Openstream.ai®
Delivering Intelligence

Heralding the Intelligent Enterprise

www.Openstream.ai

Conversational Information Retrieval using Knowledge Graphs

Pruthvi Raj Venkatesh

Rishu Kumar

K Chaitanya

P Radha Krishna

Agenda

1. Introduction
2. Problem Statement
3. Proposed Solution
4. Multimodal Knowledge Graph and Ontology
5. NLU Service
6. Multimodal Context Building
7. Dynamic Query Generator
8. Results Presentation
9. Q & A

Background

- Recent years have seen a huge increase in the popularity of information retrieval(IR) systems, which enable users to hold natural language conversations, these IR systems are called conversation agents or bots
- Due to the increased usage of conversational agents in recent years, researchers have enhanced them to adapt other modalities, such as graphics, sound, and video, to improve conversational flexibility
- Conversational agents are typically goal-oriented and use queries generated using a trained model or an intent and query mapping repository to retrieve information from backend systems
- Trained models that generate the back-end queries for question require huge dataset to ensure higher accuracy
- Most conversational agents are context aware and provide results based on the inputs provided by the user in a multi-turn conversation
- Conversational agents are developed using frameworks that provide a platform for developing and hosting conversational agents

Problem Statement

- Though Conversational IR systems effectively address users’ information requirements, there is a need for an approach that can easily adapt to diverse use cases and meet all user’s information needs without requiring the user to be aware of the backend system
- Conversational IR agents are typically goal-oriented and use predefined *intent query* mapping to retrieve information from backend systems. As the number of intent increases, the complexity of managing the system also increases
- Conversational agents that required training data have limited usability in domains with minimal or no training data
- Queries generated using trained models require additional effort to add more variations of the intent or when queries require changes.
- Every new addition of an intent required corresponding addition to the dialogue flow to support the added intent. This reduced the manageability of the system as more intents are added to the system

Openstream.ai®
Delivering Intelligence

Proposed Solution

- The approach described in this research, MIR-KG (Multimodal Information Retrieval System utilizing Knowledge Graph), uses ontologies and knowledge graphs to power the question-answering process
- The proposed approach combines transformer based intent and multimodal entity detection from user questions with procedural dynamic query creation
- The multimodal entities provided as input to the conversational agent are progressively collected into the conversation context. The dynamic query generation engine uses the context to generate multimodal knowledge graph queries
- Multimodal Knowledge Graphs (MMKG) is used to store data and ontology to represent the structural information of the MMKG database.
- Ontology is used by the Natural Language Understanding (NLU) engine to convert the user question into a logical form.
- Dynamic query builder generates the query form the logical form to be executed against the back end data source.

Proposed Solution(Contd..)

