

Project 2: Video Collection Analysis

Erica Krimmel

San José State University

LIBR 282, Prof. Horodyski

31 March 2014

Introduction

In this project, I evaluated a collection of 18 video clips provided to us by Professor Horodyski. After reviewing the content, I created standard metadata fields to describe the clips. In the analysis below I discuss the justification for these metadata fields, as well as issues related to the access and content of the collection as a whole.

Metadata Fields & Justification

Metadata is important to digital assets because it increases accessibility, retains context, and expands use (Gill, 2008). For this collection, I began defining metadata with a field that would uniquely identify each clip, **Filename**. Then I moved on to other basic and objective information: **Title**, which I standardized as “Video Clip #...;” **Length**, in the format HH:MM:SS, with seconds left off if unnecessary; **Dimensions** of the video frame; **Codec** that the video was compressed with; and **Sound**, which I intended to record the audio codec, although it turned out that none of the clips had sound.

From here, I moved on to more subjective metadata fields. **Setting** is meant to describe where the clip takes place, using pre-coordinate terms such as “Indoor–Office.” I decided to use pre-coordinate terms in order to standardize the setting metadata; the terms also offer users a sense of hierarchy, from broad (“Indoor”) to specific (“Office”), which facilitates easier browsing (NISO, 2005). Some of the setting terms also have qualifiers, for example “Building (Abandoned).” **Keywords** is used to draw out core concepts from each clip for ease of retrieval, while **Description** provides a more detailed, though not indexed, summary of the clip.

I noticed that all of the video clips appeared to be from popular films; even though I didn’t have verification of this, it seemed like valuable information to include in the metadata. Thus, I added the fields **Excerpt From**, **Year Published**, and **Recognizable Actors**. When

assigning metadata to each clip, if I recognized the film it was from I put the film's title under **Excerpt From**, and when it was released under **Year Published**. If I recognized the actors in the clip I named them under **Recognizable Actors**, in the format "Lastname, Firstname." While I knew some of the actors names on sight, others I looked up using the Internet Movie Database website. Finally, I added the field **Number of Characters** as a reference point for **Recognizable Actors**—if I couldn't figure out who one or more of the actors was, at least I could record their presence as a number.

You may view the metadata I assigned to each clip in "Appendix A: Video Clip Metadata." Most of these instances were straightforward, but for several clips I had no idea what film they were excerpted from. In these cases, I filled in the unknown fields with "unk." to acknowledge that I thought about the field, but could not offer any real metadata for it. I think filling in *something* for each field is important in metadata that will be used externally, because although the instances of "0," "unknown," and no metadata seem equivalent, they can actually mean very different things. Zero should literally mean "0," as it does when I used it for **Number of Characters** (i.e. no people were present in the clip). Meanwhile, having a blank field could mean either "0" or "unknown," which is misleading and must be avoided.

Bulk Entry & Automatically Generated Metadata

Like most file types, these video clips have embedded metadata that is generated automatically. The fields I included with automatically generated metadata are **Dimensions**, **Codec**, and **Length**. The file type is also automatically generated (".mov" for all of these clips), however I decided not to include this as a metadata field by itself because it is included in the field **Filename**. According to Dale and Waibel (2004), technical metadata, such as **Codec** and **Filename**, are critical because they "assure that the information content of a digital file can be

resurrected even if traditional viewing applications associated with the file have vanished.”

Both **Filename** and **Title** could be bulk entry metadata, as they both follow a distinct naming pattern. In the scenario of this assignment, **Filename** could be considered automatically generated because I did not have to do anything to determine it—but I think this is deceptive because at some point a human decided what to name each clip.

Bulk and automatic metadata can be excellent timesavers. As Schaffel (2006) points out, embedded metadata allows a DAM system to actively harvest information rather than passively having people contribute it. In a DAM system, bulk metadata can either be defined once and applied to many records, or defined as a pattern and applied uniquely to each record. For this assignment I couldn't make the best use of bulk and automatic metadata because my DAM system was just a Microsoft Word document, nonetheless, I definitely see the benefit for more capable systems.

