Constraining LLM-Based RDF Generation with Ontology-Guided Prompting in Urban Agriculture

Leonardo Vianna do Nascimento[1], **José Palazzo Moreira de Oliveira**[2]

[1]Instituto Federal de Educação, Ciência e Tecnologia (IFRS) – Alvorada – RS – Brazil

[2]Instituto de Informática – Universidade Federal do Rio Grande do Sul (UFRGS)
Caixa Postal 15.064 – 91.501-970 – Porto Alegre – RS – Brazil

leonardo.nascimento@alvorada.ifrs.edu.br, palazzo@inf.ufrgs.br

Abstract. *The construction and maintenance of knowledge graphs remain labor-intensive tasks, particularly when relevant information is available primarily in semi-structured textual sources. Recent advances in large language models (LLMs) offer new opportunities for automating knowledge graph population, yet generating RDF instances that comply with ontology constraints remains challenging. This paper investigates an ontology-guided prompting strategy for generating RDF knowledge graphs from textual descriptions of horticultural cultivars. The proposed approach embeds a domain ontology together with modeling patterns derived from established semantic standards and provides explicit schema definitions, anchored instances, and few-shot RDF/Turtle examples within the prompt. The method is evaluated on a dataset of 42 cultivar webpages using SHACL-based validation, precision–recall analysis, and hallucination detection. Results show that ontology-guided prompting substantially improves structural validity and semantic accuracy while significantly reducing hallucinated assertions.*

1. Introduction

Knowledge graphs have become a central infrastructure for representing structured knowledge in domains where heterogeneous information sources must be integrated and queried in a semantically meaningful way [Gong and Li 2025]. By organising entities, relations, and attributes according to explicit ontological schemas, knowledge graphs enable interoperability across datasets, support advanced reasoning capabilities, and facilitate the development of intelligent decision-support systems. In domains such as agriculture, where information about cultivars, environmental conditions, and cultivation practices is distributed across scientific publications, institutional repositories, and commercial documentation, knowledge graphs offer a promising mechanism for consolidating dispersed agronomic knowledge into machine-interpretable form.

Recent advances in large language models (LLMs) have opened new possibilities for automating knowledge graph population [Pan et al. 2024], yet applying LLMs to generate RDF knowledge graphs poses significant challenges: models frequently violate ontology constraints, introduce inconsistent vocabularies, or fabricate information absent from the source text [Ali et al. 2026] — issues particularly problematic in semantic knowledge bases where structural correctness and ontological consistency are essential for reliable querying and reasoning. A promising mitigation strategy is to embed explicit semantic constraints directly into the prompting process, incorporating ontology definitions, schema-level instructions, and structured examples that restrict the space of possible outputs, transforming ontologies from passive vocabularies into active control mechanisms shaping generation. However, systematic investigations of how such prompts can enforce ontology-compliant RDF generation remain limited, particularly in domain-specific contexts requiring the integration of multiple semantic standards.

In the agricultural domain, these challenges are particularly pronounced. Descriptions of plant cultivars typically include heterogeneous information such as morphological traits, environmental preferences, planting seasons, quantitative measurements, and geographic recommendations. Representing this information in a semantically consistent way requires integrating multiple ontologies and modelling patterns, including taxonomic relationships, quantity representations, temporal structures, and agronomic traits.

This paper investigates the use of ontology-guided prompting to enable the reliable population of knowledge graphs from semi-structured agricultural webpages. The proposed approach embeds a domain ontology for horticultural cultivars together with modelling patterns derived from widely adopted semantic standards for quantities, temporal information, and plant traits. By providing the large language model with explicit ontological context, reusable instance anchors, and few-shot examples expressed directly in RDF/Turtle, the prompt constrains the generation process so that extracted information can be serialized as ontology-compliant knowledge graph instances.

To evaluate the effectiveness of this strategy, we conducted experiments on a dataset of webpages describing horticultural cultivars published by commercial seed manufacturers. Generated RDF graphs were assessed along three complementary dimensions: structural validity with respect to SHACL constraints, semantic accuracy relative to manually constructed reference graphs, and the rate of hallucinated information unsupported by the source text. The results indicate that ontology-guided prompting substantially improves both the structural correctness and the factual reliability of LLM-generated RDF graphs when compared with vocabulary-guided extraction strategies.