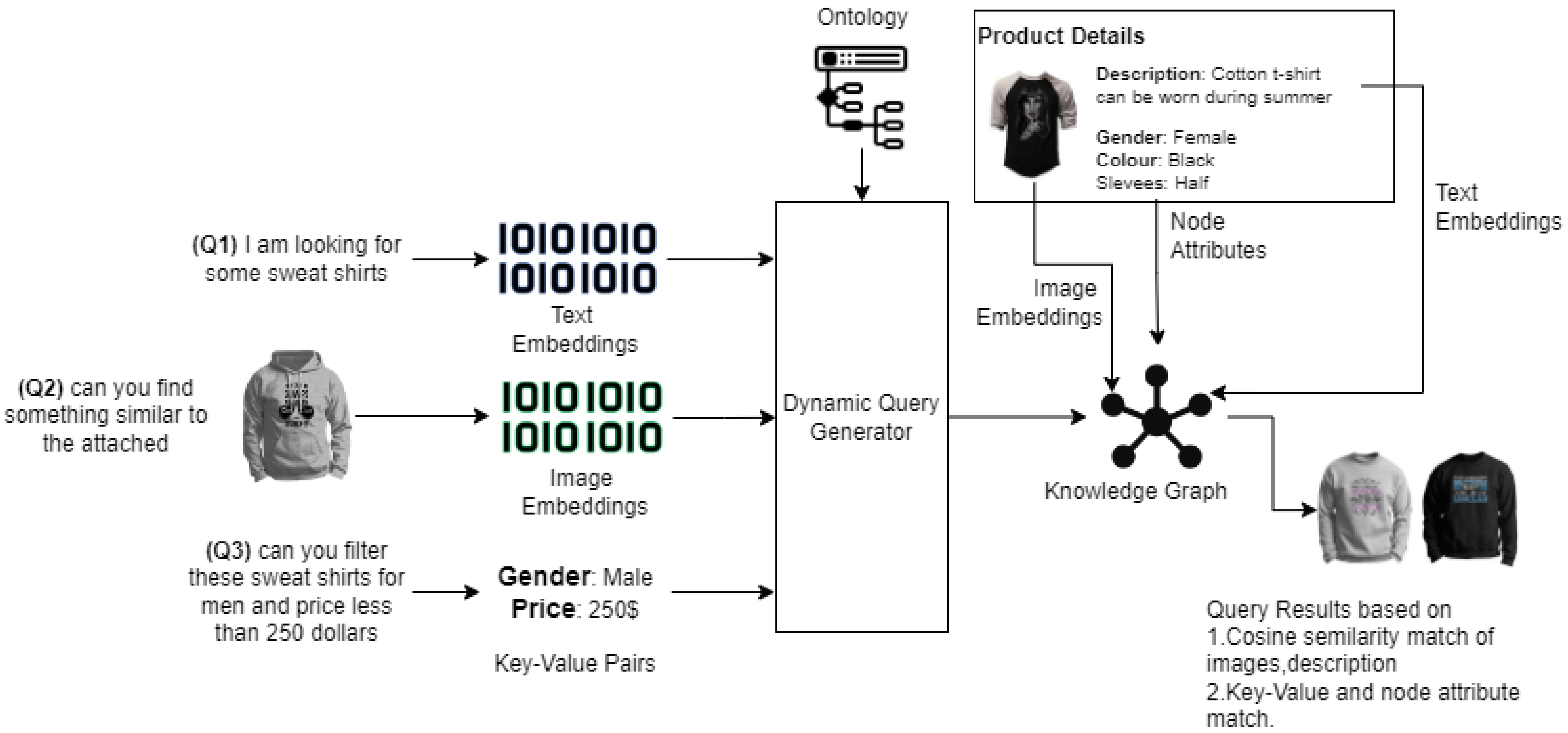


Fig 1. Multi-turn IR in Fashion Domain using proposed solution

Proposed Solution (Contd..)

