



Scott  
Polar  
Research  
Institute

# *Have you searched the Library catalogue?*

*Migrating catalogue records from pre-internet  
software to a modern library management system  
supporting MARC21*

Martin Lucas-Smith

Peter Lund



UNIVERSITY OF  
CAMBRIDGE



# Introduction



Photo credit Silvio Tanaka CC BY 2.0



Photo credit: Juan Cristobal Cobo CC BY 2.0

Aim: to create a shared catalogue in Alma  
searchable through Primo (iDiscover)



• Photo credit: Linda CC BY 2.0

Widely accessible in the Arctic...



Public Library, Tromsø

...and Antarctic



The James Clark Ross, Rothera Research Station, British Antarctic Survey

# Bibliography – Polar UDC and Muscat

- Gilbert, M; Lane, H. (2008) Forty-five numbers for snow: a brief introduction to the UDC for Polar libraries. Extensions and Corrections to the UDC, 30 (2008), pp. 23-26.  
<http://hdl.handle.net/10150/110845>
- Porter, M. and Galpin, V. (1988), "Relevance feedback in a public access catalogue for a research library: Muscat at the Scott Polar Research Institute", *Program: electronic library and information systems*, Vol. 22 No. 1, pp. 1-20. <https://doi.org/10.1108/eb046983>

# Impact

Catalogue is accessible from (almost) anywhere

Increase in visits to library by first time users and Geography undergraduates

Increase in inter-library loan requests



Photo credit: Hugh Broughton Architects Halley VI centre CC BY 2.0

# Further developments

Using Alma  
facilitated transfer  
of low use serials to  
off site store

Enabled creation of  
reading lists in  
Leganto.





THANK  
you

- Emma Coonan
- George Cronin
- Martin French
- Rick Frolich
- Pete Girling
- Janey Huber-Reacher
- Robin James
- Ann Keith
- Martin Lucas-Smith
- Frankie Marsh
- Sarah Middle
- Niamh O'Mahony
- Hilary Shibata
- Paul Taylor-Crush
- Isabella Warren
- Jeremy Wong

Editing file !r-txt

```
*q0    2017
*doc
*a Brōugger /Waldemar Christopher
*a Rolfsen /Nordahl
*t Fridtjof Nansen 1861-1896
*tt Fridtjof Nansen 1861-1896
*lang German
*e Translated into German by
*n Enzberg /Eugen von
*edn Fifth ed.
*ee original drawings by
*n Krohg /Chr.
*n Sinding /Otto
*n Werenskiold /E.
*ee photographs taken in Greenland by
*n Drygalski /Erich von
*pl Berlin
*pu Fussingers Buchhandlung
*d 1900
*p 479p., ill.
*size 24 cm
*abs Biography, first published in 1896 as "Fridtjof Nansen 1861-1893". Translator's appendix to this edition adds short account of \vFram\n expedition, 1893-96, a history of German participation in Arctic exploration, and other material
(*3)
*k Arctic regions
*k (*301)
*k North Pole
*k (*38)
*k Greenland
*k 91(091)
*k Exploration, history
*k 92[Nansen, Fridtjof]
*k Biographies: Nansen, Fridtjof
*rpl X
*loc Shelf 92[Nansen, F.]
#
```



# Muscat: 1980s technology

- 204,000 records
- c. 100 person years work
- Not MARC
- Pre-unicode
- Little validation
- No authority control
- Muscat - 1980s database

# Muscat conversion project

This screenshot shows the 'Reports' section of the Muscat conversion project interface. The top navigation bar includes links for Home, Reports, Tests, Records, Fields, Search, Post-migration, Export, Import, and Admin. A message at the top states 'Data from Muscat dated: 25th September 2019'. Below this is a dropdown menu labeled 'Switch to:'. The main content area is titled 'Welcome' and contains a paragraph explaining the purpose of the system: 'This administrative system enables Library staff at SPRI to get an overview of problems with Muscat records so that they can be prepared for eventual export to Voyager.' A heading 'Reports available' is followed by a link 'Or filter to: post-migration only'. A key at the top right indicates four categories: OK (green), Problems (red), Post-migration (blue), and Informational (grey). The main table lists various record types and their status:

Description	Problems?	%
Records without a *q0	None	-
Records without a category (*doc/*art/*ser)	None	-
Records without a *d that are not *ser and either no status or status is GLACIOPAMS	None	-
Records without a *acc	141,573	69.29%
Records without a *	None	-
Records with multiple adjacent *t values	None	-
Records with an empty *t/*tt/*tc	None	-

# The conversion process

- 6-12 hours overnight
- Data import
- Mixture of writing code, and fixing records: highly iterative
- About 200 days of coding work over 7 years.



