Knowledge Based Modelling of the Movie Domain

Mita Paul

Dr. Sibsankar Jana

ABSTRACT-

The ontology has been developed in order to represent the Movie as domain entity and the characteristics of movies has been modelled through specification and definition of subclasses, object properties, data properties and annotation properties within Protege ontology editor. This work is the further incremental development of the existing movie ontology developed by IMDb in 2009. The integration and association of related concepts have developed through Methontology approach.

Keywords: Movie Ontology, Protege, Ontology Development, Upper-Level Ontology

Mita Paul

Librarian, Karimpur Pannadevi College, Nadia, West Bengal, India
Email- mita.sumi@gmail.com

Dr. Sibsankar Jana

Associate Professor, Department of Library & Information Science
University of Kalyani, West Bengal, India
Email-sibs_jana@yahoo.com

INTRODUCTION

Ontology development is such an area where multi-faceted knowledge across fields mingles together. But the process of ontology development is not same, rather the development process varies according to knowledge architecture of the domain and application use and thus ontologies differ from each other due to existence of ontological commitments and choice of application use-cases. Knowledge structure is very much multi-dimensional and its underlying knowledge is shaped by growth of body of the knowledge within a subject, spatio-temporal factors, existence of isolates and others. The success of an ontology depends on reuse of ontological terms, property assertions, ontology integration, arrangement of terms according to filiatory sequence, identification of existent terms and placing them into hierarchical structures, items accumulation, metadata use etc. Semantic Web constitute a semantic operability between classes which can be axiomatized through object and data properties and help to integrate different concepts and information to build meaningful relations and each are uniquely identified by URI.

Film is a visual art and it's a powerful medium to connect people. Every movie has an objective message to the audience. The knowledge structure of a movie encompasses the themes, motifs, symbols, character arcs, narrative arcs, and stylistic choices that make up the film's story. It includes the relationships between characters, their motivations, and the events that drive the plot. The knowledge structure also encompasses the cultural, historical, and social context in which the movie was made, as well as its reception and impact. Understanding the knowledge structure of a movie requires a comprehensive understanding of film history, storytelling techniques, and narrative structure, as well as a deep

appreciation for the art form itself.

Most importantly, film making is an industry and cinematic form is manifested through scripts, songs, dialogues through actors. Writer writes real life situations into scripts on paper, likewise a director transforms those life situations into a series of image and sound through proper sequencing. Film making can be divided into three parts – scripting a scenario, making an environment by selecting a location and shooting sequences. The last part is the editing and production process of the movie. Film making is one area, but after production of a film it becomes an information object with different attributes and properties. All these are together inter-woven and it has potential to create data model.

REVIEW OF LITERATURE

After a brief literature review, no specific literature was found with special reference to Bengali movies. Although, the available literature are mostly circled around designing movie recommendation system using algorithms of information retrieval such as – Cosine tf-idf similarity measurement, BM25, Jaccard Co-efficient, Hidden Markov Model (HMM). Pilaszy and Tikk(2009) have created metadata to collect explicit feedbacks from the user and analyzed to rank movies depending on genres using Netflix data. Ahn and Shi(2009) have developed a simple and low-cost movie recommendation system harnessing vast cultural metadata, about movies, existing on the Web and for the description work cultural metadata that is from IMDb. Bogers (2010) have created a contextual graph of users, tag, movie, actors and genre. The contextual graph was used to compute random walk to make a ContextWALK Model for contextual recommendation of movies based on graph weight scores. Hassanzadeh, O., &Consens, M. P. (2009) have made Linked Movie Database (LinkedMDB) project providesa demonstration of the first open linked dataset connecting several major existing (and highly popular) movie web resources. The database exposed by LinkedMDB contains millions of RDF triples with hundreds of thousands of RDF links to existing web data sources that are part of the growing Linking Open Data cloud, as well as to popular movie-related web pages such as IMDb. Park, S. T., &Pennock, D. M. (2007) have built a prototype movie search engine called MAD6 (Movies, Actors and Directors; 6degrees of separation) and used collaborative filtering algorithm to generate personal item authorities for each user and combines them with item proximities for better ranking. Sejwal & Abulaish (2021) worked on designing a context driven domain ontology taking data from 1103 movies and sentiment analysis was done from different users as well as critics. Tomai, Eleni and et al (2005) addressed the issue of trip planning in the context of web services and investigated a algorithm with ontological expressions of tourism entities and geo-locations. Nešić, S., Rizzoli, A. E., &Athanasiadis, I. N. (2011) presented resource-publishing system, called AGROPub, which was developed to aid agro-environmental resource providers to annotate, publish and integrate their resources to LOD. The system comprises services and tools that enable resource providers to annotate their resources by relevant concepts from selected agro-environmental domain ontologies, to generate and publish RDF descriptions of the resources to LOD and to link the published resources to related resources from LOD. Bagosi, T. And et al. (2014) developed an OBDA framework to map relational database to RDF mapping language and tested the framework on Movie domain using OWL and Sesame API. Suganyakala, R., & Rajalaxmi, R. R. (2013) developed a framework for ontology based information retrieval system. In order to overcome the usability limitations, a query interface which required the user to enter the query in natural language is provided. A domain-specific ontology based on movies is used to develop a prototype of the proposed model which improves search accuracy. Several patents have been made for movie related recommendation system and data stream metadata. Ontology-based knowledge representation includes organizing information in a manner that allows robots to comprehend and logically analyze it. Typically, this involves establishing clear definitions, classifications, and connections within a specific field. Movie as information resource has enough opportunity to