The contributions of this work are threefold:

1. **An ontology-guided prompting strategy** that integrates domain and cross-domain ontologies to constrain RDF generation by large language models.
2. **A knowledge graph population workflow** capable of extracting agronomic information from semi-structured webpages and producing ontology-compliant RDF instances. For instance, the generated knowledge graph enables structured queries such as: '*which carrot cultivars tolerate pH between 5.5 and 6.5, have a winter harvest cycle under 90 days, and are recommended for planting in Southern Brazil in March?*' Such queries are not possible over the original semi-structured webpages.
3. **A multi-dimensional evaluation methodology** combining SHACL validation, comparison with reference graphs, and hallucination analysis to assess the reliability of LLM-generated knowledge graphs.

By demonstrating how structured prompts can operationalise semantic constraints during generation, this work contributes to the emerging intersection of knowledge graph engineering and large language models and provides practical insights for developing ontology-aware information extraction pipelines in agriculture and other data-intensive domains.

The remainder of this paper is organized as follows: Section 2 reviews related works; Section 3 presents the ontologies used in the prompt; Section 4 presents the design strategies; Section 5 describes the evaluation approaches and the obtained results; Section 6 discusses results; and Section 7 concludes.

2. Related Works

The automatic construction and population of KGs has long been a central topic in knowledge engineering and Semantic Web research. More recently, the use of LLMs to automate or semi-automate KG population from textual data has attracted growing attention [Pan et al. 2024].

[Caufield et al. 2024] propose the SPIRES method (Structured Prompt Interrogation and Recursive Extraction of Semantics), which employs zero-shot learning to extract information from text and populate knowledge bases according to predefined schemas. Similarly, [Norouzi et al. 2025] investigate the effectiveness of LLMs in populating the Enslaved.org knowledge graph, reporting that models such as GPT-4 can recover approximately 90% of the expected triples when guided by modular ontologies embedded in the prompt. These studies emphasize that performance improves significantly when the ontology schema is explicitly provided in the prompt—an approach also adopted in the present work.

From the perspective of RDF Turtle generation, [Frey et al. 2023] conducted a systematic evaluation of commercial LLMs (GPT-3.5, GPT-4, Claude 1.3, and Claude 2.0) regarding their ability to analyze, interpret, and generate knowledge graphs serialized in Turtle. Their findings indicate that newer models demonstrate improved proficiency in Turtle syntax compared to earlier versions, although challenges remain in strictly adhering to required output formats. These results underscore the importance of carefully designed prompts—such as those proposed here—that incorporate prefix declarations, class definitions, instance examples, and detailed mapping instructions to guide generation.

In the agricultural domain, [Wilson et al. 2024] propose, following a Design Science Research methodology, a semi-automated system for extracting agricultural ontological entities using LLMs such as GPT-4, with a focus on pest management information. Their study shows that combining few-shot learning with ontology context in the prompt substantially improves extraction accuracy. However, their approach does not address the generation of fully structured RDF Turtle integrating multiple ontologies, nor the extraction of semantically typed quantitative data.

Finally, [Ciatto et al. 2025] explore the use of LLMs as “oracles” for ontology instantiation with domain-specific knowledge. While demonstrating the feasibility of this strategy, they also highlight the risk of hallucinations, particularly when instances are assigned to incorrect classes. The approach proposed in this work mitigates this issue by supplying the model, within the prompt, not only with the T-box structure of the cultivar ontology but also with predefined instances of external entities (e.g., species, months, and geographic regions) and complete A-box examples in Turtle to constrain and guide generation.

Beyond the agricultural domain, ontology-guided prompting has been applied in biomedical [Ali et al. 2026], construction engineering [Zeng et al. 2026], and humanities [Norouzi et al. 2025] domains, consistently demonstrating that embedding ontological context within prompts improves structured output quality and reduces hallucinations — corroborating the hypothesis underlying the present work.

