MOVIELABS COMMON ONTOLGY

Daniel Lucas
V.P. Business Intelligence, Technology
What is an Ontology?

An Ontology for Metadata

“A set of concepts and categories in a subject area or domain that shows their properties and the relations between them.” Wikipedia

Provides a framework and vocabulary for metadata related to a creative work.
Why Should I Care...

Benefits

- Simplify the interchange of data between systems
- Map between internal silos
- Map and ingest from outside vendors
- Share data with other partners or studios
Traditional ETL

Multiple mappings from each source to every other: different terminology, normalizations and schemas every time
Multiple mappings from each source to every other: different terminology, normalizations and schemas every time
Applying an Ontology & Mapping

Single Mapping and normalization for each source to Common Ontology
Ontologies are not Schemas

Ontology

RDF Triples

SQL Schema

JSON Schema

XML Schema

RDF Graph
Driving Principles

What is Different

Existing open schemas did not cover industry use cases
Covers the major elements of title metadata
Describe international details, language, country
Extensible, either in public or private domain
Source (provenance) of every element
External identifiers and references for many fields, allowing other systems to be linked
Focus on practicality and implementation, we built a working prototype as a graph and in MongoDB
Common Ontology Classes

Creative Work
- Identifiers
- Titles
- Contributors
- Release date and consumption metrics
- Descriptive data (keywords, rating, run length, locations, cost)

Person
- Identifiers
- Names
- Birth, Death, Gender, etc.
- Work to which person has contributed

Groups
- Creative Works
- Type

Identifiers
- Titles
- Releases Dates
- Consumption
- Original Language
- Run Time
- Associated Companies
- Ratings
- Rankings
- Keywords
- Based On
- Cost
- Artwork
- Text
- Contributors
Roadmap

Current status
• White Paper published
• RDF available
• JSON schema implementation ready
• EIDR, IMDb mappings.

Future Steps
• Focus on use cases
• Extend ontology (Groups, TV, Manifestations)
• Common Metadata and other mappings
• Evangelize with industry
• Generate momentum with open-source applications

https://movielabs.com/creative-works-ontology/

NodeJS code base to create conformant JSON data
AWS Neptune Instantiation
BACKUP
Tested Implementation

Important that ontology was practical

Built prototype to test and demonstrate (MongoDB & Graph)

Includes data from over a dozen different sources

Over 250,000 film titles

All normalized, using controlled vocabulary, ISO

With mapping tools new sources can be incorporated in as little as a few hours
# Creative Work Class

<table>
<thead>
<tr>
<th>Identifiers</th>
<th>Identifiers for work in other systems</th>
<th>URI</th>
<th>Id, type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Titles</td>
<td>The name of the work</td>
<td></td>
<td>Language, Country, Type(official, original, AKA)</td>
</tr>
<tr>
<td>Releases</td>
<td>Release dates</td>
<td></td>
<td>Date Distribution(theatrical, broadcast, home), Channel(festival, wide, pvod, disc, digital), Format(dvd, 3d, iMax, 70mm)</td>
</tr>
<tr>
<td>Consumption</td>
<td>Consumption metrics</td>
<td></td>
<td>Time period(dates, weeks), Units(currency, theatres, views), Amount</td>
</tr>
<tr>
<td>Original Languages</td>
<td>Original language of work</td>
<td></td>
<td>Language</td>
</tr>
<tr>
<td>Associated Locations</td>
<td>Locations associated with the work</td>
<td>Y</td>
<td>Country, Location(coordinates), Type(filming, setting, funding, production)</td>
</tr>
<tr>
<td>Approximate Length</td>
<td>The approximate running length</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Associated Companies</td>
<td>Companies associated with the work</td>
<td>Y</td>
<td>Type (studio, production, producer)</td>
</tr>
<tr>
<td>Ratings</td>
<td>Film classification or rating</td>
<td>Y</td>
<td>Country, System</td>
</tr>
<tr>
<td>Rankings</td>
<td>User or critic rankings</td>
<td></td>
<td>Type (user, critic, aggregated)</td>
</tr>
<tr>
<td>Keywords</td>
<td>Descriptive keywords</td>
<td></td>
<td>Type (genre, subject)</td>
</tr>
<tr>
<td>Based On</td>
<td></td>
<td></td>
<td>Type (book, comic, person, character, play, musical, game)</td>
</tr>
<tr>
<td>Cost</td>
<td>Costs associated with film</td>
<td></td>
<td>Type(Production, marketing)</td>
</tr>
<tr>
<td>Artwork</td>
<td>Artwork associated with film</td>
<td>Y</td>
<td>Type(Poster,</td>
</tr>
<tr>
<td>Text</td>
<td>Descriptive text related to work</td>
<td>Y</td>
<td>Type(synopsis, review, copyright)</td>
</tr>
<tr>
<td>Contributors</td>
<td>Cast &amp; Crew associated with work</td>
<td>Y</td>
<td>Name, role, portrayal, billing order</td>
</tr>
</tbody>
</table>