

LLM-Assisted Expansion of Winery Ontologies with Kosovo Agritech Context

Eliot Bytyçi, Bekim Gashi

0000-0001-7273-9929, 0000-0003-1770-3143

University of Prishtina, Prishtina, Kosovo

Email: {eliot.bytyci, bekim.gashi}@uni-pr.edu

Arianit Kurti

0000-0003-0512-6350

Linnaeus University, Vaxjo, Sweden

Email: arianit.kurti@lnu.se

Abstract—Ontology engineering is an intensive and costly process, requiring domain expertise, ontology design skills and also manually intervening in the end. With the recent advantages in Large Language Models, new possibilities have been opened to automate or at least expedite the ontology creation, by enabling their localization first of all, extraction and alignment. Thus, this motivated us to explore development of Kosovo Winery Ontology, being a semantic framework for Kosovo's emerging wine industry. Our approach integrates three existing ontologies: W3C wine ontology, Davis Wine ontology and Wine Cloud ontology, with extension of local data context. The resulting ontology should serve as schema for the Agritech project, a project that will use precision agriculture in Kosovo. Moreover, it will serve as a case study in LLM assisted ontology engineering, demonstrating further the fostering of digital transformation in agriculture.

I. INTRODUCTION

RECENT advancements of large language models (LLMs) have transformed how researchers approach ontology engineering, substituting some of the manual phases of semantic resource creation, with automation. These tools can help now in supporting ontology design, creation of concept and relations, linking of different resources, and localization of different ontologies. Moreover, they can help in developing specific domain but also regional ontologies, that previously have lacked the involvement, such as Kosovo.

Thus, in this context, wine domain and Kosovo region as one of the wine producers not so well explored, is the ground for our explorations with LLMs and ontologies. Three main wine ontologies were localized, again through the help of LLM: W3C Wine ontology [4], Davis Wine Ontology [5] and Wine Cloud ontology [6]. W3C Wine ontology provides generic wine semantics for demonstration and interoperability. Davis Wine ontology emphasizes viticultural and enological properties, while Wine Cloud ontology introduces a cloud-based semantic framework tailored to the integration of wine data within distributed digital ecosystems.

Nevertheless, these ontologies do not capture regional, varietal and cultural properties of Kosovo, as a wine producing landscape. There are different varieties such as Vranc and Kallmet, and several micro winery regions such as Rahovec and Suhareka. These local semantics limit both interoperability with wine knowledge systems globally and it also limits digital transformation of Kosovo viticultural sector.

To address these gaps, we introduce the Kosovo Wine Ontology (KWO), constructed by merging and extending the three reference ontologies described above through LLM-assisted generation, enriched with Kosovo-specific viticultural knowledge including regional appellations and local grape varieties. KWO is designed to serve as the semantic backbone of AgriTech Kosovo, an ongoing research and innovation initiative that has been in conceptual and technical development for several years, targeting the digitalization and quality improvement of Kosovo's wine and agricultural sector. The ontology contributes directly to the project's interoperability layer, enabling structured data exchange between IoT sensors, domain databases, and analytical tools deployed across Kosovo's viticultural regions. Building on this foundation, AgriTech Kosovo is being positioned for submission to Horizon Europe funding calls — in particular under the Farm-to-Fork Strategy and agricultural innovation priorities — where KWO will serve as a core deliverable demonstrating semantic readiness and knowledge infrastructure maturity.

The remainder of this paper is structured as follows. Section 2 reviews related work on LLM-assisted ontology engineering and existing wine ontology formalizations. Section 3 presents the methodology adopted for ontology merging, LLM-assisted extension, and Kosovo-specific knowledge integration. Section 4 describes the resulting Kosovo Wine Ontology (KWO) structure, reports evaluation results, and illustrates its application within AgriTech Kosovo scenarios. Section 5 concludes the paper by summarising the key contributions, acknowledging limitations, and outlining directions for future research.

II. RELATED WORK

LLMs have been increasingly applied to automate knowledge graph and ontology construction, yet significant challenges remain in domain-specific precision, completeness, and structural correctness.

Recent evaluations show that LLM-based approaches [1], spanning zero-shot prompting, few-shot in-context learning, and fine-tuning, demonstrate strong performance on term typing tasks but struggle significantly with relation extraction on domain-specific corpora, and completeness remains a challenge compared to manually curated ontologies.

For ontology engineering specifically, authors in [2] demonstrate that LLMs using advanced prompting techniques such as

Chain-of-Thought and sub-task decomposition show effectiveness in generating OWL ontology components — performing comparably to novice ontology engineers — while noting that schema coherence and structural correctness still require expert validation.