map a huge range of data and information across domains. But, from literature review, it has been perceived that, no such work has been done with Indian cinema as well as from the perspective of LIS domain.

OBJECTIVES OF THE STUDY

- To identify the contributions of the actors in different roles in different movies
- To create a knowledge base of movies.
- To Build semantic relationships through object and data properties
- To create different relationships that can exist among the movies.
- To analyze the domain knowledge after defining declarative specification of terms collected from specific resources.
- To make association as well as dissociation between domain knowledge and operational knowledge.

Ontology Design

In this paper, a conceptual model of the Movie domain has been developed using "METHONTOLOGY" (Fernández et al.,1997)method. The developmental steps used here are presented as follows –

1. Specification – At this initial design phase, the specification phase makes the designers to define the scope, accumulating the conceptual facets and isolates from the domain of discourse and level of granularity of the description level. The conceptual model of "Movie" domain should fit as a describing feature of movie production phase, editing phase, distribution phase and the information description (just like bibliographic description)of 'Movie' crew. The crew members are – Director, Actor, Actress, Art Director, Hair Stylists, Dress Designer, Music Director, Production Manager etc.

Purpose – This conceptual model can be used to design a database for storing movie related information, coinage of standard terms, establish associations between related term showing semantic relations and also to identify identical terms and retrieval of cross-language terms.

Level of Formality – Semi-formal in nature. The level of formality in methontology is an important aspect of the field that contributes to its rigor and precision, and it helps to ensure the quality and reliability of the ontologies that are produced.

a) Short Film $\sqsubseteq$ Film $\sqcap$< 40 Minutes

b) Feature Film $\sqsubseteq$ Film $\sqcap$> 40 Minutes

c) Movie $\sqsubseteq$ Short Film $\sqcup$ Feature Film $\sqcup$ Animation $\sqcup$ Documentary

d) Film Director $\sqsubseteq$ $\exists$ Directs. Movie

e) Character $\sqsubseteq$ $\exists$ Movie(x). $\forall$Role Names(Y_1,Y_2,Y_3…,Y_i ,...Y_n) : i > 0

f) Actor $\sqsubseteq\exists$actedIn.Movie(x) $\sqcup\exists$hasRole.Character(y)

In the axiomatization of ontology, the concepts and relationships within the domain of knowledge are formalized using a formal language, such as first-order logic or description logics. The axioms used in the ontology describe the classes of entities in the domain, their properties, and the relationships between them. The axioms also specify constraints and conditions that must be satisfied in order for the ontology to be considered consistent and coherent.

Scope – Conceptual arrangement of related terms which can be categorised into several phases related movies' production and associated persons. The granularity of the knowledge organization should be deep enough to populate the qualitative information of the attributes at the instance level of movies. The ontology should be incremental by scope of further development. The persons related to movies can be categorised into different departments (Makeup department, Production Department, Music Department etc.).