Presented	MARC record	Muscat as XML	Processed version	Raw data	Muscat editing view	#20001
Field	Value					
<u>q0</u>	20001					
<u>art</u>						
<u>a</u>	Anon.					
<u>t</u>	Bericht u^uber die Ta^utigkeit des Eidg. Institutes fu^ur Schnee- und Lawinenforschung 1944/45. No. 2					
<u>j</u>						
<u>t</u>	Schweizerische Zeitschrift fu^ur Forstwesen					
<u>d</u>	1945					
<u>pt</u>	1945(12):5 p.					
<u>k</u>	551.5					
<u>kw</u>	Meteorology					
<u>k</u>	551.578.46					
<u>kw</u>	Snow cover and snow patches					
<u>k</u>	(234.3)					
<u>kw</u>	Alps					
<u>k</u>	(494)					
<u>kw</u>	Switzerland					
<u>rpl</u>	E10					
<u>loc</u>	Basement Seligman 23					

# Conversion process

## STAGE 1. Convert 200MB text dump to a database

We then had a basic key-value pair database, of 9m individual data lines, about 40 per record.



Scott  
Polar  
Research  
Institute

Field	XPath	Value
q0	/q0	20001
art	/art	
ag	/art/ag	
a	/art/ag/a	
n1	/art/ag/a/n1	Anon.
tg	/art/tg	
t	/art/tg/t	Bericht über die Tätigkeit des Eidg. Institutes für Schnee- und Lawinenforschung 1944/45. No. 2
tc	/art/tg/tc	Bericht über die Tätigkeit des Eidg. Institutes für Schnee- und Lawinenforschung 1944/45. No. 2
j	/art/j	
tg	/art/j/tg	
t	/art/j/tg/t	Schweizerische Zeitschrift für Forstwesen
d	/art/j/d	1945
pt	/art/j/pt	1945(12) :5 p.
k	/art/j/k	
ks	/art/j/k/ks	551.5
kw	/art/j/k/kw	Meteorology
k	/art/j/k	
ks	/art/j/k/ks	551.578.46
kw	/art/j/k/kw	Snow cover and snow patches
k	/art/j/k	
ks	/art/j/k/ks	(234.3)
kw	/art/j/k/kw	Alps
k	/art/j/k	
ks	/art/j/k/ks	(494)
kw	/art/j/k/kw	Switzerland
rpl	/art/j/rpl	E10
loc	/art/j/loc	
location	/art/j/loc/location	Basement Seligman 23

# Conversion

## 2. Interpret the record structure.

Used Xpath to ‘drop’ a flat-structured record into a schema.

Also turned proprietary character codes into Unicode, e.g. a<sup>u</sup> into ä



#20001

Presented MARC record Muscat as XML Processed version Raw data Muscat editing view

```
<q0>20001</q0>
<art>
  <ag>
    <a>
      <n1>Anon.</n1>
    </a>
  </ag>
  <tg>
    <t>Bericht über die Tätigkeit des Eidg. Institutes für Schnee- und Lawinenforschung 1944/45. No. 2</t>
    <tc>Bericht über die Tätigkeit des Eidg. Institutes für Schnee- und Lawinenforschung 1944/45. No. 2</tc>
  </tg>
  <j>
    <tg>
      <t>Schweizerische Zeitschrift für Forstwesen</t>
    </tg>
    <d>1945</d>
    <pt>1945(12) :5 p.</pt>
    <k>
      <ks>551.5</ks>
      <kw>Meteorology</kw>
    </k>
    <k>
      <ks>551.578.46</ks>
      <kw>Snow cover and snow patches</kw>
    </k>
    <k>
      <ks>(234.3)</ks>
      <kw>Alps</kw>
    </k>
    <k>
      <ks>(494)</ks>
      <kw>Switzerland</kw>
    </k>
    <rpl>E10</rpl>
    <loc>
      <location>Basement Seligman 23</location>
    </loc>
  </j>
</art>
```

# Conversion

## 3. Converted the hierarchy into XML.

Now have a structure easier to manipulate.

Can now determine parent-child relationships.

Also things like which author is 4<sup>th</sup>.

# Conversion

## 4. Start with basic mappings of Muscat to MARC for each field

E.g. loop though each \*k (UDC classification number) pair, and map to a 650 or 651.

Soon we had some very basic MARC records.

```
<k>
  <ks>551.5</ks>
  <kw>Meteorology</kw>
</k>
<k>
  <ks>551.578.46</ks>
  <kw>Snow cover and snow patches</kw>
</k>
<k>
  <ks>(234.3)</ks>
  <kw>Alps</kw>
</k>
<k>
  <ks>(494)</ks>
  <kw>Switzerland</kw>
</k>
```



```
650 07 #a551.5 -- Meteorology.#2fdc
650 07 #a551.578.46 -- Snow cover and snow patches.#2fdc
650 07 #aE10 -- Glaciology: snow and avalanches.#2local
651 #7 #a(234.3) -- Alps.#2fdc
651 #7 #a(494) -- Switzerland.#2fdc
```