Licensing and Legal Metadata

I had a difficult time deciding what to include in my metadata schema for licensing and legal information. In the end, I decided against including any fields related to these issues because I felt the metadata generated would be inaccurate. If I had a field for legal or licensing information, any entries would have been based purely on my guess at which film each clip came from, and although I am fairly confident in the clips I assigned films to, I would want to be 100% certain to extend licensing information also. This is not to say that this type of metadata is unimportant; on the contrary I believe if I were to go any farther with this collection it would have to include licensing and legal metadata so as not to violate any copyright.

Collection Accessibility

Initially this collection was accessible only as individual files, which wasn't ideal because users didn't know what a file contained until it was opened *and* viewed. With the addition of my metadata records, the collection is much more accessible because users can browse or search for desired clips. Of course, with metadata recorded only in text format, as it is now, search capacity is limited, but even this limited version is much more effective than a collection of files alone. In a proper DAM system, it would be even more accessible, as the metadata could be searched post-coordinately and limited by fields.

Collection Content, Classification & Taxonomy

I imagine this collection could be included as a digital reference for some sort of media company that uses themed movie clips in new media creation. Nearly every clip contains either fighting or fire, often in the form of explosions, so if this collection were included in a DAM system, I might classify it as "Major Motion Picture Video Clips" and subclassify it with the tag "violence," which I think applies to all of the 18 clips. Someone using the DAM system could easily search within this classification for more specific violence, such as "fight," "collision," or "explosion."

When creating my metadata schema, I did not use any official taxonomy because that seemed contrary to the learning purpose of this assignment. That said, if I was to apply this assignment to any sort of real-world application I would definitely research existing taxonomies that could help classify these clips. As the National Information Standards Organization (2005) states, interoperability between taxonomies is critical to harness the collective power of different collections. Creating or linking this collection to an established taxonomy would also enable better system organization and findability, and a thesaurus would help define relationships

between terms (Hedden, 2010).

Conclusion

In this project I simulated the experience of digital asset metadata creation from start to finish. I approached a collection of unknown video clips, assessed them for meaningful similarities and differences, turned this assessment into a metadata schema, and assigned metadata to each clip based on my schema. I experienced the iterative nature of metadata design, in that as I was assigning metadata to individual clips I realized that my schema needed modification. Finally, I evaluated the collection based on its accessibility, content, and metadata. Throughout this process, rationale from our course readings helped me make decisions and justifications.

References

- Dale, R. and Waibel, G. (2004). Capturing technical metadata for digital still images. *RLG DigiNews*, 8 (5). Retrieved from: <http://worldcat.org/arcviewer/2/OCC/2009/07/10/H1247256969461/viewer/file1.html>
- Gill, T., Gilliland, A., and Woodley, M. Baca, M. (ed.). (2008). *Introduction to metadata: Pathways to digital information, Online edition, version 2.1*. Los Angeles: J. Paul Getty Trust. Retrieved from http://www.getty.edu/research/publications/electronic_publications/intrometadata/pdf.
- Hedden, H. (2010). Chapter 1: What are taxonomies? In *The Accidental Taxonomist* (pp. 1–37). Medford, NJ: Information Today.
- National Information Standards Organization (NISO). (2005). Guidelines for the construction, format, and management of monolingual controlled vocabularies: an American national standard developed. Bethesda, Md.: NISO Press.
- Schaffel, B. (2006). Using DAM to drive workflow: A concept for transforming information into knowledge. *Journal of Digital Asset Management*, 2(3-4). Retrieved from <http://www.palgrave-journals.com/dam/archive/index.html>

Appendix A: Video Clip Metadata

Filename: LIBR_284_01.mov

Title: Video Clip #1

Length: 00:07

Dimensions: 647 x 486

Codec: Sorenson Video 3

Sound: none

Excerpt From: *The Matrix*

Year Published: 1999

Recognizable Actors: Weaving,
Hugo

Number of Characters: 2

Setting: Indoor–Building
(Abandoned)

Keywords: collision

Description: Two agents leap into the air, collide with rafters and wall, fall to floor.