KG-augmented LLMs — where structured ontologies guide generation — have demonstrated strong hallucination reduction, achieving up to 98% factual accuracy in closed-domain settings compared to 37–52% for ungrounded LLMs [3]. This finding motivated our two-stage validation approach combining ontological reasoning with domain expert review.

These LLM-assisted approaches are particularly relevant for domain-specific ontology construction, such as in viticulture, where multiple formalisation efforts already exist.

Wine ontologies have been formalised through multiple efforts: the W3C OWL Wine Ontology serves as a canonical reasoning demonstration in the OWL specification[4], the UC Davis Wine Ontology covers varietal designations and wine classification properties for the broader wine industry [5], and the Wine Cloud Ontology introduces IoT and sensor semantics for precision viticulture applications [6]. Building on these ontological foundations, precision viticulture systems using IoT sensors for vineyard monitoring — such as commercial deployments — report expected improvements of up to 25% in grape quality and 30% in operational cost efficiency [7].

Nevertheless, to our knowledge, no ontology captures Kosovo’s regional viticultural context, and the domain of grape quality improvement in Kosovo remains largely unexplored in the research literature — motivating the present work.

III. METHODOLOGY

The ontology construction process began with a merging phase, in which we used Perplexity Pro as the primary LLM-assisted generation environment. All three referenced ontologies — the W3C OWL Wine Ontology, the UC Davis Wine Ontology, and the Wine Cloud Ontology — were provided as contextual input within the Perplexity session, enabling the model to reason across existing formalizations during generation.

These ontologies were identified through a targeted review of existing wine ontology initiatives. The W3C OWL Wine Ontology was selected as the primary starting point due to its standardized status, OWL expressivity, and its well-established coverage of core concepts such as Winery, Region, Grape, and Vintage. Its clarity and use in reasoning demonstrations make it a solid foundation for domain-specific extensions. The UC Davis Wine Ontology builds upon this foundation with U.S.-specific wine labelling conventions, Wikidata alignment, and wine type classifications — properties that are relevant for Kosovo’s export-oriented wine production. The Wine Cloud Ontology was included for its AgriTech orientation, introducing sensor data integration, drone monitoring, and blockchain traceability semantics — all relevant to the farm-to-table value chain targeted by the present project.

These three ontologies were subsequently extended with Kosovo-specific knowledge, comprising five viticultural re-

gions and thirteen grape varieties — seven red and six white — producing a lightweight, interoperable ontology that builds upon established wine knowledge representation for a Kosovo AgriTech context. The resulting ontology components were extracted in Turtle (.ttl) and OWL formats and imported into Protégé for structural validation, class hierarchy refinement, and reasoner-based consistency checking..

Thus, this new ontology, presented in Figure 1, reuses basic wine properties and classifications from the UC Davis Wine Ontology:

- classes: Wine, RedWine, WhiteWine.
- properties: color, VarietalDesignation.

Then it takes relations from W3C Wine Ontology:

- classes: Wine, Region, Winery, WineGrape.
- relationship: madeFromGrape, locatedIn, hasMaker.

Then again from Wine Cloud ontology taking modern precision agriculture concepts such as:

- classes: Parcels, Disease, WeatherSensor.
- properties: observes.

Then some concepts are aligned, such as:

- UC Davis: Wine and W3C: Wine, and
- W3C: Parcels and Wine Cloud: Region.

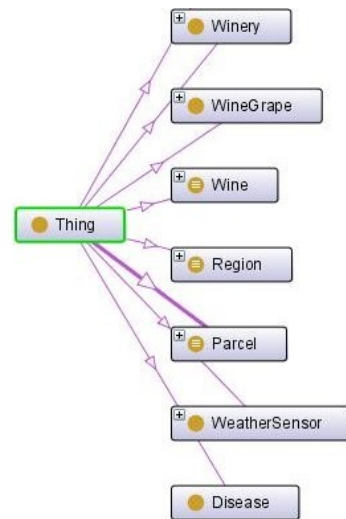


Fig. 1. Kosovo Wine Ontology.