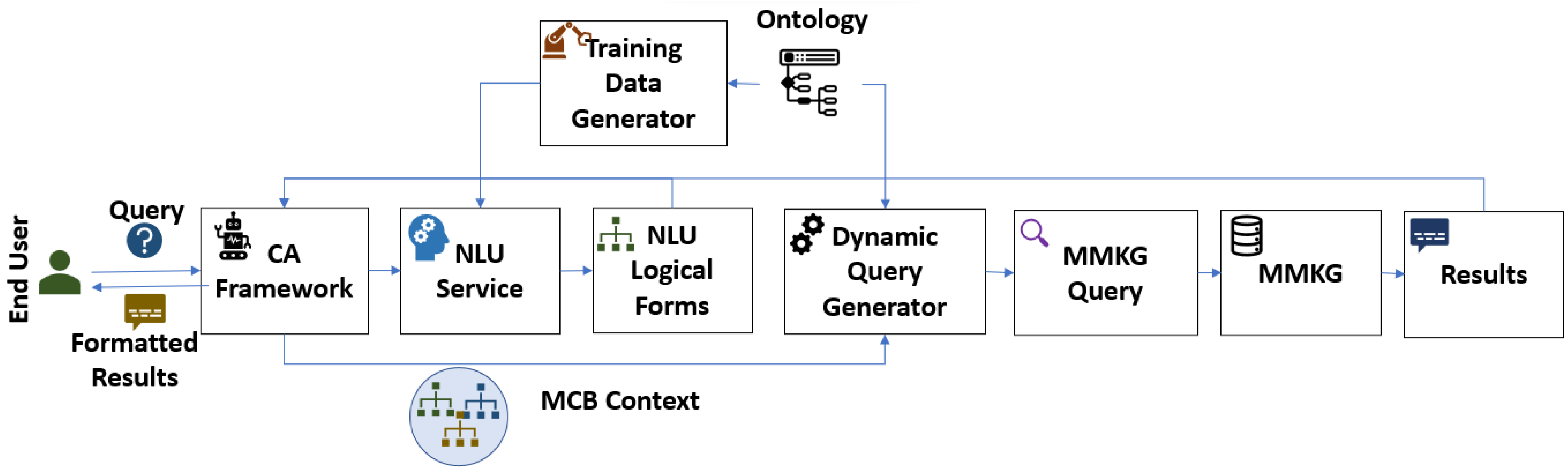
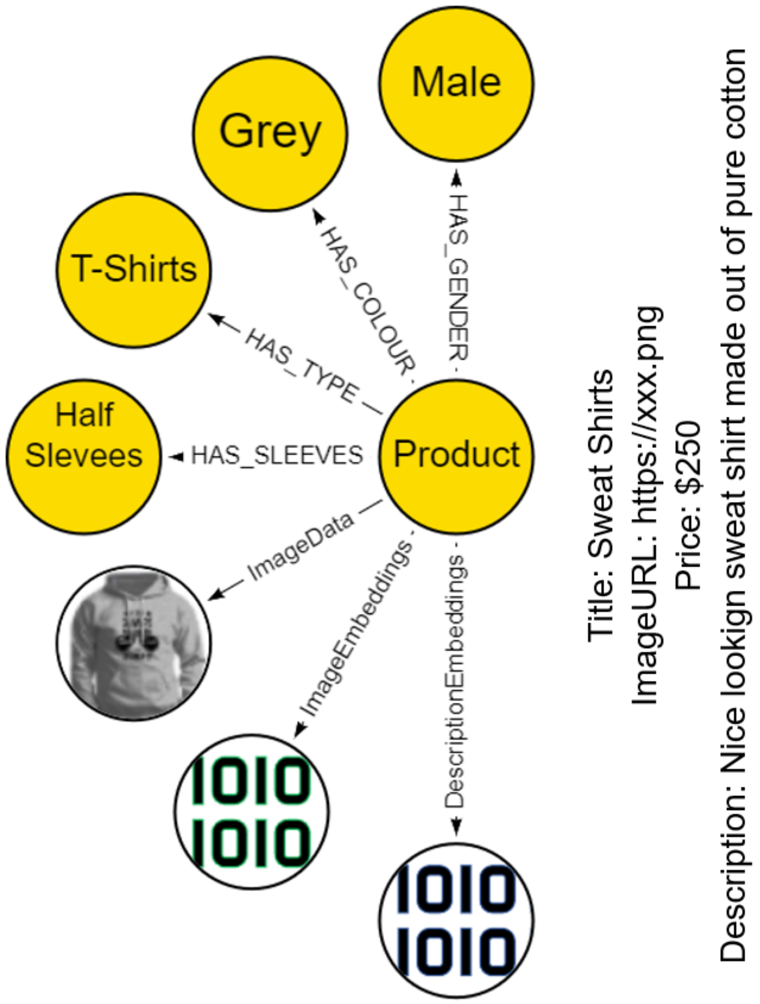


Fig 2. QnA Framework

Openstream.ai® Delivering Intelligence

Multimodal Knowledge Graph and Ontology



(a) - Multimodal Knowledge Graph

```
"nodes": {
  "Product": {
    "Title": "Text",
    "ImageURL": "Text",
    "Price": "Numeric",
    "ImageData": "Image",
    "Description": "Text",
    "ImageEmbeddings":
"Embeddings",
    "Description
Embeddings": "Embeddings"
    },
```

```
"Color": {
   "Title": "Text"
 },
 "Gender": {
   "Title": "Text"
 },
 "Sleeves": {
   "Title": "Text"
 },
 "Type": {
   "Title": "Text"
 }
```