2. Knowledge Acquisition - Though this step is independent of first step, but this step can be done in parallel with the specifying the scopes and formality. For this work, two separate types of sources have been used. IMDb(International Movie

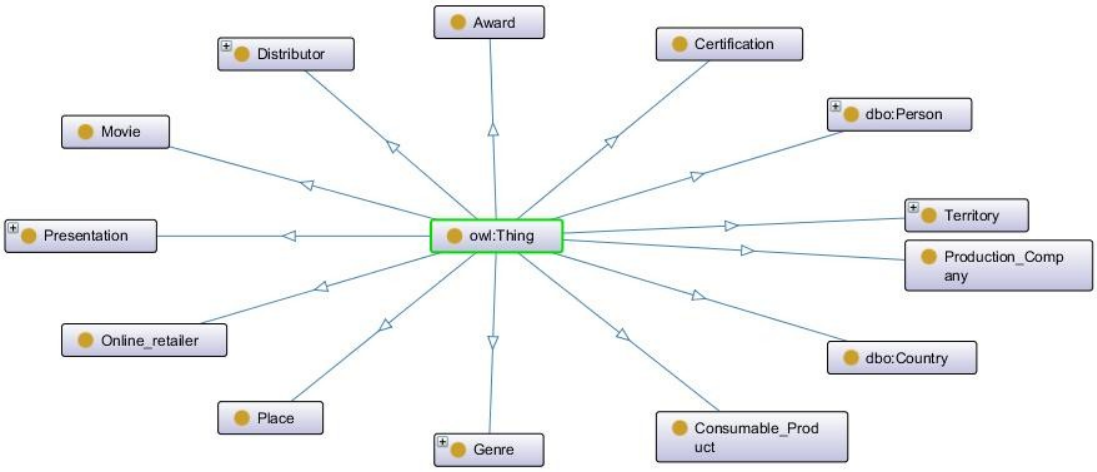

Figure:1 Ontology Classes of the Movie Ontology

Table:1 Class-Property-Instance or Attributes of Movie Domain

Class Attributes	Instance Attributes	Property Attributes
Persons	Instances of Actor, Director etc.	Movie belongs to genre
Award	Award names - Academy Awards, Golden Globe etc.	Movie is directed by directors
Genre	Action, Adult, Adventure, Romance etc.	Actors act in movie
Distributor	Name of Distributors for movies	A producer produces a movie
Country	Name of countries	A movie is released in a year
Movie	Movie Names	Movie production needs several departments e.g. Production, Makeup, Editing, Sound, Camera, Direction, Animation, Cinematography, Art etc.

Database) has developed one ontology which present the knowledge representation of movies which was helpful for storage of related information of the movies. Another source of information is the credit crawls which appear at the end of the movie which entails the role of staffs involved in all processes of a movie. The second source of information carries a variety of information snippets and many of them being ignored by standard movie databases. Another point of concern is the lack of usage of standard terms often poses ambiguity. Sometimes it so happens, over usage of non-standard terms becomes a standard one. Another direction is newer version of a term replaces the older terms. For example – Hair Dresser, Hair Stylist; Makeup Artist, Makeup

Stylists, Makeup Director etc. Incorporation of latest technologies into movie production also needs to be integrated within the existing knowledge storage and representation system. In movie domain, there is an absence of standard reference tools (e.g. glossary, dictionary, thesaurus etc).