Scott  
Polar  
Research  
Institute

# Conversion

5. Extend this principle,  
by create macros for each  
MARC field conversion.

A 520 3# \$a{ //abs | macro:excludeNoneValue | macro:dotEnd(true)}

110 2# \$aL.S. Navigational Consulting Services Ltd.  
245 10 \$aReport on examination of navigation equipment for Arctic LNG carriers. Book 1 / \$cL.S.  
Consulting Services Ltd.  
260 ## \$a[S.l.] : \$bL.S. Navigational Consulting Services Ltd, \$c[n.d.]  
300 ## \$1 volume (unpaged) ; \$c28 cm.  
520 3# \$aShips on route through Davis Strait, Baffin Bay, Barry Channel to Melville Island.  
650 07 \$a656.6 -- Sea transport.\$2udc  
650 07 \$a656.61.052 -- Seamanship.\$2udc  
650 07 \$aL -- Communication and transportation.\$2local

Basically each {...} block is  
a placeholder to be filled  
in. E.g. this macro says:

Take //abs contents, and  
put into 520 3# \$a{...}  
and add a dot at end.



```
</pu>L.S. Navigational Consulting Services Ltd</pu>  
</pg>  
<d>[n.d.]</d>  
<p>unpaged, ill.</p>  
<size>28 cm.</size>  
<abs>Ships on route through Davis Strait, Baffin Bay, Barry Channel to Melville Island</abs>  
<k>  
    <ks>(*60)</ks>  
    <kw>Arctic Ocean and adjacent waters</kw>  
</k>  
<k>  
    <ks>(*62)</ks>  
    <kw>Canadian Arctic waters</kw>  
</k>
```

Parser definition: \*

```
LDR {/*|macro:generateLeader}
A 001 SPRI-{/q0}
A 005 {/*|macro:migrationDatetime}
A 007 {/*|macro:generate007}
A 008 {/*|macro:generate008}
AR 020 ## /*|isbn|macro:validisbn}
AR 022 ## †a{/*|issn}
A 035 ## †aSPRI-{/q0}
A 035 ## †a{/*|macro:voyagerBelievedMatch(absolute)}
A 040 ## †aukCU-P‡beng‡eaacr
A 041 i{'0'/*|macro:languages041(true)}# †a{./|macro:languages041}
A 100 {/*|macro:authorsField(100)}
A 110 {/*|macro:authorsField(110)}
A 111 {/*|macro:authorsField(111)}
E 240 10 †a{//tg/to|macro:stripLeadingArticle240}{//tg/to|macro:ifElse{//tg/lto,,,})†l{//tg/lto}
A 242 i{'0'/*|macro:indicator1xxPresent(l)}i{/*|tg/tt|macro:nfCount(English)} †a{/*|tg/tt|macro:colonSplit(b)}
A 245 {/|macro:generate245}
AR 246 3# †a{/*|tg/ta|macro:colonSplit(b)}
A 250 ## {/|macro:generate250}
E 260 ## {/|macro:generate260}
A 300 ## †a{./|macro:generate300}
A 310 ## †a{/*|freg|macro:ucfirst}
AR 490 0# {/ts|macro:generate490}
A 500 ## {/|macro:generate500analytics}
AR 500 ## †a{/*|notes/note|macro:generate5xxNote(500)}
AR 505 {/notes/note|macro:generate505Note}
A 520 3# †a{/*|abs|macro:excludeNoneValue|macro:dotEnd(true)}
A 530 ## †a{'Also issued online'} †u{'true'}|macro:ifNotXpathValue{//location[1],Digital Repository}|macro:ge
AR 533 ## †a{/*|notes/note|macro:generate533Note}
AR 538 ## †a{/*|notes/note|macro:generate5xxNote(538)}
A 541 0# {/*|macro:generate541}
A 546 ## †a{/*|macro:multipleValues(/root/*|tg/lang)|macro:commaAnd|macro:exceptExactly(English)|macro:prepe
A 546 ## †a{/*|macro:transliterationProcessNote}
AR 561 1# †a{/*|notes/note|macro:generate5xxNote(561)}
```

# Conversion

## 6. Full macro definition. Mapping each MARC field from Muscat fields.

Each MARC line ‘looks up’  
data from the Muscat  
XML and processes it  
through a macro.

<https://github.com/camunigeog/muscat-conversion/blob/master/tables/marcparser.txt>

# Conversion

## 7. Each macro has quite complex logic.

Receives the value from the Macro definition, then builds the result.