(b) – Ontology Nodes

```
"relationtrees": {
    "paths": [
      "(P:Product)-[HT:HAS_TYPE]-(T:Type)",
      "(P:Product)-[HC:HAS_COLOR]-(C:Color)",
      "(P:Product)-[HG:HAS_GENDER]-(G:Gender)",
      "(P:Product)-[HS:HAS_SLEEVES]-(SL:Sleeves)"
    ]
  }
```

(c) – Ontology Relations

Fig 3. Knowledge Graph and Ontology

NLU Service

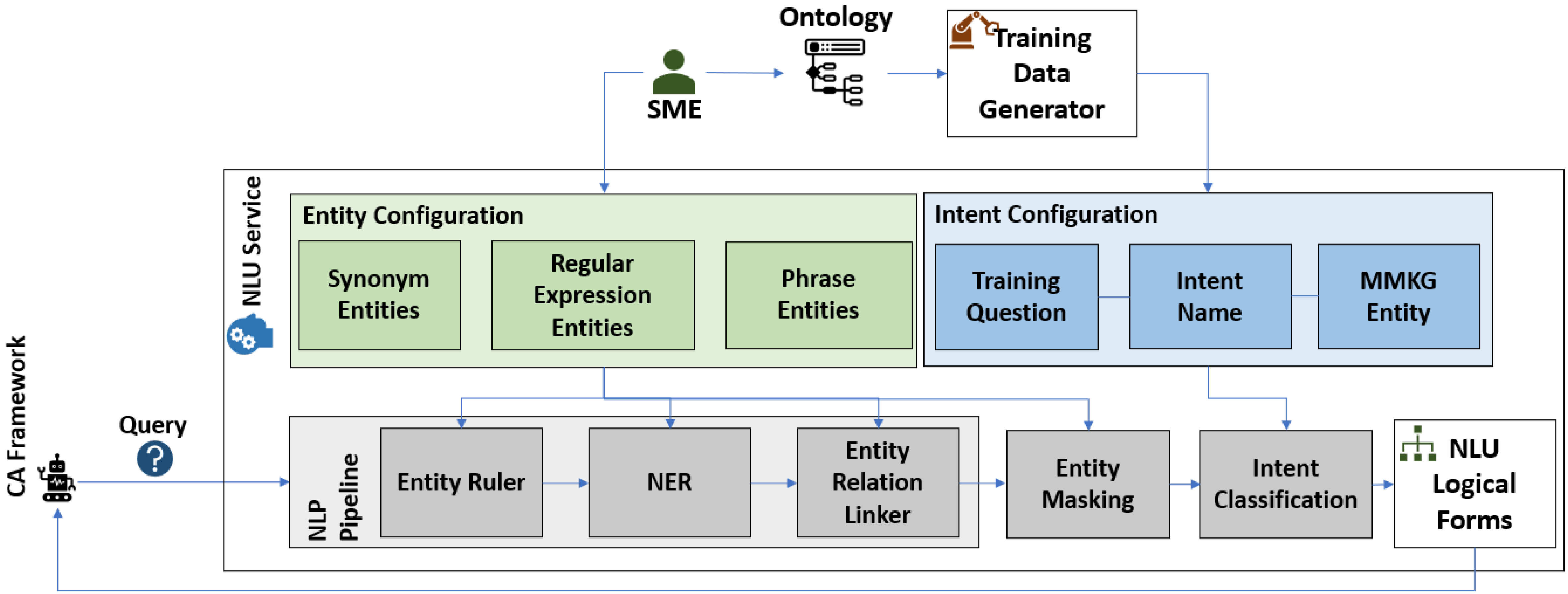


Fig 4. NLU Service Pipeline

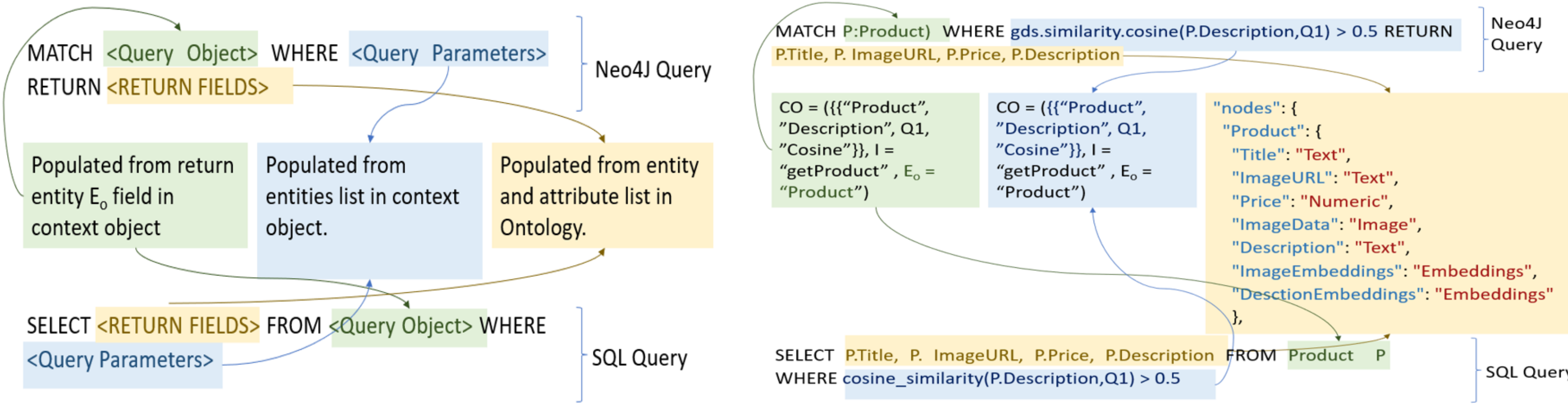
Multimodal Context Building(MCB)

Table 1

MCB Context Generation

Turn	CA Response	User Question/Context	NLU Entities and Intents identified	Context Object State
1	Please let us know what you are looking for.	(Q1) = I am looking for some sweat shirts	ER = {}, I = “getProduct” , E_o = “Product”	CO = ({{“Product”, ”Description”, Q1, ”Cosine”}}, I = “getProduct” , E_o = “Product”)
2	These are the results, Do you want to provide more filters.	(Q2) = can you find something similar to the attached (user attaches a image of the shirt)	ER = {{“Product”, ”ImageURL”, “http://....png”, ”Cosine”}}	CO = ({{“Product”, ”Description”, Q1, ”Cosine”}, {“Product”, ”ImageURL”, “http://....png”, ”Cosine”}}, I = “getProduct” , E_o = “Product”)
3	These are the results, provide more context to get more filtered results.	(Q3) = can you filter these sweat shirts for men and price less than 250 dollars	ER = {{“Gender”, ”Title”, “men”, ”=”}, {“Product”, ”Price”, “250”, ”<”}}	CO = ({{“Product”, ”Description”, Q1, ”Cosine”}, {“Product”, ”ImageURL”, “http://....png”, ”Cosine”}, {“Gender”, ”Title”, “men”, ”=”}, {“Product”, ”Price”, “250”, ”<”}}, I = “getProduct” , E_o = “Product”)

