ELSA Knowledge Graphs for Animal Treatment Recommendations

Varsha Kalidas[1,*], André Gomes Regino[2,3], Anderson Rossanez[2], Julio Cesar dos Reis[2], Tarek Alskaif[1] and Ricardo da Silva Torres[1]

[1]*Wageningen University and Research, the Netherlands, e-mail: varsha.kalidas@wur.nl, tarek.alskaif@wur.nl, ricardo.dasilvatorres@wur.nl*

[2]*Institute of Computing, University of Campinas, Brazil, e-mail: andre.regino@ic.unicamp.br, anderson.rossanez@ic.unicamp.br, jreis@ic.unicamp.br*

[3]*DIMEC, Center for Information Technology Renato Archer – Brazil, e-mail: aregino@cti.gov.br*

Abstract

Despite the potential of Knowledge Graphs (KGs) in enhancing treatment recommender systems (RSs), there is a significant research gap in developing ontologies and KGs that effectively capture ethical, legal, and societal aspects (ELSA) in healthcare RSs. This study explores the use of Large Language Models (LLMs) to assist in generating KGs that incorporate ELSA for a dairy-cow treatment RS. Our approach involves the generation of five KGs: for cows (the users); treatments (the items); ethical considerations; legal requirements, and social preferences. The KG generation process is semi-automatic, with LLMs aiding in generating structured tuples based on predefined ontologies. Researchers manually curate the information to ensure accuracy and relevance. Our study focuses on the ontological modeling of the KGs, demonstrating how entities and relationships are defined to capture the complex interplay between veterinary treatments and ELSA compliance. Use-inspired scenarios illustrate the generation of KGs based on the ontologies tailored to Dutch dairy production. The resulting KGs provide a foundation for developing dairy treatment RSs that consider ELSA. Our results contribute to integrating domain expertise with AI tools to create knowledge structures supporting informed and ELSA-compliant livestock management decision-making.

Keywords

Knowledge graphs, Large Language Models (LLMs), Ethical, Legal, Societal Aspects (ELSA), Recommender Systems, Animal Treatment

1. Introduction

Treatment recommender systems (RSs) [1] have emerged as powerful tools in healthcare, leveraging data-driven approaches to suggest personalized therapies for patients. These systems enhance clinical decision-making and improve patient outcomes by analyzing vast medical data and patient information to provide tailored treatment recommendations.

In recent years, the integration of Ethical, Legal, and Societal Aspects (ELSA) [2] has gained significant relevance in the development of trustworthy RSs [3, 4]. ELSA considerations ensure that artificial intelligence (AI)–driven recommendations align with ethical concerns, legal requirements, and societal preferences. This is particularly crucial in healthcare applications, where treatment decisions can profoundly impact patients' lives. Ensuring ELSA compliance in RSs has become increasingly important, especially with the introduction of legislative frameworks such as the EU AI Act and data protection regulations like the GDPR. Compliance protects users and fosters trust in healthcare RSs.

Knowledge Graphs (KGs) have emerged as a powerful tool for representing complex, interconnected information in a structured manner [5]. They capture diverse entities and relationships, making them

KG-STAR@ECMS25: Knowledge Graphs for Responsible AI Workshop in conjunction with ESWC25, June 1, 2025, Portoroz, Slovenia

*Corresponding author.

varsha.kalidas@wur.nl (V. Kalidas); andre.regino@ic.unicamp.br (A. G. Regino); anderson.rossanez@ic.unicamp.br (A. Rossanez); jreis@ic.unicamp.br (J. C. d. Reis); tarek.alskaif@wur.nl (T. Alskaif); ricardo.dasilvatorres@wur.nl (R. d. S. Torres)

0009-0006-4957-0743 (V. Kalidas); 0000-0001-9814-1482 (A. G. Regino); 0000-0001-7103-4281 (A. Rossanez); 0000-0002-8538-7261 (J. C. d. Reis); 0000-0002-1780-4553 (T. Alskaif); 0000-0001-9772-263X (R. d. S. Torres)

particularly suitable for modelling the multifaceted nature of healthcare and ELSA. A key advantage of KGs is their machine-readable format. It enables computational systems to store data and reason over interconnected data, perform complex inferences, and extract meaningful insights through semantic analysis. KGs are constructed relying on Resource Description Framework (RDF) triples through data extraction, transformation, and loading. RDF triples consist of subject-predicate-object statements and can be collected from various sources such as websites, existing semantic markups, or unstructured text documents [6]. Recently, Large Language Models (LLMs) [7] appeared promising in the data transformation process of KG construction [8, 9] due to their advanced natural language understanding and prompt-informed generation capabilities [10].

Ontologies are crucial computer artifacts in modelling domain knowledge, such as ELSA, for creating KGs. They provide a formal, shared domain conceptualization, allowing for consistent knowledge representation [11, 12]. This semantically structured approach enables advanced reasoning and decision-making in RSs. Ethical principles, legal requirements, and social preferences or norms can be modelled using ontologies to create KGs. Moreover, existing ontologies can be leveraged from domains such as healthcare, ethics, and legal studies to accelerate KG development and ensure alignment with established standards.

Despite the potential of KGs in enhancing ELSA-compliant treatment RSs, there is a significant research gap in developing comprehensive ontologies and KGs that effectively capture the complex interplay between ELSA and treatments in healthcare decision-support systems. Addressing this gap would involve creating robust ELSA-focused ontologies that assist RSs in providing ELSA-compliant treatment recommendations. One challenge refers to the encoding of domain-specific ELSA into ontologies. This task requires aligning existing ontologies (concepts and relationships) from various fields (*e.g.*, medical and legal) or creating new ones to address existing gaps. Also, domain-specific ELSA knowledge is often scattered across different sources and encoded into unstructured documents (*e.g.*, legislation). Social aspects are primarily intangible and mutable [13], making it challenging to find relevant information digitally or otherwise. The encoding of evolving ELSA knowledge further poses challenges. In this case, created ontologies should enable the dynamic update of KGs as ELSA standards change over time.

This study proposes incorporating ELSA into treatment RSs by leveraging LLMs to generate KGs. In the context of treatment RSs for Dutch dairy production, we model ontologies to represent various aspects, such as animal health and regulatory compliance. We explore ontologies to define concepts related to cow health, treatment options, and their associated ethical and legal aspects. Use-inspired scenarios illustrate the generation of KGs based on the ontologies tailored to Dutch dairy production. In particular, we present and discuss challenges and lessons learned related to a comparative study addressing different strategies involving different LLM models and ontology-based prompt construction approaches for structuring ELSA-based knowledge into KGs.

The contributions of this article are threefold:

- Introduce domain-specific ELSA ontologies in healthcare RSs and validate the approach using a Dutch dairy production use case.
- Introduce use-inspired scenarios by illustrating KGs created based on developed ontologies.
- Demonstrate a comparative study including different LLMs and prompt design construction strategies for generating ELSA-related RDF triples in our KGs.

Our findings suggest that the integration of domain-specific ELSA ontologies with LLMs is a promising direction for structuring relevant ELSA knowledge into the implementation of KG-based RSs [14, 15]. These findings provide valuable insights to developers, stakeholders, and practitioners interested in creating ELSA-compliant RS. Particularly in dairy production, such RSs recommend effective treatments and ensure compliance with veterinary regulations, animal welfare standards, and ELSA in dairy farming.

