

Topic Maps for Cultural Heritage Collections

Conal Tuohy

Senior Developer

New Zealand Electronic Text Centre

www.nzetc.org



Site search Search

NZETC NEW ZEALAND ELECTRONIC TEXT CENTRE
HOME ADVANCED SEARCH ABOUT

Browse Collection By...

- Authors
- Works
- Subjects...
 - Autobiography, Biography, Journals, Correspondence
 - Contemporary Māori and Pacific Islands
 - Historical Māori and Pacific Islands
 - Language
 - Literary Criticism and History
 - Literature
 - New Zealand History
 - Science and Natural History
- Projects

Welcome to the New Zealand Electronic Text Centre's free online archive of New Zealand and Pacific Islands texts and heritage materials, which offers an ever expanding, fully searchable, set of images and full-text books, manuscripts and journals.

We are always keen to hear suggestions of other rare or significant texts which we should make available through the collection. Please contact the [NZETC director](#).

The NZETC also works with partners in the cultural heritage and ePublishing sector on a wide variety of digitisation and digital content delivery projects. More information on the Centre's mission and activities can be found in our [mission statement](#).

Moko; or Māori Tattooing



The New Zealand Electronic Text Centre is pleased to announce a new collection of historical texts on Ta Moko as the result of a project which tries to take into account sensitivities around the digitisation of mātauranga Māori and textual taonga.

The project was centred on the digitisation of an annotated copy of the 1896 publication "Moko; or Maori Tattooing" by Horatio Gordon Robley. An additional six texts have also been digitised to provide contextual information about Ta Moko, mokamokai, Robley himself and his art.

"Tattooing on the face of **TE PEHI KUPE**, drawn by himself." from "Moko; or Maori Tattooing"

Home | Advanced Search | About
© 2007 Victoria University of Wellington | Contact us | Conditions of use
Victoria
UNIVERSITY OF WELLINGTON
A Member of the 21st Century

Website visitor statistics (daily)

- around 9k visitors
- around 70k hits
- around 30k web pages
- > 1GB traffic

Website content statistics

- 75k web pages
 - 50% represent digitised documents
 - books, magazines, letters
 - articles, chapters, sections
 - illustrations
 - The other 50% are about *things*
 - people, organisations
 - places, ships,
 - literary works
 - even a few animals!
- 3.5M hyperlinks

Resource-centric vs subject-centric systems

- “Resource-centric” systems focus narrowly on digital resources
 - a catalogue of digital items
 - everything else is peripheral or secondary
- “Subject-centric” systems can accommodate anything of interest:
 - information resources
 - abstract concepts,
 - or physical things