Overall, the above review suggests that, to the best of our knowledge, no existing approach simultaneously combines: (a) a cultivar ontology grounded in best-practice ontologies; (b) the integration of the Plant Trait Ontology for modeling plant characteristics; and (c) LLMs guided by structured prompts and Turtle examples to automate the extraction of data from commercial seed manufacturers' webpages into valid, interoperable RDF Turtle. The integration of these three dimensions constitutes the main and distinguishing contribution of the present work with respect to the state of the art.

3. Ontologies and Semantic Standards Employed

The proposed approach relies on the integration of multiple ontologies and semantic standards to ensure consistent, interoperable, and formally structured knowledge graph generation. Rather than adopting a single domain-specific vocabulary, the modeling strategy combines a domain ontology for agricultural cultivars with established upper-level, quantitative, temporal, and trait ontologies.

At the core of the modeling framework lies *COnto (Cultivar Ontology)*, a domain ontology developed by the authors to represent cultivars and their agronomic characteristics, whose RDF/Turtle serialization is publicly available online[1]. COnto defines the primary conceptual entities involved in horticultural knowledge representation. The central class, *Cultivar*, is described by properties covering identification (SKOS *prefLabel*), textual description (RDFS *comment*), taxonomic classification (*hasSpecies*), and agronomic relationships (*providesProduct*, *hasTrait*, *hasPreference*, and *hasManagement*). Associated classes include *PlantSpecies* for taxonomic species; *Product* for edible outputs, linked to harvest information via *hasHarvest*; *AnnualCropHarvest* for seasonal harvest timing; *CultivarPreference* for environmental and soil requirements; and *PlantingManagement*, *PlantingSeason*, and *PlantingMethod* for modeling planting practices, seasonality, and regional recommendations.

Several properties involve quantitative measurements, including planting spacing, root dimensions, temperature tolerance, soil pH ranges, and time to harvest. To represent such information in a standardized manner, the approach adopts the *Quantities, Units, Dimensions and Data Types Ontologies* (QUDT)[2], instantiating explicit *Quantity* individuals with structured quantity values, unit-aware modeling, and interval-based representations using minimum and maximum inclusive bounds. Temporal information is structured using the *W3C Time OntologyW3C Time Ontology*[3], with temporal entities explicitly instantiated and linked to planting season descriptions to enable formal temporal reasoning.

Morphological characteristics such as root shape, color, thickness, and length are

[1]`https://github.com/lvnascimento/sbbd2026/blob/main/conto.ttl`
[2]`https://www.qudt.org/`
[3]`https://www.w3.org/TR/owl-time/`

modeled using the *Plant Trait Ontology (TO)*[4], which provides controlled identifiers for phenotypic attributes, promoting semantic consistency with plant science vocabularies and facilitating integration with agronomic datasets. Regional planting recommendations are represented using the *Descriptive Ontology for Linguistic and Cognitive Engineering (DUL)*[5], particularly the *Place* concept and part–whole relations, enabling hierarchical geographic reasoning and structured spatial queries.

4. Prompt Design for Ontology-Guided RDF Generation

The prompt proposed in this work aims to instruct an LLM to perform, in a single interaction, two chained tasks: (i) to extract structured agronomic information from seed manufacturers' webpages; and (ii) to serialize this information as valid RDF instances in Turtle format, strictly compliant with the COnto ontology and the external ontologies integrated into it. The adopted approach follows a few-shot prompting strategy with full ontological context: rather than merely requesting data extraction, the prompt provides the model with the complete semantic framework required for the output to be directly usable as linked data.

The prompt was constructed following an iterative three-phase process: domain analysis of seed manufacturer webpages to identify the types of information to be extracted and the most appropriate ontological standard for each (quantities with units, regional seasonality, morphological traits); definition of six functional components — task definition and source restriction, class definitions and property mappings, external ontology mappings, few-shot Turtle examples, predefined shared instances, and prefix declarations — each corresponding to a strategy identified in the literature (ontology-based prompting, few-shot exemplification, instance anchoring); and iterative refinement through short cycles of RDF generation, content analysis, violation identification, and component adjustment, from which specific instructions such as the prohibition on duplicating quantity individuals emerged. The prompt components are described in next subsections.