Dynamic Query Generator



(a) Dynamic Query Generation Rules **(b) Dynamic Query Generation for Q1**

Fig 5 Dynamic Query Generator

Results

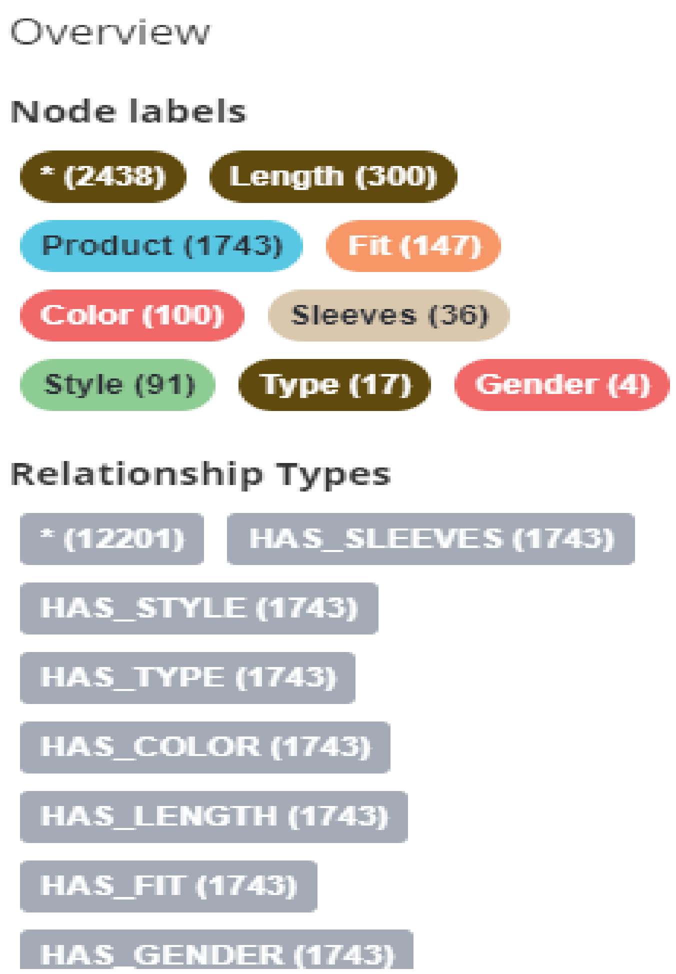


Fig 6. Node Types and Relations

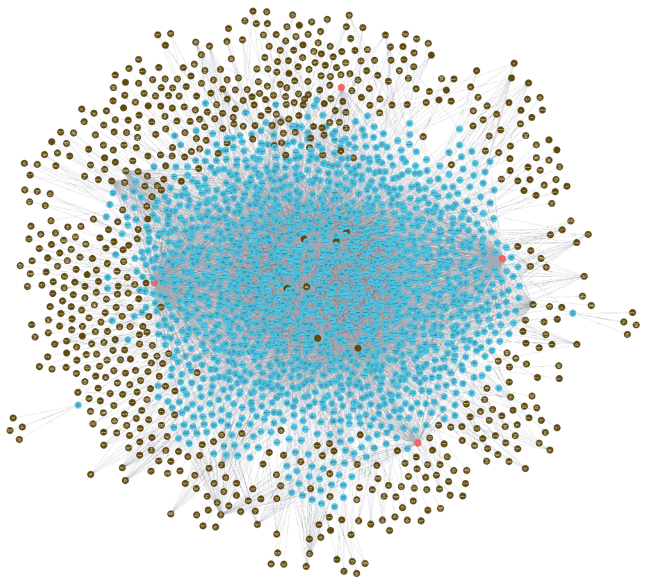


Fig 7. Zoomed out view of Neo4J Database

Actions	Stats
Number of ingested JSON from MMD Website	1743
Number of Noe4J Nodes	2438
Number of Noe4J Relations	12201
KG Data load time	~10 mins
Embedding Generation Time	2 secs