This article is organized as follows: Section 2 describes related work and introduces relevant background concepts; Section 3 describes the proposed pipeline for ELSA KG construction, highlighting

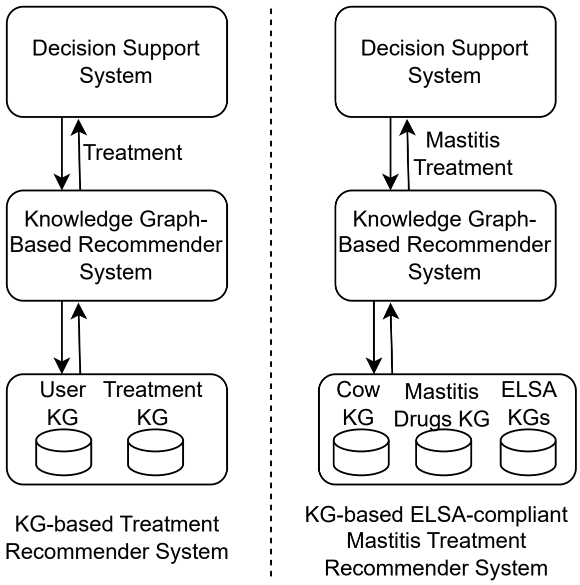


Figure 1: Conceptual view of a Treatment RS (left) vs. a Mastitis Treatment RS (right).

the modelling of ontologies derived from a dairy production case; Section 4 presents the validation procedures and results of the modelling; Section 5 discusses the findings. Section 6 draws conclusion remarks and points out directions for future work.

2. Background and Related Work

2.1. KG-based Recommender Systems Including for Mastitis Treatment

Fig. 1 presents the conceptual view of a treatment recommendation system (left) and a mastitis treatment recommendation system (right). Both systems rely on KG-based recommender systems, which explore KGs in encoding user and treatment (item) properties and their previous interaction [14]. The reader may refer to [14, 16] for an overview of KG-based recommendation systems. In particular, our study focuses on modelling and creating ELSA-related KG aiming for their use in a KG-based mastitis treatment recommender system, a problem overlooked in the literature.

Bovine mastitis is a painful mammary gland inflammation due to physical trauma or microorganism infections. Despite the advances concerning diagnostic procedures, mastitis is still the most frequent and costly disease in dairy farms, with significant negative impacts on profitability and animal welfare [17]. Mastitis is also the primary driver of antibiotic usage in dairy animals. It demands judicious and responsible decision-making regarding its use in livestock production in the face of emerging pathogen resistance [18]. Searching for suitable treatments demands the identification of their efficacy considering multiple variables and dimensions (*e.g.*, animal welfare, operational costs, estimated profits, and environmental conditions). Currently, farmers often rely on their past experiences when they make decisions [13]. In our formulation, the goal is to recommend mastitis treatments (*e.g.*, drugs) to farmers and veterinarians, taking into account domain-specific ELSA (*e.g.*, ethical issues related to animal welfare or existing regulations related to the use of antibiotics).

2.2. ELSA in Recommender Systems

In recent years, the integration of trustworthy principles or ELSA, such as Safety, Robustness, Non-discrimination, Fairness, Explainability, Privacy, Environmental Well-being, and Accountability, has gained significant relevance in the development of trustworthy RSs [3, 4, 19]. However, as noted in the survey by Ge *et al.* [4], research into trustworthy RSs often has a narrow focus on singular aspects, such as fairness or privacy. The study by Wang et al. [3] is one of the first that presents a conceptual framework to support the construction of trustworthy RSs.

Research on building RSs with a broad and comprehensive focus on all critical aspects is still limited [4]. Existing approaches have a global or overall RS-level view when integrating ELSA and do not focus on local or individual recommendations. Additionally, the practice seems to operationalize an aspect of trustworthiness by translating it to a mathematical metric [20], such as a fairness score. Unlike existing approaches, our solution presents a practical way of integrating multiple domain-specific ELSA into an RS by encoding the ELSA into KGs instead of metrics. In addition, we focus on ELSA at the recommendation level.

2.3. KG Creation based on Large Language Models (LLMs)

Bosselut *et al.* [21] developed Commonsense Transformers to generate common-sense knowledge graphs, thereby demonstrating the ability of pre-trained language models to generate common-sense descriptions in a loosely structured format. In another venue, Melnyk *et al.* [8] extracted KGs from the New York Times and TekGen datasets in a two-stage approach, where a pre-trained LLM extracts nodes and an edge construction head defines the edges. Zhang and Soh [9], in turn, demonstrated the generation of triples for large texts even when the ontology is unavailable or too large to fit into the context window of an LLM. They used the WebNLG, REBEL, and Wiki-NRE datasets. In another venue, Zhang *et al.* [10] fine-tuned LLMs to generate triples on public datasets, such as WebNLG, SKE, DocRed, FewRel, and KELM.

In contrast to the work of Bosselut *et al.* [21], we generate triples in a canonical format. Different from Melnyk *et al.* [8], we extract the entities and relations using an LLM. Similar to Zhang and Soh [9], we experiment with generating triples with or without the inclusion of the ontology. Unlike Zhang *et al.* [10], we do not fine-tune the LLM; we use it in its original pre-trained form by exploring in-context learning. Distinct from all these initiatives, we originally focused on creating KGs related to veterinary healthcare.

In this work, we follow the study by Regino *et al.* [22], who proposed an LLM-based approach for triple generation toward creating e-commerce KGs. Their approach comprised four steps: (i) identification of text source type, (ii) extraction of relevant information from the text source, (iii) generation of RDF triples based on identified information, and (iv) validation of generated triples. LLMs are employed in the first three steps. Different from their study, we focus on the LLM-guided sentence simplification and triple-generation problems.

3. ELSA KG Generation and Ontology Modelling

This section introduces the proposed method for ELSA KG generation and the ontological modeling of KGs.

3.1. Conceptual View

Fig. 2 presents our proposed LLM-based pipeline to create ELSA-related triples inspired by the work of Regino *et al.* [22]. The pipeline comprises three components: (1) Sentence Simplification, (2) LLM-based Triple Generation, and (3) Triple Curation.

Figure 3 illustrates how an input text document is processed by each component of the pipeline. It shows the intermediary output generated by each component for the given example text and shows the resulting output of curated triples. The Sentence Simplification component composes simpler sentences comprising a single subject, object, and predicate relations. Qualifiers for the predicate are included in the simpler sentence to improve the efficiency of triple extraction. This way, instead of querying the model twice to generate the predicate and qualifier relations, it is efficient to generate these relations with one query. The Sentence Simplification component receives the unstructured text documents,

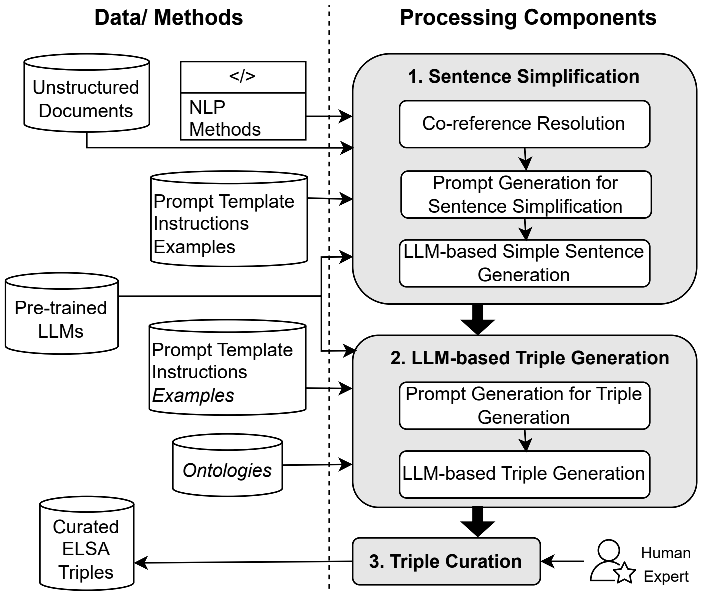


Figure 2: Illustration of the Triple Generation pipeline: The Sentence Simplification component transforms unstructured text documents into simpler sentences using NLP methods and LLMs. Then, the LLL-based Triple Generation component receives the simple sentences and generates triples using LLMs. Finally, the generated triples are curated by an expert to yield Curated Triples.

including ELSA, NLP methods, a prompt template, instructions, examples, and a pre-trained LLM as input. It produces a list of simpler sentences as output.

