1 Introduction

Metadata is information in a structured format that describes a resource. One of the key components for delivering findNZarticles is the quality and consistency of the metadata contributed.

This document defines the metadata standards that partners must meet, and it is split into three sections:

1. Section Two provides an explanation of what data formats will be currently accepted for participation in findNZarticles.
2. Section Three provides an overview of the standards and is intended to give the contact person within a partner organisation a basic understanding of what is required to comply with the guidelines for the purposes of the findNZarticles.
3. Section Four provides some examples of the above mentioned data formats.

2 Data Acceptance

findNZarticles is based on both centralised and distributed databases. This document specifies the preferred formats for the provision of indexing data from contributing sites to a central repository. Three alternative formats are given. These formats have been chosen so as to make it as easy as possible for libraries to contribute records to the central MARC/XML database. For those organisations that are unable to generate MARC or MARCXML records we will accept two text formats, Dublin Core [DC] and RIS, which are often available as an output format from library applications, such as Endnote or Reference Manager.

Data in any of the following formats will be accepted

1. Data extracted and converted from the local format to MARC or MARC XML format based on an agreed mapping document
   or
2. Where either the tools or expertise to generate MARC are not available, data in RIS text format is acceptable.
   or
3. For libraries already working with Dublin Core, data tagged with Simple or Qualified DC tags will be an acceptable format for converting to MARC/XML.
If your indexing data is in another format that can be expressed as delimited text you may be able to easily convert it to MARC using a free utility such as MARC Edit <http://oregonstate.edu/~reeset/marcedit/> to translate text to MARC.

3 Standards

3.1 MARC and MARCXML

MARC is the acronym for MAchine-Readable Cataloging. It defines a data format that emerged from a Library of Congress-led initiative that began thirty years ago. It provides the mechanism by which computers exchange, use, and interpret bibliographic information, and its data elements make up the foundation of most library catalogues used today. MARC became USMARC in the 1980s and MARC21 in the late 1990s.

The record structure of MARC is an implementation of ISO 2709, also known as ANSI/NISO Z39.2. MARC records are comprised of three elements: the record structure, the content designation, and the data content of the record. The record structure implements national and international standards (e.g. Z39.2, ISO2709.) The Content Designation is the codes and conventions established to identify explicitly and characterize data elements within a record and support their manipulation. The content of data elements in MARC records is defined by standards outside the formats such as AACR and Library of Congress Subject Headings.

MARCXML is an XML schema based on the fairly common MARC21 flavour among the bibliographic MARC standards. It was developed by the US Library of Congress and adopted by it and others as a means of easy sharing of, and networked access to, bibliographic information.

3.1.1 Specifications for key fields

773 Host Item Entry field

This field is provided in order to enable the user to locate the physical piece that contains the component part or subunit being described. Thus, only those data elements required to assist in the identification of the host item need to be included in the field. Information in subfield $g that points to the exact location of the component part within a bibliographic item is necessary.

The 773 is made up of a number of subfields stored in the following order

$t title of source journal

$g date, volume, issue/number, and pagination information

$h size
In order to provide the easy parsing of MARC/MARCXML data for the creation of OpenURLs, the format of the 773$g subfield is important.

Ideally data will be in the following format:

[Date]; [volume]. [number/issue]: [page(s)]

**Date**

Publication date information is contained in the 773$g from the start of the subfield until the first semi-colon “:” character.

Ideally, the publication date should be given in the form:

- YYYY-MM-DD, or
- YYYY-MM, or
- YYYY.

as appropriate.

If date fields are in readable text, these should be in the format of:

[day], space, [month as 3 letter abbreviation or spelled out in full], space, [year, as YYYY]

e.g. 28 Dec 2005

**Volume, issue number, pagination**

Ideally displayed as [lower case] v, dot, [volume number], space, [lower case] n, dot, [number], colon, space, [lower case] p, dot, [pagination]

e.g.

v.12 n.67: p.122-138

v.8 n.1: p11,13

**Examples**

$\text{t} \text{Otago Daily Times}, \text{g} \text{28 Dec 2005; p.21} \text{x} \text{0114-426X}$

$\text{t} \text{Listener, g 25 Dec 1989; v.125 n.2599:sup.p.54} \text{x} \text{0110-5787}$

or in MARC XML

<MR773 ind1="0">
3.2 Dublin Core Metadata Element Set

The Dublin Core metadata element set is a standard for cross-domain information resource description. It provides a simple and standardised set of conventions for describing things online in ways that make them easier to find. Dublin Core is widely used to describe digital materials such as video, sound, image, text, and composite media like web pages. Dublin Core is defined by NISO Standard Z39.85-2001.

The Dublin Core standard includes two levels: Simple and Qualified. Simple Dublin Core comprises fifteen elements; Qualified Dublin Core includes three additional elements (Audience, Provenance and RightsHolder), as well as a group of element refinements (also called qualifiers) that refine the semantics of the elements in ways that may be useful in resource discovery.

