IPTC Video Metadata Hub: Moving to Adoption

Pam Fisher
The Media Institute
p.fisher@cs.ucl.ac.uk
IPTC Autumn Meeting: Video Day
IPTC Video Metadata Hub

• ‘Hub’ rather than stand-alone standard, supporting all video metadata communities
  – Acknowledges prior investments in standard creation and adoption
  – Unifies vertical ‘silos’ of support

• Represents the metadata community with the leading track record of adoption: IPTC

• **Now comes the interesting part: implementation and adoption!**
  – Three areas to discuss:
    • Automated metadata extraction
    • Cross-collection vocabularies
    • Open tools
Video Metadata Communities Abound

- Differentiated by key purpose:
  - Workflow and working practice
    - (e.g. IMF, DPP, NABA)
  - Asset management
    - (e.g. pbCore, EBUcore)
  - System operations
    - (e.g. SMPTE, ISO-MPEG)
  - Search and discovery
    - (e.g. EIDR, ISAN, ContentID, Teletrax)

- Until now, vertical silos:
  - Broadcast contribution
  - Stock footage sales
  - Digital cinema
  - News contribution
  - Social media / user generated

- Might include any or all of 3 broad sets of data:
  - Semantic (IPTC: describing a/v content)
  - Rights (IPTC: rights)
  - Engineering (IPTC: technical)
Moving Image Demographics

• 200m+ unique hours of professionally-produced or professionally-accessioned moving image content around the world today
  – Many billions of instances / versions / caches / etc
  – Less than 10% originals are digitally mastered (both traditional and ‘born digital’)
  – Less than 5-21% rights-ready (upper figure: Screen Digest/FOCAL)
  – Annual growth? (ISAN say 6m hours annually)

• Video now dominates the Internet
  – IP traffic worldwide is 68% video and growing (excluding BitTorrent, P2P) – will reach 90% by 2020 (Cisco VNI: http://bit.ly/CiscoVisNetIndex)
    • Some interesting subsets: VR 4x growth 2015, 61x by 2020; 19.7% of digital ad spending is video

• Overall, minimal metadata coverage within any of the vertical silos
  – Plus, challenges to accuracy and sustainability - example
Unique IDs, opportunities and issues

Unique, searchable IDs are marching forward (ISAN, EIDR, et al). Currently, much confusion and work required by rights holders (versus distributors & aggregators)

- Example: EIDR record entered and maintained by Rovi, versus ITV shot log (partial data shown)

### EIDR entry:

**BASE OBJECT DATA**

<table>
<thead>
<tr>
<th>Field</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>EIDR ID</td>
<td>10.5248/AAAB-2ICA-40DA-ABSB-26EF-F</td>
</tr>
<tr>
<td>Structural Type</td>
<td>Abstraction</td>
</tr>
<tr>
<td>Mode</td>
<td>Audio/Visual</td>
</tr>
<tr>
<td>Title</td>
<td>Sea Otter Saga</td>
</tr>
<tr>
<td>Ext. Original title</td>
<td>Survival - Saga Of The Sea Otter</td>
</tr>
<tr>
<td>Main Title</td>
<td>Survival</td>
</tr>
<tr>
<td>Caption</td>
<td>&quot;Sea otter floating on back among kelp fronds, flat stone on abdomen&quot;</td>
</tr>
<tr>
<td>Query Fld</td>
<td>VAR CU sea otter floating on back among kelp fronds, flat stone on abdomen</td>
</tr>
<tr>
<td>Narrative</td>
<td>&quot;VAR CU sea otter floating on back among kelp fronds, flat stone on abdomen</td>
</tr>
<tr>
<td>Media Form</td>
<td>Seq</td>
</tr>
<tr>
<td>Colour Designator</td>
<td>Col</td>
</tr>
<tr>
<td>Source</td>
<td>gbs/V02/GBSO0000000362</td>
</tr>
<tr>
<td>Creation Date</td>
<td>09 February 2004</td>
</tr>
<tr>
<td>Transmission Date</td>
<td>25 January 1973</td>
</tr>
<tr>
<td>Program Number</td>
<td>71/10</td>
</tr>
<tr>
<td>Filming Composing Date</td>
<td>01 January 1971</td>
</tr>
<tr>
<td>Duration</td>
<td>00:00:07:00</td>
</tr>
<tr>
<td>Tape Number</td>
<td>GBS00000000495</td>
</tr>
<tr>
<td>Mrl Sum Dir</td>
<td>gbs/V02/GBSO0000000362</td>
</tr>
<tr>
<td>Quality Notes</td>
<td>No.2532 Ends At 57.22</td>
</tr>
<tr>
<td>Collection Name</td>
<td>SLA</td>
</tr>
<tr>
<td>Owner</td>
<td>Granada</td>
</tr>
<tr>
<td>Country Of Origin</td>
<td>UK</td>
</tr>
<tr>
<td>Time Code In</td>
<td>00:02:52:00</td>
</tr>
<tr>
<td>Time Code Out</td>
<td>00:02:59:00</td>
</tr>
</tbody>
</table>