First, in the Co-reference Resolution sub-component, co-references in the document are resolved using NLP methods. Then, in the Prompt Generation for Sentence Simplification sub-component, a prompt for sentence simplification is formulated using the prompt template, instructions, and examples. The LLM-based Simple Sentence Generation sub-component uses this prompt to query the input LLM to generate simpler sentences. In Fig. 4 (left), the prompt template, including an example sentence, illustrates an example prompt used for sentence generation. Due to space considerations, only one example is provided in the figure.

The LLM-based Triple Generation component receives these simpler sentences, pre-trained LLMs, a prompt template, instructions, examples, and ontologies as input. It produces a set of triples as output.

In the Prompt Generation for Triple Generation sub-component, a triple-generation prompt is formulated using the prompt template, instructions, examples, and ontologies. Note that the examples and ontologies are optional. The LLM-based Triple Generation subcomponent uses this prompt to query the input LLM to generate the triples. In Fig. 4 (right) the prompt template, including an example sentence, illustrates an example prompt used for triple generation. Due to space considerations, only one example is provided in the figure.

The Triple Curation component is responsible for the curation of the triples. An expert reviews the generated triples and makes suitable modifications, making this a semi-automatic pipeline. This step can also be made automatic by employing a machine-learning-based curation component. Finally, the curated triples are stored in a triple store for further use in generating KGs.

3.2. Ontological Modelling of Knowledge Graphs

An ontology formally represents computer-interpretable knowledge within a specific domain, consisting of concepts or entities and their relationships [11, 12]. It refers to a directed-graphical structure where the nodes represent classes or entities and the edges represent the relationships or properties between the entities. The relationships describe how entities relate to each other. The definition of entities and relationships enables a shared understanding of a domain, enabling data integration, knowledge

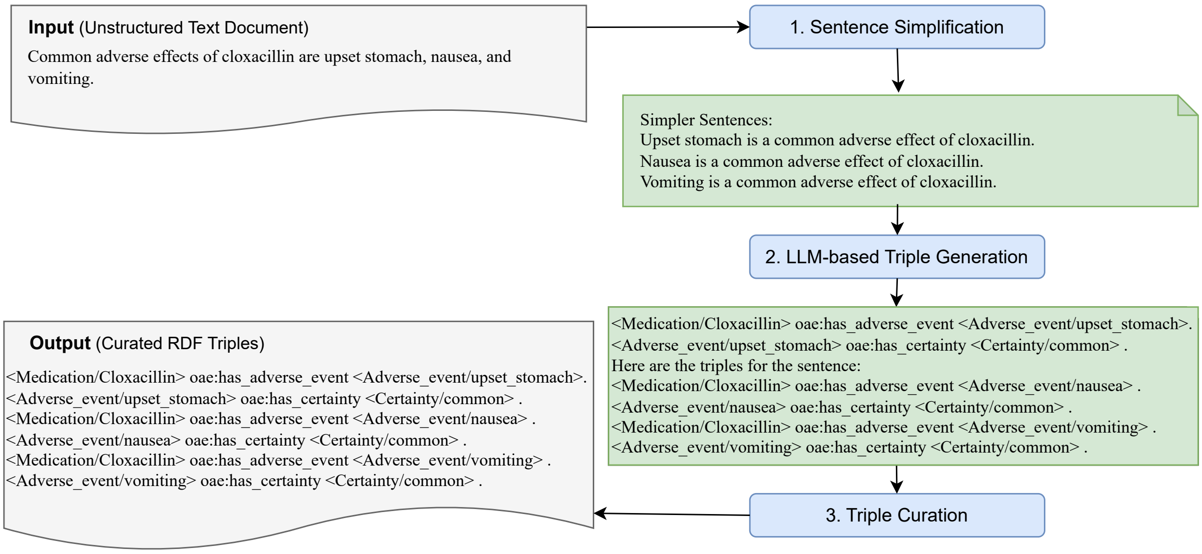


Figure 3: Illustration of the processing of input text documents in the form of unstructured text documents to curated ELSA triples, including the various intermediary outputs from each component of the processing pipeline.