Filename: LIBR_284_02.mov

Title: Video Clip #2

Length: 00:03

Dimensions: 647 x 486

Codec: Sorenson Video 3

Sound: none

Excerpt From: *James Bond*

Year Published: unk.

Recognizable Actors: Brosnan,
Pierce

Number of Characters: 2

Setting: Indoor–Train Tunnel

Keywords: sports car, 007, confetti,
machine gun

Description: Brosnan tosses box out over empty train track, mounted machine gun shoots and it explodes into confetti.

Filename: LIBR_284_03.mov

Title: Video Clip #3

Length: 00:02

Dimensions: 647 x 486

Codec: Sorenson Video 3

Sound: none

Excerpt From: *RoboCop*

Year Published: 1987

Recognizable Actors: Weller, Peter
Number of Characters: 2

Setting: Outdoor–Industrial

Keywords: gun, explosion, robot

Description: Villain shoots at RoboCop in industrial complex, explosion ensues.

Filename: LIBR_284_04.mov

Title: Video Clip #4

Length: 00:02

Dimensions: 647 x 486

Codec: Sorenson Video 3

Sound: none

Excerpt From: *RoboCop*

Year Published: 1987

Recognizable Actors: Weller, Peter
Number of Characters: 1

Setting: Indoor–Office

Keywords: breaking through
window, robot, explosion

Description: Explosion causes RoboCop to fall backward through window, breaking glass and landing in empty office setting.

Filename: LIBR_284_05.mov

Title: Video Clip #5

Length: 00:02

Dimensions: 647 x 486

Codec: Sorenson Video 3

Sound: none

Excerpt From: *The Fifth Element*

Year Published: 1997

Recognizable Actors: Willis, Bruce

Number of Characters: 1

Setting: Indoor–Atrium

Keywords: explosion, fire, furniture

Description: Fire is already burning when an explosion shakes the room, as Willis takes cover beneath a table.

Filename: LIBR_284_06.mov

Title: Video Clip #6

Length: 00:02

Dimensions: 647 x 486

Codec: Sorenson Video 3

Sound: none

Excerpt From: *The Fifth Element*

Year Published: 1997

Recognizable Actors: Willis, Bruce

Number of Characters: many

Setting: Indoor–Atrium

Keywords: explosion, shooting star, fire

Description: Groups of Characters are huddled in atrium when a shooting star causes an explosion and Willis takes cover under a table.

Filename: LIBR_284_07.mov

Title: Video Clip #7

Length: 00:01

Dimensions: 647 x 486

Codec: Sorenson Video 3

Sound: none

Excerpt From: *The Matrix*

Year Published: 1999

Recognizable Actors: Reeves, Keanu; Fishburne, Laurence

Number of Characters: 2

Setting: Indoor–Studio (Karate)

Keywords: fight, karate

Description: Reeves and Fishburne karate fight in a studio.

Filename: LIBR_284_08.mov

Title: Video Clip #8

Length: 00:01

Dimensions: 647 x 486

Codec: Sorenson Video 3

Sound: none

Excerpt From: *Planet of the Apes*

Year Published: 2001

Recognizable Actors: Wahlberg, Mark

Number of Characters: 4

Setting: Outdoor–Jungle

Keywords: apes, captive, explosion

Description: Wahlberg is being held captive by three apes when an explosion disrupts them.

Filename: LIBR_284_09.mov

Title: Video Clip #9

Length: 00:04

Dimensions: 647 x 486

Codec: Sorenson Video 3

Sound: none

Excerpt From: *The Matrix*

Year Published: 1999

Recognizable Actors: Reeves, Keanu; Weaving, Hugo

Number of Characters: 2

Setting: Indoor–Building (Abandoned)

Keywords: fight, anti-gravity, collision

VIDEO COLLECTION ANALYSIS

10

Description: Weaving grabs and punches Reeves, who flies backward through a shelf.

Filename: LIBR_284_10.mov

Title: Video Clip #10

Length: 00:03

Dimensions: 647 x 486

Codec: Sorenson Video 3

Sound: none

Excerpt From: *iRobot*

Year Published: 2004

Recognizable Actors: unk.