4.1. Task and Source Definition

Figure 1 presents the first section of the prompt. The original prompt was written in Portuguese and is translated into English here.

The prompt begins by specifying the task in natural language: the model must access a provided URL and extract from it a explicitly enumerated set of agronomic attributes. The instruction that the information must be extracted "solely and exclusively from the referenced website, without using any other sources" is crucial: it prevents the model from filling gaps with its parametric knowledge, thereby reducing the risk of hallucinations that could contaminate the knowledge base.

This component also provides the definitions of the Brazilian geographic regions adopted in the model: Region 1 (all other states), Region 2 (São Paulo and Minas Gerais), and Region 3 (the Southern states). This enables the model to correctly interpret regional planting indications found in the sources and map them to the corresponding location instances in the ontology.

[4]`https://planteome.org/node/1`
[5]`http://ontologydesignpatterns.org/wiki/Ontology:DOLCE+DnS\_Ultralite`

I would like you to extract some information about the cultivars described on the following websites:

[LINKS TO WEB PAGES]

The information to be extracted includes: cultivar name, species name, description, name of the harvested product, cultivar characteristics (fruit/root color, leaf color, flower color, root or fruit dimensions (diameter and/or length), root or fruit shape, time to harvest (in days) in both summer (if available) and winter (if available), preferred ambient temperature and soil pH, planting months in different regions, and spacing between plants and rows.

I want you to extract this information solely and exclusively from the referenced website, without using any other sources. Consider Region 3 as the Southern states of Brazil, Region 2 as comprising São Paulo and Minas Gerais, and Region 1 as all other Brazilian states.

Figure 1. First section of the prompt.

4.2. COnto Ontology Classes Definition and Properties Mapping

Figure 2 details the second part of the prompt. The prompt reproduces, in commented RDF Turtle syntax, the definitions of the main classes of the CONTO ontology relevant to the task. Including the class definitions constitutes an *ontology-based prompting* strategy [Zeng et al. 2026]: by presenting the class hierarchy and its semantics, the prompt reduces ambiguity regarding how each type of information should be represented and in which class it should be instantiated.

Model the information below as RDF Turtle using the classes defined in the Conto ontology, specified below:

[CONTO ONTOLOGY TURTLE DEFINITION]

For individuals of the class Cultivar, use the following properties:

[DESCRIPTION OF Cultivar CLASS PROPERTIES]

For individuals of the class Product, use the following properties:

[DESCRIPTION OF Product CLASS PROPERTIES]

A single cultivar may have more than one product, such as roots, fruits, leaves, and flowers. Be aware that the websites mentioned above may provide details about additional products of the cultivar, such as edible flowers.

For harvest information, use the class AnnualCropHarvest for annual crops, with the property timeToHarvest_Summer to specify the time to harvest in summer and timeToHarvest_Winter to specify the time to harvest in winter. In both cases, use an individual of the QUDT Quantity class, following the example below **[QUANTITY EXAMPLES]**

For cultivar preferences, use subclasses of CultivarPreference. For each preference, associate a Quantity corresponding to a value range, as previously shown for root diameter, but linked to appropriate quantity kinds and units. Some useful quantity kinds may include: `quantitykind:Acidity`, `quantitykind:Temperature`. Some useful units may include: `unit:PH`, `unit:DEG_C`.

Regarding planting, create an individual of the PlantingManagement class in Conto with the following properties (all defined in Conto):

[DESCRIPTION OF PlantingManagement PROPERTIES]

For an individual of the PlantingSeason class, include the following properties:

[DESCRIPTION OF PlantingSeason PROPERTIES]

Figure 2. Second section of the prompt.

For each main class, the prompt explicitly specifies which properties must be used to generate instances and what each represents. Figure 3 presents an example of the specification of Product class properties.

- rdfs:label for the product name.

- hasHarvest from the Conto ontology to relate a product to its harvest information.

Figure 3. An example of properties definition for the Product class.

This explicit specification of properties plays a role analogous to that of a relational database schema: it reduces the model's degrees of freedom and makes the output structurally predictable, thereby facilitating the subsequent validation and integration of the generated triples.