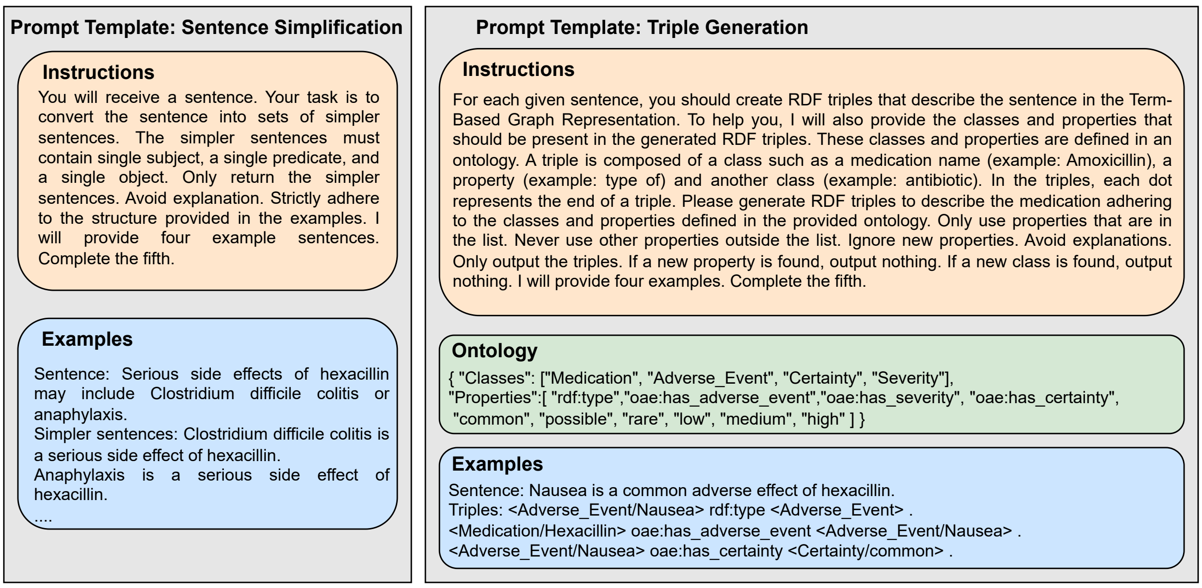
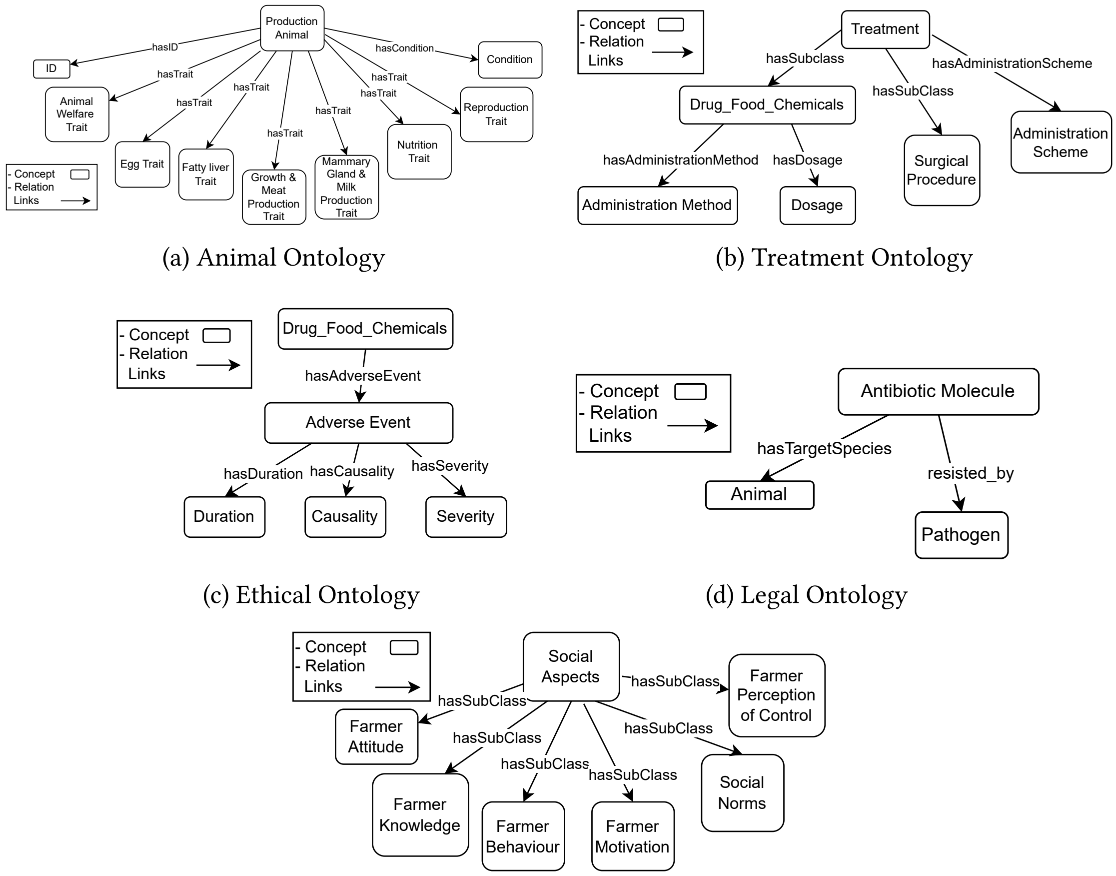


Figure 4: Illustration of the Prompt templates for LLM-based Sentence Simplification (left) and Triple Generation (right) tasks.

sharing, and automated reasoning. Ontologies can be reused across different domains and applications, allowing knowledge models to be adapted and extended to solve diverse challenges.

We elaborate on the ontologies modelled for the animals, treatments, and ethical, legal, and social aspects, relying on existing ontologies adapted and reused. They are explored for developing KGs in a treatment RS for mastitis in Dutch dairy cattle. We present a high-level conceptual model rather than a comprehensive ontology for clarity and brevity organized in the main modules concerning "Animal (Users)" (cf. Section 3.2.1), "Treatment (Items)" (cf. Section 3.2.2), and "Ethical, Legal, and Social" (cf. Section 3.2.3). Specific details are provided in the appropriate sections.



(a) Animal Ontology

(b) Treatment Ontology

(c) Ethical Ontology

(d) Legal Ontology

(e) Social Ontology

Figure 5: (a) Top-level classes and properties of the Production Animal ontology. (b) Top-level classes and properties of the Treatment ontology. (c) Top-level classes and properties of the ontology for the ethical aspects. (d) Top-level classes and properties of the ontology for the legal aspects. (e) Top-level classes of the ontology for the social aspects.

3.2.1. Animal (Users)

The "users" in our context are production animals in precision production systems because the treatment recommendations are provided for them. Farmers or veterinarians may also use the system itself. The essential data for an RS would be the animal-identifying characteristics such as the animal ID, age, weight, and historical health conditions. We start by using all production animal trait classes (animal welfare trait, fatty liver trait, egg trait, growth and meat production trait, mammary gland and milk production trait, nutrition trait, and reproduction trait) from the Animal Trait Ontology for Livestock (ATOL) by Golik et al. [23] available at Bioportal.[1] This ontology was selected because it provided a comprehensive class set describing different traits of production animals. This enables the extension of the system to other animals in the future. Next, we add the `Condition` class from the Disease-Treatment Ontology proposed by Khoo et al. [24] to represent additional conditions or patient attributes that may influence the treatment efficacy. This class includes the age, medical history, and other diseases that may be present. Finally, since production animals are likely to have a unique identifier, this is added as an RDF property `hasID`. The final ontology comprises 2356 classes and 6 property types. The ontological model for a production animal is presented in Fig. 5a.

[1]https://bioportal.bioontology.org/ontologies/ATOL (As of January 2025).

3.2.2. Treatment (Items)

The treatment model is modelled by re-using the `Treatment` class of the Disease-Treatment Ontology [24]. This ontology was used for its ability to systematically encode disease-treatment information into distinct classes (e.g., disease, treatment, condition, effect, and evidence), enabling precise representation of information from medical abstracts, similar to our case using textual information about medications. The ontological model for a treatment is presented in Fig. 5b. It has information about the drug dosage and administration, surgical procedures, and administration scheme as properties. Drug information is used from the `Drug Product` class from the Drug Ontology available at Bioportal.[2] The final ontology comprises 4 classes and 6 property types.

3.2.3. Ethical, Legal, and Social Knowledge Graphs

Ethical KG

Beneficence, non-maleficence, autonomy, and justice constitute the four main principles in clinical ethics [25, 26]. Beneficence is the obligation of the physician to act for the benefit of the patient [25]. Non-maleficence is the obligation of a physician not to harm the patient [25]. The principle of autonomy is interpreted as all persons having intrinsic and unconditional worth and therefore allowed the power to make rational decisions and moral choices, and each allowed to exercise his or her capacity for self-determination [27]. It does not extend to persons who can not act autonomously [25]. Since the use case involves animals, this principle is not extended to cows since they cannot act autonomously. A production animal cannot be expected to choose which treatment it would prefer. Therefore, autonomy is not part of the ontology. Justice is interpreted as fair, equitable, and appropriate treatment [25]. This principle can be addressed by other means, such as using fair recommendation algorithms and fair training data. Hence, it is not part of the ontology.

Practically, the application of beneficence and non-maleficence can be illustrated when a physician weighs the benefits against the burdens of all possible treatments and chooses the ones that are the best course of action for the patient, and refrains from using those that are inappropriately burdensome. A physician's obligation and intention to relieve the suffering of a patient by the use of appropriate drugs override the foreseen but unintended harmful effects or outcomes [25, 28]. Specifically for animals, the application of beneficence and non-maleficence corresponds to ensuring freedom from discomfort and freedom from pain, injury, or disease (see Five Freedoms for Animal Welfare[3]).