E.g. this shows the LDR.  
\*form=Map becomes  
LDR position 6 = e

```
2098
2099 # Macro for generating the Leader (LDR); see: https://www.loc.gov/marc/bibliographic/bdleader
2100 private function macro_generateLeader ($value)
2101 {
2102     # Start the string
2103     $string = '';
2104
2105     # Positions 00-04: "Computer-generated, five-character number equal to the length of
2106     $string .= '____';           // Will be fixed-up later in post-processing, as at 1
2107
2108     # Position 05: One-character alphabetic code that indicates the relationship of the record
2109     $string .= 'n';           // Indicates record is newly-input
2110
2111     # Position 06: One-character alphabetic code used to define the characteristics and content
2112     #!#M If merging, we would need to have a check that this matches
2113     switch ($this->form) {
2114         case 'Internet resource':
2115         case 'Microfiche':
2116         case 'Microfilm':
2117         case 'Online publication':
2118         case 'PDF':
2119             $value06 = 'a'; break;
2120         case 'Map':  
             ↓
2121             $value06 = 'e'; break;
2122         case 'DVD':
2123         case 'Videorecording':           // E.g. /records/9992/ (test #385)
2124             $value06 = 'g'; break;
2125         case 'CD':
2126         case 'Sound cassette':
2127         case 'Sound disc':
2128             $value06 = 'i'; break;
```

LDR 00938nam#a2200000#a#4500  
001 SPRI-1102  
005 20200904200401.0  
007 ta  
008 200904nuuuuuuuuuxx#a###|####|0||#0#eng#d  
035 ## [#aSPRI-1102](#)  
040 ## [#aUkCU-P](#)[#beng](#)[#eaacr](#)  
110 2# [#aL.S. Navigational Consulting Services Ltd.](#)  
245 10 [#aReport on examination of navigation equipment for Arctic LNG carriers. Bo](#)  
[ting Services Ltd.](#)  
260 ## [#a\[S.l.\]](#) :[#bL.S. Navigational Consulting Services Ltd,](#) [#c\[n.d.\]](#)  
300 ## [#a1 volume \(unpaged\) ;](#) [#c28 cm.](#)  
520 3# [#aShips on route through Davis Strait, Baffin Bay, Barry Channel to Melville](#)  
650 07 [#a656.6 -- Sea transport.](#) [#2udc](#)  
650 07 [#a656.61.052 -- Seamanship.](#) [#2udc](#)  
650 07 [#aL -- Communication and transportation.](#) [#2local](#)  
651 #7 [#a\(\\*60\) -- Arctic Ocean and adjacent waters.](#) [#2udc](#)  
651 #7 [#a\(\\*62\) -- Canadian Arctic waters.](#) [#2udc](#)  
852 7# [#2camdept](#) [#bSCO](#) [#cSPRI-SHF](#) [#h\(\\*60\) : 656.61.052](#) [#9Create 1 item record](#) [#0Migrate](#)  
917 ## [#aUnenhanced record from Muscat, imported 2019](#)  
948 3# [#a20200904](#)

# Conversion

## 8. Eventually end up with a full MARC record.

Generated live. (Change logic and refresh shows change.)

Iteration and testing with the OPAC.

Thanks to Rick Frohlich whose work was critical



Scott  
Polar  
Research  
Institute

[www.spri.cam.ac.uk/library/catalogue/records/1102/](http://www.spri.cam.ac.uk/library/catalogue/records/1102/)

# Reports and tests

#985	✓	500 Ensure whole string in /art/in genuine analytic ends with dot	2905	500	Dumont d'Urville.	500 ## †aIn: Voyage au Pôle Sud et dans l'Océani années 1837-1838-1839-1840 sous le commandement Jules Sébastien César Dumont d'Urville.
#986	✓	Leader position 06 has 't' for theses	61792	LDR	/^.{6}t/	LDR 01035ntm#a2200000#a#4500
#987	✓	300 †a process [unpaged] as unpaged	24258	300	‡a1 volume (unpaged)	300 ## †a1 volume (unpaged)
#988	✓	300 †a unpaged for Map appears as map	155387	300	‡a1 map ;‡c	300 ## †a1 map ;‡c23 cm.
#989	✓	300 †a unpaged for Microfiche/Microfilm appears as microform	60971	300	‡a1 microform.	300 ## †a1 microform.
#990	✓	In 852, where both *kg and location=Periodical are present, *kg takes priority in location matching over title-based matching	43303	773	‡wSPRI-23052	773 0# †7nnas †tANARE Report. Series B. ‡wSPRI-2
#991	✓	Library of Congress Cyrillic, created from generated Cyrillic (from BGN/PCGN), replaces original BGN/PCGN entered in Muscat	1043	245	Trudy Murmanskoī nauchno-promyslovoī ékspeditsii	245 00 †6880-01‡aTrudy Murmanskoī nauchno-promys
#992	✓	Prevent multiple italic sections being swallowed in transliterated protected string handling - Cyrillic version (see also #46)	17256	880	‡aBelukha или beluga	880 10 †6245-02/(N‡aBelukha или beluga /‡cK.K. Ч
#993	✓	880 italic protection - Cyrillic version (see also next test)	1664	880	‡alabed'-klikun (Cygnus Cygnus) на	880 10 †6245-02/(N‡alabed'-klikun (Cygnus Cygnus) на
#994	✓	880 italic protection - Cyrillic version (see also previous test)	1664	245	‡aLebed'-klikun (Cygnus Cygnus) на	245 10 †6880-02‡aLebed'-klikun (Cygnus Cygnus) на

- Extensive reports
- Data cleanup work
- 1,071 tests
- Test for every conversion feature

## Import

Imports can be done using this form.

