TamilInfoG: A Strategic Framework for Tamil Knowledge Centric Infographics recommendation using Generative Semantic Intelligence

Prakhar Langer, Gerard Deepak
School of Computer Engineering, Manipal Institute of Technology Bengaluru, Manipal Academy of Higher Education, Manipal, Karnataka - 575 104, India.
Dayanand Sagar Academy of Technology and Management
Email IDs: *prakhar.mitblr2022@learner.manipal.edu, gerard.deepak.christuni@gmail.com*

***Abstract*—Tamil knowledge-based infographics recommendation is offered in a frame of strategies suggested in this paper. The hybridity frame integrates the Tamil-to-English knowledge generation and the English-to-Tamil knowledge translation and then it transforms this knowledge into increment of knowledge. Various views of the dataset are adopted, whereby firstly the dataset is classified with BERT and Tamil BERT to come out with classified instances. Nilamani and labels are then extracted in English and in Tamil. The Tamil terms are fed through a Tamil large language model constructed on Google Gemma 7B and English terms are translated into Tamil by using application programming rivets and fed into the model. The framework also uses tailor made Web 3.0 crawled infographics and these types of infographics consist of infographic images with Tamil text and labels added to them hence facilitating retention and reflectiveness of central knowledge. Queries are received in Tamil and English to take care of multilingual access and bilingual situational Latent Semantic Indexing (LSI) model is used. The Resulting terms in the LLLs of both languages are then translated into the Tamil language using a cross-lingual entity-linking API (Linked Open Data API) and they are, in turn incorporated into the structure. In the case of incremental knowledge aggregation, the framework uses the quantitative semantic reasoning. Particularly, adaptive pointwise mutual information measures, the Lance-and-Williams index and the index of Lloyd are employed at different stages of the pipeline. Differentials are used in setting thresholds and as much as it is necessary step-deviance measures are used. In order to have the best results a spiral optimization strategy simulating the processes in nature is undertaken to find the best set of solutions. Furthermore, the contributions of Google, Knowledge Graph API, instances of categorized dataset, and context-feedback of entities provided by the Web 3.0 are also considered in the framework. Tamil annotations are enriched with the help of relevant labels and terms of these sources that are collectively aggregated with the help of the corresponding quantitative reasoning schemes. These augmented annotations are then accounted to particular tests and then used to spur the process of a recommendation. Lastly, the suggested framework is performing better in terms of Tamil infographics recommendation, with a precision of 97.89%, an FDR of 0.03, and an F-measure of 98.42%, which proves its effectiveness in producing the knowledge graph reliant on the semantic web in a context of the knowledge of the Tamils.**

***Index Terms*—Tamil BERT, Knowledge Graph, Semantic Web, Web 3.0, Infographics Recommendation, Cross-lingual Entity Linking, Quantitative Semantic Reasoning, Spiral Optimization Algorithm, Natural Language Processing (NLP)**

I. Introduction

The swift shift in Web 2.0 to Web 3.0 has incorporated a sea change in the expression, conveying, and promoting information. Web 3.0 focuses on semantic intelligence, personalization, as well as knowledge based presentation of content and as such recommendation systems are required that not only recalls the content but also contextualizes it and tailors it to different linguistic-cultural contexts. As a form of visual communication, infographics also play a special role in the context since the medium incorporates formatted knowledge, images, and written word to pass more complex information efficiently. Systems that can cleverly suggest knowledge-based infographics are becoming essential in multilingual cultures like in India that still require regional languages to be useful in engaging digitally. In this changing environment, Tamil infographics recommendation becomes highly significant. Tamil is a classical language that has a rich literary and cultural heritage and has had a wider use in digital ecosystems. Recent recommendation platforms are, however, not as deep semantically and cannot generate as well as they would have to support Web 3.0 scenarios that are Tamil specific. The drive to this study is that it resolved this shortcoming by implementing technologies of Generative Semantic Intelligence fused with Tamil large-language models, BERT challenges categorization, and quantitative semantic reasoning. The integration of statistical and semantic similarity-based algorithms will allow incorporating knowledge in an incremental way with accuracy and favor. Since the need to maintain a strategic and AI-oriented, and linguistically pluralistic recommendation system is a long-standing challenge in question, the developed framework represents a rather timely solution that will comply with the overarching necessity to prioritize Indian and regional languages in the Web 3.0 era. The contributions of the framework are varied. It solely uses the dataset per se as its input in classification via BERT and then incorporates the classified ones into the end solution set to add domain richness. In addition to that, the system will provide bilingual search schemes of Tamil and English and therefore facilitate indexing of semantics of multilingual query entries. A tailored crawler is used for gathering enriched Tamil infographics with related text and labels, which form an initial solution space subsequently further enriched using feedback. The model also incorporates a Tamil LLM based on Google's Gemma 7B to facilitate text summarization, captioning, and the addition of high-density auxiliary knowledge in Tamil. To enhance semantic reasoning, the system employs Lloyd's index, the Lance and Williams index, and adaptive control with empirically set step-deviance thresholds. Lloyd's index also serves as an objective criterion for the spiral optimization algorithm, a nature- inspired meta-heuristic for guaranteeing optimum solution generation. Together, these contributions set the framework as being the first of its type in driving Tamil knowledge-based infographics recommendation powered by semantic intelligence for Web 3.0.

