Domain Ontology and Character-Topic Relationship in Srimad-Bhagavatam

Apurba Paul and Dipankar Das

Abstract **The Srimad-Bhagavatam** is a foundational text in Hindu mythology, characterized by intricate narratives and profound philosophical teachings. This paper addresses the analytical challenges of the text by developing tailored ontologies: **`SBC-Ontology`** for characters and **`SBT-Ontology`** for topics, which together create a structured framework for deeper understanding. Advanced transformer models, including BERT, DistilBERT, and RoBERTa, were utilized to analyze character-topic relationships, with an ensemble approach(F1-score 0.79) demonstrating superior performance in capturing nuanced relationships. Despite the text's inherent complexities and limited data availability, this research significantly contributes to digital humanities, providing innovative tools for exploring and interpreting the rich narratives of **The Srimad-Bhagavatam**.

Keywords Ontology · Character-topic relationship · Srimad-bhagavatam · Knowledge representation · Transformer-based models

1 Introduction

Analyzing complex ancient texts like **The Srimad-Bhagavatam**[1] presents significant challenges due to their intricate narratives and profound philosophical concepts. As a cornerstone of Hindu philosophy, this text spans twelve cantos filled with narratives, hymns, and discourses on Lord Vishnu and his avatars.

[1] https://prabhupadabooks.com/sb.

A. Paul (✉)
Institute of Engineering and Management, University of Engineering and Management, Kolkata, India
e-mail: apurba.saitech2021@gmail.com

A. Paul · D. Das
Jadavpur University, Kolkata, India

J. K. Mandal et al. (eds.), *Recent Advances in Artificial Intelligence and Smart Applications*, Lecture Notes in Networks and Systems 1380,
https://doi.org/10.1007/978-981-96-5822-0_35

To tackle these challenges, this research focuses on two key methods: *domain ontology development* and *character-topic relationship* analysis. The Domain Ontology specifies concepts within the domain, identifying entities and revealing familial relationships, thematic links, and constraints [1]. Such ontologies provide valuable insights, aiding both human scholars and intelligent software agents in understanding the characters and themes in the text [2]. Meanwhile, Character-Topic relationship analysis employs computational techniques to uncover thematic connections between characters and topics, thereby enhancing scholarly comprehension and accessibility.

This study identifies critical gaps in existing ontologies for the Srimad-Bhagavatam, emphasizing the absence of comprehensive domain-specific datasets for Hindu mythology, which limits character representation and interaction analysis.

The absence of comprehensive domain-specific datasets for Srimad-Bhagavatam is the critical gap identified by this study. Precisely this gap limits character representation and interaction analysis of Hindu mythology. It also highlights the lack of robust character-topic datasets, impacting the accuracy of association analyses.

By identifying characters, relationships, and related subjects, the research seeks to create a thorough domain ontology that will improve comprehension of the text's philosophy and story. With the use of sophisticated transformer models like BERT, RoBERTa, and DistilBERT, it analyzes character-topic relationships.

Key contributions include the creation of specialized datasets for Character and Topic Ontologies, the development of a Character-Topic Relationship dataset for in-depth analysis, and the implementation of transformer models.

A visual representation of the research framework is demonstrated in Fig. 1. The figure highlights two main components of this research. The first component focuses