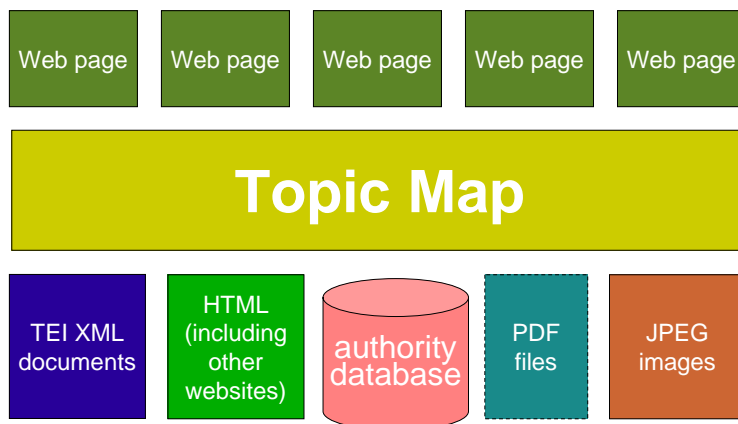
Information Architecture goals

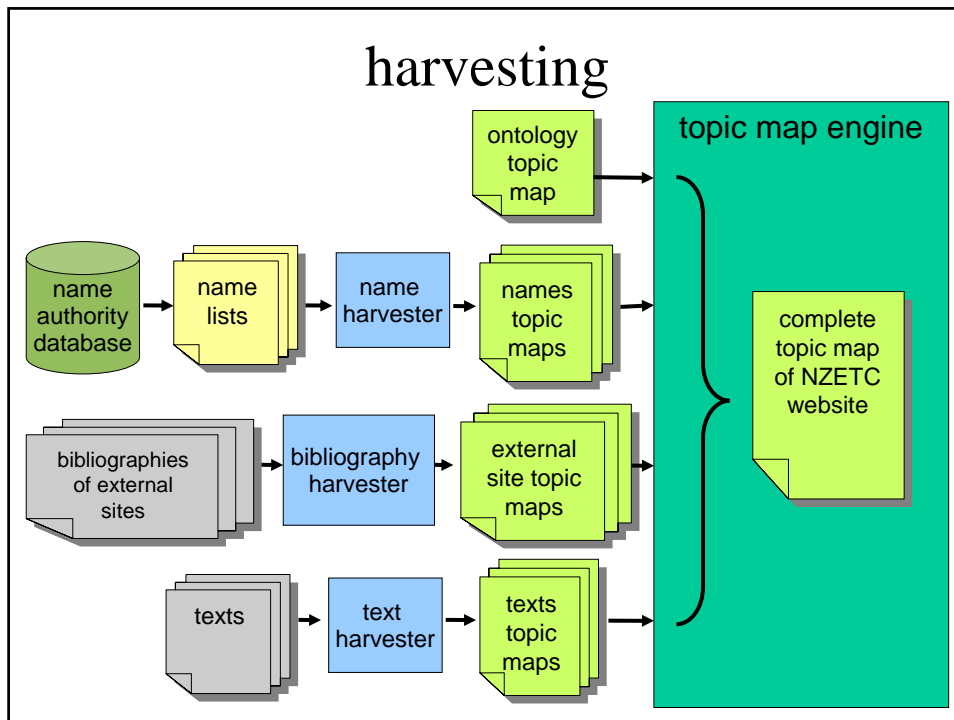
Information resources must be transformed, chunked, and linked together into a navigable web

- Need to present information *in context* on every page
- Need an explicit model of the entire website logical structure. Not just a sitemap, but an *ontological model*
- Need to build the model automatically



Topic Map layer above the digital resources

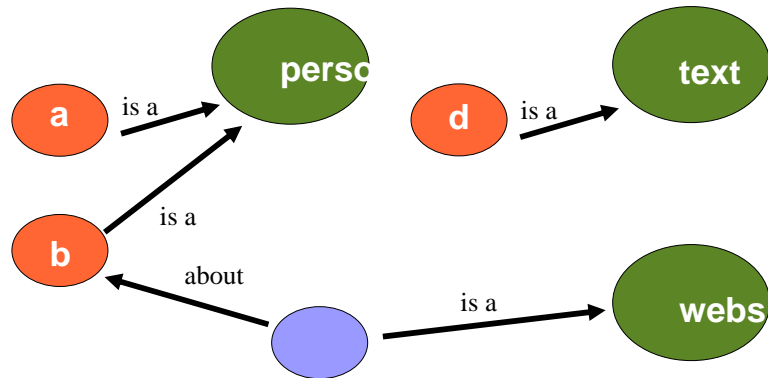




Entity Authority

- We built an authority file of entities of interest.
- We've developed a specialised database for this purpose, which we call “Entity Authority Tool Set” (EATS) to manage names and identifiers (a PSI server).
- In our digitised documents we tag every mention of these entities with their identifier. Our taggers search in EATS for a name, and select from the possible matches.

