Terminology for new media art

James M Turner, professeur

Outline

Introduction
- The general problem of terminology
- Past and present
- DOCAM's terminology committee

The general problem
- The nature of languages
- Waltz, Markey, Panofsky
- Representation and management

The nature of languages
- Tool for human communication
- Basic instrument for thinking
- National, regional, artificial, slang, secret
- Organic : evolves over time
- The need to manage all this
- Things to think about :

Waltz, Markey, Panofsky
- David Waltz : language is like a game of charades
- We use a simplified version to represent the world, something like shorthand
- But each world is like an indexing term, which refers to a much broader reality
Markey

The Gombrich fallacy:
- The message received is the same as the message transmitted (Markey 1981, 20)
- For each person, any word can have particular meaning, for example:
  - mother, dog, coriander, métro

Panofsky

Another level of translation, of representation, the expression of a deeper reality:
- when a man lifts his hat upon greeting a lady, a social interaction is taking place

And so on

Terminology, indexing, representation, communication, it’s all semiotics

The essential problem in information science: how to represent things so they can best be managed

Representation and management

With the information explosion (early 20th century) came the need to manage all this information

This is done by representation: extract the essential reality, which then points to a much broader universe

Tools that developed: cataloguing, indexing, classification

Cataloguing, indexing, classification

The catalogue record represents the object at hand:

Past and present

Cataloguing, indexing, classification

Thesauruses

Cuts, deterioration

The new world of information
Once upon a time...

The library in miniature
Authors, titles, subjects

Introduction
Past and present
Cataloguing, indexing, classification

Indexing
Quickly now: what's this finger called?

Book indexes
Many more indexes in English-language books than in French-language books -- the tradition is much stronger
A piece of intellectual work, a kind of synthesis of the contents
In the 1960s: KWIC, KWOC, KWAC
In the 1980s: emphasis on generating automatically, a kind of concordance

KWOC

Picture indexing
Pre-iconographic: woman, wheel, sword, cushion, palm leaf
Iconographic: Saint Catherine of Alexandria, Christian martyr, 4th C., Caravaggio
Iconological: the wheel: an instrument of torture; the palm: sacrifice, victory; the sword: the fight for truth; the cushion: nobility
Classification

- Emerged from the information explosion, the industrial revolution, compulsory education
- Development of book classifications (Cutter, Dewey, CDU, Ranganathan, Bliss)
- An attempt to find a place for each concept in the world of knowledge

Stabilising, education

- With time, methods developed for cataloguing, indexing, classification become stable
- Libraries, archives, and museums are the storehouses of knowledge, and they play an important role in education

Thesauruses (thesauri)

- “Controlled vocabulary” is a way to manage synonyms and multiple meanings
- Another vocabulary management tool: the thesaurus
- From a long tradition (Roget 1852), development of a tool for managing specialised vocabulary
- In information science, usually specific to an area of knowledge

Art, images

- Art & Architecture Thesaurus (Getty Research Institute, <http://www.getty.edu/research/conducting_research/vocabularies/aat/>)
Cuts, deterioration

- In the 1990s, lots of budget cuts
- Many collections bought and sold
- This sometimes involves important changes (e.g. Corbis buying the Bettman Archive)
- Often, “simplifying” tools built with much care
- Deterioration of practices, cutting positions, cheaper but untrained personnel

Introduction
Past and present

The new world of information

- Meanwhile, the World Wide Web arrives on the scene (November 1994)
- Suddenly, connectivity (online access to many collections)
- Anarchy reigns: each collection has its own cataloguing policy, indexing vocabulary

Introduction
Past and present

What now?!

- A great mass of documentation on the web
- Norms, layers of software to make it work, e.g. search engines, Z39.50
- Some attempts to classify web pages
- “Indexing” done by search engines
- The semantic web, Web 2.0

Introduction
Past et présent
The new world

The semantic web

- An attempt to impose order
- Elaborate metadata, layers of metadata, metadata wrappers
- Dublin Core (DC), Resource Description Framework (RDF)
- Long-term, large-scale development
- In the long term, we can hope for information structured as it is created

Introduction
Past et présent
The new world

Web 2.0

- Web services like YouTube, MySpace, flickr, del.icio.us, Facebook, blogs, and so on, have created a new environment
- This involves indexing (tagging) by users, uncontrolled vocabulary, multiple indexers
- Folksonomy: collaboratively creating and managing tags to annotate and categorise content (Wikipedia)

Introduction
Past et présent
The new world

Interesting facts

- Inter-indexer consistency only reaches about 50%
- The huge mass of documentation available requires automated techniques, it’s a question of economics
- Systems don’t necessarily perform very well, but some of them are no worse than what was there before
The DOCAM terminology committee

Why a new tool?