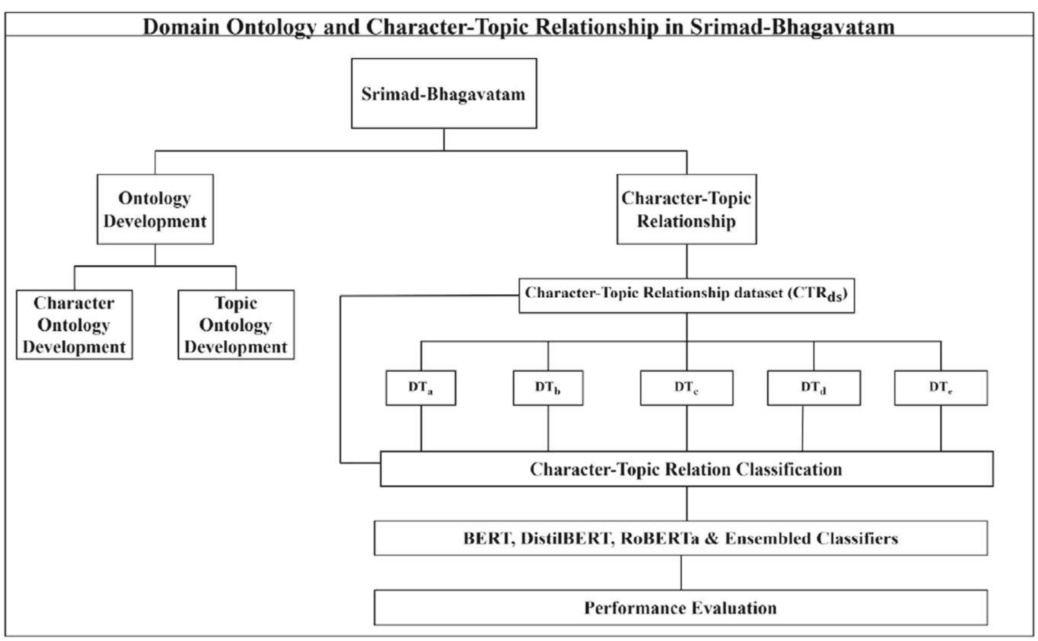

Fig. 1 Proposed research framework

on the development of domain ontologies, including Character Ontology and Topic Ontology. The second component involves utilizing the CTR_{ds} dataset to create five distinct sub-datasets (DT_a, DT_b, DT_c, DT_d, DT_e). The performance of various models, including BERT, DistilBERT, RoBERTa, and an ensemble method, is then evaluated using these sub-datasets along with CTR_{ds}.

The paper is organized as follows: it begins with the *Related Work*, followed by the *Dataset Preparation* and *System Framework* sections. This is followed by *Result Analysis*, *Error Analysis*, and *Discussion & Observations*. The paper concludes with the *Conclusion & Future Scope* section.

2 Related Work

Research on domain ontology and character-topic relationships in religious texts, particularly the Srimad-Bhagavatam, has garnered attention for enhancing understanding through structured frameworks. However, studies specifically addressing the ontological representation of characters and topics in these texts are limited. Varadarajan et al. [3] introduced GENOME, a methodology for improving the ontological modeling of epics. Syamili et al. [4] developed an ontology for Greek mythology, while Juan-Antonio [5] created the Ontology of Greek Mythology (OGM) to enhance semantic representation on the Semantic Web.

Watrobski [6] discussed ontology learning methods, and Espinoza-Arias et al. [7] examined smart city ontologies, emphasizing their lack of interoperability. Noy et al. [8] traced the evolution of ontologies, highlighting initiatives like RDF for knowledge encoding. Raja Mohan et al. [9] presented a Protege-based methodology for Indian medicinal plant ontology, while Hadaiq et al. [10] described an automated wayang ontology construction method.

Kulthida et al. [11] merged digital humanities and folktale studies in the Greater Mekong Subregion, while Sattar et al. [12] reviewed ontology methodologies in Interlocking Institutional Worlds. Franziska et al. [13] proposed a technique to organize myth variants using shallow ontologies. Ibukun et al. [14] developed a domain ontology for Nigerian history through a semi-automated approach, and Wolff et al. [15] suggested a method for rapid ontology construction using text mining.

Vlasenko et al. [16] developed a domain ontology for monitoring the condition of equipment in designing intelligent decision-support systems. Haridy et al. [17] proposed a tourism ontology development methodology beneficial for e-tourism. Pan et al. [18] presented a semi-automatic ontology development framework to build an energy management ontology and evaluated the data coverage.

3 Preparation of Datasets

Preparing datasets for studying domain ontology and character-topic relationships in the Srimad-Bhagavatam requires a structured approach to capture the thematic and narrative complexities of this ancient scripture. Paul et al. [19] developed an extensive dataset for the Pouranic topic classification challenge based on Indian mythology. From this dataset, we selected the Srimad-Bhagavatam for our domain ontology development and character-topic relationship tasks. The procedures for creating the datasets for these tasks are detailed in the following sections.

3.1 Datasets for Domain Ontology

We separated our work into two distinct subtasks for the domain ontology: character and topic ontology development, and we then prepared the dataset accordingly.