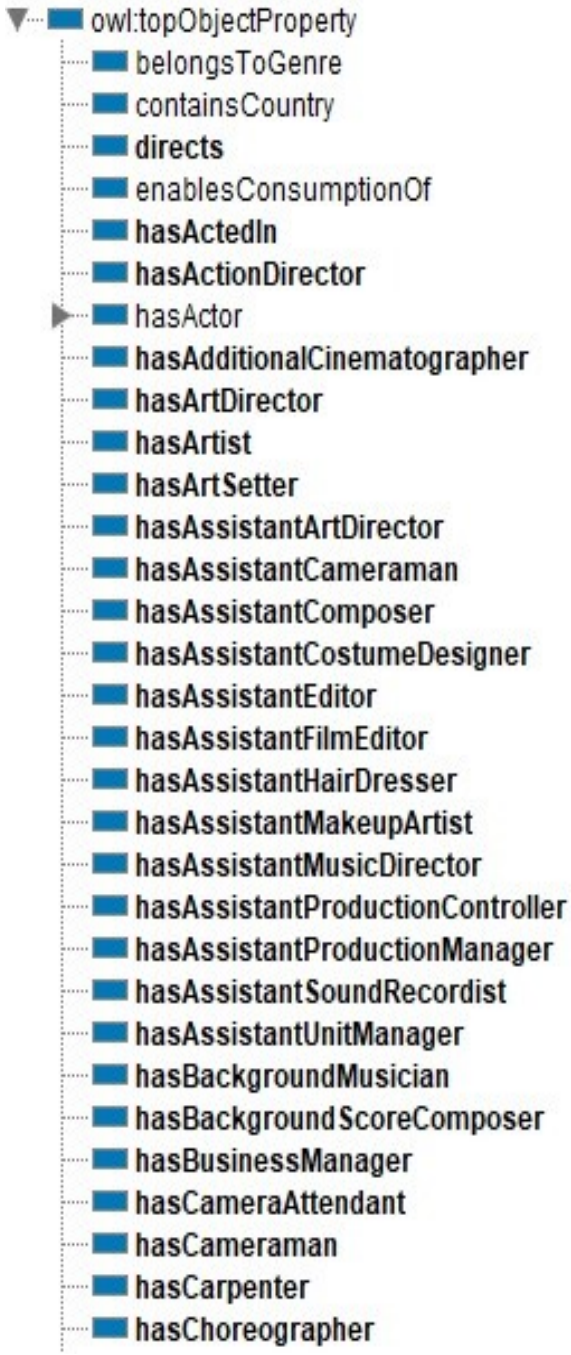

Figure 2: List of Object Properties mapping **d**omain and Range1

Though, there exist several popular websites dedicated to movies only(IMDb, TMDB, TheTVDB, letterboxd, Rotten Tomatoes).

Non-existence of standard terms increase the challenge of using a prevalent term to designate the semantics of a concept. In this step, all the collected

Figure 3: List of Object Properties mapping domain and Range1

Figure 4: List of Object Properties mapping Domain and Range3

Table 2: Object Properties of Movie Ontology and their defined Domain and Range

Property Name	**Domain**	**Range**
belongsToGenre	Movie	Genre
containsCountry	Territory	dbo:Country
Directs	Director	Movie
enablesConsumptionOf	Consumable_Product	Movie
hasActionDirector	Movie	Action_Director
hasActor	Movie	dbo:Actor

hasAdditionalCinematographer	Movie	Additional_Cinematographer
hasArtDirector	Movie	Art_Director
hasArtist	Movie	Artist
hasArtSetter	Movie	Art_Setter
hasAssistantArtDirector	Movie	Assistant_Art_Director
hasAssistantCameraman	Movie	Assistant_Cameraman
hasAssistantComposer	Movie	Assistant_Composer
hasAssistantCostumeDesigner	Movie	hasAssistantCostumeDesigner
hasAssistantEditor	Movie	Assistant_Editor
hasAssistantFilmEditor	Movie	Assistant_Film_Editor
hasAssistantHairDresser	Movie	Assistant_Hair_Dresser
hasAssistantMakeupArtist	Movie	Assistant_Makeup_Artist
hasAssistantMusicDirector	Movie	Assistant_Music_Director
hasAssistantProductionController	Movie	Assistant_Production_Controller
hasAssistantProductionManager	Movie	Assistant_Production_Manager
hasAssistantSoundRecordist	Movie	Assistant_Sound_Recordist
hasAssistantUnitManager	Movie	Assistant_Unit_Manager
hasBackgroundMusician	Movie	Background_Musician
hasBackgroundScoreComposer	Movie	Background_Score_Composer
hasBusinessManager	Movie	Business_Manager
hasCameraAttendant	Movie	Camera_Attendant
hasCameraman	Movie	Cameraman
hasCarpenter	Movie	Carpenter
hasChoreographer	Movie	Choreographer
hasCinematographer	Movie	Cinematographer
hasColor	Movie	Colorinfo
hasComposer	Movie	Composer