Data from Muscat dated: 25th September 2019

1. Create the exports as follows:

Open a Muscat terminal and type the following. Note that this can take a while to create.

Be aware that you may have to wait until your colleagues are not using Muscat to do an export, as exporting may lock Muscat access.

n-voyager\_export

2. Upload the export files to this website, using this form. Note that this can take several minutes, so please be patient.

muscatview file:  No file chosen

rawdata file:  No file chosen

3. Run the import using the form below. This will reset the data in this system.

Select export files dated: \*

Import type: \*

## Conversion

### 9. Loop through all 204,307 records.

We ran this overnight, took between 6-12 hours!

# Conversion

## 10. Run through MarcEdit, and Bibcheck

Final stage of import.  
Converts MARC text  
to .mrc format, using  
MarcEdit (not UI).

Used Bibcheck, to  
check for problems.

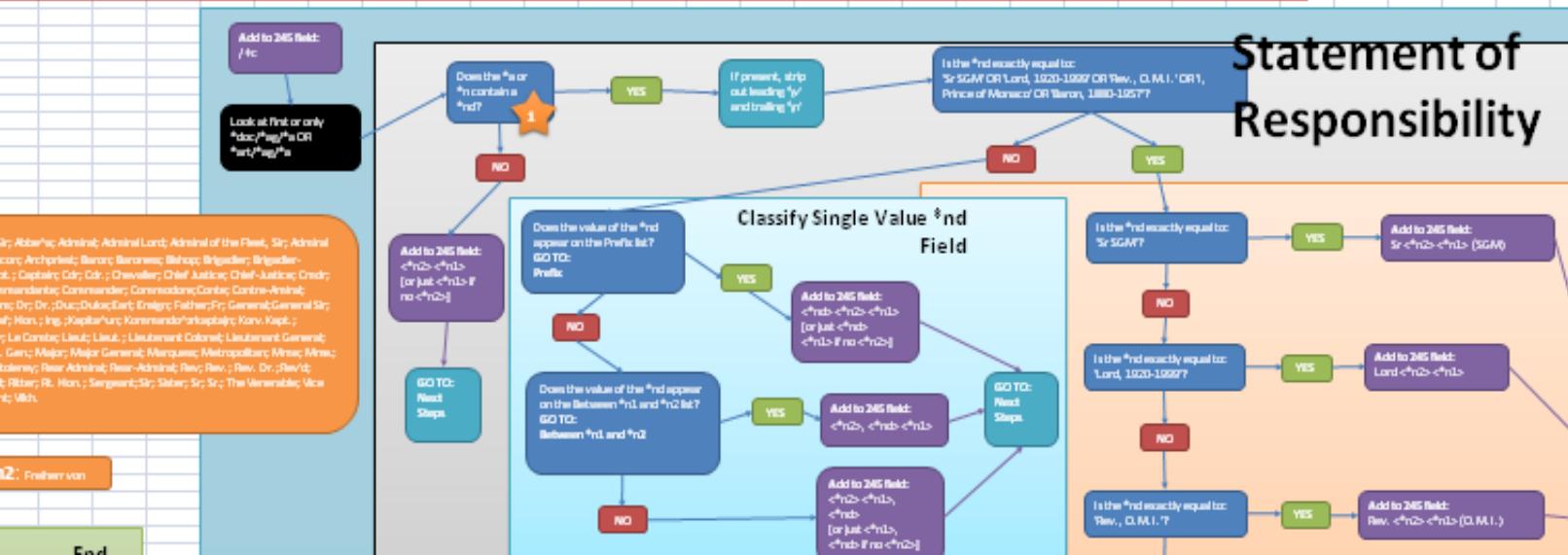
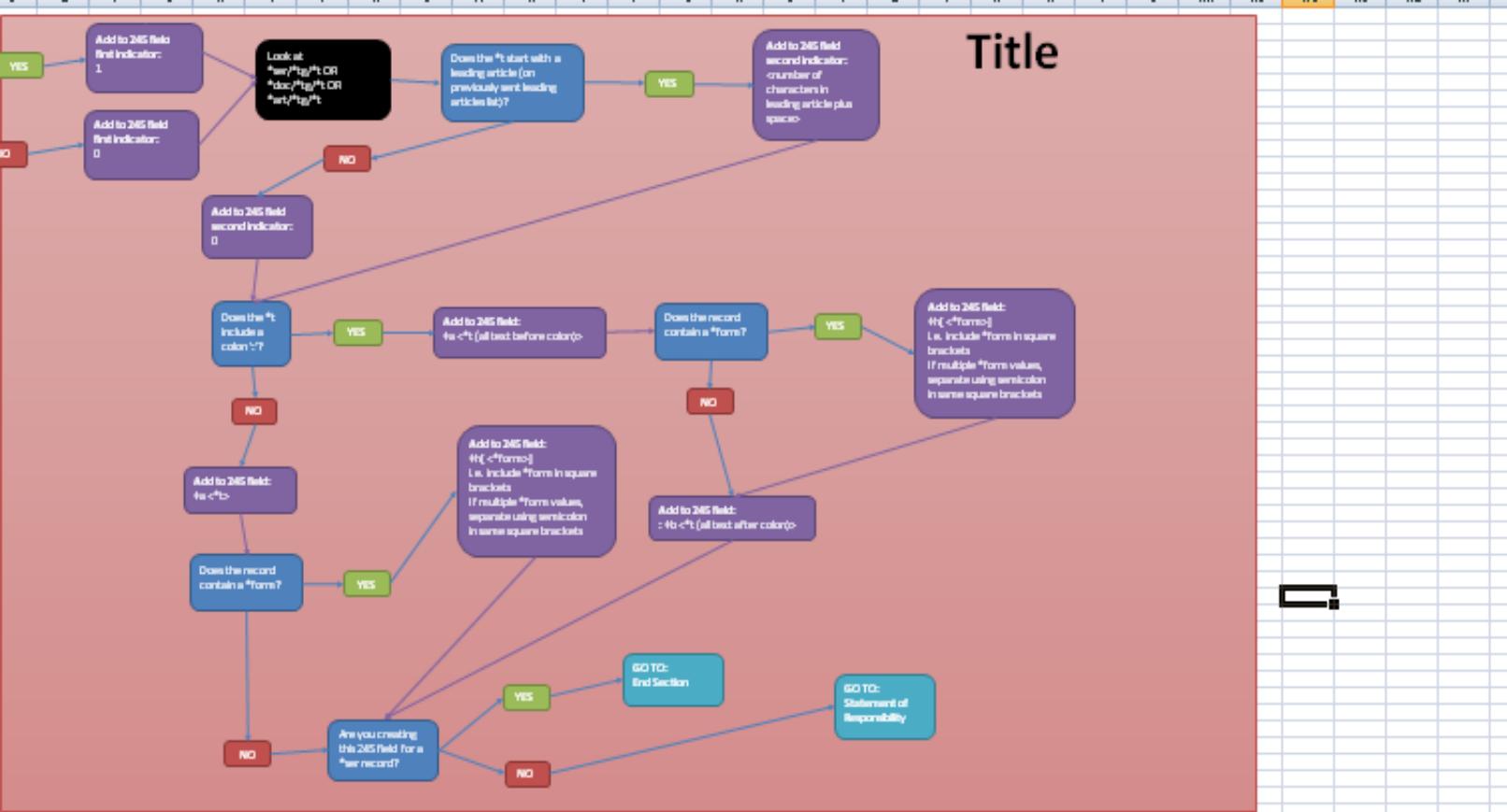
### Export MARC21 output