Tools already exist for media art, variable media, unstable media
So why create another one?
Because of all these things we’ve looked at (synonyms, multiple meanings, semiotics, representation)

Some plans

Some of the work we hope to accomplish:

- build a glossary
- build a thesaurus
- build an ontology
- Adopting SKOS (Simple Knowledge Organization System)
- Many discussions later ...

Context

An emerging discipline
All work is experimental
Each project has particular needs
DOCAM does too

Glossary

Description
Characteristics
Work to date
Structure
Take [two]

Description
A tool for managing terminology
By definition, only offers definitions (one or many) for terms
Support DOCAM researchers, the professional community
A complement to existing tools and projects
Characteristics

- Bilingual
- Useful
- Easy access

Bilingual

- French and English
- Finding sources in each language, not translating the definitions
- Transcreation: establish an equivalent in the other language, without translating necessarily

Useful

- Current
- Rapid revisions possible, if necessary
- Respond to user needs

User needs

- A call out to DOCAM committees to send terms
- If necessary, multiple definitions:
  - For example, maybe the definition of a piece of technology is not the same for a cataloguer as for a preserver

Easy access

- Available on the web
- So the most current version is the one online
- Using World Wide Web Consortium (W3C) standards to foster longevity
  - XHTML, CSS2, SKOS
- XHTML version online, new version uses PHP

Work to date

- Search for existing tools
- Search for terminology
- Studying data structures
- Organising, coding the data
- Always some hesitation about the structure, because of interoperability concerns (thesaurus, ontology, systems elsewhere)
**Structure**

- **Term in French**
- **Equivalent in English**
- **See also (other names for this term)**
- **Context (area: film, preservation, )**
- **Definition 1, 2, 3...**
- **Reference (source for each definition)**
- **Source of the request**
- **Creation, modification date of the record**

---

**Example of a record**

- **Decision: database, or XHTML pages?**
- **Old version: XHTML pages**
- **New version: a database, using PHP**
- **An important decision, because the glossary is the basis for the thesaurus**

---

**Thesaurus**

- **Semantic networks**
- **Bilingual**
- **Online**
- **Intellectual work**
- **Partnering**
- **Plans**

---

**Semantic networks**

- **For each term, identify other terms in the glossary that are somehow related**
- **What is the relationship?**
- **Decide which term to use**
- **Identify equivalent in the other language**

- **Note of application:** A chair or a seat has many names. It is possible to see a chair by turning a wheel of a chair.
- **Note:** Similarly, a chair has many names. It is possible to see a chair with a seat.
Bilingual

- Bilingual = unilingual twice
- For many terms, no direct equivalent
- In this example, these equivalents (EP) don’t exist in English

Other aspects

- In general, parallel bilingual thesauruses only work well in specialised situations (e.g. medecine, sciences), otherwise they are messy
- This is for the same reason automatic translation doesn’t work very well
- Different contexts, different cultures, different realities -- there isn’t necessarily an exact equivalent

A few examples

- Fleuve
- Rivière
- Boulangerie
- Pâtisserie
- Bakery

Different reality

Online

- A hypertext version is desirable, so users can visit the nodes of the semantic network
- Of course, accessibility from anywhere is also a consideration

Intellectual work

- As with building an index, building a thesaurus is an intellectual task
- Humans need to do this work, at least for the moment
- There is lots of work going on, but computers can’t yet determine relationships very well
Partnering

- Insofar as it’s possible, we’d like to avoid reinventing the wheel
- The need to study existing thesauruses and develop partnerships
- Some work with V2, Institute for Unstable Media, Rotterdam (<http://www.v2.nl/>)
- Their thesaurus: >V2_Archive

Other possibilities

- Rhizome (<http://rhizome.org>) >Archives
  >TextBase (or >ArtBase to get artists’ names)
- Of course, there are limits to the type of collaboration
- Furthermore, nobody else is using SKOS yet

Ontology?

- A nice dream, we’ll see how we do
- If we get that far, the glossary and thesaurus will certainly contribute
- Maybe a structure like V2, PADI
- Perhaps a longterm project for DOCAM, the FDL, or...

Some thoughts

- User needs vary greatly
- It’s intellectual work, but ultimately, all tools for organising knowledge need to be practical
- Our goal: offer tools that will help users find their way in the sea of available terminology
- Artists, curators, researchers, preservationists can use the tools as they see fit

References


That’s it

james.turner@umontreal.ca
http://mapageweb.umontreal.ca/turner