### Extra Object Metadata

- **Parent**: 10.5248/F6EO-FD44-1FA2-EO2A-3EAC-W
- **Sequence Info**
  - Distribution Number: 12
  - House Sequence: 112

**Catalogue entry for example shot:**

```
"title_all": "Survival - Saga Of The Sea Otter -",
"title_2": "Saga Of The Sea Otter",
"main_title": "Survival"
"caption": "Sea otter floating on back among kelp fronds, flat stone on abdomen",
"qry_fld": "VAR CU sea otter floating on back among kelp fronds, flat stone on abdomen | Sea otter floating on back among kelp fronds, flat stone on abdomen | Survival | Saga Of The Sea Otter | |
"narrative": "VAR CU sea otter floating on back among kelp fronds, flat stone on abdomen",
"media_form": "Seq", "colour_designator": "Col",
"media_resource_locator": "10.5248/AAAB-2ICA-40DA-ABSB-26EF-F",
"creation_date": "09 FEBRUARY 2004",
"transmission_date": "25 JANUARY 1973", "program_number": "71/10",
"filming_composing_date": "01 JANUARY 1971",
"duration": "00:00:07:00",
"tape_number": "GBSO0000000495", "mrl_sum_dir": "gbs/V02/GBSO0000000362",
"quality_notes": "No.2532 Ends At 57.22",
"collection_name": "SLA",
"owner": "Granada",
"country_of_origin": "UK",
"time_code_in": "00:02:52:00",
"time_code_out": "00:02:59:00",
"medium": "Film",
"film_roll_number": "0688",
"film_guage": "16mm",
```

**Total IMDB entry:**

Survival

**Saga of the Sea Otter** (7 Feb. 1973)

TV Episode - Documentary
Automated Metadata Extraction: Historic

Automated metadata extraction has been in academic (1990) and commercial (1995) development for over 20 years

Automated Metadata Extraction: Now

• Consolidation in new ‘deep’ learning methods (DCNN); robust new methods (VLAD, VLAC, VVLAD); performance advances (e.g. during real-time transit)
• More info: [http://resources.tmi.io/IBC2015_VideoClarity.pdf](http://resources.tmi.io/IBC2015_VideoClarity.pdf)
Why do we need this (1)?

- Discover semantic meaning; e.g. ImageNet (valuable ‘ground truth’):
  - Major human tagging initiative, 14m images, organized by WordNet nouns
  - See Professor Li Fei Fei’s excellent TEDx talk here: http://bit.ly/FFL-IM
Why do we need this (2)?

- Anti-piracy
- De-duplication of file systems
- With visual quality analysis – compare and characterize all instances of a work
- Displaces earlier fingerprinting technology
Why do we need this (3)?

- Because visual quality matters!
  - Just when you think it can’t get worse...
Cross-collection Vocabularies

• Significant work to be done
  – Aim: simplicity and adoptability!
• Coding format example:
  – How many? (pbCore: 243; Netflix IMF: 2)
• Significant growth in complexity must be represented:
  – color space (Rec.2020, ACES, HDR)
  – Object-based media, integrative media (W3C MSE, EME, Timed Text 2)
  – Production formats and profiles as SMPTE standards (GoPro Cineform VC-5, Apple ProRes VC-6)
  – Visual quality metrics, human perception
Open Tools

• Open toolsets are needed to accelerate adoption
  – Example: IMF-DPP mapping, re-formatting tool wipes out format conflict in broadcast contribution (Netflix, others)
No substitute for action:

• Moving to the next stage:
  – Open toolsets
  – Vocabulary effort
  – Genuine experiments

• Thank you!
  – Pam Fisher (p.fisher@cs.ucl.ac.uk)