#### Downloads

Category:	Serials ( /ser ):	Monographs & articles ( /doc , /art/in , /art/j ):
Migrate to Alma, with item record(s): 56,797 records	<a href="#"> MARC21 data (text)</a> <a href="#"> MARC21 data (binary .mrc)</a> 3,179 serials records	<a href="#"> MARC21 data (text)</a> <a href="#"> MARC21 data (binary .mrc)</a> 53,618 monographs & articles records
Migrate to Alma: 112,539 records	<a href="#"> MARC21 data (text)</a> <a href="#"> MARC21 data (binary .mrc)</a> 0 serials records	<a href="#"> MARC21 data (text)</a> <a href="#"> MARC21 data (binary .mrc)</a> 112,539 monographs & articles records
Suppress from OPAC, with item record(s): 21,761 records	<a href="#"> MARC21 data (text)</a> <a href="#"> MARC21 data (binary .mrc)</a> 147 serials records	<a href="#"> MARC21 data (text)</a> <a href="#"> MARC21 data (binary .mrc)</a> 21,614 monographs & articles records
Suppress from OPAC: 252 records	<a href="#"> MARC21 data (text)</a> <a href="#"> MARC21 data (binary .mrc)</a> 0 serials records	<a href="#"> MARC21 data (text)</a> <a href="#"> MARC21 data (binary .mrc)</a> 252 monographs & articles records
Ignore record: 12,958 records	<a href="#"> MARC21 data (text)</a> <a href="#"> MARC21 data (binary .mrc)</a> 6 serials records	<a href="#"> MARC21 data (text)</a> <a href="#"> MARC21 data (binary .mrc)</a> 12,952 monographs & articles records