II. Motivation

The motivation to suggest this framework stems from the increasing demand for knowledge-based schemes in Tamil infographics recommendation within the realm of Web 3.0. The suggestion of Tamil infographics implies strong impacts in the long term in this dynamic digital ecosystem. The framework offered can respond to this need by relying on Generative Semantic Intelligence and is implemented with the help of the combination of Tamil language model with BERT architectures and the support of quantitative semantic reasoning. It is a synthesis methodology that brings together statistical methodologies with semantic similarity-based techniques to facilitate good precision and relevance. With the internet becoming Web 3.0-based rather than Web 2.0-based, the demand to have such strategic, semantically oriented and generative artificial intelligence methods which can support advanced-level recommendation systems is growing more imperative. Specifically, the increased attention to Indian and other regional language has gained traction, which is why the solutions that are language-based are not only important but also necessary. Based on this, the framework, as described below, furthers the proposal on Tamil infographics in tandem with the overall purpose of Web 3.0.

III. Contribution

The framework proposed in this study offers a number of innovative contributions to the field of infographic recommendation that is knowledge-centric, i.e. revolves around Tamil, in the Web 3.0 environment. To achieve domain richness, the framework directly uses the dataset and classifies it through the main BERT library and finally consolidates the instances that have been classified as part of the final solution set. The framework also absorbs the bilingual retrieval systems in the English language and the Tamil language. This method is unique in its application of guided language model in the course of preprocessing and hence eases the indexing of semantics on bilingual inputs during the querying stage. In addition, improved Tamil pictures and visual information with Tamil writing and labels are gathered through a specially created crawler. The retrieved entities form a starting solution, which is then stored and enriched within a feedback loop that is undertaken by the entire solution set hence improving quality of recommendations. Another important strength is the embedding of a Tamil large-language model on the platform of Google Gemma 7B that are used to produce summaries of text in Tamil, captions extractors, and fine-tune the inclusion of high-density auxiliary knowledge into the framework. Furthermore, semantic-oriented reasoning is performed using Lloyd's index, the Lance and Williams index, and adaptive measures with differential step- deviance thresholds designed for varied situations, each empirically determined. Lloyd's index is further used as an objective function within the spiral optimization algorithm, which serves as a strategic nature-inspired meta-heuristic to yield the optimal solution set. Col- lectively, these contributions make the framework not only novel but also the first of its kind in advancing Tamil knowledge-centric infographics recommendation with semantic intelligence for Web 3.0.

IV. Organization

The subsequent part of the paper is structured as follows: The second portion offers a brief overview of the related work. The third part explains the architecture of the suggested system. The fourth segment mainly concentrates on the execution and efficacy evaluation of the technique. The final part presents the conclusions inferred from the study.

