mailto:jc.dauphin@gmail.com

INNZ metadata - *Source: National Library of New Zealand Te Puna Matauranga o Aotearoa,*
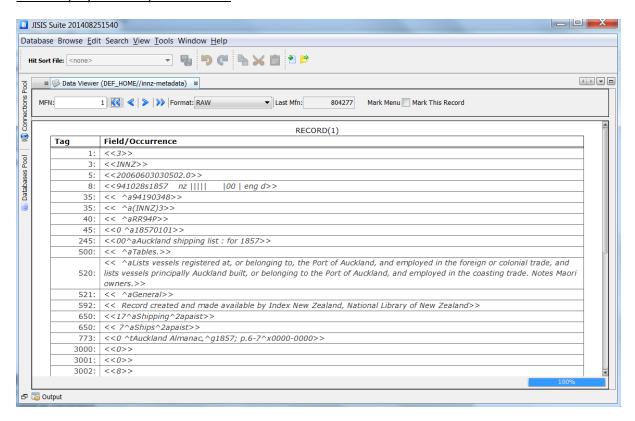
licensed by the Department of Internal Affairs for re-use under the Creative Commons Attribution 3.0 New Zealand Licence.

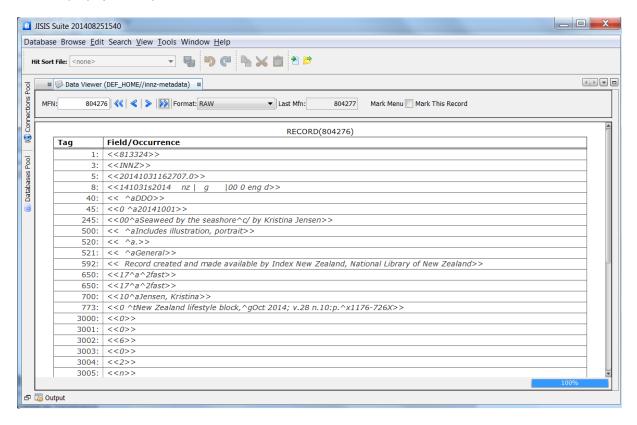# Testing J-ISIS Import of MARC Data with the INNZ metadata MARC file

The first import attempt failed due to an empty field occurrence near record 400 000. The J-ISIS Import source code was changed to skip empty field occurrences making the Import process more robust. The second import attempt was successful and 804 276 records were imported (I am not sure this is the right number of records).

1142213 milliseconds (~19 Minutes) to import ISO file on a Windows 7 64-bits computer with 4 GB RAM and Intel Core 2 CPU 2.50 GHz
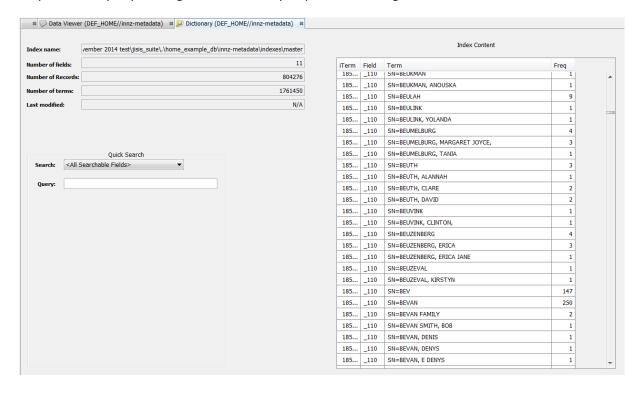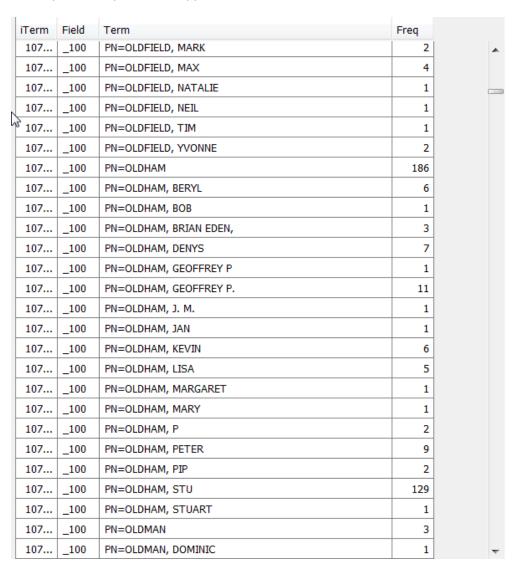
J-ISIS Display of 1st imported record

J-ISIS Display of last imported record

Indexing was quite slow but succeeded and 1 761 450 terms were indexed. I forgot to remove the stopwords. Anyway looking at the dictionary is quite interesting:
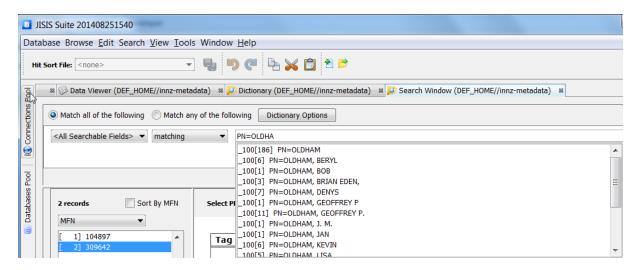
Terms prefixed by PN= are supposed to be Personal Names

iTerm	Field	Term	Freq
107...	_100	PN=OLDFIELD, MARK	2
107...	_100	PN=OLDFIELD, MAX	4
107...	_100	PN=OLDFIELD, NATALIE	1
107...	_100	PN=OLDFIELD, NEIL	1
107...	_100	PN=OLDFIELD, TIM	1
107...	_100	PN=OLDFIELD, YVONNE	2
107...	_100	PN=OLDHAM	186
107...	_100	PN=OLDHAM, BERYL	6
107...	_100	PN=OLDHAM, BOB	1
107...	_100	PN=OLDHAM, BRIAN EDEN,	3
107...	_100	PN=OLDHAM, DENYS	7
107...	_100	PN=OLDHAM, GEOFFREY P	1
107...	_100	PN=OLDHAM, GEOFFREY P.	11
107...	_100	PN=OLDHAM, J. M.	1
107...	_100	PN=OLDHAM, JAN	1
107...	_100	PN=OLDHAM, KEVIN	6
107...	_100	PN=OLDHAM, LISA	5
107...	_100	PN=OLDHAM, MARGARET	1
107...	_100	PN=OLDHAM, MARY	1
107...	_100	PN=OLDHAM, P	2
107...	_100	PN=OLDHAM, PETER	9
107...	_100	PN=OLDHAM, PIP	2
107...	_100	PN=OLDHAM, STU	129
107...	_100	PN=OLDHAM, STUART	1
107...	_100	PN=OLDMAN	3
107...	_100	PN=OLDMAN, DOMINIC	1

Here we can see that there are some data entry inconsistencies such as "OLDHAM, GEOFFREY P" and "OLDHAM, GEOFFREY P." which are considered as different terms.

Search Suggestions works but takes about 1 second to be displayed

Searching for OLDHAM