3.2.1 Specifications
As no one has contributed indexing data to findNZarticles in Dublin Core this information is currently in a separate “draft” document nl_cims-#240609-Draft Dublin Core metadata guidelines for contributors to findNZarticles. Once a library has contributed some data in Dublin Core and it has been successfully added to the central repository this section will be updated.

3.3 RIS Format

The RIS file format is a tagged format for expressing bibliographic citations. It is supported by a number of databases and reference managers, such as Endnote or Reference Manager.

3.3.1 Specifications

Unicode Compliance

Please ensure the submitted file is UTF-8 (Unicode) compliant.

Elements
For a complete list of RIS elements, see the Reference Manager web site at http://www.refman.com/support/risformat_intro.asp

Use separate lines for each tag
Each tag and its contents must be on a separate line, preceded by a “carriage return/line feed” (ANSI 13 10).
Tag order
Except for the first tag of each reference, which must be “TY - “ and the last tag of each reference, which must be “ER - ,” the tags within each reference can be in any order.

Characters allowed in fields
The characters allowed in the reference ID fields can be in the set “0” through “9,” or “A” through “Z.” The characters allowed in all other fields can be in the set from “space” (character 32) to character 255 in the IBM Extended Character Set. Note, however, that the asterisk (character 42) is not allowed in the author, keywords or periodical name fields.

How to handle long fields
If the information following any one tag is more than 70 characters long, it is allowable (though not necessary) to insert a carriage return/line feed at the end of 70 characters, and continue on the next line.

4 Examples

4.1 Example MARC Records

```
00868nan 2200217z |
45000100170000005001700007008004100024033001300065040000800780450013000
862450043000992502720014250000800414521000120042260000310043465001800465
650002800483700001900511400006000530773006000590-452439-20050222133439.0-050
22s2005 nz | g |00 | eng d- a20050222- aDDS-0 a20050217-00aAt home on the range - or on the stage- aMeets the singer/songwriter, who will sing his composition 'Otago my home' accompanied by the Southern Sinfonia at the Last Night at the Proms concert, Dunedin Town Hall, 18 Feb 2005. Discusses some of his protest songs provoked by environmental challenges in Otago.- all- aGeneral-14aCurtis, Marjorie(Cardrona)-17aMusic-2apaist-17aPerforming arts2apaist-14aCook, Marjorie- tOtago Daily Times,g17 Feb 2005; P.29;h38cmx0114-426X-0 tOtago Daily Times,g17 Feb 2005; P.29;h38cmx0114-426X-0 00779caj 2200229z
450000100170000005001700007008004100024033001300065040000800780450013000
862450043000992501500016950000100003195210001200329600004400341650001800385
650002000430009925000170004237000210044040004400461773004400505-438245-2004091
0160336.0-040910s2004 nz | g |00 0 eng d- a20040910- aDDL-0 a20040801-00- aAngela Annabell, 1929-2000; Composed music of the colonial period- aPays tribute to the musicologist and brings to light her previously unpublished article on composed NZ music of the 19th-century colonial period.- aNotes- aGeneral-14aAnnabell, Angelag(1929-2000, Auckland)-17aMusic2apaist-17aPerforming arts2apaist-14aPalmer, Jill-14aAnnabell, Angela- tCrescendo,gAug 2004; 68:6-8x0111-8994-0 tCrescendo,gAug 2004; 68:6-8x0111-8994-000 00895can a2200229z| 450
001 452439
005 20060604014634.0
008 050222s2005 nz | g |00 | eng d
033 __ |a 20050222
035 __ |a (INNZ)452439
040 __ |a DDS
045 0 |a 20050217
245 00 |a At home on the range - or on the stage
500 __ |a Ill
```
520 __ |a Meets the singer/songwriter, who will sing his composition 'Otago my home' accompanied by the Southern Sinfonia at the Last Night at the Proms concert, Dunedin Town Hall, 18 Feb 2005. Discusses some of his protest songs provoked by environmental challenges in Otago.
521 __ |a General
600 14 |a Curtis, Martin |g (Cardrona)
650 17 |a Music |2 apaist
650 17 |a Performing arts |2 apaist
700 14 |a Cook, Marjorie
773 0_ |t Otago Daily Times, |g 17 Feb 2005; p.29 |h 38cm |x 0114-426X

Or

= LDR 00000nam 22000000a 45e0
=008 050819s999994||\x|x|\\\\\\\\000\0\und\d
=245 \$a Illustrations of indigenous New Zealand plants in "Curtis's Botanical Magazine"=500 \$a 16 refs; 2 tables.
=500 \$a Matthews 1977
=520 \$a Illustrations of 135 New Zealand indigenous plants have appeared in "Curtis's Botanical Magazine", and these are tabulated, together with any subsequent taxonomic judgement. (auth)
=653 \$a ILLUSTRATIONS; INDIGENOUS TAXA; PLANTS; FLOWERS; LEAVES
=700 \$a Matthews, B.A., $g DSIR, Botany Division, Lincoln
=773 \$t New Zealand Journal of Botany$x0028-825X$y NZJBAS$g NZJBot 15(4): p. 767-773