V. Literature Review

Literature on semantic intelligence and knowledge-centric recommendation has grown significantly with the emergence of Web 3.0 and generative AI. Hemashree et al. [1] discussed a systematic literature review on tagging multimedia content in Web 3.0 based on semantic artificial intelligence, with a special emphasis on the growing importance of semantic reasoning in multimedia indexing and recommendation. Deepak et al. [2] pushed this idea further with OntoInfoG++, a semantic fusion technique specifically tailored for infographics recommendation, showing the worth of ontology-based knowledge fusion for higher relevance. Opoku [3] examined infographic design problems in the digital age, emphasizing the role of contextual and user-oriented design factors that have a direct impact on the effectiveness of recommendation systems. In a similar vein, Prathibha and Deepak [4] introduced StrategicVideoRec, combining BERT and fact-based semantics for video recommendation, drawing parallels between infographic recommendation and semantic video pipelines. Aljamel et al. [5] followed this path by introducing domain knowledge-oriented evaluation methodology for intelligent information retrieval, focusing on knowledge-aware indexing as a means to increased precision.

From a more general semantic web point of view, Chaudhuri et al. [6] envisioned knowledge-centric methodologies in robotics, highlighting semantic reasoning's cross-domain nature and ontology learning. Al Khatib et al. [7] performed a wide-ranging survey for patient-centric knowledge graphs, pointing out challenges and methods for domain-specific graph building, which mirrors the activity of developing Tamil-centric knowledge bases. Guo et al. [8] gave a baseline overview of knowledge graph-based recommender systems, outlining architectures utilizing semantic fusion and embedding strategies. Zhou et al. [9] followed by enhancing conversational recommender systems based on knowledge graph-based semantic fusion, while Lully et al. [10] highlighted the importance of improving explanation in recommendations using graph-based reasoning. Wang et al. [11] also added by providing knowledge graph convolutional networks, providing a connection between graph learning and recommendation precision.

The significance of ontology and knowledge maps has also been underlined in education and content provision. Pei et al. [12] provided ontology-based construction of curriculum knowledge maps, showing how semantic enrichment supports recommendation pipelines. Deldjoo et al. [13] and Zhang et al. [14] moved in the direction of generative AI, presenting how generative models and agents redefine context and personalization in recommender systems. Hong et al. [15] presented a survey of multi-objective recommendation during the generative AI era, describing new approaches for balancing accuracy, diversity, and fairness. Nawara and Kashef [16] expanded this by thoroughly surveying LLM-based recommender systems, from discriminative to multimodal frameworks.

In the general semantic web environment, Hassan and Rashid [17] surveyed AI-based natural language processing and ontology learning algorithms, illustrating the manner in which semantic intelligence facilitates strong knowledge modeling. At the optimization level, Karandikar et al. [18], Maureira et al. [19], and Naghavipour et al. [20] also focused on hybrid metaheuristics and adaptive approaches for knowledge mapping and service composition, further verifying the tendency to interweave metaheuristic optimization into semantic frameworks. Lastly, benchmark datasets like USAR [21], RTAI [22], and TVAI [23] also offer empirical basis for recommender system testing, allowing standardized comparisons in models. All these studies collectively provide a robust basis for the suggested framework, which combines Tamil-centric language models, bilingual semantic indexing, knowledge graph augmentation, and optimization-based reasoning to push forward infographics recommendation under the Web 3.0 model.

VI. Proposed System Architecture

The structure of the proposed Tamil knowledge-based infographic recommendation system based on the hybridized semantic system of intelligence is shown in Fig.1. The system is designed to work on both user-based searches and a curated list of Tamil-influenced images, which is a source of foundation. Such images can have Tamil inscriptions or writing in English as it is widely seen in Web 3.0 repositories whereby the metadata is commonly in English. The framework is designed to elaborate on this attempt of processing bilingual information to facilitate semantic reasoning and knowledge accommodation.

The system makes two primary streams of processing. The second step is to run the dataset through term and annotation extraction, in which there are Tamil terms recognized with the help of iNLTK, and annotations are translated with the help of a Lingvanex-based English-to-Tamil translation pipeline. Simultaneously, the dataset is categorized with narVidhai, a tool of BERT classification to come up with contextual instances that facilitate aesthetic organization of the semantic comprehension in a different perspective of classification. On the Tamil text, a Tamil LLM 7B is fine-tuned and specialized using the Google Gemma 7B and Alpaca Tamil datasets to produce highly contextual text summaries. These digests give important keys and captions in the Tamil language, which are enhanced terminations to be processed further. This rich data is further added with a Web3.0 crawler which gathers Tamil labeled infographics and