4.3. Mapping of External Ontologies Classes

One of the central technical contributions of the prompt is the detailed instruction on how to use classes and properties from external ontologies to represent specific cultivar characteristics:

- The prompt maps each type of characteristic to a specific TO term — “to:0002710" (root shape), “to:0002628" (fruit shape), “to:0000065" (root color), “to:0002620" (exocarp color), “to:0002619" (mesocarp color), “to:0000537" (flower color), “to:0000326" (leaf color), “to:0000306" (root thickness), and “to:0000227" (root length). For each term, an instance is created and labeled with an appropriated description.
- For quantitative attributes (dimensions, harvest cycles, spacing, temperature preferences, and soil pH), the prompt instructs the use of the QUDT structure: “qudt:Quantity", “qudt:QuantityValue", “qudt:hasQuantityKind", “qudt:hasUnit", “qudt:minInclusive", and “qudt:maxInclusive". Recommended quantity kinds (“quantitykind:Diameter", “quantitykind:Length", “quantitykind:Time", “quantitykind:Temperature", “quantitykind:Acidity") and units of measurement (“unit:CentiM", “unit:DAY", “unit:DEG_C", “unit:PH") are provided for each type of parameter;
- For representing planting seasons, the prompt instructs the use of “time:hasTime", linking to predefined instances of "time:Instant" corresponding to each month of the Gregorian calendar;
- The property “conto:hasLocation" is instructed to point to predefined instances of “dul:Place" representing geographic regions.

This layer of the prompt is particularly relevant because it requires the model to reason about the mapping between attributes extracted from natural language text (e.g., “length of 17 to 19 cm”) and multi-level semantic structures involving chained instances of QUDT classes.

4.4. Complete Examples in RDF/Turtle (Few-Shot Examples)

For each type of complex structure—quantity intervals, time-to-harvest instances, and cultivar preferences—the prompt provides at least one complete and correct example in Turtle. These examples cover the instantiation of quantities and value ranges (using the

QUDT Quantity class with the *minInclusive* and *maxInclusive* properties), as well as examples of instances of classes defined in the TO for modeling characteristics. Such examples serve as *format anchors* that the model can follow when generating analogous instances for the cultivar under analysis.

The prompt also includes a final instruction—"*Do not duplicate quantity individuals when certain properties share the same values*"—which serves as a quality-control directive. This guideline encourages the reuse of instances, particularly for quantities (which often recur across similar cultivars), thereby avoiding redundancy in the generated A-box and promoting normalization of the resulting knowledge graph.

4.5. Anchoring Instances and Prefix Specification

To ensure consistency and reusability across separate generation sessions, the prompt provides a set of predefined instances that the model must reference rather than create: twelve named individuals of the OWL Time *Instant* class associated with Gregorian calendar months; individuals of the DUL *Place* class linked via *hasPart* to the corresponding Brazilian states; and individuals of COnto *PlantSpecies*, each annotated with scientific name (SKOS *prefLabel*), regional common names (SKOS *altLabel*), and a textual description. This anchoring strategy constrains the model's space of choices for species and location associations, preventing the creation of spurious instances inconsistent with the ontology's T-box and promoting A-box coherence across executions involving different cultivars.

The prompt concludes with the complete declaration of namespace prefixes to be used in Turtle serialization. Explicit prefix specification prevents the model from inventing or varying namespaces across different generations — a problem documented in the literature on Turtle generation by LLMs [Frey et al. 2023].

5. Evaluation

This section presents a comprehensive evaluation of the proposed ontology-guided prompting strategy. The assessment is structured along three complementary dimensions: (i) structural validity, (ii) semantic accuracy, and (iii) hallucination rate. The goal is to measure not only syntactic correctness but also ontology compliance and factual reliability.