In the context of diseases in production animals, the RS is intended to recommend drugs for the benefit of the animal (beneficence). However, drugs may cause adverse events (AEs). By giving preference to drugs with the least AEs or ones with less severe AEs, the system can be made to be as non-maleficent as possible. Therefore, AEs and severity are included in the ontology. The drug information from the Drug Ontology used for the treatment KG is also used for this KG. For AEs, the Ontology of Adverse Events is available at Bioportal.[4] We reused the sub-classes `Adverse event`, `Causality`, `Duration`, and `Severity`. The `Adverse event` class provides a rich taxonomy of several AEs. The `Causality`, `Duration`, and `Severity` classes provide sub-classes to define the certainty of occurrence, duration (acute or chronic), and severity (six grades from GO, indicating an AE with sign and symptom within normal limits, to G5, indicating an AE that results in death) of AEs. The final ontology comprises 9113 classes and 4 property types. The ontological model for the ethical aspects is presented in Figure 5c.

Legal KG

The rules and regulations on veterinary medicinal products are established at the European level and set out in the Veterinary Medicinal Products Regulation.[5] The regulation is binding for all European

[2]https://bioportal.bioontology.org/ontologies/DRON?p=classes&conceptid=http%3A%2F%2Fpurl.obolibrary.org%2Fobo%2FDRON_00000005 (As of January 2025).
[3]https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=LEGISSUM:animal_welfare (As of February 2025).
[4]https://bioportal.bioontology.org/ontologies/OAE (As of January 2025).
[5]https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX%3A32019R0006&from=EN (As of January 2025).

Union member states and, therefore, applicable to the Netherlands. For our purposes, we found two aspects that could be modelled for use in the KG: the marketing authorization for veterinary medicinal products and the susceptibility of antimicrobial resistance (AMR) of a veterinary medicinal product. Only veterinary medicinal products that a Member State grants marketing authorisation to should be used in that Member state. The marketing authorisation also specifies the intended target species for the product. This information was added as an additional `has_target_species` property. Information about AMR would ensure the prudent use of products such as antimicrobials, avoiding their routine use and restricting the use of antimicrobials critical for preventing or treating life-threatening infections in humans. The Antibiotic Resistance Ontology (ARO) [29], available at Bioportal[6] was reused to model the AMR of drugs. The `antibiotic molecule` class was reused. A property relation `resisted_by` was added to indicate that an antibiotic is resisted by a pathogen. The final ontology comprises 662 classes and 2 property types. The ontological model for the legal aspects is presented in Figure 5d.

Social KG

For modelling the social aspects, since these are very context-specific, we looked into the available literature on social factors related to mastitis management on dairy farms. The key factors at play are farmer attitudes, knowledge, behaviour, social norms, and perception of control [30, 31, 32]. Farmer attitudes relate to the problem level bulk milk somatic cell count (BMSCC) used by the farmer for diagnosing mastitis, preferences for reducing the BMSCC, aspects of mastitis that the farmer finds annoying, such as increased labour or uncertainty of recovery [32]. Farmer knowledge relates to the knowledge level of the farmer concerning mastitis and the main sources of mastitis information [32]. Farmer behaviour relates to the particular farm management practices employed by the farmer, herd size, diagnosing practices, and information exchange sources [31, 32]. Social norms are factors that may influence farmers, such as extending antibiotic treatments for 'being a good farmer', perceived positive reference groups with whom the farmer identifies and regularly communicates with, such as other farmers, the herd veterinarian, and other farm advisors [33]. Different farmers have different motivations, such as avoiding penalties, achieving premiums, improving farm efficiency, or basic economic motivations [34]. Further, the costs of mastitis are different for different farmers and cannot be generalized. Even for similar cases of mastitis in an animal, the farm economics that influence management decisions are different [35]. The perception of control refers to how comfortable farmers feel about managing a disease. Some farmers may feel anxious because they believe they lack control over the situation, while others might experience frustration due to aspects that are challenging to manage. Furthermore, some may feel extremely anxious and attribute their struggles to bad luck [31]. The final ontology comprises 8 classes and 0 property types. The corresponding ontological model is shown in Figure 5e.

3.3. Knowledge Graph Generation using Large Language Models (LLMs)

In our approach to extracting RDF triples from natural language texts, LLMs were leveraged using zero-shot and few-shot prompting [36], along with the predefined ontologies. The instructions, ontology, and input sentences were provided for zero-shot prompting. For few-shot prompting, five to seven example sentences were provided in the prompting for each of the treatment, ethical, and legal KGs, along with the corresponding RDF triples that cover a wide range of the entity and relationship types relevant to that KG.

The list of ontology classes and properties was provided to the LLM to ensure alignment with the intended KG structure. A prompt template was constructed that included the examples (in the case of few-shot prompting) and the ontology, followed by the input text from which triples were to be extracted. The LLM was instructed to generate RDF triples for the new text in a format consistent with the provided examples and using only the specified ontology statements. Algorithm 1 formalizes our designed procedure.

[6] https://bioportal.bioontology.org/ontologies/ARO (As of January 2025).

Algorithm 1: Extraction of Knowledge Graph Triples using LLMs

Function *ExtractTriples*

Data: an Ontology $\mathcal{O}$ associated with KG, a pre-trained LLM model $\mathcal{L}$, a prompt base $\mathcal{B}$ that includes the instructions (and examples in the case of few-shot prompting) for the KG, and the input text document $\mathcal{I}$ consisting of a list of sentences

Result: Triples $\mathcal{T}$ associated with the input document

$\mathcal{T} \leftarrow \emptyset$

foreach $s \in \mathcal{I}$ **do**

$\mathcal{P} \leftarrow constructPrompt(\mathcal{B}, \mathcal{O}, s)$

$\mathcal{T} \leftarrow \mathcal{T} \bigcup getTriples(\mathcal{P}, \mathcal{L})$

return $\mathcal{T}$

4. Validation

This section describes the methodology used to validate the developed ontologies using real-world data and the results we achieved.

4.1. Methodology

The validation was guided by practical considerations and data availability, focusing on those aspects of the ontologies for which we could either generate or acquire data for input to component 1 of the Triple Generation pipeline (Fig. 2). Therefore, the KGs in this section may have fewer entities and relationship types than those described in the models in Section 3.2. For the sentence simplification and triple generation sub-components, state-of-the-art open-source LLMs were used. In particular, for the triple-generation sub-component, we perform a comparative analysis of two types of LLMs: general-purpose LLMs and LLMs specialized for reasoning tasks. The quality of the generated triples in terms of correctness and completeness was inspected by the researchers based on their knowledge. The state-of-the-art open-source general-purpose LLM (as of the 10th of February 2025), the llama-3.3-70b-versatile model[7], was used for the triple generation task. Additionally, another model specialized in reasoning tasks, the deepseek-r1-distill-llama-70b model [8], was used to compare the two models. The distilled version is also more efficient in terms of cost and energy. Therefore, the comparison would show whether a reasoning-specialized model would be more suited for the task. The temperature hyperparameter was set to 0.6 for both models to ensure a balance between deterministic and creative output. Table 1 provides a brief description of the models compared for triple generation.

Further, we also compare the performance of these models in three cases: prompts with only instructions, prompts with instructions and ontologies, and prompts with instructions, ontologies, and examples.

4.2. Data

For the cow production animal (user) KG, cow health data can be gathered from electronic veterinary health records (EVHRs) or milking robots. For the treatment, ethical, and legal KGs, we first compiled a list of drugs that are approved to be used to treat mastitis in the Netherlands[9] The resulting list comprised eleven single-compound drugs. For the sake of simplicity of KG construction, combinations of multiple drugs were excluded. Then the relevant drug information was collected from publicly

[7]https://console.groq.com/docs/model/llama-3.3-70b-versatile (As of April 2025).

