The IFLA/MIC localisation project

Coming together: building community through Moving Image Collections (MIC)
2006.10.12 16:00-17:30

James M Turner

Faculté des arts et des sciences
École de bibliothéconomie et des sciences de l'information
Outline

- Background
- Localisation of web sites
- Work to date on the IFLA/MIC project
- Next steps
Background

- Context
- Chronology
- Project team
- Project goal
- Project objectives
MIC: “a window to the world’s moving images”

But users can access MIC only in English

AMIA’s efforts to reach out into the international community

Some preliminary work by students, some French-language content now
Chronology

2003 IFLA Berlin: the Section on Audiovisual and Multimedia takes on a section project: advising AMIA on how to internationalise the MIC site

2004 IFLA Buenos Aires: report on a needs survey and proposal to develop a collaborative project

2005 IFLA Oslo: application to IFLA for funding

2006: Funding obtained (9500€)

2006: Planning meetings in Montréal
Project team

James Turner, École de bibliothéconomie et des sciences de l’information, Université de Montréal (research coordinator project leader)

Samira Sambaíno, Instituto Universitario Asociación Cristiana de Jóvenes, Montevideo, Uruguay (research coordinator)

Marwa El Sahn, Biblioteca Alexandrina, Alexandria, Egypt (research coordinator)
Help make MIC available to a broader user base by opening it to people who speak languages other than English
Project objectives

- Advise AMIA and MIC on needs of the international community
- Build a kit to create localised versions of the MIC site
- Test the kit with French, Spanish, Arabic
Localisation of web sites

- What it means
- Some examples
- Work on localisation
What it means

- Also called internationalisation, globalisation, i18n, g11n, l10n

“the process of adapting technical media products... into a form where they are... acceptable in countries outside the original target market”

Translating is not enough because users don’t connect to cultural aspects
A related concept is transcreation, a kind of interpretation of the text. Word-for-word translation often does not get the message across very well. It’s better to re-write the text in the target language. The content remains, but the text sounds/reads like it was created in the target language.
Some examples

- Date formats
- Colours
- Giving examples
- Direction of writing
- Order of names
- Representation of numbers
- Money
Date formats

- Several calendars used around the planet
- Even countries using the same calendar have different ways of expressing dates
June 28, 2004

United States   6/28/2004
Europe          28/6/2004
International System (SI) 2004/06/28
Two-digit year = Evil

United States  06/07/04

Europe  07/06/04

International System (SI)  04/06/07
Colours

- Red = celebration or bloodshed
- White = death or purity
- Green = hope or solidarity

A need to consider the symbolic value of colours
Giving examples

- Dublin Core metadata example of a film:
  - *Citizen Kane* a good bet in the U.S.
  - *La Strada* better in Italy
  - *Der blaue Engel* in Germany
Direction of writing

- Roman alphabet left to right
- Hebrew, Arabic right to left
- Chinese right to left, top to bottom, now left to right
- (Luckily, the MIC site already uses XHTML and the Unicode character set, which includes a declaration of writing direction)
Order of names

- Firstname Lastname (North America)
- Lastname Firstname (Much of Europe, Asia, Africa)
- Many systems for children’s last names (consider Spain, Iceland, Québec)
Representation of numbers

- Various systems of punctuation for grouping:
  - 3,224,758 / 3.224.758 / 3 224 758

- Confusion about large numbers:
  - in the U.S. a billion means a thousand million
  - in the U.K. a billion means a million million
The $ sign is used to represent several different dollars (e.g. U.S., Canada, Australia)

The $ sign also used to represent several different national pesos and other currencies

How to write amounts of money changes, even within the same country, e.g. Canada:

$35,244.87 (English-language contexts)

35 244,87$ (French-language contexts)
Work on localisation

- A new area of study and research
- The work to date in the business community
- Study of localisation of cultural web sites barely begun
- At UdeM, a few projects and a web site
Work to date on the IFLA/MIC project

- Planning meetings
- Test pages identified
- Localisation kit
Planning meetings

- 5 days in Montréal, July 28 - August 1
- Localisation training
- Review of MIC site map, selection of test pages
- Planning the localisation kit
- List of tasks, sequencing, timeline
- Job description for student assistants
Test pages identified

- Visit of MIC site map to understand the structure
- Choice of Archivists portal as most relevant to IFLA’s goals, institutional needs
- Establishing priorities
Priorities

- Priority 1: Cataloguing and Metadata portal
- Priority 2: Preservation portal
- Also Priority 1: general MIC site information and searching information
- Within these, each page assigned Priority 1, 2, 3

Work to date
Test pages identified
URLs

- URLs for all pages to be used in the project gathered to include in kit
- Student assistants will access each page, get the source code, put French, Spanish, or Arabic text directly between the existing tags
- Also, identify localisation issues for each page

Work to date
Test pages identified
Localisation kit

Almost ready, out for review and final comments

Table of contents:

- Introduction to the kit
- Localisation
- Detailed steps in building a test version in another language
- Other information
Next steps

- Hire students to test the kit
- Quality control by research coordinators
- Preparation of XHTML files for MIC team
- Transfer of files to MIC team
- Install test site (hopefully)
- After testing, update and install kit so further language versions can be built (hopefully)
The End

james.turner@umontreal.ca

http://mapageweb.umontreal.ca/turner