5.1. Experimental Setup

Although the dataset comprises 42 webpages, it represents a substantial volume of structured agronomic information and constitutes a meaningful sample for evaluating the proposed approach. Each webpage provides a detailed description of a cultivar, including multiple categories of agronomic attributes such as cultivar identification, botanical species, textual description, harvested product, morphological traits (e.g., color, shape, and dimensions), environmental preferences (temperature and soil pH), planting management parameters (plant and row spacing), planting season across geographic regions, and time-to-harvest information. When transformed into RDF according to the ontology modeling patterns adopted in this work, each webpage typically generates dozens of interconnected triples involving multiple classes and relations across several ontologies. Consequently, the full dataset results in a knowledge graph composed of hundreds to thousands of RDF triples, including complex structures such as quantity intervals, temporal entities, and trait instances. From the perspective of evaluation, the unit of analysis is therefore not merely the number of webpages but the number and diversity of structured assertions extracted from them. Moreover, the selected webpages originate from commercial seed manufacturers and follow a relatively consistent genre of agronomic description commonly used in horticultural documentation. As a result, the dataset captures a representative range of attributes and modeling patterns required for cultivar representation, including quantitative measurements, categorical traits, and regional planting recommendations. A RDF file was manually annotated to produce a gold-standard RDF graph from the selected dataset using COnto and the integrated ontologies[6].

Two prompting configurations were evaluated:

- *Baseline (Vocabulary-Guided)* - The baseline configuration, referred to as vocabulary-guided prompting, provides the LLM with the COnto ontology vocabulary — including class and property identifiers — but omits the structural components that characterize the proposed approach: no modeling patterns for external ontologies (QUDT, W3C Time, Plant Trait Ontology, DUL) are specified, no few-shot RDF/Turtle examples are included, no shared anchor instances for months, geographic regions, or plant species are predefined, and no prefix declarations are enforced. In this configuration, the model is expected to infer how to structure and serialize the extracted information solely from the vocabulary identifiers, without explicit guidance on instantiation patterns or output format. This setup isolates the contribution of the ontology-guided components by keeping the underlying LLM, the source webpages, and the extraction task identical across both configurations.[7]
- *Proposed Ontology-Guided Prompt* - The full prompt design incorporating explicit modeling patterns for quantities (QUDT), temporal entities (W3C Time), trait alignment (Plant Trait Ontology), and structural constraints[8].

All experiments were conducted using the same LLM client[9] and configuration to isolate the effect of prompt design. It is important to note that the present experiments are intended as a feasibility validation of the proposed prompting strategy rather than a large-scale industrial benchmark. The consistent and expressive results obtained — including full SHACL conformity and near-perfect precision and recall — indicate that the dataset, despite its moderate size, was sufficient to demonstrate the effect of ontology-guided prompting with statistical clarity. Large-scale evaluations involving broader cultivar datasets, additional seed suppliers, multiple domains, and diverse LLMs are planned as future work.

5.2. Structural Validation with SHACL

Structural validity was assessed using SHACL shapes derived from COnto and the integrated ontologies[10]. The validation process was implemented using the *pySHACL* engine

[6] `https://github.com/lvnascimento/sbbd2026/blob/main/cultivares_groundtruth.ttl`

[7] `https://github.com/lvnascimento/sbbd2026/blob/main/baseline_prompt.txt`

[8] `https://github.com/lvnascimento/sbbd2026/blob/main/prompt.txt`

[9] `https://claude.ai/`

[10] `https://github.com/lvnascimento/sbbd2026/blob/main/shacl.ttl`

with RDFS inference enabled. This dimension evaluates how effectively each prompting strategy enforces ontology-compliant structure.

For each generated RDF graph, we computed:

- *Conformity*: whether graphs fully conforming to all SHACL constraints.
- *Violation Count*: total number of SHACL validation errors.

5.3. Semantic Accuracy

Semantic accuracy was measured by comparing generated RDF graphs against manually constructed gold-standard graphs. This evaluation dimension measures the factual correctness and completeness of the generated knowledge graphs.

Let:

- G_{gen} be the set of generated triples.
- G_{gold} be the set of gold-standard triples.

We compute:

$$Precision = \frac{|G_{gen} \cap G_{gold}|}{|G_{gen}|} \quad (1) \qquad Recall = \frac{|G_{gen} \cap G_{gold}|}{|G_{gold}|} \quad (2)$$

Before comparison, graphs were normalized to account for namespace expansion and literal normalization.

