

### Analysis

The objects I chose to examine for this metadata creation assignment were taken from my personal DVD collection. While I don't aspire to any claims that this collection of approximately 200 films is of much worth to anyone besides myself, I can list any number of ways in which it has an enduring personal value. Looking back over these discs, I find it possible to chart my own growth and development as a film lover - a path that certainly runs parallel to the academic and professional choices that led me to where I am today, in the MIAP program. I don't know that a collector's edition of *The Social Network* has any direct connection to an interest in archiving, but I can't say they're entirely unrelated either. More than anything I suppose my collection is a fantastic record of personal taste: these are the films that made enough of an impression on me that I specifically sought them out for purchase. As someone who loves to discuss film, recommend viewing and loan movies to friends and acquaintances, I found the idea of cataloging my DVDs an interesting and useful experiment in self-reflection. I also opted specifically to address just my DVDs, as opposed to including films I have acquired via digital download or have available streaming, because of the personal care and attention that has already gone into arranging those physical objects. My efforts to alphabetize and otherwise aesthetically organize my DVD collection, contained either in one of several plastic-sleeved carrying cases or in their original packaging, made this a more natural, contained collection to begin cataloging.

Given the general informality of this collection, in identifying the qualities I wanted to record, I tended to focus mainly on descriptive and physical/technical metadata. Considering my familiarity with the collection and its (I believe) intuitive arrangement on my shelf, I didn't consider assigning identifiers necessary - the title information would be enough for me (or anyone else who happened to be using these records) to find the desired disc. I did, however, want the flexibility to list possible alternative or translated titles for any given film, particularly as I have a number of foreign films in my collection that might be known to someone else by a different title than what I would personally search for. For the other descriptive metadata elements that I considered, I tried to keep in mind that a major use of my collection has been to share films with other people, and that therefore I probably wanted a system in place to help further recommendations by identifying trends in the film's makers or subject matter. Elements such as the film's director and major starring actors are easily identifiable qualities that most people base their taste on, e.g. "I'm a big Alfred Hitchcock fan" or "I think Tilda Swinton is a really interesting actress." Cross-referencing my collection based on these names could lead to other desirable titles. Similarly, a movie's genre and topical contents are great ways to generate suggestions: just as Netflix could guide you from *Se7en* to *Memories of Murder* by way of "critically-acclaimed crime thrillers," so could I do with my collection. A film's original release language could also help here (I don't think anyone would argue the French, for instance, have their

own style), but I also wanted that field just out of curiosity for how many non-English titles were in my collection. Running time is also generally crucial to helping myself or others decide what to watch – I might only have a limited window for viewing and can only consider those films that are 100 minutes in length or less. Finally, I wanted the film's original release date – again, a simple descriptive field that can actually describe a lot about the circumstances in which a film was made, and therefore its content and style. Especially for those with a little more knowledge about film history, a film's release date is excellent shorthand for background information that doesn't fit well into a cataloging record.

These above elements are for the most part about the general work contained on a DVD – they are details that one could acquire from IMDB or other sources and are not particular to my collection itself. So I was also interested in obtaining metadata about these DVDs as unique, physical objects, to use in assessing both the exact quantity and quality of my collection. For instance, while I was not especially concerned with the studio that originally produced the film, I did want to record the company that manufactured and distributed the actual disc. This would be more for my personal benefit than anyone else looking at the record, as I'm generally familiar with the quality of various company's digital transfers and packaging; I would know, for instance, that a specialty outfit like The Criterion Collection is likely going to have a more satisfactory release of obscure titles than a major studio's perfunctory attempt to capitalize on their back-catalog. The publishing date of the DVD, similarly, could offer information about the disc's quality (and also double as preservation metadata, since it would help identify older discs that are more likely to break down). I would want to note the number of discs in my collection associated with each title, another sort of shorthand field that would help identify whether there are a considerable number of special features available for any given film (and help control the total number of objects in the collection). The condition of the disc(s) would also be crucial – I want to be able to identify films in my collection that may not be playable anymore, as candidates for replacement or “de-accession.” Finally, I wanted to note the type of packaging or container that the disc was kept in, whether it was in a sleeved case or its original box – both as a way of locating the disc, and as administrative metadata for organizing the collection (indicating whether I might need to buy a new sleeved case, for instance).