IV. ONTOLOGY AND EVALUATION

The Kosovo Winery Ontology (KWO) is constructed by merging three foundational ontologies — the W3C Wine Ontology, the UC Davis Wine Ontology, and the Wine Cloud Ontology — into a unified knowledge graph extended with Kosovo-specific entities. Integration is achieved via OWL equivalence axioms and shared namespace prefixes, ensuring cross-ontology reasoning. Specifically, *ucd:Wine owl:equivalentClass w3c:Wine* aligns the UC Davis and

W3C class hierarchies, while *wc:Parcel owl:equivalentClass w3c:Region* bridges the Wine Cloud vineyard model with the W3C regional structure.

Core classes drawn from each source ontology are as follows: from W3C, *w3c:Wine*, *w3c:Region*, *w3c:WineGrape*, and *w3c:Winery*; from UC Davis, *ucd:RedWine* and *ucd:WhiteWine* as subclasses of *ucd:Wine*; and from Wine Cloud, *wc:Parcel*, *wc:WeatherSensor*, and *wc:Disease* to represent IoT-monitored vineyard infrastructure. Key object properties — *w3c:hasMaker*, *w3c:locatedIn*, and *w3c:madeFromGrape* — enable traceability across the production chain.

Kosovo-specific extensions introduce a *kos:namespace* containing local subclasses, individuals, and properties. Notable additions include *kos:Kosovo* as a *w3c:Region* instance, local grape varieties such as *kos:Vranac* and *kos:SheshiZi* as instances of *w3c:WineGrape*, and the region *kos:Rahovec* modelled as a *kos:KosovoRegion*. These extensions are validated in Protégé using the HermiT reasoner, which infers, for example, that *kos:RahovecVranac rdf:type ucd:RedWine*.

LLMs were employed to extract Kosovo-specific relations from local agricultural corpora and regulatory documents, populating the *kos:namespace* automatically. Extracted triples — such as *kos:RahovecWine w3c:locatedIn kos:Kosovo* — were validated against the merged schema before insertion. Ontology labels were further localized into Albanian using LLM-based translation, with *rdfs:label* annotations added for multilingual accessibility.

KWO is designed as the semantic backbone for future Agritech deployments in Kosovo. The integration of Wine Cloud’s IoT classes enables sensor-driven queries, such as identifying parcels at risk of downy mildew via *wc:Disease* instances linked to *wc:WeatherSensor* observations. Blockchain-based provenance can be anchored to *kos:Region* parcels, supporting tamper-proof traceability for Rahovec Vranac export compliance. The following SPARQL query illustrates a representative use case:

Listing 1: Example SPARQL query retrieving Kosovo wines made from Vranac grapes

```
SELECT ?wine
WHERE {
    ?wine w3c:madeFromGrape kos:Vranac .
    ?wine w3c:locatedIn kos:Kosovo .
}
```

KWO is evaluated along two dimensions: structural completeness, measuring the coverage of wine domain concepts relative to the three source ontologies; and logical coherence, verified through the absence of unsatisfiable classes under HermiT reasoner execution in Protégé.

The structural completeness evaluation demonstrates that KWO successfully integrates all three source ontologies while extending coverage with Kosovo-specific entities. In Table 1 we have summarized the structural metrics of KWO as reported by the Protégé ontology metrics panel. The merged ontology comprises 21 classes, 8 object properties enabling

rich semantic relationship, 5 data properties, and 26 individuals, producing 103 logical axioms and 252 total axioms. The 2 EquivalentClass axioms confirm the cross ontology alignment (*ucd:Wine* $\equiv$ *w3c:Wine* and *wc:Parcel* $\equiv$ *w3c:Region*). The 115 annotation assertions reflect extensive multilingual labelling in English and Albanian, demonstrating the LLM-assisted ontology localization contribution of this work. Logical coherence was validated using the HermiT reasoner in Protégé 5.6.5, with zero unsatisfiable classes detected across axioms.

A key risk in LLM-assisted ontology engineering is the generation of plausible but factually incorrect triples — a well-documented failure mode in KG construction. To mitigate possible hallucinations, all LLM-generated triples were validated in two stages: first against the merged OWL schema using the HermiT reasoner to detect logical inconsistencies, and second via manual spot-checking by a domain expert against Kosovo wine documents.

V. CONCLUSIONS

This paper presented the Kosovo Winery Ontology (KWO), a domain-specific knowledge graph constructed by integrating three established wine ontologies — the W3C Wine Ontology, the UC Davis Wine Ontology, and the Wine Cloud Ontology — and extending them with Kosovo-specific entities, relations, and multilingual labels. By leveraging LLMs for relation extraction, corpus annotation, and ontology localization, we demonstrated that the knowledge graph engineering pipeline can be substantially automated while maintaining logical coherence and domain coverage.

