Enhancing Knowledge Graphs with Large Language Models: Contributions to E-commerce Question Answering Systems

André Gomes Regino
aregino@cti.gov.br
Instituto de Computação, Universidade Estadual de Campinas (UNICAMP) and Centro de Tecnologia da Informação Renato Archer
Campinas, SP, Brazil

Julio Cesar dos Reis
jreis@ic.unicamp.br
Instituto de Computação, Universidade Estadual de Campinas (UNICAMP)
Campinas, SP, Brazil

ABSTRACT

E-commerce platforms demand structured and dynamic knowledge to deliver accurate, humanized, and scalable customer support. This Ph.D. Thesis introduces QART, a novel framework that integrates Large Language Models (LLMs) to automatically extract high-quality RDF triples from question-answer (Q&A) dialogues, enabling population of domain-specific Knowledge Graphs. The method combines semantic summarization, few-shot triple generation, and automated consistency validation to align outputs with existing ontologies. Experimental results across multiple LLMs demonstrate up to 20% gains in precision over baseline approaches and high accuracy in semantic validation tasks (F1 = 0.98). The solution has been validated in an industrial environment, showing potential to reduce human effort, enhance recommendation quality, and improve customer satisfaction through more informed and context-aware answers.

KEYWORDS

Electronic Commerce, Semantic Web, Generative Artificial Intelligence, Large Language Models

1 INTRODUCTION

In the dynamic environment of e-commerce, customer satisfaction increasingly depends on timely, accurate, and personalized responses to product-related inquiries [6]. Traditional question-answering systems often fall short due to their reliance on manual responses, which introduce delays and inconsistencies. These limitations are especially evident when customers abandon purchases after receiving late or no answers to critical product questions [5]. To address these gaps, the integration of Knowledge Graphs (KGs) emerges as a promising solution for structuring and reusing question-answer (Q&A) knowledge. By capturing and organizing interactions between customers and sellers, KGs can automate responses, reduce repetitive queries, and enhance the efficiency of digital platforms [5]. The Ph.D. research, conducted in partnership with the Brazilian AI startup GoBots, explores how combining LLMs [3] with KGs can transform e-commerce support into a more scalable, humanized, and intelligent process, ultimately improving customer experience and operational efficiency.

Despite the potential of KGs and LLMs, several challenges hinder their effective application in real-world e-commerce scenarios. Extracting structured RDF triples from informal, ambiguous, and error-prone customer interactions [1] is a complex task that requires accurate intent recognition, entity linking, and contextual understanding. Moreover, ensuring that newly generated triples are semantically consistent with existing knowledge (without introducing redundancy or contradictions) adds another layer of complexity to KG evolution. These tasks are compounded by the need to align extracted knowledge with predefined ontologies and to support real-time, human-like interactions. Without ontologies, the LLMs can generate triples that might not fit the existing knowledge present in the KG.

This thesis develops and evaluates a structured approach that leverages LLMs to update KGs with e-commerce Q&A texts. The solution addresses challenges such as syntactic errors in natural language e-commerce texts and their corresponding semantic representation through techniques for sentence simplification, ontology-aligned triple generation, and semi-automated syntactic and semantic validation. Additionally, it explores using structured triples to enhance user experience by generating more humanized responses and delivering more accurate product recommendations.

Our method advances the state of the art by offering a scalable and domain-specific framework for transforming e-commerce text (natural language) into structured knowledge (RDF triples). Unlike prior studies such as Seq2RDF [2], which is limited to generating a single triple per sentence and lacks domain specificity, or KGen [4], which is semi-automatic and tailored to biomedicine, our approach fully automates RDF triple generation and insertion in an e-commerce context using LLMs and ontology parts (classes and properties).