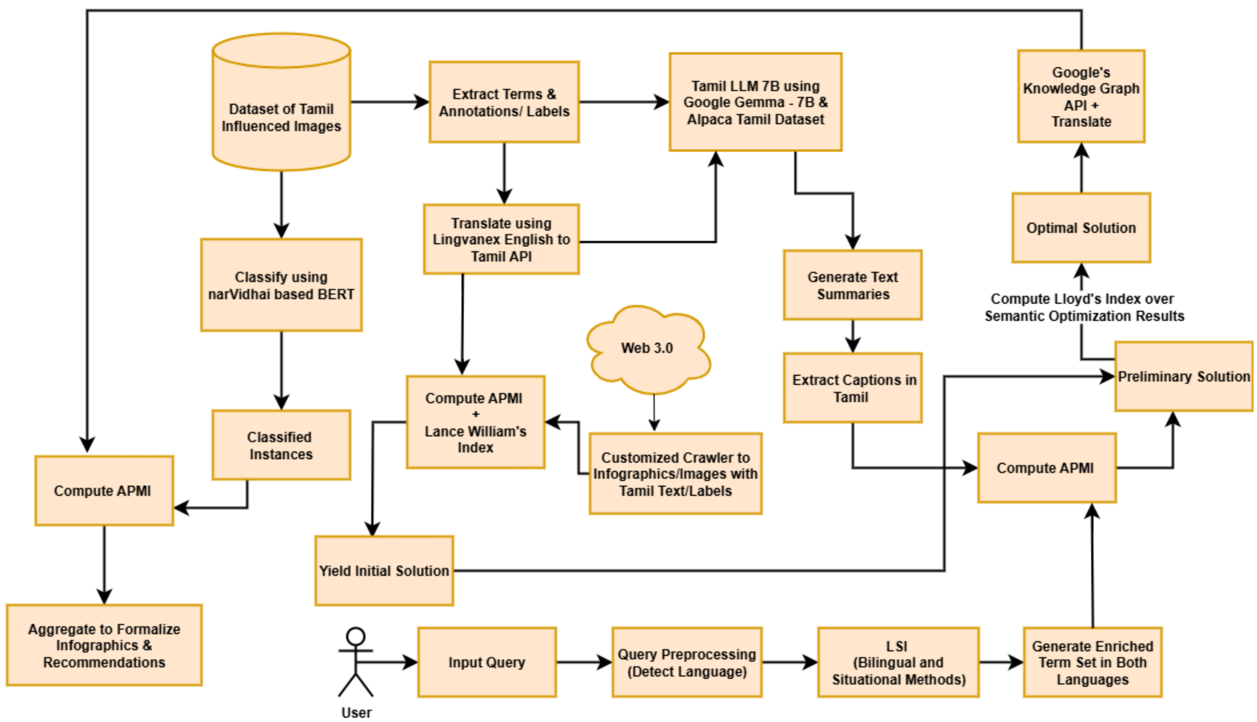

Fig. 1. Architecture of the proposed Tamil knowledge-centric infographic recommendation system.

photographs thus upgrading the auxiliary knowledge base with varied contextual information.

The queries made by the users (Tamil or English) are transferred along a bilingual query preprocessing pipeline. It contains language detection by iNLTK and NLTK, tokenization, stop-Word elimination, named-entity recognition and Latent Semantic Indexing by using bilingual and situational modeling. Using this contextual indexing, the system can produce enriched sets of terms in both languages in accordance with the underlying topic models.

Semantic matching A processing step of computing the computed Adaptive Pointwise Mutual Information between Tamil LLM captions and extracted annotations drives semantic matching. The extensions of PMI and NPMI are also calculated in two phases as the APMI that is a strict empirical threshold of 0.70 so that only the most informative term associations are preserved. This provides a set of preliminary solutions that will be the most semantically useful words to the user query. The initial solution is then optimised with Spider Optimization algorithm which is a metaheuristic that is based on web traversal behaviour. The algorithm has been designed to cover and semantically cohesive hypernyms by making sure that every term in the set of solutions is accessed at least twice. Lloyds Index, used as the cost function when optimising this project, will have a threshold of 0.15 and will bring the algorithm towards globally optimal solutions. Alternative preliminary solutions based on Web 3.0 crawled material are also considered on the basis of the Lance Williams Index which has a threshold of 0.108 and then reintroduced into the optimisation loop to make a compromise in terms of overall term coherence. The final optimal solution set produced from both query driven and data driven paths is passed through Google's Knowledge Graph API for knowledge enrichment. Entities are translated, mapped, and enriched using the Lingvanex Tamil API. To further personalize recommendations, classified instances from the BERT based stream are incorporated with a relaxed APMI threshold of 0.80, since they align closely with the already optimized output. Finally, the enriched entities are mapped back to both the dataset and Web 3.0 crawled infographics. Using relevance-based merging techniques, these infographics are composed into a single, contextually unified recommendation. Users can interact with this infographic set, request alternative permutations, or finalize a recommendation based on the presented combinations.