[8]https://console.groq.com/docs/model/deepseek-r1-distill-llama-70b (As of April 2025).

[9]https://www.knmvd.nl/app/uploads/sites/4/2024/10/240815-formularium-melkvee-versie-1.10.pdf (See Page 25 for mastitis (in Dutch) – As of Dec. 2024).

Table 1
Brief description of the models compared for triple generation.

LLM Model	Description
llama-3.3-70b-versatile	• An LLM based on Llama-3.3-70B-Instruct, fine-tuned using supervised fine-tuning and reinforcement learning with human feedback to align outputs with human preferences for helpfulness and safety. Optimized for a wide range of tasks such as coding, reasoning, math, general knowledge tasks, instruction following, and tool use. Offers high performance across various benchmarks while maintaining efficiency. • Cost-wise, cheaper than models like GPT-4 but still expensive compared to distilled models [37]. • The dense parameter structure also consumes more energy compared to distilled models [38].
deepseek-r1-distill-llama-70b	• A distilled LLM based on Llama-3.3-70B-Instruct, using outputs from DeepSeek R1. Leverages knowledge distillation to retain robust reasoning capabilities in smaller, more agile architectures. Delivers exceptional performance on benchmarks in mathematical and logical reasoning. • Lower cost due to reduced computational requirements from distillation [37]. • Highly energy-efficient due to compact architecture and optimized inference [38].

available data sources, [10] [11] [12] and stored in text format. Next, co-references in the sentences were resolved using the spacy NLP library [39]. Lastly, sentences were simplified to contain a single subject, predicate, and object by the Sentence Simplification sub-component as described in Section 2.

For the treatment KG, information about drug family classification, and mode of administration was collected. For the ethical KG, information about drug adverse effects along with their severity was collected. We used three grades (High, Medium, and Low) for the severity since data at the level specified in the ontology could not be found easily. For the legal KG, information about drug antimicrobial resistance and intended target species was collected. The requirement for marketing authorization is implicitly met since information about only the approved drugs is collected.

4.3. Results

This section first provides examples of KG models using a dairy production system's use case, followed by a comparison of the results obtained for the generated triples using different LLMs with different prompt formulations.

4.3.1. Use case

In this section, we provide examples of the KGs that can be generated using the RDF triples extracted using LLMs, thereby demonstrating how the ontological models can be used in the context of a treatment RS considering the practical availability of data. In the figures accompanying each example KG, note that classes from the ontology are referred to as concepts, while nodes containing information about specific instances are referred to as entities. The edges that represent predicate relationships between entities are called relation type links, whereas the edges connecting a node to concepts in the

[10] www.medlineplus.gov (As of Dec. 2024).
[11] www.wikipedia.com (As of Dec. 2024).
[12] www.noahcompendium.co.uk (As of Dec. 2024).

ontology are referred to as relation links. For simplicity, each entity is connected to the top-level class in the ontology through a single relation link. However, the specific class or property associated with the entity may be located deeper in the ontology and can be reached by traversing multiple levels from the top-level class. For instance, in the ATOL Ontology, the `Milk Somatic Cell Count` property can be reached by the following path: `Mammary Gland & Milk Production Trait -> Mammary Gland Production Trait -> Milk Trait -> Milk Quality Trait -> Milk Composition Trait -> Milk Cell Count -> Milk Somatic Cell Count`.

For the bovine KG, due to the limited availability of EVHRs, synthetic cow profiles were created manually using mastitis diagnostic knowledge from veterinary literature [40]. An example profile is shown in Fig. 6a. For the treatment KG, due to the limited availability of veterinary treatments, the full treatment ontological model was not used. Instead, the drug name, family, and mode of administration were used to build the KG since these were available on public sources. An example KG for the drug Ampicillin is shown in Fig. 6b. For the ethical KG, the adverse events and severity information were used and an example KG for the drug Ampicillin is shown in Fig. 6c. For the legal KG, the AMR and target species information was used and an example KG for the drug Ampicillin is shown in Fig. 6d. A conceptual view of the social KG for different farmers evolving through different times is shown in Fig. 6e. Those KGs would encode, for a given farm, the treatment history for different cows. The social KG was not generated due to two reasons. The first is that the ontological model could be improved further but needs more detailed research by domain experts, which is out of the scope of the present researchers. The second is that it is difficult to obtain such social data at dairy farms.

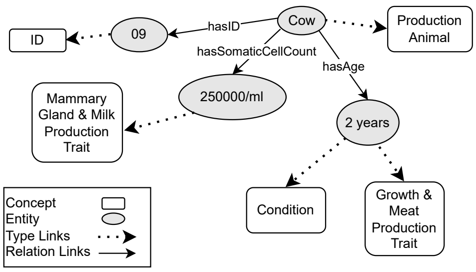
4.3.2. Knowledge Graph Generation

The results of using these LLMs to extract triples using only instructions (without ontologies and examples), zero-shot prompting, and few-shot prompting are shown in Tables 2, 3, and 4, respectively. Note that for the deepseek-r1-distill-llama-70b model, only the final output after the thinking token is displayed.

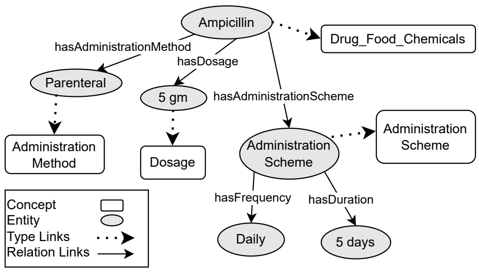
When generating triples without ontologies or examples, interestingly, the llama-3.3-70b-versatile model generated triples adhering to the instructed triple format, whereas the deepseek-r1-distill-llama-70b model was unable to do so. As shown in Table 2, for the input text *"Upset stomach is a common adverse effect of cloxacillin"*, the llama-3.3-70b-versatile model first generated the namespace definitions for the generated classes and properties (*http://example.org/medical_condition/UpsetStomach, http://example.org/relationship/commonAdverseEffectOf, http://example.org/drug/-Cloxacillin*), and then generated the triples *"<http://example.org/medical_condition/UpsetStomach> <http://example.org/relationship/commonAdverseEffectOf> <http://example.org/drug/Cloxacillin> ."* The deepseek-r1-distill-llama-70b model, in turn, generated the triples *"upset_stomach - adverse_effect_of - cloxacillin." and "upset_stomach - common - true"*.

When using zero-shot prompts with ontologies, both models struggled with adhering to the provided classes and properties and the required triple format, and generated additional classes and properties. As shown in Table 2, for the input text *"Pain occurs at the injection site"*, the llama-3.3-70b-versatile model generated the triples *_:b1 rdf:type Adverse_Event. , _:b1 oae:has_severity low. , _:b1 oae:has_certainty common. , _:b2 rdf:type Medication. , and _:b2 oae:has_adverse_event _:b1* ., while the deepseek-r1-distill-llama-70b model generated *<Medication> rdf:type Medication. , <Medication> oae:has_adverse_event <Adverse_Event>. , <Adverse_Event> rdf:type Adverse_Event. , and <Adverse_Event> common "Pain".* Here, the entities *"_:b1", "_:b2"* and the relation *"common"* are incorrect.