2 METHOD

Our solution introduces QART (Question and Answers to RDF Triples, see Figure 1), a novel framework that transforms e-commerce Q&A text into RDF triples for updating a domain-specific KG. The framework is composed of three key stages: (1) summarization of noisy Q&A interactions into factual statements; (2) triple generation using few-shot prompting with LLMs; and (3) semantic validation of the generated triples, including syntax checking, URI duplication, and ontology consistency. Each stage is supported by prompt-based LLMs tailored to the structure and vocabulary of the domain ontology. Beyond the triple pipeline, we demonstrate the practical utility of QART by integrating the enriched KG into two real-world applications: (i) a humanized answer generator that retrieves relevant
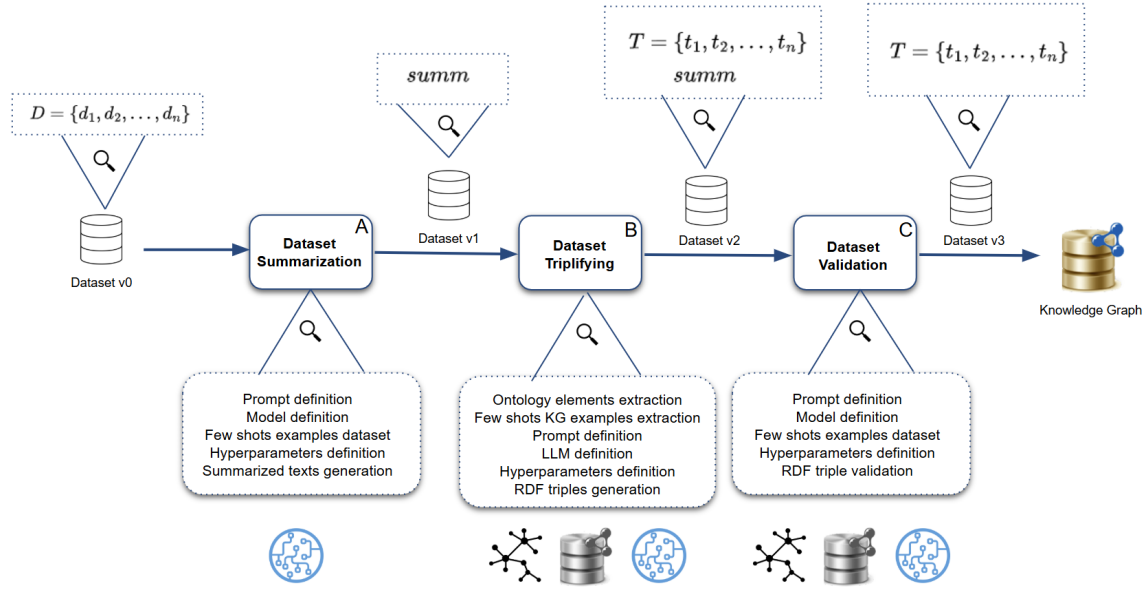
Figure 1: QART framework: a three-stage solution to KG population from Q&A text using LLMs.

knowledge and reformulates responses with contextual empathy, and (ii) a product recommendation engine that leverages structured relations for more precise and explainable suggestions.

3 EVALUATION

We evaluated QART by analyzing each of its core components in isolation. In the summarization stage, we compared fine-tuning and few-shot prompting across different LLMs and found that few-shot prompting consistently outperformed fine-tuning, requiring fewer resources while achieving higher quality summaries (BLEU score gains of up to 15%). In the triple generation stage, few-shot learning again outperformed zero-shot and multilingual baselines, with Portuguese-specific models achieving results on par with LLaMA. Sentences involving compatibility relations produced the highest F1 scores (≈ 0.92). Finally, in the validation stage, LLMs were tested on their ability to detect syntactic, URI-related, and semantic errors. Top-performing models reached F1 scores of 0.98 for syntactic and URI errors, and 0.90 for semantic consistency checks.

These results confirm the effectiveness of QART not only in automating KG population but also in ensuring triple quality and enabling direct value creation in end-user applications.

4 ADHERENCE TO WEBMEDIA

The thesis aligns with the themes of WebMedia by addressing key areas such as E-commerce, Artificial Intelligence, Deep Learning, Ontologies, Semantic Web, Natural Language Processing, and Recommender Systems. By proposing a framework that automates the structuring of user interactions into KGs using LLMs, the work contributes to the scalability and intelligence of digital services. Socially, it improves customer experience through more humanized and context-aware responses, supporting fairer and more informed consumption decisions.

5 CONTRIBUTION

One of the main contributions of the Ph.D. thesis is the development of **QART**, and methods for generating humanized answers and compatibility-based product recommendations from structured data. QART has been adopted by **GoBots**, a Brazilian AI company, and applied through a formal **industry-academia partnership** with UNICAMP, demonstrating real-world impact. The work also led to an **international collaboration** with Wageningen University & Research (WUR), where the proposed techniques were extended to domains such as animal welfare. The results of the thesis were disseminated through **15 scientific papers** and a **hands-on tutorial**[1].

REFERENCES
[1] Xusen Cheng, Ying Bao, Alex Zarifis, Wankun Gong, and Jian Mou. 2021. Exploring consumers' response to text-based chatbots in e-commerce: the moderating role of task complexity and chatbot disclosure. *Internet Research* 32, 2 (2021), 496–517.

[2] Yue Liu, Tongtao Zhang, Zhicheng Liang, Heng Ji, and Deborah L McGuinness. 2018. Seq2rDF: An end-to-end application for deriving triples from natural language text. In *CEUR Workshop Proceedings*, Vol. 2180. CEUR-WS.

[3] Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, Ilya Sutskever, et al. 2019. Language models are unsupervised multitask learners. (2019).

[4] Anderson Rossanez, Julio Cesar Dos Reis, Ricardo da Silva Torres, and Hélène de Ribaupierre. 2020. KGen: a knowledge graph generator from biomedical scientific literature. *BMC medical informatics and decision making* 20, 4 (2020), 1–24.

[5] Diogo Teles Sant'Anna, Rodrigo Oliveira Caus, Lucas dos Santos Ramos, Victor Hochgreb, and Julio Cesar dos Reis. 2020. Generating Knowledge Graphs from Unstructured Texts: Experiences in the E-commerce Field for Question Answering. In *Advances in Semantics and Linked Data: Joint Workshop Proceedings from ISWC 2020.* 56–71.

[6] Snezhana Sulova. 2018. Integration of Structured and Unstructured Data in the Analysis of E-commerce Customers. *International Multidisciplinary Scientific GeoConference: SGEM* 18, 2.1 (2018), 499–505.

---

[1] Available at https://bit.ly/sbbd_tutorial