The text encompasses 1,714 characters, categorized into 17 types, including *God*, *Demigod*, *Human*, and *Animal*. The *God* category has four sub-categories: *Godhead*, *Goddess*, *Lord's-spiritual-energy*, and *Bhagavan*. The *Human* category includes 15 sub-categories, such as *King*, *Queen*, and *Sage*. In the *Animal* category, sub-types include *Elephant* and *Horse*. Examples of character classifications include: `Durga` ***rdf:type*** `Goddess`, `Narada` ***rdf:type*** `Devarsi`, and `Bharata` ***rdf:type*** `Maharaja`. The text identifies 24 family relation types, such as *father of* and *sister of*.

The Srimad-Bhagavatam is divided into 12 cantos. Each canto contains a varying number of chapters, some with fewer chapters and others with more extensive ones. These topics encompass stories, dialogs, teachings, and philosophical discussions. For example, the Second canto, titled *The Age of Deterioration*, contains 10 topics, out of which the first two topics are *The First Step in God Realization* and *The Lord in the Heart*.

3.2 Datasets for Character-Topic Relationship

We extracted sentences from The Srimad-Bhagavatam for each character and computed sentiment attributes-*polarity*, *subjectivity*, *positivity*, and *negativity*-using the *NaiveBayesAnalyzer* of TextBlob.[2]

These metrics were aggregated to reveal relationships and patterns among characters based on their appearances in various topics. The **PouranicTopic** dataset was utilized to tag sentences, resulting in the CTR_{ds} dataset, which contains 45,941 observations. For instance, Table 1 analyzes a sentence related to Ramacandra from

[2] https://textblob.readthedocs.io/en/dev/.

Table 1 Sample observation of CTR_{ds}

Character	Canto	Sentence	Pol	Subj	Pos	Neg	Senti	VP	Topic
Ramacandra	Liberation	All the brahmanas who were	1	0	0.62	0.37	pos	'brahmanas', 'engaged'	Lord Ramacandra Rules the World

Pol=Polarity; Subj=Subjectivity; Pos=Positivity; Neg=Negativity; Senti=Sentiment; VP=Verb or a particle attached to a verb

the Liberation section, noting a positive polarity (1), objective subjectivity (0), a positivity score of 0.62, and a negativity score of 0.37. The associated verbs are *brahmanas* and *engaged*, with the topic being Lord Ramacandra Rules the World. We constructed five datasets (DT_a, DT_b, DT_c, DT_d, DT_e) from the CTR_{ds} to evaluate various features on character-topic relationships. Each dataset includes different features, with the classifiers BERT, DistilBERT, RoBERTa, and an ensemble model tested on them.

Table 2 summarizes the CTR_{ds} dataset, highlighting the range of Polarity scores (–1 to +1), Subjectivity (0 to 1), Positivity (0.002 to 1), and Negativity (0.002 to 0.99). The dataset comprises 38,296 positive and 7,645 negative sentences. Notable topics include *Gajendras Prayers of Surrender*, *The Killing of the Demon Trnavarta*, *The Descendants of Ajamidha*, *The Gopis Songs of Separation*, and *Five Queens Married by Krsna*.

Table 2 Descriptive statistics of the CTR_{ds} dataset

Pol range	Subj range	Pos range	Neg range	Total pos	Total neg	Frequent top 5 topics
–1 to +1	0 to 1	0.002 to 1	0.002 to 0.99	38296	7645	Gajendras Prayers of Surrender The killing of the demon Trnavarta The descendants of Ajamidha The gopis songs of separation Five queens married by Krsna The liberation of king Jarasandha The sages teachings at Kuruksetra

Pol range: Polarity score range; Subj range: Subjectivity score range; Pos range: Positivity range; Neg range: Negativity score range; Total pos/neg: Total positively/negatively classified sentences

4 System Framework

This section outlines the system framework for exploring domain ontologies, focusing on character and topic ontologies, along with character-topic relationships in The Srimad-Bhagavatam. The framework consists of two main components: Domain Ontology and Character-Topic Relationship.

4.1 Domain Ontology Development

The Character and Topic ontologies for The Srimad-Bhagavatam were developed following Stanford University's guidelines [8]. The Character ontology, referred to as SBC-Ontology, and the Topic ontology, called SBT-Ontology, are illustrated in Fig. 2.

To clarify motivation, competency questions in natural language define the scope and coverage of both ontologies, ensuring they address all relevant queries. Examples include: *Who is Abhimanyu? Who is Ramacandra? What is Srimad-Bhagavatam? What topics are available in each canto of Srimad-Bhagavatam?*

In the case of existing ontology, we found mythological ontologies, notably the Mahabharata ontology[3] from IIT Madras.