Fig 9. Data load Statistics

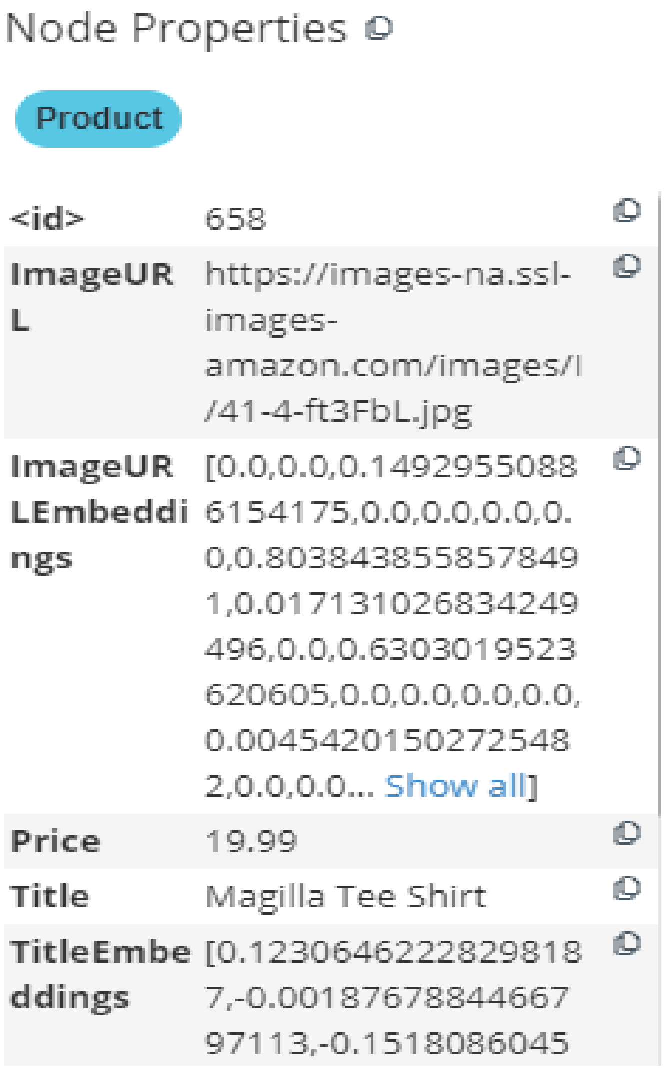


Fig 8. Node Properties of Product Node

Results

Table 2

MHRED-attn Vs MIR-KG Response

<table>
<tr><td></td><td rowspan="4">MIR-KG Response
The product details are
Title : Men Flex Fury Athletic Black Wolf Running Shoes
Description : A pair of black walking shoes, has regular Styling, lace-ups detail, Mesh upper, Cushioned footbed and Textured and patterned outsole. It has a feet tip and rubber sole.
Price : 25$
Image : </td></tr>
<tr><td>User Query: What is the sole material and tip in the 4^{th} result?</td></tr>
<tr><td>Gold Response: The running shoes in the 4^{th} image has feet tip
rubber sole material.</td></tr>
<tr><td>Baseline Response (MHRED-attn): The running shoes in the 4th image has rubber sole.</td></tr>
</table>

Summary

- Conversational agents enable users with information retrieval by querying the underlying systems.
- Traditional state-of-the-art transformer-based systems were trained with a huge dataset to achieve high accuracy. The traditional approach is time-consuming and requires a huge effort and computing resources.
- This paper presents a novel predictive and rule-based approach that reduces training time and supports multimodal conversations using a multimodal knowledge graph.
- We introduced the context-aware multimodal approach, where a context object is progressively built during the multi- turn conversation.
- The context gathered along with ontology is used to query the database with all the previously recorded entities.
- Domain ontology defines entities, attributes, and rules and drives the dynamic query generation module, which helps create conversational IR systems for a new domain with minimal effort.
- We compared our results with a Multi-headed Hierarchical Encoder-Decoder with attention approach and found that the proposed approach gives a more detailed response to user queries.