# Authors: 1xx/7xx/245

- Flow chart for each field
- Emulated SoR
- 2<sup>nd</sup>+ 1xx mutates to 7xx
- Contributor terms
- (880 handling also!)



```
651 #/ #a(*/ ) -- Antarctic regions.#2udc
651 #7 #a(*78) -- Sub-Antarctic islands.#2udc
651 #7 #a(*785) -- Heard and McDonald Islands.#2udc
773 0# #7nnas #tANARE Report. Series B. #wSPRI-23052
852 7# #2camdept#bSCO#cSPRI-SHF#h(*785) : 595.7#9Create 1 item record#0Migrate
917 ## #aUnenhanced record from Muscat, imported 2019
949 24 #--201901217
```

Year All ▾ Volume All ▾ Description All ▾

[Scott Polar Research Institute Library](#) > [Serials](#) > [148 B-D](#)

Related holding

Holdings: 1953-64

[Scott Polar Research Institute Library](#) > [Shelf](#) > [\(\\*785\):595.7](#)

Details

Title	The insects of Heard Island / K.G. Brown.
Author	K. G. Brown >
Publisher	Melbourne : Antarctic Division, Dept. of External Affairs, 1964.
Creation Date	1964

## Analytics – parent/child relationships

- Extensive 773/852 logic
- Pamphlet handling
- Processing order
- Language selection
- Name alignment with UL

Russian	M.N.	<u>М.Н.</u>	M.N.
Russian	Bogatov	<u>Богатов</u> 701	Bogatov
Russian	Otchet ob ekspeditsiy	Отчет об экспедиций	Otchet ob ékspeditsii
=	Imperatorskoy	Императорской	Imperatorskoj
French	Akademiy Nauk na Novuyu Zemlyu letom 1896 goda = Compte rendu de l'expédition, envoyée par l'Académie Impériale des Sciences à Novaïa Zemlia en été 1896	Академий Наук на Новую Землю летом 1896 года = Compte rendu de l'expédition, envoyée par l'Académie Impériale des Sciences à Novaïa Zemlia en été 1896	Akademii Nauk na Novuiu Zemliu letom 1896 goda = Compte rendu de l'expédition, envoyée par l'Académie Impériale des Sciences à Novaïa Zemlia en été 1896
Russian	V.V.	<u>В.В.</u>	V.V.
Russian	Amnikolapodobnyye Gastropoda iz presnykh vod Dal'nego Vostoka SSSR	Амниколаподобные Gastropoda из пресных вод дальнего Востока СССР	Amnikolapodobnye Gastropoda iz presnykh vod Dal'nego Vostoka SSSR
Russian	Zoologicheskiy Zhurnal	Зоологический Журнал	Zoologicheskii Zhurnal
Russian	Pushkin	<u>Пушкин</u> 15449	Pushkin
Russian	A.F.	<u>А.Ф.</u>	A.F.

# Russian / Transliteration

- Highly complex area
- 140,000 new Cyrillic strings
- Muscat: Roman charset
- BGN/PCGN 1947
- Non-reversible scheme!
- 30+ stages to get to 880s
- Language identification
- Italics: names, etc.
- Parallel titles, etc.

# Reports

Reports define conditions which a record must or must not match.

Enables problems with the records to be tackled over time.

Plenty of work to fix these for years to come!

Report for p not followed by a dot in *p / *pt	None	-
*p values containing colon-space rather than space-colon-space	None	-
*ser records with multiple *r	None	-
/art/j records with no *kg in the Pamphlets	18,159	8.89%
*t with explicit slash also having *e	43	0.02%
Records containing a field with an empty dash, with a SPRI location	None	-
Records containing a field with an empty dash, without a SPRI location	None	-
Records with an invalid *con syntax	None	-
Records with names for transliteration in other languages (e.g. Yakut, Chinese, etc.) for upgrading	16	0.01%
Records with location = ??	495	0.24%
Records with *to without *lto defined, having whitelisted English *to in a non-Russian record	None	-
Records with *lto present but no *to	None	-
Records with more than one *to/*lto	None	-
Records with *nt=None which have not had / upgrading for 245	10	0.01%
Non-Russian records with Cyrillic *nt = BGNRus/LOCRus	33	0.02%
Non-Russian records with all Cyrillic *nt for types never supported	21	0.01%
Records with *nt=LOCRus that need diacritics adding	23	0.01%
Records with a suspected wrong language due to leading article mismatch in 245	None	-
Records whose title is in round brackets	None	-
*to/*ts records ending with a dot	None	-



# Project close

- Data was finalised end September 2019, and set to University Library.
- Plans for journal article about the conversion project.
- Code is open source on Github (and will have a DOI soon):  
<https://github.com/camunigeog/muscat-conversion>
- Documentation at:  
<https://github.com/camunigeog/muscat-conversion/tree/master/doc>
- Ongoing long-term work to fix up legacy record data