Following research by [20, 21], we extracted and categorized characters and topics in The Srimad-Bhagavatam for Knowledge Acquisition and Glossary of Terms. The SBC-Ontology includes characters, genders, and relationships, while SBT-Ontology addresses cantos, topics, and hierarchies.

For classes and hierarchy, SBC-Ontology categorizes terms into classes such as Animal, Celestial-Beings, and God, while SBT-Ontology includes Canto and Topic classes. For a detailed illustration of these classifications, refer to Fig. 3.

The key properties in SBC-Ontology include *hasFather* and *isFatherOf*, while SBT-Ontology features *hasSentence* and *isBelongsToCanto* (Fig. 3).

The Knowledge Production phase integrates data to enhance semantic depth, with *MythoDomainEntity* as the top-level class in SBC-Ontology, defined via the *subClassOf* property.

The SBC-Ontology employs annotation and object properties to represent complex relationships, including lineage and parent-child connections. It comprises 3,530 axioms (2,551 logical), defines 86 classes, and includes 797 individuals, showcasing a detailed semantic structure. The SBT-Ontology organizes concepts into classes like Mythology and Canto, linking them through properties such as *hasTopic*. It contains 1,457 axioms (1,258 logical), 8 classes, and 451 individuals.

The structure of the ontologies is validated by the domain experts using the popular tool named Protégé[4] and finally populated the ontology.

[3] https://sites.google.com/site/ontoworks/about-us.

[4] https://protege.stanford.edu/.

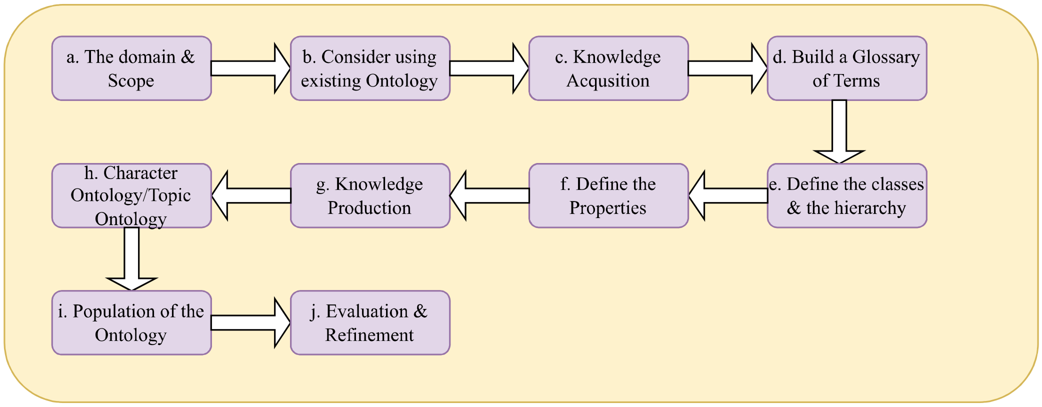

Fig. 2 Proposed model framework for ontology development

SBC-Ontology	
Class Hierarchy	**Object Properties**
owl: Thing	owl : topObjectProperties
MythoDomainEntity	hasWife
Hill	hasHusband
Hindu-Mythological-Characters	hasFather
Animal	hasMother
Celestial-Beings	isWifeOf
Deity	isHusbandOf
Devarsi	isFatherOf
Devotee	isMotherOf
Emperor	isDevoteeOf
SBT-Ontology	
Class Hierarchy	**Object Properties**
owl: Thing	owl : topObjectProperties
Canto	hasSentence
Mythology	hasTopic
Person	isBelongsToCanto
Sentence	isBelongsToMythology
Topic	isBelongsToTopic

Fig. 3 Class hierarchy and object properties of **`SBC`** and **`SBT-Ontology`**

The Ontology evaluation assesses quality and correctness [22], with experts manually reviewing the ontologies for accuracy.

4.2 *Character-Topic Relationship*

Understanding the relationship between characters and topics in The Srimad-Bhagavatam involves analyzing how characters engage with and contribute to various topics throughout the text. We first identified the top 5 characters and topics by frequency to highlight their significance in The Srimad-Bhagavatam. Table 3 illustrates

Table 3 Top 5 frequent characters and topics

Characters	Freq.	Topics	Freq.
Krsna	6445	Krsna, the supreme personality	792
Supreme personality of godhead	1574	Descendants of Ajamidha	641
Balarama	1017	Advent of lord Krsna	531
Vasu	956	Diti vows to kill Indra	509
Brahma	951	Pastimes of lord Ramacandra	421

that Krsna is the most frequently mentioned character, while *Krsna, the Supreme Personality* is the most common topic.