hasConceptDesigner	Movie	Concept_Designer
hasContactLensTechnician	Movie	Contact_Lens_Technician
hasCostumeCoordinator	Movie	Costume_Coordinator
hasCostumeDesigner	Movie	Costume_Designer
hasCraneOperator	Movie	Crane_Operator
hasDirector	Movie	Film_Director_Department
hasDraughtsman	Movie	Draughtsman
hasDressman	Movie	Dressman
hasEditor	Movie	Film_Editorial_Department
hasElectrician	Movie	Electrician
hasFilmEditor	Movie	Film_Editor
hasFilmLocation	Movie	Film_Location
hasFinanceManager	Movie	Finance_Manager
hasFoleyArtist	Movie	Foley_Artist
hasFoleyRecordingEngineer	Movie	Foley_Recording_Engineer
hasGenre	Movie	Genre
hasGraphicDesigner	Movie	Graphic_Designer
hasGreensman	Movie	Greensman
hasHairDresser	Movie	Hair Dresser ≡ Hair Stylist
hasHDCameraOperator	Movie	HD_Camera_Operator
hasLightAttendant	Movie	Light_Attendant
hasLyricist	Movie	Lyricist
hasMakeupArtist	Movie	Makeup_Artist
hasMakeupStylist	Movie	Makeup_Stylist
hasModeller	Movie	Modeller
hasMusicDirector	Movie	Music_Director
hasMusician	Movie	Musician

hasPainter	Movie	Painter
hasPlasterer	Movie	Plasterer
hasPostProductionAssociates	Movie	Post_Production_Associates
hasPostProductionHead	Movie	Post_Production_Head
hasProducer	Movie	Producer
hasProductionAssistant	Movie	Production Assistant
hasProductionController	Movie	Production Controller
hasProductionDesigner	Movie	Production Designer
hasProductionManager	Movie	Production Manager
hasProductionSupervisor	Movie	Production Supervisor
hasPropMaker	Movie	Prop Maker
hasProstheticMakeupArtist	Movie	Prosthetic Makeup Artist
hasProstheticMouldMaker	Movie	Prosthetic Mould Maker
hasReleaseTime	Movie	
hasReleasingCountry	Movie	dbo:Country
hasScreenplayer	Movie	Screenplayer
hasScriptWriter	Movie	Script Writer
hasSetDecorator	Movie	Set Decorator
hasSinger	Movie	Singer
hasSoundDesigner	Movie	Sound Designer
hasSoundmix	Movie	Sound Mix
hasSoundMixer	Movie	Sound Mixer
hasSoundMixingEngineer	Movie	Sound Mixing Engineer
hasSoundRecordist	Movie	Sound Recordist
hasStillPhotographer	Movie	Still Photographer
hasSupplier	Movie	Supplier
hasTitleDesigner	Movie	Title Designer

hasTranslatedMovie	dbo:Language	Movie
isTranslatedTo (inverse of hasTranslatedMovie)	Movie	dbo:Movie
hasTrolleyOperator	Movie	Trolley Operator
hasUnitProductionManager	Movie	Unit Production Manager
isActorIn	dbo:Actor	Movie
isAssistantArtDirectorOf	Assistant Art Director	Movie
isAwardedWith	Movie	Award
isCertifiedWith	Movie	Certification
isColorOf	Colorinfo	Movie
hasColor	Movie	Colorinfo
hasComposer	Movie	Composer
isComposerOf (inverse of hasComposer)	Composer	Movie

Table 3: Sample Competency Questions for Ontology Evaluation Process

Competency Question	Subject-Predicate-Object	Evaluative Decision
Q1.What are the types of genres of the movies?	Movie ->hasGenre -> Genre	Yes
Q2.What was the budget of any movie?	No such property defined	No
Q3.What is the box-office collection of a movie?	No such property defined	No
Q4.Who is the director of a movie?	Movie ->hasDirector -> Director	Yes
Q5.Who is the hair-stylist of a movie?	Movie ->hasHairStylist ->HairStylist	Yes
Q6.When was the movie released?	Movie ->wasReleasedIn -> Year	Yes
Q7.What were the members working in production department in a movie?	RDFS :subclassOfProduction_Department Production Manager Production Controller Production Designer	Yes

	Production Assistant Unit Production Manager Assistant Unit Manager Assistant Production Manager	
Q8.Who is the art director in a movie?	Movie ->hasArtDirector ->ArtDirector	Yes
Q9.What is the running time of a movie?	No such Property defined	No
Q10.What are the character names of the actors acting in a movie?	No such Property defined	No
Q11.Who is the choreographer of a movie?	Movie ->hasChoreographer -> Choreographer	Yes
Q12.What are those movies acted by an actor X?	Actor ->hasActedIn -> Movie	Yes
Q13.What is the quality of the movie?	No such Property defined	No
Q14.Name the movies which fall under Horror/War/Romantic/ Comedy category?	Movie ->hasGenre -> Genre	Yes
Q15.Who are the co-actors of the main actor in a movie?	Actor ->hasActedIn -> Movie	Yes
Q16.What are the short films existing in the knowledge base?	No such Property defined	

terms were categorized into classes and instances by analyzing their attributive features.