5.4. Hallucination Analysis

Hallucination was defined as the presence of RDF triples that are not supported by the source webpage content. We distinguish three categories:

- *Ontological Hallucination* - Use of predicates or classes not defined in COnto or integrated ontologies.
- *Factual Hallucination* - Assertions not supported by the webpage text.
- *Numerical Hallucination* - Fabricated quantitative values or ranges.

The hallucination rate is computed as:

$$HallucinationRate = \frac{|G_{unsupported}|}{|G_{gen}|} \quad (3)$$

where $G_{unsupported} \subseteq G_{gen}$ denotes triples not supported by the source text.

5.5. Results

Table 1 summarizes the comparative results across the three prompting strategies.

Table 1. Results obtained after prompts evaluation.

Prompt	SHACL Conf.	Viol. Count	Precision	Recall	Halluc. Rate
Baseline	No	188	0.35	0.10	0.65
Our Approach	No	54	0.97	0.96	0.03

Preliminary observations indicate that:

- The ontology-guided prompt significantly improves SHACL conformity.
- Precision increases as structural constraints reduce ambiguous predicate usage.
- Hallucination rates decrease due to vocabulary grounding and modeling restrictions.
- The baseline exhibits the highest structural violations and predicate variability.

6. Discussion

The results demonstrate that ontology-guided prompting substantially enhances both structural compliance and factual reliability of RDF generation. The approach integrates three complementary strategies from the literature. First, ontology-based prompting: rather than merely instructing the model on what to extract, the prompt supplies the complete semantic schema — classes, properties, hierarchies, and external ontologies — within which extraction must be articulated, significantly improving triple precision [Norouzi et al. 2025]. Second, few-shot exemplification: complete Turtle examples for complex structures serve as syntactic and semantic anchors, reducing formatting errors and increasing structural conformity [Brown et al. 2020, Zeng et al. 2026]. Third, anchoring of shared instances: predefined IRIs for months, regions, and species prevent the fragmentation of the graph caused by redundant instance creation across separate generation sessions, promoting A-box coherence for SPARQL queries.

Several limitations should be acknowledged. The current prompt was manually designed to ensure transparency and reproducibility. As ontologies evolve, however, maintaining prompt alignment introduces maintenance overhead. Future work could address this through automatic prompt generation from OWL/RDFS specifications, template-based synthesis tied to ontology versioning pipelines, and dynamic retrieval of relevant ontology fragments at inference time. Prompt sensitivity to LLM version is also a relevant practical concern; while the strategy is model-agnostic in principle, systematic evaluation across multiple LLMs and providers would be necessary to characterize its robustness under different generation configurations.

The baseline was deliberately designed to isolate the effect of ontological grounding under controlled conditions. Comparisons against traditional Information Extraction pipelines — rule-based systems, BERT-based relation extraction, or OpenIE frameworks — constitute a complementary research direction that would require additional annotation infrastructure and output format harmonization beyond the scope of this feasibility study.

The main scalability challenge lies in prompt length and token consumption, which grow with ontology complexity and the number of few-shot examples. Several mitigation strategies are available: ontology modularization, retrieval-augmented ontology injection, embedding-based schema selection, and reduced few-shot sets for simpler structures. The present study did not conduct a systematic cost analysis per triple generated; the approach relies on a single LLM call per webpage, with cost dominated by prompt length rather than output volume, meaning the effective cost per triple decreases with source information density. A rigorous cost-benefit comparison against smaller fine-tuned models remains an important direction for future work.

The current dataset is drawn from a limited number of commercial seed manufacturers. Since the prompt encodes constraints at the ontological schema level rather than at the level of any specific source format, the approach is in principle source-agnostic.

Broader validation across manufacturers from different countries, languages, and documentation traditions is planned. Cross-source integration challenges — conflict resolution, provenance management, confidence estimation, and semantic reconciliation — fall outside the scope of this initial study and are identified as future research directions, with potential extensions including provenance metadata, trust scoring, and SHACL-based cross-source consistency rules.