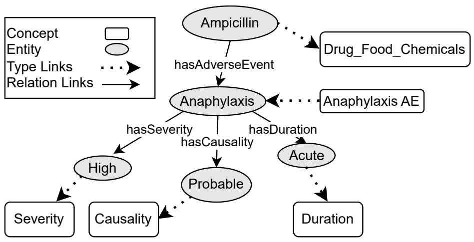
When using few-shot prompts, both models could generate similar results. When the generated triples were accurate, the performance of both models was similar. Table 4 presents for the input texts *"Upset stomach is a common adverse effect of cloxacillin"* and *"Fever is an adverse reaction to cephapirin"*, both the models generate the same complete and correct triples. When the models encountered a sentence structure different from the provided examples, there was a difference in performance with the llama-3.3-70b-versatile model generating triples with additional properties. The deepseek-r1-distill-llama-70b model, in turn, generated only the main entity. For the input text *"Pain occurs at the injection site"*, the



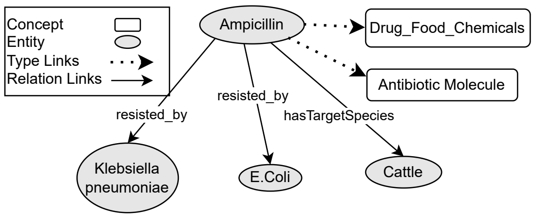
(a) Cow (User) KG



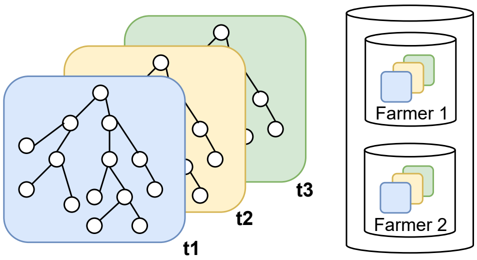
(b) Treatment (Item) KG



(c) Ethical KG



(d) Legal KG



(e) Social KG

Figure 6: (a) KG example of a cow with ID number 09 aged 2 years with a somatic cell count of 250000/ml. (b) KG example of a 5 mg Ampicillin treatment administered parenterally daily for 5 days. (c) KG example for ethical aspects showing an adverse event for the drug Ampicillin. (d) KG example for the legal aspects showing the AMR and target species for the drug Ampicillin. (e) A Conceptual View of KGs that evolve through time for different farmers.

llama-3.3-70b-versatile model generated non-existing properties related to certainty and severity, while the deepseek-r1-distill-llama-70b model only generated the triple for the adverse event related to pain. Although there is an example of not generating triples when there are no side effects, both models fail to follow this example. The llama-3.3-70b-versatile model tended to provide explanations, while the deepseek-r1-distill-llama-70b model tended to extract the main entity, as in the case of the generated triples for the input text *"Danofloxacin has no side effects"*. Furthermore, although there are instructions to not provide explanations, the llama-3.3-70b-versatile model tended to ignore this instruction and generated explanatory text or other conversational phrases such as *"Here are the triples:"*.

To illustrate the size of the output generated by the models under different prompting strategies, consider an input document consisting of two sentences *"Adverse events of procainebenzylpenicillin include pain at the site of injection, blood clotting problems, seizures, and allergic reactions. procaine penicillin can cause severe adverse effects such as seizures and CNS abnormalities."*. The llama-3.3-70b-versatile model generates 21, 33, and 16 triples for instruction-only, zero-shot, and few-shot prompts respectively. The deepseek-r1-distill-llama-70b model generates 12, 23, and 19 triples for instruction-only, zero-shot, and few-shot prompts respectively.

Table 2
Comparison of RDF Triples Generated only instructions (without ontologies or examples).

Prompt: For each given sentence, create RDF triples that describe the sentence in the Term-Based Graph Representation. Avoid explanations. Only output the triples.		
Input Text	**Generated Triples**	
	llama-3.3-70b-versatile	**deepseek-r1-distill-llama-70b**
Upset stomach is a common adverse effect of cloxacillin.	(http://example.org/medical_condition/UpsetStomach, http://example.org/relationship/commonAdverseEffectOf, http://example.org/drug/Cloxacillin) <http://example.org/medical_condition/UpsetStomach> <http://example.org/relationship/commonAdverseEffectOf> <http://example.org/drug/Cloxacillin> .	upset_stomach – adverse_effect_of – cloxacillin. upset_stomach – common – true.
Fever is an adverse reaction to cephapirin.	(http://example.org/medical_condition/Fever, http://example.org/relationship/adverseReactionTo, http://example.org/drug/Cephapirin) <http://example.org/medical_condition/Fever> <http://example.org/relationship/adverseReactionTo> <http://example.org/drug/Cephapirin> .	1. Fever rdf:type AdverseReaction 2. Fever isAdverseReactionTo cephapirin
Pain occurs at the injection site.	(http://example.org/symptom/Pain, http://example.org/relationship/occursAt, http://example.org/location/InjectionSite) <http://example.org/symptom/Pain> <http://example.org/relationship/occursAt> <http://example.org/location/InjectionSite> .	Triple 1: :Pain :occurs :InjectionSite . Triple 2: :Pain :type :Discomfort . Triple 3: :InjectionSite :location "injection site"@en .
Danofloxacin has no side effects.	(http://example.org/drug/Danofloxacin, http://example.org/relationship/hasSideEffect, "None") <http://example.org/drug/Danofloxacin> <http://example.org/relationship/hasSideEffect> "None".	- Danofloxacin -> has -> "no side effects"

Table 3
Comparison of RDF Triples Generated with Ontologies using Zero-shot Prompts.

Prompt: For each given sentence, you should create RDF triples that describe the sentence in the Term-Based Graph Representation. To help you, I will also provide the classes and properties that should be present in the generated RDF triples. These classes and properties are defined in an ontology. Please generate RDF triples to describe the medication adhering to the following classes and properties defined in this ontology: Classes: ["Medication", "Adverse_Event", "Certainty", "Severity"] Properties: ["rdf:type", "oae:has_adverse_event", "oae:has_severity", "oae:has_certainty", "common", "possible", "rare", "low", "medium", "high"]. Just use properties that are in the list. Never use other properties outside the list. Ignore new properties. Avoid explanations. Only output the triples. If a new property is found, output nothing. If a new class is found, output nothing.		
Input Text	**Generated Triples**	
	llama-3.3-70b-versatile	**deepseek-r1-distill-llama-70b**
Upset stomach is a common adverse effect of cloxacillin.	cloxacillin rdf:type Medication. upset_stomach rdf:type Adverse_Event. cloxacillin oae:has_adverse_event upset_stomach. upset_stomach oae:has_severity low. upset_stomach oae:has_certainty common.	1. cloxacillin rdf:type Medication 2. cloxacillin oae:has_adverse_event upset_stomach 3. upset_stomach rdf:type Adverse_Event 4. upset_stomach oae:has_certainty common
Fever is an adverse reaction to cephapirin.	Fever rdf:type Adverse_Event. Cephapirin rdf:type Medication. Cephapirin oae:has_adverse_event Fever.	:medication1 rdf:type Medication. :medication1 common "cephapirin". :medication1 oae:has_adverse_event :adverse_event1 . :adverse_event1 rdf:type Adverse_Event. :adverse_event1 common "Fever".
Pain occurs at the injection site.	_:b1 rdf:type Adverse_Event. _:b1 oae:has_severity low. _:b1 oae:has_certainty common. _:b2 rdf:type Medication. _:b2 oae:has_adverse_event _:b1 .	<Medication> rdf:type Medication. <Medication> oae:has_adverse_event <Adverse_Event>. <Adverse_Event> rdf:type Adverse_Event. <Adverse_Event> common "Pain".
Danofloxacin has no side effects.	Danofloxacin rdf:type Medication. Danaofloxacin oae:has_adverse_event "no side effects". "no side effects" rdf:type Adverse_Event. "no side effects" oae:has_certainty common. "no side effects" oae:has_severity low.	<med:Danofloxacin> rdf:type oae:Medication. <med:Danaofloxacin> oae:has_adverse_event <ae:NoSideEffects>. <ae:NoSideEffects> rdf:type oae:Adverse_Event.

