

Pawarisa Nipawattanapong

Access to Moving Image Collections (CINE-GT 1803)

Data Mapping Project

Fall 2011

Professor Linda Tadic

Data Mapping Project

In this project, I select three data standards to create a crosswalk as follows: MARC 21, Dublin Core, and PB Core.

MARC - leader, Directory, and variable fields are the main components of MARC 21 bibliographic record. It is a data format and set of related standards used by libraries to encode and share information about books and other material they collect.

Strengths:

- Thoroughly divide information

MARC has many numeric, graphic symbols, and alphabet letters for representing data causing we can thoroughly divide information.

- Widely adopted by library communities

MARC is popular to use in libraries. It was developed by Library of Congress in the 1960s, and is still widely used today.

Weaknesses:

- Complicated to understand

MARC is quite complicated more than other formats. There are many numeric, graphic symbols, and alphabet letters to represent each field and element. Therefore, we have to interpret the codes before reading information. Example, I would like to share the information about name of producer, publisher, distributor, or manufacturer of a book, I have to encode it to 264\$b as the usage rules.

Dublin Core – its purpose is to provide a very simple set of metadata. It consists of fifteen elements for use in resource description: title, subject, description, type, source, relation, coverage, creator, publisher, contributor, rights, date, format, identifier, and language. Each Dublin Core element is optional and may be repeated.

Strengths:

- Simplicity

Dublin Core is easy to create a Dublin Core record for any kind of information resource because it is unnecessary to encode. Although people with no formal training in this data format, they can describe web-based information resources quickly and easily.

- Extensibility

These simple elements can be further defined to create more detailed metadata. Example, The description element of Dublin Core may be able to refer to display text, summary, etc., and date/time and place of an event note

Weaknesses:

- Inconsistency

Sometimes the simple elements become a problem because there are no cataloging rules that determine how data will be entered in the fields, it means that there is no consistency across different uses of Dublin Core.

PB Core - the metadata standard was created for use by public broadcasters and related communities. It is used for television, radio, and web activities. It consists of a standard for cataloging and describing media objects in three general ways: intellectual content (title, subject, description, and genre), intellectual property (creator, distributor, and publisher), and instantiation contains (format, media type, duration, file size, data rate, aspect ratio, frame rate, and many other aspects of the media object).

Strengths:

- Free software

PB Core is free to all and there are database tools already in use including variety of open source software applications.

- PB Core provides a wide range of detail, from very simple to extremely specific, therefore we can be confident that PBCore standard has the consistency and reliability for indexing, retrieving, and managing data for immediate applications and over time.

Weaknesses:

- Complexity

PB Core has greater complexity than Dublin Core. It contains more technical fields and the ability which have to train and use the experience of users in this data format.

All data standards not only have advantages but also have disadvantages. Therefore, we have to adopt the most appropriate format in order to obtain the most effective work.