The three key contributions of this work are: (1) a merged, OWL-compliant winery ontology grounded in internationally recognised schemas; (2) an LLM-assisted pipeline for extracting and localising Kosovo-specific knowledge from agricultural corpora; and (3) a deployable semantic resource supporting agritech applications, including IoT-based vineyard monitoring, supply chain traceability, and blockchain-anchored provenance for Kosovo wine exports.

Our approach has several limitations that warrant acknowledgment. First, the Kosovo agricultural corpus is relatively small, potentially limiting the diversity of extracted relations. Second, while we believe we have reduced hallucinations through our two-stage validation, this still requires manual expert review, limiting full automation. Finally, the LLM extraction precision reported here reflects Perplexity Pro performance; smaller open-source models may yield different trade-offs between accuracy and deployment cost.

Future work will focus on three directions: (1) expanding the Kosovo agricultural corpus with ongoing regulatory updates and winery data to improve relation extraction coverage; (2) integrating dynamic sensor data from Wine Cloud-compatible IoT deployments in pilot vineyards in the Rahovec region through Agritech Project; and (3) applying the ontology as the semantic layer in a future EU-funded Agritech project targeting precision viticulture in Kosovo. We also plan to explore Graph-RAG approaches for knowledge-grounded

TABLE I
KWO STRUCTURAL METRICS

Axiom Type	Count	Description
ClassAssertion	26	Each individual typed to its class
ObjectPropertyAssertion	11	Relations between individuals (such as <code>madeFromGrape</code> , <code>observes</code>)
DataPropertyAssertion	24	Literal values (such as <code>color</code> , <code>exportStatus</code>)
AnnotationAssertion	115	Multilingual labels (EN/SQ) and Wikidata links
ObjectPropertyDomain/Range	8 / 8	Full domain and range coverage on all object properties

question answering over KWO, enabling end-users such as winemakers and agricultural inspectors to query the ontology in natural language. Finally, we will investigate fine-tuning smaller open-source LLMs on our validated Kosovo wine corpus to reduce dependency on proprietary models while maintaining extraction quality.

GENERATIVE AI DISCLOSURE STATEMENT

During the preparation of this work the author(s) used Perplexity in order to assist in ontology generation, grammar and spelling check. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

For example: During the preparation of this work the author(s) used ChatGPT in order to improve the language and clarity of the manuscript. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

ACKNOWLEDGMENT

This publication is supported by ITC Conference grant from COST Action CA23147 GOBLIN – Global Network on Large-Scale, Cross-domain and Multilingual Open Knowledge Graphs, supported by COST (European Cooperation in Science and Technology, <https://www.cost.eu>).

REFERENCES

- [1] H. Babaei Giglou, J. D'Souza, and S. Auer, "LLMs4OL 2024 Overview: The 1st Large Language Models for Ontology Learning Challenge," in *Open Conference Proceedings*, vol. 4, 2024. [Online]. Available: <https://www.tib-op.org/ojs/index.php/ocp/article/view/2473>
- [2] M. J. Saeedizade and E. Blomqvist, "Navigating Ontology Development with Large Language Models," in *The Semantic Web – ESWC 2024 (Lecture Notes in Computer Science)*, Springer, 2024. [Online]. Available: https://doi.org/10.1007/978-3-031-60626-7_8
- [3] S. El-Sappagh *et al.*, "Ontology-grounded knowledge graphs for mitigating hallucinations in LLM-based clinical question answering," *J. Biomed. Inform.*, vol. 126, p. 104171, 2026. [Online]. Available: <https://pubmed.ncbi.nlm.nih.gov/41610815/>
- [4] M. K. Smith, C. Welty, and D. L. McGuinness, "OWL Web Ontology Language Guide (Wine Ontology Example)," World Wide Web Consortium (W3C), W3C Recommendation, 2004. [Online]. Available: <https://www.w3.org/TR/owl-guide/>
- [5] UC Davis Library, "Wine Ontology Project," GitHub repository, 2016. [Online]. Available: <https://github.com/UCDavisLibrary/wine-ontology>
- [6] CheckSem, "The Wine Cloud Ontology," Online ontology resource, 2023. [Online]. Available: <https://ontology.winecloud.checksem.fr/>
- [7] Libelium, "Smart Wine: IoT Technology Allows Predictive Control of Vineyard Management," Libelium Case Studies, 2025. [Online]. Available: <https://www.libelium.com/libeliumworld/success-stories/smart-wine-libeliums-iot-technology-allows-predictive-control-of-vineyard-management/>