VII. Implementation and Performance Evaluation

The performance of the proposed Tamil and knowledge-centric infographic recommendation frame-

Model	Average Precision %	Average Recall %	Average Accuracy % (P+R)/2	Average F-Measure % (2*P*R)/(P+R)	FDR
USAR [21]	90.78	91.84	91.31	91.31	0.10
RTAI [22]	91.22	93.01	92.12	92.10	0.09
TVAI [23]	92.44	93.79	93.12	93.11	0.08
Proposed TamilInfoG	97.89	98.97	98.43	98.42	0.03

Fig. 2. Comparison of Performance of the proposed TamilInfoG with other approaches

work, referred to as Tamil InfoG, was evaluated using a comprehensive set of semantic and recommendation-specific metrics. The USAR model, designed as a user-centric semi-automated infographic authoring and recommendation tool, integrates human-in-the-loop elements like interactive interfaces for novice designers and domain experts. Despite incorporating some auxiliary knowledge through visual group entities, USAR lacks machine-driven semantic reasoning and quantitative analysis, resulting in significantly lower performance Precision: 90.78%, Recall: 91.84%, and FDR: 0.1. The RTAI model, a retrieval-based system with neural network sampling from learned distributions, uses recursive neural networks for mapping infographic labels. However, it is limited by its shallow auxiliary knowledge density and absence of deep semantic analysis, yielding intermediate results with Precision: 91.22%, Recall: 93.01%, and FDR: 0.09. The TVAI model, which automates infographic generation from natural language statements, showed relatively stronger performance but still lacked contextual enrichment and semantic reasoning. It achieved Precision: 92.4%, Recall: 93.7%, and FDR: 0.0, placing it below Tamil InfoG in all aspects. Tamil InfoG outperformed these models due to its deeply knowledge-centric architecture. The dataset of Tamil-influenced images was processed from multiple perspectives: direct classification using narVidhai-based BERT, extraction and translation of labels using the Lingvanex API, and generation of Tamil text summaries via the Tamil NLM7B model fine-tuned on Google's Gemma 7B and Alpaca datasets. Additionally, the model integrates Web 3.0-based crawling to enrich infographic text and labels and employs bilingual query preprocessing (Tamil and English) using INLTK. The structure of the proposed Tamil knowledge-based infographic recommendation system based on the hybridized semantic system of intelligence is shown in Fig.1. The system is designed to work on both user-based searches and a curated list of Tamil-influenced images, which is a source of foundation. Such images can have Tamil inscriptions or writing in English as it is widely seen in Web 3.0 repositories whereby the metadata is commonly in English. The framework is designed to elaborate on this attempt of processing bilingual information to facilitate semantic reasoning and knowledge accommodation.

The system makes two primary streams of processing. The second step is to run the dataset through term and annotation extraction, in which there are Tamil terms recognized with the help of iNLTK, and annotations are translated with the help of a Lingvanex-based English-to-Tamil translation pipeline. Simultaneously, the dataset is categorized with narVidhai, a tool of BERT classification to come up with contextual instances that facilitate aesthetic organization of the semantic comprehension in a different perspective of classification. On the Tamil text, a Tamil LLM 7B is fine-tuned and specialized using the Google Gemma 7B and Alpaca Tamil datasets to produce highly contextual text summaries. These digests give important keys and captions in the Tamil language, which are enhanced terminations to be processed further. This rich data is further added with a Web3.0 crawler which gathers Tamil labeled infographics and photographs thus upgrading the auxiliary knowledge base with varied contextual information.