Topic distribution task aims to identify the most frequent topics linked to specific characters. Understanding how topics are distributed across the text allows us to see which topics are most relevant to particular characters. The character ***Krsna*** appears most frequently in the following top 5 topics: *Krsna, the Supreme Personality of Godhead* (106 times), *The Descendants of Ajamidha* (90 times), *The Advent of Lord Krsna-Introduction* (85 times), *Five Queens Married by Krsna* (74 times), and *The Story of the Syamantaka Jewel* (66 times).

Now we will discuss various classification models for the Character-Topic Relationship. The Character-Topic Relationship Model leverages advanced natural language processing techniques to analyze and predict relationships between characters and topics in mythological texts. We utilized several SOTA models: BERT [23] for contextual understanding, DistilBERT[24] for faster training, and RoBERTa [25] for improved pretraining. Each model was optimized with the standard configurations The evaluation of the models was performed using categorical cross-entropy loss along with the Adam optimizer. To enhance performance, we implemented an ensemble classifier that combines outputs from BERT, DistilBERT, and RoBERTa through majority voting, leveraging individual strengths to improve overall Precision, Recall, and F1-score.

5 Result Analysis

Table 4 provides a detailed comparison of performance metrics across different models on various datasets. The models evaluated include BERT, DistilBERT, RoBERTa, and an Ensemble approach. Each row in the table represents a different dataset, denoted as DT_a, DT_b, DT_c, DT_d, DT_e, and CTR_{ds}.

The BERT classifier demonstrates stable performance, with Precision (Pr) scores varying between 0.65 and 0.70 across datasets, and F1-Scores (F1s) ranging from 0.64 to 0.67. DistilBERT outperforms BERT, with Pr scores from 0.70 to 0.75 and F1 scores (F1s) between 0.69 and 0.73, indicating robust performance. RoBERTa yields similar results, with Precision (Pr) scores ranging from 0.67 to 0.72 and F1-Scores

Table 4 Models performance across all datasets

	BERT			DistilBERT			RoBERTa			Ensemble		
	Pr	Rc	F1s	Pr	Rc	F1s	Pr	Rc	F1s	Pr	Rc	F1s
DT_a	0.70	0.65	0.67	0.75	0.70	0.72	0.72	0.68	0.70	**0.80**	**0.78**	**0.79**
DT_b	0.68	0.64	0.66	0.73	0.69	0.71	0.70	0.66	0.68	**0.78**	**0.76**	**0.77**
DT_c	0.69	0.67	0.68	0.74	0.71	0.73	0.71	0.69	0.70	**0.79**	**0.77**	**0.78**
DT_d	0.66	0.63	0.64	0.71	0.68	0.69	0.68	0.65	0.66	**0.76**	**0.74**	**0.75**
DT_e	0.67	0.64	0.65	0.72	0.69	0.70	0.69	0.66	0.67	**0.77**	**0.75**	**0.76**
CTR_{ds}	0.65	0.63	0.64	0.70	0.68	0.69	0.67	0.65	0.66	0.75	0.73	0.74

$DT_a = \{Character, Canto, Sentence, VP\}$
$DT_b = \{Character, Canto, Sentence, Pol, Senti, VP\}$
$DT_c = \{Character, Sentence, VP\}$
$DT_d = \{Character, Canto, VP\}$
$DT_e = \{Character, Canto, Sentence, Pos, Senti, VP\}$
$CTR_{ds} = \{Character, Canto, Sentence, Pol, Subj, Pos, Neg, Senti, VP\}$

(F1s) between 0.66 and 0.70. The Ensemble approach excels across all datasets, achieving the highest metrics with Pr scores from 0.75 to 0.80 and F1-Scores (F1s) between 0.74 and 0.79, particularly standing out in the CTR_{ds} dataset.The Ensemble model consistently outperforms the individual models across all datasets. The highest performance is observed with DT_a, which includes the features Character, Canto, Sentence, and VP, achieving the best F1-Scores (F1s) across the board, particularly when using the Ensemble approach. DT_a emerges as the most effective dataset, while the Ensemble model is identified as the best-performing model.