3. Integration

Integration step involves the search of an ontology which is in existence and investigate the possibility of using terms or organization of knowledge structure or using properties of the existing ones. In this study, another ontology designed by IMDb has been integrated into our ontological framework to use the standard terms defined by IMDb(http://www.movieontology.org/2009/10/01/movieontology.owl#).

4. Implementation

The ontology in discourse have been developed with Protege ontology editor.

5. Evaluation

After development of the ontology, the most important step is to evaluate them with competency questions. Competency question are the user-centric interrogation to justify the scope of the ontology. Competency questions are a set of queries that are designed to test the coverage, expressiveness, and reasoning capabilities of the ontology. These questions are formulated to test the ontology's ability to answer specific questions and to check if the ontology can handle complex reasoning tasks. The answers to these questions provide insight into strengths and weaknesses of the ontology in discourse, and can help identify areas that need improvement.

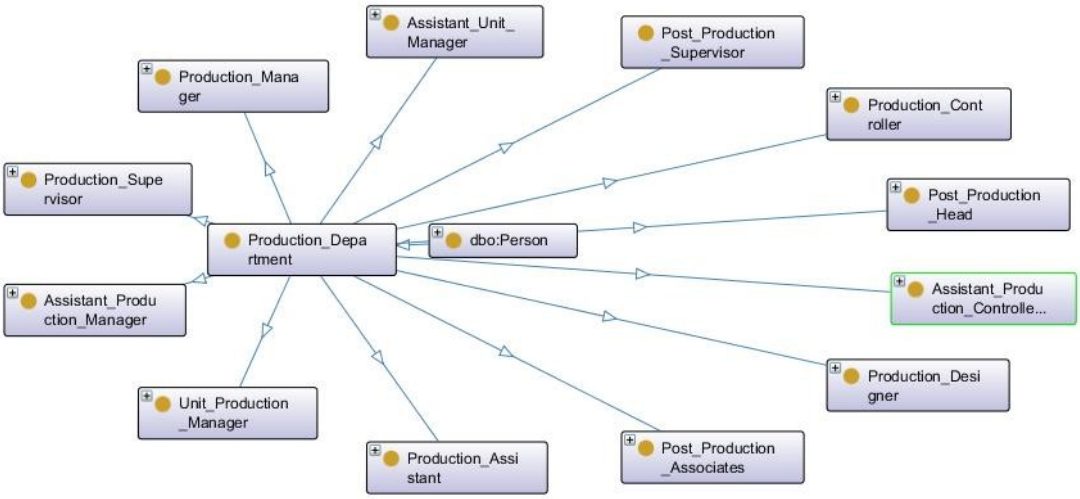

Figure 5: Underlying production associates of Production Department as retrieved from the query (Q7)