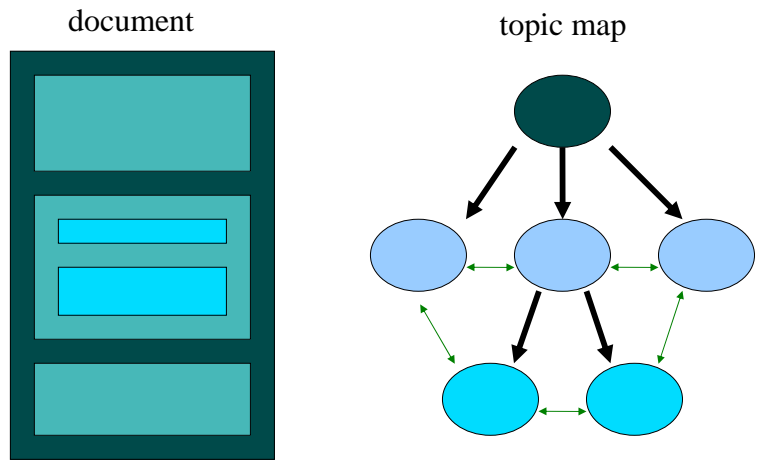
“authority” topic maps



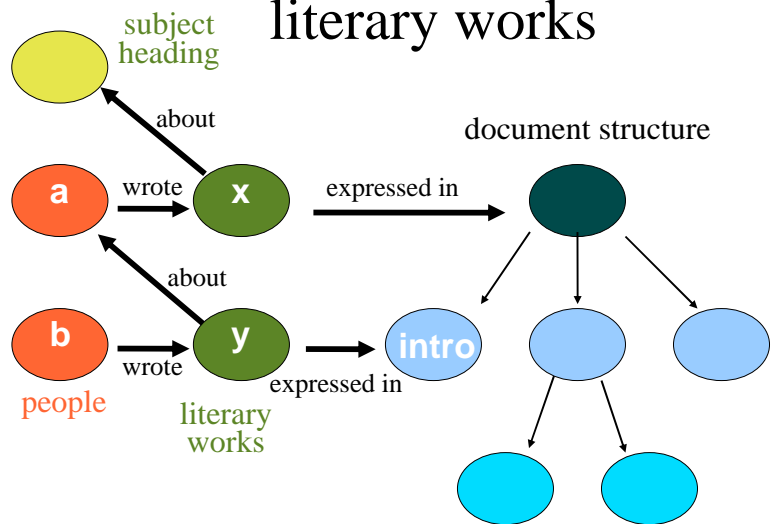
Text Encoding for Interchange (TEI)

- Bibliography
- Subject classification
- Textual structure
- Cross-references
- External references
- Commentary
- *etc.*

a document's internal structure



literary works



Multiple editions of a single work

Search | Site search

NZETC NEW ZEALAND ELECTRONIC TEXT CENTRE
HOME ADVANCED SEARCH ABOUT

Browse Collection By...

Authors

Works

Subjects...

Autobiography; Biography; Journals; Correspondence

Contemporary Māori and Pacific Islands

Historical Māori and Pacific Islands

Language

Literary Criticism and History

Literature

New Zealand History

Science and Natural History

THE TREATY OF WAITANGI (TEXT)

Digitised Editions of this Text in Our Collection
Mentioned In

Digitised Editions of this Text in Our Collection

- [English](#)
- The Treaty of Waitangi
- The Treaty of Waitangi. — 6th February, 1840.
- Maori Version

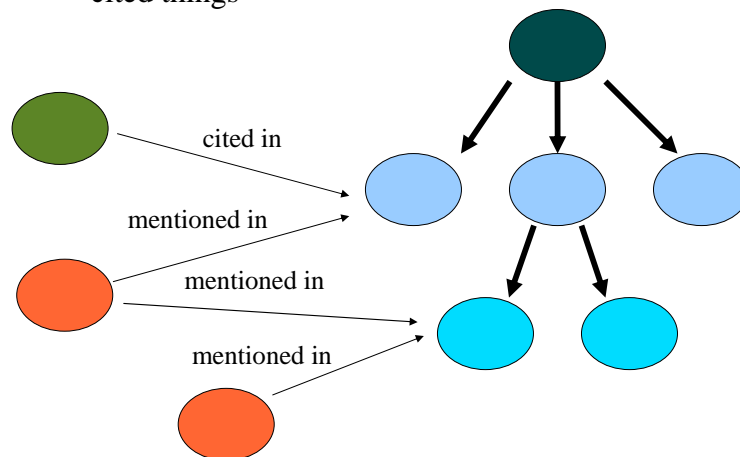
Mentioned in

- Book & Print in New Zealand : A Guide to Print Culture in Aotearoa
 - Printed works in Māori to the 1850s
 - Index
- Extracts from a Diary during Heke's War in the North in 1845
 - Introduction

mentions, depictions, citations

mentioned, depicted and cited things

document structure



NZETC NEW ZEALAND ELECTRONIC TEXT CENTRE
HOME ADVANCED SEARCH ABOUT

KIMBLE BENT
(PERSON)


Soldier, deserter, Pakeha-Maori, character.

[Subject of Texts](#)
[Image Gallery](#)
[Mentioned In](#)
[External Links](#)

Subject of Texts

- Monday's Warriors
- The Adventures of Kimble Bent

Image Gallery



Mentioned in

- Hero Stories of New Zealand
 - The Defence of Turuturu-Mokai Redoubt — The "Rorke's Drift" of New Zealand
- Kōtare 1998, Volume One, Number One

Topic map statistics

- 126k topics
- 126k occurrences
- 242k associations
- 1M roles
- 115k base names
- 69k variant names (sort names)

Benefits

- Easier to provide links and contextual information
- Easy to pull together information from a variety of sources
- Implicit topics of interest are made explicit
- Improved our own understanding of our collection
- Easier to find information on the site
- Google searches work better

questions?

- only easy questions please
- contact me: conal@nzetc.org