6 Error Analysis

We analyzed the maximum number of misclassified topics across all models. The topic *The Descendants of Ajamidha* had the highest misclassification with 56 errors, often incorrectly predicted as *Krsna, the Supreme Personality of Godhead*, *Lord Parasurama Destroys the Worlds Ruling Class*, or *Parasurama, the Lord's Warrior Incarnation*. The topic *Diti Vows to Kill King Indra* followed with 44 misclassifications, frequently confused with *The Pastimes of the Supreme Lord, Ramacandra*, *Krsna, the Supreme Personality of Godhead*, or The Advent of Lord Krsna-Introduction. Additionally, *Krsna, the Supreme Personality of Godhead*" faced 42 misclassifications, commonly confused with *Diti Vows to Kill King Indra*, *The Dynasties of the Sons of Yayati*, or *The Descendants of Ajamidha*. The topic *Lord Krsna Shows the Universal Form Within His Mouth* had 40 misclassifications, often predicted as *Diti Vows to Kill King Indra*, *Krsna, the Supreme Personality of Godhead*, or *The Advent of Lord Krsna-Introduction*. Lastly, *The Killing of the Demon Putana* was misclassified 38 times, with predictions including *Krsna, the Supreme Personality of Godhead*, *Diti Vows to Kill King Indra*, or *The Descendants of Ajamidha*.

7 Discussion and Observations

The paper presents a framework for analyzing character-topic relationships in the Srimad-Bhagavatam, a key text in Hindu mythology. It introduces two core ontologies: SBC-Ontology for categorizing 1,714 characters into 17 groups (e.g., Gods, Humans, Animals) and identifying 24 family relations, and SBT-Ontology which organizes the text into 12 Cantos, mapping character interactions within specific topics.

The study utilizes a Character-Topic Relationship dataset (CTR_{ds}) comprising 45,941 observations. Transformer models such as BERT, DistilBERT, and RoBERTa are employed, with an ensemble model achieving the highest F1-score of 0.79. To minimize the error rate, a domain-specific transformer-based embedding model can be introduced in future work. The analysis also highlights Krsna as the most mentioned character and establishes a strong correlation, indicating that characters associated with positive topics have minimal negative interactions.

8 Conclusion and Future Scope

This research offers a structured approach to understanding the complex narratives and themes in the Srimad-Bhagavatam by developing domain-specific ontologies. This study provides a detailed framework for categorizing the vast array of characters, their relationships, and the topics they engage with across the scripture. After that using the advanced transformer models like BERT, DistilBERT, and RoBERTa, this research analyzes character-topic relationships. With an F1-score of 0.79, the ensemble model performed better than the others, indicating that it could capture intricate interactions between topics and characters.

This research opens avenues for further exploration, including expanding the dataset to include other significant Hindu texts, such as the Mahabharata and the Ramayana, to develop a more extensive ontology and to find better relationships between characters and topics. Integrating these ontologies and character-topic relationships with global knowledge graphs like DBpedia or Wikidata could enhance model interoperability beyond Hindu mythology.

References

1. Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition, 5*(2), 199–220. https://doi.org/10.1006/knac.1993.1008
2. Jiang, L. P. (2012). Designing family ontology with the Protégé OWL plugin. In *Materials Science and Information Technology II* (pp. 836–840). Trans Tech Publications Ltd. https://doi.org/10.4028/www.scientific.net/AMR.532-533.836