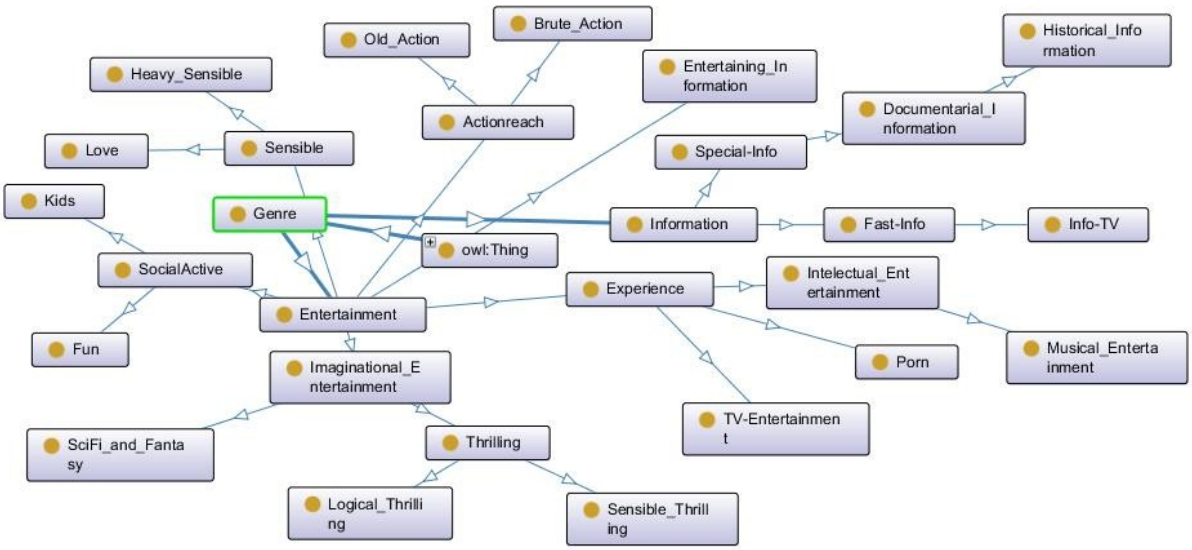

Figure 6: Different types of Genre categories of the movies as retrieved from query (Q1)

FINDINGS AND CONCLUSION

The structure of a movie typically consists of three parts:

1. The set-up or exposition: introduces the setting, characters, and conflict, setting the stage for the rest of the story.
2. The rising action: builds tension and conflict as the story progresses, leading to the climax.
3. The resolution or falling action: concludes the conflict and ties up loose ends, leading to the conclusion or ending of the story.

Some filmmakers also include additional elements such as a subplot, flashbacks, and foreshadowing to add depth and complexity to the story. However, the basic three-act structure remains the foundation of most movie narratives.From knowledge organization perspective, movie ontology has been

created by associating classes, properties and instances by obeying 3 part structure of the movie. The ontology has been developed with the help of Methontology by defining property roles among the entities and classes and specification of instances within classes. The ontology has been evaluated with competency questions. The open area of further development is to crosswalk among terms with cross language assertion of the defined terms within the knowledge base.

REFERENCES

1. Ahn, S., & Shi, C. K. (2009). Exploring movie recommendation system using cultural metadata. In *Transactions on Edutainment II* (pp. 119-134). Springer, Berlin, Heidelberg.
2. Bagosi, T., Calvanese, D., Hardi, J., Komla-Ebri, S., Lanti, D., Rezk, M., ... & Xiao, G. (2014, August). The ontop framework for ontology based data access. In *Chinese Semantic Web and Web Science Conference* (pp. 67-77). Springer, Berlin, Heidelberg.
3. Bogers, T. (2010, January). Movie recommendation using random walks over the contextual graph. In *Proceedings of the 2nd Intl. Workshop on Context-Aware Recommender Systems*.
4. Sejwal, V. K., & Abulaish, M. (2021). CAMO: A context-aware movie ontology generated from L OD and movie databases. *Multimedia Tools and Applications*, *80*(5), 7247-7269.
5. Fernández-López, M., Gómez-Pérez, A., &Juristo, N. (1997). Methontology: from ontological art towards ontological engineering.
6. Hassanzadeh, O., &Consens, M. P. (2009, April). Linked Movie Data Base. In *LDOW*.
7. Lugmayr, A., Niiranen, S., &Kalli, S. (2004). Digital interactive TV and metadata. *Signals and Communication Technology, Springer, New York, NY*.
8. Nešić, S., Rizzoli, A. E., &Athanasiadis, I. N. (2011, October). Publishing and linking semantically annotated agro-environmental resources to LOD with aGROPub. In *Research Conference on Metadata and Semantic Research* (pp. 478-488). Springer, Berlin, Heidelberg.
9. Park, S. T., &Pennock, D. M. (2007, August). Applying collaborative filtering techniques to movie search for better ranking and browsing. In *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 550-559). ACM.
10. Pilászy, I., &Tikk, D. (2009, October). Recommending new movies: even a few ratings are more valuable than metadata. In *Proceedings of the third ACM conference on Recommender systems* (pp. 93-100), ACM.
11. Suganyakala, R., &Rajalaxmi, R. R. (2013, February). Movie related information retrieval using ontology based semantic search. In *2013 International Conference on Information Communication and Embedded Systems (ICICES)* (pp. 421-424). IEEE.
12. Tomai, E., Spanaki, M., Prastacos, P., &Kavouras, M. (2005, October). Ontology assisted decision making–a case study in trip planning for tourism. In *OTM Confederated International Conferences" On the Move to Meaningful Internet Systems"* (pp. 1137-1146). Springer, Berlin, Heidelberg.