7. Conclusion

This work investigated how large language models (LLMs) can be systematically guided to generate ontology-compliant RDF knowledge graphs from semi-structured textual sources. Focusing on horticultural cultivar descriptions published on seed manufacturers' webpages, we proposed an ontology-guided prompting strategy that embeds explicit schema definitions, modeling patterns, and anchored instances directly within the prompt.

To assess the reliability of the generated knowledge graphs, we introduced a multi-dimensional evaluation framework combining SHACL-based structural validation, comparison with manually constructed reference graphs, and analysis of hallucinated assertions. The experimental results obtained from a dataset of 42 cultivar webpages indicate that ontology-guided prompting substantially improves the structural validity and semantic accuracy of the generated RDF graphs while significantly reducing unsupported statements. These findings suggest that embedding ontological constraints within prompt design can effectively transform LLMs from generic text generators into controlled generators of semantically consistent knowledge graph instances.

Beyond the specific agricultural application investigated here, the proposed methodology illustrates how prompt engineering can operationalize semantic constraints during the generation process, thereby supporting the integration of large language models into knowledge graph engineering workflows. In this sense, the approach contributes both a practical strategy for ontology-guided information extraction and a reusable evaluation methodology for assessing the quality of LLM-generated RDF graphs.

Future work will focus on expanding the dataset of cultivars and seed suppliers. It also will explore ontology modularization, prompt compression, and retrieval-augmented ontology injection strategies to assess the robustness and scalability of the approach at an industrial scale, including an analysis of token consumption trade-offs relative to output quality.

An additional direction of interest concerns the integration of the generated knowledge graph into a retrieval-augmented generation architecture. In such a setting, verified RDF instances produced by the ontology-guided pipeline could serve as a dynamic retrieval source for subsequent extraction prompts, providing corpus-specific few-shot examples and constraining the generation space for new cultivars.

Acknowledgments. This study was supported by National Council for Scientific and Technological Development (CNPq) through CNPq/MCTI Nº 10/2023 grant n. 402086/2023-6.

References

Ali, M., Taha, Z., and Morsey, M. M. (2026). Ontology-grounded knowledge graphs for mitigating hallucinations in large language models for clinical question answering. *Journal of Biomedical Informatics*, 175:104993.

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., et al. (2020). Language models are few-shot learners. *Advances in neural information processing systems*, 33:1877–1901.

Caufield, J. H., Hegde, H., Emonet, V., Harris, N. L., Joachimiak, M. P., Matentzoglu, N., Kim, H., Moxon, S., Reese, J. T., Haendel, M. A., et al. (2024). Structured prompt interrogation and recursive extraction of semantics (spires): A method for populating knowledge bases using zero-shot learning. *Bioinformatics*, 40(3):btae104.

Ciatto, G., Agiollo, A., Magnini, M., and Omicini, A. (2025). Large language models as oracles for instantiating ontologies with domain-specific knowledge. *Knowledge-based systems*, 310:112940.

Frey, J., Meyer, L.-P., Arndt, N., Brei, F., and Bulert, K. (2023). Benchmarking the abilities of large language models for rdf knowledge graph creation and comprehension: how well do llms speak turtle? In *ISWC: Workshop Deep Learning for Knowledge Graphs*.

Gong, R. and Li, X. (2025). The application progress and research trends of knowledge graphs and large language models in agriculture. *Computers and electronics in agriculture*, 235:110396.

Norouzi, S. S., Barua, A., Christou, A., Gautam, N., Eells, A., Hitzler, P., and Shimizu, C. (2025). Ontology population using llms. In *Handbook on Neurosymbolic AI and Knowledge Graphs*, pages 421–438. IOS Press.

Pan, S., Luo, L., Wang, Y., Chen, C., Wang, J., and Wu, X. (2024). Unifying large language models and knowledge graphs: A roadmap. *IEEE Transactions on Knowledge and Data Engineering*, 36(7):3580–3599.

Wilson, S., Ginige, A., and Goonatilake, J. (2024). Design science research approach for ontology development in agriculture: Utilising advances of llm for automated entity extraction.

Zeng, C., Hartmann, T., and Ma, L. (2026). Ontology-based prompting with large language models for inferring construction activities from construction images. *Advanced Engineering Informatics*, 69:103869.