

OLAC CAPC
Moving Image Work-Level Records Task Force
Final Report and Recommendations
July 30, 2010

Summary of the Task Force's Reports

Task Force Members:

Kelley McGrath (chair; subgroup 1 and 4 leader)
Susannah Benedetti (subgroup 2)
Karen Gorss Benko (subgroup 3)
Lynne Bisko (subgroup 4)
Greta de Groat (subgroup 1)
Scott M. Dutkiewicz (subgroup 4)
Ngoc-My Guidarelli (subgroup 2)
Jeannette Ho (subgroup 2 leader)
Nancy Lorimer (subgroup 1)
Scott Piepenburg (subgroup 3)
Thelma Ross (subgroup 3 leader)
Walt Walker (subgroup 3)

Advisors to the Task Force:

David Miller
Jay Weitz
Martha Yee

Pt. 1 Definition of a Moving Image Work and Examination of Work Boundaries

Definition:

After much discussion, we decided on a top level that we called Work/Primary Expression (WPE). The primary expression is usually, but not always, the original public release version. This was a practical compromise that we felt would let us efficiently consolidate into one record all the information associated with one WPE (both work and history of the primary expression) that we would like to re-use in association with any new expression or manifestation.

We also discussed moving image works that are performances of existing works. In this case, we found the FRBRoo (http://cidoc.ics.forth.gr/frbr_inro.html) categories of “performance work,” which they defined for us as “designed to cover the sets of concepts that pertain to the elaboration of live performances and of performances that take place with the sole purpose of being recorded (e.g., in movies, studio recordings of music, etc.)” and “recording work,” which they said was “intended to apply to sets of concepts

that pertain to the elaboration of any kind of recording. The type of the thing recorded is not taken into consideration: it can be birdsong, the changing aspect of the Empire State Building over eight hours in an Andy Warhol movie, or anything” very helpful. From this perspective, we thought that we could define a cinematographic work of a performance as a new recording work based on a particular performance or performances. So each new cinematographic or recording work of a performance of Romeo and Juliet would require a new WPE, as well as a link to the original play and would incorporate the characteristics of a particular performance or performances.

Boundaries:

FRBR doesn’t provide absolute guidance on how to make the decision as to when one work becomes a new work.

“Because the notion of a work is abstract, it is difficult to define precise boundaries for the entity. The concept of what constitutes a work and where the line of demarcation lies between one work and another may in fact be viewed differently from one culture to another.” (FRBR Report)

We attempted to find a practical, but flexible method that will work for most situations. We settled on some heuristics, rather than hard-and-fast rules:

- Degree of commonality among versions. For example, the theatrical release and the director’s cut or the edited-for-airline-viewing version of movies generally share most characteristics in common.
- Extent to which there are primary and derivative versions where the derivative versions are expressions of the primary version. So in the examples given above, we would consider the theatrical release to be the primary expression and the others to be modified expressions of the original one.
- Extent to which a new version can substitute for the original version. For example, we considered the Spanish and English versions of the 1931 *Dracula* to be separate but related WPEs.

In the full report, we provide a number of additional examples that we hope will illuminate these principles.

Pt. 2 Core Attributes and Relationships for Moving Image WPE Records

We listed data elements for inclusion in moving image WPE records and designated each as either Core, Recommended, or Optional. These are generalizations and our priorities might not apply to all situations. It should be possible to expand the model to incorporate new types of data as needed.

We also gave a list of commonly-occurring relationships and commonly-occurring roles. Although we listed the primary creator(s) and contributor(s) as core, we also recognized that what roles are primary depends on the type of work. Therefore, we listed some examples of different types of moving image works and the primary roles commonly associated with them.

Pt. 3a Operational Definitions

This section deals with operational definitions and guidance for five common characteristics of moving image WPEs and expanded considerably from our initial intentions.

In addition to providing operational definitions for the five elements, we gave some guidance on identifying these characteristics. We also realized that our structure would benefit from additional data about data (sometimes called parametadata or meta-metadata). Some of this meta-metadata elaborates on the type of data recorded (e.g., TitleStatus, which identifies a particular title as preferred or non-preferred and TitleLanguage which identifies the language of the title). Other meta-metadata identifies the source of the data in the main element (whose MARC record or what manifestation or what reference source was consulted). This type of granular data about the provenance of information should improve error-checking and enable assessments about the quality of data to be more easily performed without re-verifying the data.

Pt. 3b Data Sources for Information about Moving Image Works

This section examines a number of primary and secondary sources that include information about moving images works and provides brief annotations of coverage, assessment of reliability, and a list of elements covered in each source. Print, and free and subscription online sources are included. Information about these sources is being entered into a relational database in the hope that it can be made available on the web in a flexible manner (e.g., by element covered, types of works covered, level of reliability, online vs. print). We provided some examples of what entries might look like, but the number of sources and elements covered will need to be expanded before implementation.

Pt. 4 Extracting Work-Level Information from Existing MARC Manifestation Records

This part consisted of a pilot project that attempted to extract the same five sample WPE elements used in part 3 from existing MARC bibliographic records. In all cases, we tried a multi-pronged approach as the data we wanted was not always in the same place in all records. Where possible, we used data that is straightforward for a computer to interpret, such as identifying the authorized form of directors' name where relator codes or terms have been used.

We also undertook some more error-prone and processing-intensive approaches. For director, this meant that we parsed 245 \$c (statement of responsibility) and 508 (credits note) into individual statements by using the standard space-semicolon-space punctuation as break points. We then went through all the individual statements looking for the root “direct” and attempted to match the other words in that statement to any of the names in fields. Because this data exists in many records whereas use of relator codes and terms is fairly infrequent, this matching approach had an overall higher success rate. However, there were a number of complications, including variant forms of names that didn’t match (some of which could be fixed by trying to match on cross-references in the name authority record), the need to compile a list of non-English terms for director (which we didn’t do for the pilot project), and the need to identify and exclude other terms that include “direct,” such as “director of photography” or “art direction.”

Although this process is unlikely to yield complete and accurate WPE information in all cases, the success rate is high enough that we think it is a reasonable strategy for initially populating WPE records.

We also gave suggestions for how catalogers can encode these sample WPE characteristics in such a way that they will be more easily extractable by automated means, such as using relator terms or codes with authorized name entries to identify the role(s) that the entity played.

Part 4 also includes an appendix comparing our extracted data with several external sources. We found that, in general, our data agreed with the external sources. In some cases, the external reference sources disagreed among themselves. The main failings of the MARC extraction process were lack of the relevant WPE data in the MARC bibliographic record or our inability to extract the WPE data rather than incorrect data in the MARC record.