3. Varadarajan, U., Bagchi, M., Tiwari, A., & Satija, M. P. (2022). GENOME: A GENeric methodology for ontological modelling of epics. ArXiv. https://api.semanticscholar.org/CorpusID:247158154
4. Syamili, C., & Rekha, R. (2017). Developing an ontology for Greek mythology. *The Electronic Library 36*, 119–132. https://api.semanticscholar.org/CorpusID:31343095
5. Pastor-Sánchez, J.-A., Kontopoulos, E., Saorín, T., Bebis, T., & Darányi, S. (2021). Greek mythology as a knowledge graph: From Chaos to Zeus and beyond. ArXiv https://api.semanticscholar.org/CorpusID:235657768
6. Watrobski, J. (2020). Ontology learning methods from text–an extensive knowledge-based approach. In *International Conference on Knowledge-Based Intelligent Information & Engineering Systems*. https://api.semanticscholar.org/CorpusID:226255735
7. Espinoza-Arias, P., Poveda-Villalón, M., García-Castro, R., & Corcho, Ó. (2018). Ontological representation of smart city data: From devices to cities. In *Applied Sciences*. https://api.semanticscholar.org/CorpusID:116614841
8. Noy, N. F., & McGuinness, D. L. (2001). Ontology development 101: A guide to creating your first ontology. In *Stanford Knowledge Systems Laboratory Technical Report KSL-01-05*.
9. Mohan, A. R., & Arumugam, G. (2010). Developing Indian medicinal plant ontology using OWL and SWRL. In *International Conference on Data Engineering and Management*. https://api.semanticscholar.org/CorpusID:21166742
10. Sanabila, H. R., & Manurung, R. (2014). Automatic wayang ontology construction using relation extraction from free text. In *LaTeCH@EACL*. https://api.semanticscholar.org/CorpusID:16152128
11. Tuamsuk, K., Chansanam, W., & Kaewboonma, N. (2018). Ontology of folktales in the Greater Mekong Subregion. *International Journal of Metadata, Semantics and Ontologies*, *13*, 57–67. https://api.semanticscholar.org/CorpusID:67868859
12. Sattar, A., Salwana, E., Nazir, M., & Ahmad, M. (2020). Comparative analysis of methodologies for domain ontology development: A systematic review. In *International Journal of Advanced Computer Science and Applications*. https://api.semanticscholar.org/CorpusID:222176816
13. Pannach, F., & Blaschke, T. (2023). Modeling and comparison of narrative domains with shallow ontologies. In *International Conference on Language, Data, and Knowledge*. https://api.semanticscholar.org/CorpusID:264439793
14. Afolabi, I. T., Daramola, O. J., & Adio, T. A. (2014). *Developing domain ontology for Nigerian history*. https://api.semanticscholar.org/CorpusID:55595742
15. Gersberg, I. W., & Ebecken, N. F. F. (2014). A simple strategy to start domain ontology from scratch. *International Journal of Advanced Computer Science and Applications*. https://api.semanticscholar.org/CorpusID:18753697
16. Vlasenko, L. O., Lutska, N. M., Zaiets, N. A., Shyshak, A. V., & Savchuk, O. V. (2022). Domain ontology development for condition monitoring system of industrial control equipment and devices. *Radio Electronics, Computer Science, Control*. https://api.semanticscholar.org/CorpusID:248456848
17. Haridy, S., Ismail, R. M., Badr, N. L., & Hashem, M. (2023). An ontology development methodology based on ontology-driven conceptual modeling and natural language processing: Tourism case study. *Big Data and Cognitive Computing*, *7*(101). https://api.semanticscholar.org/CorpusID:258855924
18. Pan, Z., Gao, Y., Ponci, F., & Monti, A. (2023). Semi-automatic ontology development framework for building energy data management. *IEEE Access, 11*, 111991–112003. https://api.semanticscholar.org/CorpusID:264069350
19. Paul, A., Seal, S., & Das, D. (2024). Transformer-based pouranic topic classification in Indian mythology. S$\bar{a}$dhan$\bar{a}$. https://doi.org/10.1007/s12046-024-02555-3
20. Paul, A., & Das, D. (2017). Identification of character adjectives from Mahabharata. In *Proceedings of the International Conference Recent Advances in Natural Language Processing, RANLP 2017* (pp. 569–576). https://doi.org/10.26615/978-954-452-049-6_074
21. Paul, A., & Das, D. (2017). A deep dive into identification of characters from Mahabharata. In *Proceedings of the 14th International Conference on Natural Language Processing (ICON-2017)* (pp. 447–455). https://aclanthology.org/W17-7555

22. Hlomani, H., & Stacey, D. A. (2014). *Approaches, methods, metrics, measures, and subjectivity in ontology evaluation: A survey*. https://api.semanticscholar.org/CorpusID:51371006
23. Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)* (pp. 4171–4186).https://doi.org/10.18653/v1/N19-1423
24. Sanh, V., Debut, L., Chaumond, J., & Wolf, T. (2019). DistilBERT, a distilled version of BERT: Smaller, faster, cheaper and lighter. ArXiv. https://api.semanticscholar.org/CorpusID:203626972
25. Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). RoBERTa: A robustly optimized BERT pretraining approach. ArXiv. https://api.semanticscholar.org/CorpusID:198953378