In choosing a data structure standard for this collection, then, I required a schema that was strong on descriptive data for moving images while also allowing for at least a decent amount of flexibility in collecting physical metadata. I ultimately settled on the MODS schema developed by the Library of Congress. While there are structural standards that have been more specifically created for cinematographic works (e.g. PB Core and EBU Core, the FIAT data set, CEN.BT 372), these schemas put a greater emphasis on the instantiation of the work than was strictly necessary for my collection. Certainly in production environments, or in large moving image collections that might contain more than one version of the same general work, this emphasis would be useful, but in dealing with a collection where all the items were of the same

format (DVD), it seemed overkill. As a personal collection, this project similarly did not really require much of a need for legal and rights information. Some of these other schema also did not even easily allow for the listing of numerous contributors, as would be necessary for cataloging various actors' names in large ensemble casts. On the other hand, my desire for some technical and administrative data (number of discs, packaging/container, etc.) required a slightly more complex system than the simplified catch-all elements of Dublin Core. MODS seemed the best option, considering its fairly thorough (and repeatable) descriptive metadata elements; in particular the ability to distinguish origin information for both the original release date of the film and the release date of the DVD was a benefit, since those were two of my key desired fields. The MODS <physicalDescription> category also hit generally the right amount of specificity, and was clearly intended for communities and users working with analog/physical formats as I was. The free-text <note> subelement here was also good for my desired “disc condition” field – I did not want to be hampered by controlled vocabulary for this element, instead seeking the ability to put the extent of damage in my own words.

I was, however attracted by the ability in MODS to easily insert controlled vocabulary authorities into the subject and genre fields. Since I was particularly interested in cross-referencing my collection in order to find thematic trends (for improved recommendations), it was essential that these fields stay consistent in their content and values. Primarily to maintain consistency with my Library of Congress-developed structural standard, I decided to use the LOC's Moving Image Genre-Form Guide for the genre field, and LOC Subject Headings to fill the subject field. A large benefit for these controlled vocabularies lies in their familiarity – anyone used to navigating academic or public library catalogs would hopefully recognize the terms used in these fields. These are terms that have been developed over the years based on the same kind of general indexing that I was hoping to do with my own collection; using the LOC Subject Headings seemed a good way to ensure that I could easily leap from, say, *12 Angry Men* to *Runaway Jury*. Meanwhile, MIGFG in particular was appealing since it was clearly tailored with an eye on specifically film history – possible terms like “film noir” and “animation” are clearly suited to the cinematic community where other genre-form guides are designed for a broader swath of media. Certainly for a non-professional collection such as the one I was dealing with, this combination of MODS and LOC controlled vocabularies seemed the best route to ensure an easily accessible level of detail and clarity.

### Data Mapping

Desired Elements	MODS Field
Title	<titleInfo> / <title>
Director	<name> / <namePart> and <name> / <role> / <roleTerm>
Actor(s)	<name> / <namePart> and <name> / <role> / <roleTerm>

Original Release Date	<originInfo type="publication"> / <dateIssued>
DVD Distributor	<originInfo type="distribution"> / <publisher>
DVD Distribution Date	<originInfo type="distribution"> / <dateIssued>
Original Language	<language>
Running Time	<physicalDescription> / <extent>
Number of Discs Associated with Film	<physicalDescription> / <extent>
Condition of Disc(s)	<physicalDescription> / <note type="condition">
Packaging/Container	<physicalDescription> / <note type="organization">
Genre	<genre>
Subject/Content Topics and Themes	<subject>