The queries made by the users (Tamil or English) are transferred along a bilingual query preprocessing pipeline. It contains language detection by iNLTK and NLTK, tokenization, stop-Word elimination, named-entity recognition and Latent Semantic Indexing by using bilingual and situational modeling. Using this contextual indexing, the system can produce enriched sets of terms in both languages in accordance with the underlying topic models.

Semantic matching A processing step of computing the computed Adaptive Pointwise Mutual Information between Tamil LLM captions and extracted annotations drives semantic matching. The extensions of PMI and NPMI are also calculated in two phases as the APMI that is a strict empirical threshold of 0.70 so that only the most informative term associations are preserved. This provides a set of preliminary solutions that will be the most semantically useful words to the user query. The initial solution is then optimised with Spider Optimization algorithm which is a metaheuristic that is based on web traversal behaviour. The algorithm has been designed to cover and semantically cohesive hypernyms by making sure that every term in the set of solutions is accessed at least twice. Lloyds Index, used as the cost function when optimising this project, will have a threshold of 0.15 and will bring the algorithm towards globally optimal solutions. Alternative preliminary solutions based on Web 3.0 crawled material are also considered on the basis of the Lance Williams Index which has a threshold of 0.108 and then reintroduced into the optimisation loop to make a compromise in terms of overall term coherence.

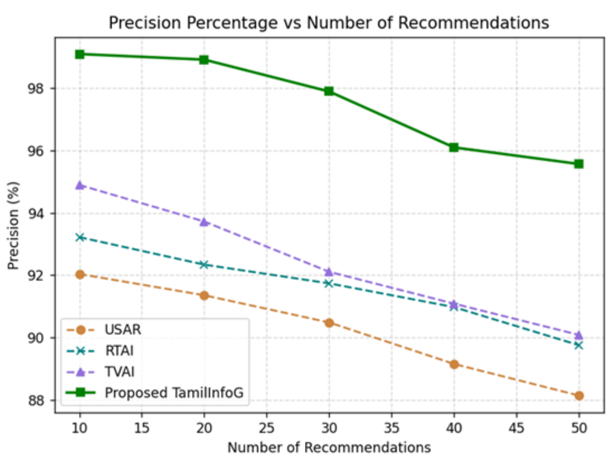

Fig. 3. Precision Vs Percentage Vs No. of Recommendations

VIII. Conclusion

The presented research provides a strategic framework of Tamil knowledge-based infographic recommendation, which makes it one of the first and most developed projects in the field. The structure combines both the semantic artificial intelligence and generative AI, where quantitative semantic reasoning controls the selection of the differentiating entities and gradual knowledge augmentation. The dataset is categorized through Tamil-BERT, and Tamil language architecture (7B) based on Gemma framework by Google creates textualization summaries, gathers captions, and relays auxiliary knowledge across the system.

The query-processing unit is fed with input in the Tamil language and in English; therefore, this makes the system support bilingual situational latent semantic indexing, which makes the system particularly new. Also, the tailored crawled infographics, enhanced with the Tamil text and labels, are added to the first set of solutions the final solution set may be improved with the feedback mechanisms. This looping enrichment cycle makes sure that the recommendation system is dynamic, adaptive and semantically rich, as it goes through its lifecycle.

The use of contentious information measures like the LanceWilliams index and the index of Lloyd is applied to strengthen the process of quantitative semantic reasoning under the support of the various thresholds as well as the step deviation metrics. These process guarantee efficient differentiation entity selection. A nature-inspired meta-heuristic known as spiral optimization algorithm is placed in strategic use to find the best solution solution in the initial solution space.

Its framework is supported by the inclusion of the Knowledge Graph API of Google, cross-lingual entity linking over Tamil-English translation, and BERT-classified instances, as well as entities retrieved through a crawler, which contribute to increasing the core capabilities of the system to formalize and suggest the infographics in Tamil. Generally, the proposed framework is more superior when compared to Tamil infographics framework as it comes out as a best-in-class model of Tamil infographic recommendation that meets the requirements of the Web 3.0. Through a combination of semantic web metadata, generative AI, bilingual processing, and incremental knowledge augmentation, the framework shows significant gains on recall, accuracy, and F -measure. This makes the model a trailblazer that does not only outdo the approaches based on a baseline but also provides a strong ground on the development of regionally centered language-centric knowledge systems in the Web 3.0 age.