4.2 Example MARC XML Record

```xml
<MARC>
  <MRleader>01472caj a2200241z 4500</MRleader>
  <MR001>444083</MR001>
  <MR005>20060603192549.0</MR005>
  <MR008>041116s2001 nz | g |00 0 eng d</MR008>
  <MR033>
    <MR033a>20041116</MR033a>
  </MR033>
  <MR035>
    <MR035a>(INNZ)444083</MR035a>
  </MR035>
  <MR040>
    <MR04oa>DDL</MR04oa>
  </MR040>
  <MR045 ind1="0">20010101</MR045a>
  <MR245 ind1="0" ind2="0">Digitisation, knowledge and metadata : the National Library of
  New Zealand as a case study</MR245a>
  <MR245>
    <MR500>Notes; refs</MR500a>
  </MR500>
  <MR520>
    <MR520a>Backgrounds the National Library of NZ's development of a
    metadata standards framework to guide its cataloguing/indexing/description activities as it
    increases its presence online. Describes the initial goal of the Library with respect to
```
metadata, its growing appreciation of the breadth and depth of metadata initiatives worldwide, and its decision to focus on resource discovery metadata as the initial output of a continually evolving metadata standards framework.</MR520>
</MR521>
<MR521a>Research</MR521a>
</MR521>
<MR650 ind1="1" ind2="7">
<MR650a>Information management</MR650a>
<MR6502>apaist</MR6502>
</MR650>
<MR650 ind1="1" ind2="7">
<MR650a>Standards</MR650a>
<MR6502>apaist</MR6502>
</MR650>
<MR650 ind2="7">
<MR650a>Cataloguing</MR650a>
<MR6502>apaist</MR6502>
</MR650>
<MR700 ind1="1" ind2="4">
<MR700a>Knight, Steve</MR700a>
<MR700g>(National Library of New Zealand)</MR700g>
</MR700>
<MR711 ind1="2" ind2="4">
<MR711a>‘Digitisation and Knowledge : Perspectives from Aotearoa/New Zealand’; Auckland University of Technology Centre for New Media Research, 14 Feb 2001</MR711a>
</MR711>
<MR773 ind1="0">
<MR773t>Access : Critical Perspectives on Cultural and Policy Studies in Education</MR773t>
<MR773g>2001; v.20 n.2:p.45-51</MR773g>
<MR773x>0111-8889</MR773x>
</MR773>
</MARC>

4.3 Example RIS records

TY - JOUR
ID - 4489
T1 - Four new species of Dorylaimida: (Nematoda) from New Zealand
A1 - Ahmad,W.
A1 - Sturhan,D.
A1 - Wouts,W.M.
Y1 - 2003/03/
RP - NOT IN FILE
SP - 43
EP - 55
JF - New Zealand journal of zoology
VL - 30
IS - 1
U1 - 2002/2003
U2 - Refereed
U3 - A
N2 - Four new nematode species belonging to the families Aporcelaimidae, Qudsianematidae and Leptonchidae (Dorylaimida) are described from native vegetation in New Zealand. Makatinus silvaticus n. sp. is 2.3-2.7 mm long and has characteristically narrow lateral hypodermal chords, a lip region set-off by a slight depression, a 2527 mum long odontostyle, and a pharyngo-intestinal junction with a disc. The females have a longitudinal vulva and characteristically shaped vaginal sclerotisation; the males have
large spicules and 11-12 ventromedian supplements. Hulqus zelandicus n. sp. females are 0.9-1.0 mm long and have a set-off lip region, a 7-9 mum long odontostyle, wine-glass-shaped amphids with small apertures, a mono-opisthodelphic gonad, a posteriorly directed vagina and a long filiform tail; males are unknown. Proleptonchus attenuatus n. sp. is 1.5-1.9 mm long and slender and has a set-off lip region, a 7.5-8.5 mum long odontostyle, and an elongate pharyngeal bulb; the female has a pseudomonodelphic gonad, the posterior branch being reduced to a vestigial oviduct and traces of an ovary; the males have short spicules. Capilionchus capitatus n. sp. is 1.7-2.2 mm long and slender and has a set-off lip region with a prominent labial disc, a 9.0-9.5 mum long, attenuated odontostyle with fine lumen and aperture, and a pharyngeal bulb set-off by a constriction; the female has a pseudomonodelphic gonad, the posterior branch being reduced to a vestigial oviduct and traces of an ovary; the males have dorylaimoid spicules and five ventromedian supplements.

TY - JOUR
ID - 4235
T1 - Adelie penguins and environmental change
A1 - Ainley,D.G.
A1 - Ballard,G.
A1 - Emsue,S.D.
A1 - Fraser,W.R.
A1 - Wilson,P.R.
A1 - Woehler,E.J.
Y1 - 2003/04/18/
RP - NOT IN FILE
SP - 429
JF - Science
VL - 300
IS - 5618
U1 - 2002/2003
U2 - Refereed
U3 - A
ER -