classification (DDC) http://www.oclc.org/dewey/
- Universal Decimal Classification http://www.udcc.org/

Frequently used subject headings systems in OAI context are

- Library of Congress Subject Headings (LCSH)
- Schlagwortnormdatei (SWD)

Besides this, OAI metadata contain discipline-related classification codes from schemes such as the Mathematics Subject Classification (MSC) and the Medical Subject Headings (MeSH) but also different local classification information.

Currently, services based on this information have serious problems to extract the information from the delivered data in an appropriate way. The first step to improve the situation should focus on making the used technique and classification scheme transparent to the service provider.

DRIVER recomends that the repository should transport the information related to the usage of classification and subject headings in the description element of the Identify response. When a classification is used for structuring the repository via sets, the classification part should be repeated in the subject element.

Best practice is to transport the classification in the element subject "URI-field" using an authoritative namespace in order to support recognizing the classification scheme. Based on this information service providers can use it for establishing services as classification browsing. This includes substituting classification codes by English terms, translating terms to different languages or doing a merge of classification codes using mapping rules.

It is recommended to use an URI when using classification schemes or controlled vocabularies especially when codified schemes are used DDC or UDC. Service providers can recognise encoding schemas more easily when the schema is "URI-fied" by an authority namespace. When the classification scheme is codified, use a human readable text of the code, preferably in English, directly below the codified element. For example:

```
<dc:subject>info:eu-repo/classification/ddc/641</dc:subject>
<dc:subject>Anatomy</dc:subject>
```

If no specific classification scheme is used we recommend the Dewey Decimal Classification (DDC). The first 1000 terms are called the Dewey Decimal Classification Summary and can be downloaded at http://www.oclc.org/dewey/resources/summaries/ if one agrees with the following terms and conditions: http://www.oclc.org/research/researchworks/ddc/terms.htm