References

[1] M. Hemashree, S. Banerjee, and R. Rashmi. A systematic review of semantic ai-based tagging for multimedia content in web 3.0 environments. *IEEE Access*, 2025.

[2] G. Deepak, A. Vibakar, and A. Santhanavijayan. Ontoinfog++: A semantic approach leveraging knowledge fusion for infographic recommendation. Publication venue not specified, 2023.

[3] D. Opoku. Examining challenges in infographic design in the modern web era. *Journal of Emerging Technologies and Innovative Research*, 2021.

[4] M. Prathibha and G. Deepak. Strategicvideorec: Integrating bert with fact-driven semantics for video recommendation. In *Innovations in Bio-Inspired Computing and Applications*, pages 39–49. Springer, Cham, 2023.

[5] A. Aljamel et al. Optimizing smart information retrieval through domain knowledge centric methods. *IEEE Access*, 7:4167–4183, 2018.

[6] S. R. Chaudhuri et al. Conceptualizing robotics using a knowledge-centric perspective. In *Proceedings of the 12th Innovations in Software Engineering Conference*, pages 1–11, 2019.

[7] H. S. Al Khatib et al. A review on patient-focused knowledge graphs: techniques, challenges, and use cases. *Frontiers in Artificial Intelligence*, 7:1388479, 2024.

[8] Q. Guo et al. Knowledge graph-enhanced recommender systems: a survey. *IEEE Transactions on Knowledge and Data Engineering*, 34(8):3549–3568, 2020.

[9] K. Zhou et al. Semantic fusion with knowledge graphs for conversational recommenders. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 1006–1014, 2020.

[10] V. Lully et al. Utilizing knowledge graphs to explain recommendations. *Procedia Computer Science*, 137:211–222, 2018.

[11] H. Wang et al. Applying knowledge graph convolutional networks to recommendation. In *Proceedings of the World Wide Web Conference*, pages 3307–3313, 2019.

[12] P. Pei, D. Xuejing, and Z. Deqing. Ontology-based curriculum knowledge mapping. In *Proceedings of the 2nd International Conference on Robotics, Intelligent Control and AI*, pages 259–265, 2020.

[13] Y. Deldjoo et al. Generative model-based recommendation: concepts and challenges, 2024.

[14] A. Zhang et al. Generative agents for recommender systems. In *Proceedings of the 47th ACM SIGIR Conference on Information Retrieval*, pages 1807–1817, 2024.

[15] Z. Hong et al. Survey on multi-objective recommendation using generative ai, 2025.

[16] D. Nawara and R. Kashef. Exploring llm-driven recommender systems: from classification to multi-modal generation. *IEEE Access*, 2025. (to be published).

[17] B. A. Hassan and T. A. Rashid. Learning semantic web ontologies via ai-driven natural language algorithms, 2021.

[18] A. Karandikar et al. A multi-objective hybrid optimization approach for self-healing concrete using ai-enhanced metaheuristics. *Asian Journal of Civil Engineering*, pages 1–13, 2025.

[19] C. Maureira et al. Adaptive macro-level knowledge modeling for the aec-ai domain. *IEEE Access*, 9:110842–110879, 2021.

[20] H. Naghavipour et al. Quality-aware service composition using hybrid metaheuristics: A structured mapping review. *IEEE Access*, 10:12678–12701, 2021.

[21] Xuan-Son Vu, Thanh-Binh Nguyen, Tai-Nhan Nguyen, Son-Tung Tran, Quang-Huy Nguyen, Hoang-An Nguyen, Van-Tam Nguyen, and Minh-Triet Nguyen. Usar: An interactive user-satisfactory-aware app recommendation system. arXiv preprint arXiv:2108.11914, 2021.

[22] S. Annam, B. E. Reddy, V. Vagavolu, S. Koneru, and C. V. Jawahar. Rtai: A real-time framework for tagging personal images based on aesthetics and content. In *2020 IEEE Region 10 Conference (TENCON)*, pages 798–803, 2020.

[23] J. Park, G. Kim, S. Oh, S. Lee, J. Kim, and Y. Han. Tvai: A novel task-oriented conversation model for video annotation. In *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 7365–7369, 2019.