Number of Characters: many

Setting: Indoor–Warehouse

Keywords: robots, attack, breaking door

Description: Human guards are flung out of the way as horde of robots breaks through double doors to attack.

Filename: LIBR_284_11.mov

Title: Video Clip #11

Length: 00:09

Dimensions: 647 x 486

Codec: Sorenson Video 3

Sound: none

Excerpt From: unk.

Year Published: unk.

Recognizable Actors: unk.

Number of Characters: 2

Setting: Outdoor–Residence

Keywords: house fire, cyclist, explosion

Description: Cyclist is watching a house on fire as man runs out from inside and house explodes in flames.

Filename: LIBR_284_12.mov

Title: Video Clip #12

Length: 00:03

Dimensions: 647 x 486

Codec: Sorenson Video 3

Sound: none

Excerpt From: *Terminator 3*

Year Published: 2003

Recognizable Actors:

Schwarzenegger, Arnold; Loken, Kristanna

Number of Characters: 2

Setting: Indoor–Warehouse

Keywords: breaking ceiling, fall, debris

Description: Schwarzenegger and Loken character fall clutching each other through the roof of the warehouse landing on table with much debris.

Filename: LIBR_284_13.mov

Title: Video Clip #13

Length: 00:03

Dimensions: 647 x 486

Codec: Sorenson Video 3

Sound: none

Excerpt From: unk.

Year Published: unk.

Recognizable Actors: unk.

Number of Characters: 2

Setting: Indoor–Office; Outdoor–Building

Keywords: explosion, run away,

Description: Man is inside his office when an explosion knocks him off his feet, meanwhile a man with a scarred face runs away outside the building.

VIDEO COLLECTION ANALYSIS

Filename: LIBR_284_14.mov

Title: Video Clip #14

Length: 00:03

Dimensions: 647 x 486

Codec: Sorenson Video 3

Sound: none

Excerpt From: unk.

Year Published: unk.

Recognizable Actors: unk.

Number of Characters: 2

Setting: Indoor–Museum

Keywords: robot, flames, escape

Description: Robot fires flame shooter at man, who escapes through another room while carrying two guns.

Filename: LIBR_284_15.mov

Title: Video Clip #15

Length: 00:00:30

Dimensions: 647 x 486

Codec: Sorenson Video 3

Sound: none

Excerpt From: unk.

Year Published: unk.

Recognizable Actors: 0

Number of Characters: 0

Setting: Outdoor–Jungle

Keywords: explosion, jungle

Description: Explosion happens in the distance, obscured from the viewer by jungle vegetation.

Filename: LIBR_284_16.mov

Title: Video Clip #16

Length: 00:03

Dimensions: 647 x 486

Codec: Sorenson Video 3

Sound: none

Excerpt From: *The Matrix*

Year Published: 1999

Recognizable Actors: Reeves, Keanu; Fishburne, Laurence

Number of Characters: 2

Setting: Indoor–Studio (Karate)

Keywords: fight, karate

Description: Reeves kicks Fishburne, propelling him backward into a beam, which Fishburne breaks.

Filename: LIBR_284_17.mov

Title: Video Clip #17

Length: 00:07

Dimensions: 647 x 486

Codec: Sorenson Video 3

Sound: none

Excerpt From: *Terminator 2*

Year Published: 1991

Recognizable Actors:

Schwarzenegger, Arnold; Patrick, Robert

Number of Characters: 2

Setting: Indoor–Building–Hallway

Keywords: fight

Description: Schwarzenegger and Patrick fight, pushing each other into and eventually through the walls.

Filename: LIBR_284_18.mov

Title: Video Clip #18

Length: 00:08

Dimensions: 647 x 486

Codec: Sorenson Video 3

Sound: none

Excerpt From: *The Fifth Element*

Year Published: 1997

Recognizable Actors: Willis, Bruce; Tucker, Chris

Number of Characters: 2

Setting: Indoor–Building

Keywords: gun, fall

VIDEO COLLECTION ANALYSIS

12

Description: Willis fires a ring of shots at the ceiling, causing the floor where Tucker is laying to give out and fall at Willis' feet.