



Using simple Dublin Core to describe eprints

by Andy Powell, Michael Day and Peter Cliff
UKOLN, University of Bath
Version 1.2

Introduction

This document provides some recommendations for the use of simple [Dublin Core metadata](#) [1] to describe eprints in eprint archives.

These recommendations are not primarily targeted at the end-users of eprint archives. Rather, they are intended to guide 'best-practice' for the maintainers of eprint archives in order that such systems can be configured to maximise the benefits of a shared approach to metadata disclosure. The intention is to facilitate more consistent results when 'cross-searching' and browsing metadata records gathered from multiple eprint archives. Nonetheless, we would hope that these guidelines to have some impact on the cataloguing guidelines and help sub-systems offered to end-users of institutional eprint archives.

As can be seen from the recommendations for dc:type (below), this document uses a fairly wide working definition of 'eprint':

"an electronic copy of an academic paper" [2].

These recommendations are specific to the use of simple (unqualified) Dublin Core. The intention is to develop a separate set of guidelines for the use of qualified Dublin Core (including the use of element refinements and encoding schemes) to describe eprints.

These recommendations draw on three documents:

- [DCMI Elements and Element Refinements - a current list](#) [3]
- [Using Dublin Core](#) [4]
- [The RDN Cataloguing Guidelines](#) [5]

General Recommendations

All eprint archives must support the 'oai_dc' record format (this is mandated by the [OAI-PMH](#) [6]).

All eprint archives are strongly recommended to support the following minimal set of DC elements within the 'oai_dc' record format.

- [dc:title](#)
- [dc:creator](#)

- [dc:subject](#)
- [dc:description](#)
- [dc:publisher](#)
- [dc:date](#)
- [dc:type](#)
- [dc:identifier](#)
- [dc:language](#)
- [dc:relation](#)

The values of these elements must **not** contain any HTML (or XML) markup. They may contain [LaTeX](#) [7] commands if desired but it is worth remembering that there is no mechanism for explicitly indicating that LaTeX is being used.

The Elements

This section lists each of the Dublin Core elements. For each element, an eprint-specific recommendation is provided followed by the authoritative definitions and comments from the Dublin Core Metadata Initiative.

(*) indicates that an element is part of the minimal set.

dc:title (*)

Eprint-specific

Recommendation: The title of the eprint.

Preserve the original wording, order and spelling of the eprint title. Only capitalize proper nouns. Punctuation need not reflect the usage of the original. Subtitles should be separated from the title by a colon. For example:

```
<dc:title>Initial sequencing and analysis of the human genome</dc:title>
<dc:title>The new nationalism and the old history: perspectives on
the West German Historikerstreit</dc:title>
```

If necessary, repeat this element for multiple titles.

DC Label: Title

DC Name: title

DC Definition: A name given to the resource.

DC Comment: Typically, a Title will be a name by which the resource is formally known.

dc:creator (*)