Scott  
Polar  
Research  
Institute

# Sample record: iDiscover

	Title	Russian Arctic seas = Moria Rossijskoj Arktiki / Natalya Marchenko.
TOP	Author	Natalya. Marchenko >
	Publisher	Berlin : Springer, 2012.
SEND TO	Identifier	ISBN : 9783642221255
GET IT	Creation Date	2012
DETAILS	Format	:274 : ill., maps ; 27 cm.
LINKS	Description	Gathers accounts of seagoing accidents in Russian Arctic Seas including Chukchi, Kara, Laptev and East Siberian Seas, mainly by ice pilots. Describes physical environment, climate, bathymetry, routes, navigational issues and vessel development. Comprises nearly 100 detailed accident reports of incidents caused by heavy ice conditions from 1900 to present. Classifies accidents by type and cause. Reports first full navigational description of Russian Arctic Seas available to English-speaking readers.
	Subjects	<a href="#">629.12 -- Vessels &gt;</a> <a href="#">629.124.791 -- Vessels for use in ice &gt;</a> <a href="#">629.124.791.2 -- Icebreakers &gt;</a> <a href="#">656.6 -- Sea transport &gt;</a> <a href="#">656.61.052 -- Seamanship &gt;</a> <a href="#">551.326 -- Floating ice &gt;</a> <a href="#">330.34 -- Economic development &gt;</a> <a href="#">502 -- Environmental issues &gt;</a> <a href="#">93 -- History &gt;</a> <a href="#">93"19" -- Twentieth century &gt;</a> <a href="#">93"20" -- Twenty-first century &gt;</a> <a href="#">91 -- Geography, exploration and travel &gt;</a> <a href="#">L -- Communication and transportation &gt;</a> <a href="#">(*501) -- Russia (Federation) &gt;</a>

# Sample Alma Record

- Leader 02765nam a2200517 a 4500
- 001 991006442493603606
- 005 20190930020543.0
- 007 ta
- 008 190930s2012 gw ab | |0||0 eng d
- 020 ##\$a9783642221255
- 035 ##\$aSPRI-204602
- 040 ##\$aUkCU-P\$beng\$eaacr
- 041 0#\$aeng\$arus
- 100 1#\$aMarchenko, Natalya.
- 245 10\$6880-01\$aRussian Arctic seas = Моря Rossii skoï Arktiki /\$cNatalya Marchenko.
- 260 ##\$aBerlin :\$bSpringer,\$c2012.
- 300 ##\$a:274 :\$bill., maps ;\$c27 cm.
- 520 3#\$aGathers accounts of seagoing accidents in Russian Arctic Seas including Chukchi, Kara, Laptev and East Siberian Seas, mainly by ice pilots. Describes physical environment, climate, bathymetry, routes, navigational issues and vessel development. Comprises nearly 100 detailed accident reports of incidents caused by heavy ice conditions from 1900 to present. Classifies accidents by type and cause. Reports first full navigational description of Russian Arctic Seas available to English-speaking readers.
- 541 0#\$cBA Russia;\$aBooks2anywhere US, Amazon\$d20130906\$eR0031\$h54.70 GBP.\$5UkCU-P
- 546 ##\$aln English and Russian.
- 546 ##\$aRussian transliteration entered into original records using BGN/PCGN 1947 romanisation of Russian; Cyrillic text in MARC 880 field(s) reverse transliterated from this by automated process; BGN/PCGN 1947 text then upgraded to Library of Congress romanisation.

# Sample Alma Record (continued)

- 650 07\$a629.12 -- Vessels.\$2udc
- 650 07\$a629.124.791 -- Vessels for use in ice.\$2udc
- 650 07\$a629.124.791.2 -- Icebreakers.\$2udc
- 650 07\$a656 continued.6 -- Sea transport.\$2udc
- 650 07\$a656.61.052 -- Seamanship.\$2udc
- 650 07\$a551.326 -- Floating ice.\$2udc
- 650 07\$a330.34 -- Economic development.\$2udc
- 650 07\$a502 -- Environmental issues.\$2udc
- 650 07\$a93 -- History.\$2udc
- 650 07\$a93"19" -- Twentieth century.\$2udc
- 650 07\$a93"20" -- Twenty-first century.\$2udc
- 650 07\$a91 -- Geography, exploration and travel.\$2udc
- 650 07\$aL -- Communication and transportation.\$2local
- 651 #7\$a(\*501) -- Russia (Federation).\$2udc
- 651 #7\$a(\*60) -- Arctic Ocean and adjacent waters.\$2udc
- 651 #7\$a(\*68) -- Arctic Ocean, seas adjacent to Russia.\$2udc
- 651 #7\$a(\*682) -- Chukchi Sea.\$2udc
- 651 #7\$a(\*683) -- East Siberian Sea.\$2udc
- 651 #7\$a(\*684) -- Laptev Sea.\$2udc
- 651 #7\$a(\*685) -- Kara Sea.\$2udc
- 880 10\$6245-01/(N\$aRussian Arctic seas = Моря Российской Арктики /\$cНаталья Марченко.
- 916 ##\$a148457 -- 2014/06/30 -- NOM
- 917 ##\$aUnenhanced record from Muscat, imported 2019
- 948 3#\$a20190930 \$bIW; NOM