Openstream.ai®
Delivering Intelligence

References

1. MMD: Towards Building Large Scale Multimodal Domain-Aware Conversation Systems https://amritasaha1812.github.io/MMD/.
2. Xiangru Zhu, Zhixu Li, Xiaodan Wang, Xueyao Jiang, Penglei Sun, Xuwu Wang, Yanghua Xiao, and Nicholas Jing Yuan, "Multimodal Knowledge Graph Construction and Application: A Survey", 22, doi: https://doi.org/10.48550/arXiv.2202.05786
3. H.Jiang, B.Yang, L.Jin and H.Wang, "A BERT-Bi-LSTM-Based Knowledge Graph Question Answering Method," 2021 International Conference on Communications, Information System and Computer Engineering (CISCE), 2021, pp.308-312, doi: 10.1109/CISCE52179.2021.9445907.
4. L.Ma, P.Zhang, D.Luo, M.Zhou, Q.Liang and B.Wang, "Answer Graph-based Interactive Attention Network for Question Answering over Knowledge Base," 2020 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking (ISPA/BDCloud/SocialCom/SustainCom), 2020, pp.521-528, doi: 10.1109/ISPA-BDCloud-SocialCom-SustainCom51426.2020.00091.
5. Y.Li, J.Cao and Y.Wang, "Implementation of Intelligent Question Answering System Based on Basketball Knowledge Graph," 2019 IEEE 4th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), 2019, pp.2601-2604, doi: 10.1109/IAEAC47372.2019.8997747.
6. S.Hu, L.Zou, J.X.Yu, H.Wang and D.Zhao, "Answering Natural Language Questions by Subgraph Matching over Knowledge Graphs (Extended Abstract)," 2018 IEEE 34th International Conference on Data Engineering (ICDE), 2018, pp.1815-1816, doi: 10.1109/ICDE.2018.00265.
7. X.Dai, J.Ge, H.Zhong, D.Chen and J.Peng, "QAM: Question Answering System Based on Knowledge Graph in the Military," 2020 IEEE 19th International Conference on Cognitive Informatics & Cognitive Computing (ICCI*CC), 2020, pp.100-104, doi: 10.1109/ICCICC50026.2020.9450261.
8. S. Zoghbi, G. Heyman, J. C. Gomez, and M-F. Moens. Fashion Meets Computer Vision and NLP at e-Commerce Search. International Journal of Computer and Electrical Engineering (IJCEE), Vol. 8, No 1, pp. 31–43, February 2016.
9. K. Laenen, S. Zoghbi, and M-F. Moens. Web Search of Fashion Items with Multimodal Querying. In Proceedings of WSDM 2018: The Eleventh ACM International Conference on Web Search and Data Mining, Marina D
10. S. Bell and K. Bala. Learning Visual Similarity for Product Design with Convolutional Neural Networks. ACM Transactions on Graphics (TOG), vol. 34 , No 4, pp. 1-10, July 2015.
11. J.-H. Hsiao and L.-J. Li. On Visual Similarity based Interactive Product Recommendation for Online Shopping. 2014 IEEE International Conference on Image Processing (ICIP), pp. 3038-3041, 2014.
12. B. Zhao, J. Feng, X. Wu, and S. Yan. Memory-Augmented Attribute Manipulation Networks for Interactive Fashion Search. IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2017). pp. 6156–6164. 2017
13. X. Han, Z. Wu, P. X. Huang, X. Zhang, M. Zhu, Y. Li, Y. Zhao, and L. S. Davis.Automatic Spatially-Aware Fashion Concept Discovery. 2017 IEEE International Conference on Computer Vision (ICCV), pp. 1472-1480, 2017

Openstream.ai
Delivering Intelligence

References

14. C. R. Sapna, M. Anagha, K. Vats, K. Baradia, T. Khan, S. Sarkar, and S. Roychowdhury. Recommendence and fashionsence online fashion advisor for offline experience. ACM International Conference Proceeding series, pp. 256–259, 2019
15. A. Paranjape, A. See, K. Kenealy, H. Li, A. Hardy, P. Qi, K. R. Sadagopan, N. M. Phu, D. Soylu, and C. D. Manning. Neural generation meets real people: Towards emotionally engaging mixed-initiative conversations. Stanford NLP, 3rd Proceedings of Alexa Prize, arXiv:2008.12348, 2020
16. Vision Transformer - https://en.wikipedia.org/wiki/Vision_transformer
17. Sentence Embedding - https://en.wikipedia.org/wiki/Sentence_embedding
18. Bot Framework SDK - https://docs.microsoft.com/en-us/azure/bot-service/bot-service-overview?view=azure-bot-service-4.0
19. Synthetic Data - https://en.wikipedia.org/wiki/Synthetic_data
20. Language Processing Pipelines · spaCy - https://spacy.io/usage/processing-pipelines