5. Discussion

We demonstrated the construction of ontologies for different ELSA in the domain of dairy treatment recommendation. The ontology construction process in this context involved carefully mapping and integrating relevant concepts from established ontologies, such as those related to the production of animals and medical treatments. For including ethical considerations (*e.g.*, animal well-being), legal requirements (e.g., EU regulations on antibiotic use), and social implications (*e.g.*, dairy farmer practices), new classes and properties were added as necessary, based on the knowledge of domain experts. This

Table 4
Comparison of RDF Triples Generated with Ontologies using Few-shot Prompts.

Prompt: For each given sentence, you should create RDF triples that describe the sentence in the Term-Based Graph Representation. To help you, I will also provide the classes and properties that should be present in the generated RDF triples. These classes and properties are defined in an ontology. Please generate RDF triples to describe the medication adhering to the following classes and properties defined in this ontology: Classes: [“Medication”, “Adverse_Event”, “Certainty”, “Severity”] Properties: [“rdf:type”, “oae:has_adverse_event”, “oae:has_severity”, “oae:has_certainty”, “common”, “possible”, “rare”, “low”, “medium”, “high”]. Just use properties that are in the list. Never use other properties outside the list. Ignore new properties. Avoid explanations. Only output the triples. If a new property is found, output nothing. If a new class is found, output nothing. I will provide four examples. Complete the fifth.

Simpler Sentence: Nausea is a common adverse effect of hexacillin.
Triples: <Adverse_Event/Nausea> rdf:type <Adverse_Event>
<Medication/Hexacillin> oae:has_adverse_event <Adverse_Event/Nausea>.
<Adverse_Event/Nausea> oae:has_certainty <Certainty/common>.

Simpler Sentence: Seizure is a severe adverse event.
Triples: <Adverse_Event/Seizure> rdf:type <Adverse_Event> .
<Adverse_Event/Seizure> oae:has_severity <Severity/High>

Simpler Sentence: Pain may occur.
Triples: <Adverse_Event/Pain> oae:has_certainty <Certainty/possible>.

Simpler Sentence: Hexacillin has no side effects.
Triples:

Input Text	**Generated Triples**	
	llama-3.3-70b-versatile	**deepseek-r1-distill-llama-70b**
Upset stomach is a common adverse effect of cloxacillin.	<Adverse_Event/Upset_stomach> rdf:type <Adverse_Event>. <Medication/Cloxacillin> oae:has_adverse_event <Adverse_Event/Upset_stomach>. <Adverse_Event/Upset_stomach> oae:has_certainty <Certainty/common>.	<Adverse_Event/Upset_stomach> rdf:type <Adverse_Event>. <Medication/Cloxacillin> oae:has_adverse_event <Adverse_Event/Upset_stomach>. <Adverse_Event/Upset_stomach> oae:has_certainty <Certainty/common>.
Fever is an adverse reaction to cephapirin.	<Adverse_Event/Fever> rdf:type <Adverse_Event>. <Medication/Cephapirin> oae:has_adverse_event <Adverse_Event/Fever>. <Adverse_Event/Fever> oae:has_certainty <Certainty/common>.	<Adverse_Event/Fever> rdf:type <Adverse_Event>. <Medication/Cephapirin> oae:has_adverse_event <Adverse_Event/Fever>. <Adverse_Event/Fever> oae:has_certainty <Certainty/common>.
Pain occurs at the injection site.	<Adverse_Event/Pain> rdf:type <Adverse_Event>. <Adverse_Event/Pain> oae:has_certainty <Certainty/common>. <Adverse_Event/Pain> oae:has_severity <Severity/low>.	<Adverse_Event/Pain> rdf:type <Adverse_Event>.
Danofloxacin has no side effects.	No triples, as there is no information to create triples with the provided properties and classes.	<Medication/Danofloxacin> rdf:type <Medication>.

approach facilitates a shared understanding of a domain based on multi-stakeholder deliberations. Reusing existing ontologies promotes standardization within the field, enabling better communication and knowledge sharing among stakeholders, including farmers, veterinarians, policymakers, and researchers. This standardization is crucial for further developing KG-based RSs to navigate the complex landscape of ELSA considerations in dairy farming.

Through the proposed semi-automatic approach, we demonstrated that structured knowledge in the form of RDF triples can be extracted from unstructured text while adhering to the ontological model, thereby facilitating the generation of KGs. Both models can generate similar results when using few-shot prompts, in line with the findings by Brown *et al.* [36]. Our experiments revealed that the generated triples are not satisfactory in their correctness and completeness when an input sentence structure differs from the provided examples, even when there are clear instructions and associated ontologies. We noted that the expertise of the researchers was relied upon when evaluating the triples. This is a study limitation, which can be addressed later when more data is available. Therefore, our preliminary experiments indicate the potential of using pre-trained LLMs, ontologies, and few-shot prompting to extract RDF triples automatically. Further experiments with more robust datasets and metrics may help establish the usefulness of our approach. We point out that different temperature settings were not experimented with, and this is left for future work.

To the best of our knowledge, our study represents the first effort to leverage LLMs and ELSA KGs for recommendations in the context of animal treatment. Unlike existing approaches, our study provided a practical method for integrating multiple domain-specific ELSA into a recommendation system by encoding ELSA into KGs rather than relying solely on traditional metrics.

Our approach focuses on ELSA at the recommendation level. In contrast to the methods proposed by Bosselut *et al.* [21] and Zhang *et al.* [10], which employed transformer-based neural networks or fine-tune existing LLMs, our solution primarily utilized zero-shot and few-shot learning techniques. We employed an LLM to extract the components of the KGs. Our results demonstrated through use cases and KG generation, offer alternative means of presenting findings, unlike most state-of-the-art studies, which primarily adopt a quantitative evaluation approach.

Developing ELSA KGs leverages opportunities to incorporate domain-specific ELSA considerations into creating RSs, particularly in animal healthcare. These ELSA-compliant RSs could explore created KGs from various perspectives. One potential approach involves implementing ranking or re-ranking procedures that consider the connections between recommended items and ELSA restrictions. The re-ranking methods based on complex network measurements, as proposed in [41], could be further developed to address ELSA relationships. Another promising avenue is to pair ELSA-compliant explanations with recommended items, enhancing the users' understanding of the ELSA risks and constraints associated with recommendations.

6. Conclusion

The literature lacks Knowledge Graphs (KGs) related to ELSA aspects and veterinary healthcare. Our investigation provided a preliminary model of ELSA ontologies relevant to this domain. Our method produced KGs that comply with these modeled ontologies and must play a key role in developing KG-based treatment RSs. We investigated and demonstrated how LLMs are suited to generate RDF triples for the ELSA-related KGs. We found that LLMs produce better results when examples are provided via few-shot prompting. Particularly, since the deepseek-r1-distill-llama-70b model had comparable effectiveness to the llama-3.3-70b-versatile model, it is much more efficient, cost- and energy-wise, to use the distilled version [42, 38, 37]. Future work will focus on investigating fully automatic triple-generation pipelines. Moreover, in the context of veterinary RSs, other ontologies, such as the Environment Ontology of Livestock, could be incorporated into the RSs by including information about farm systems. Another direction concerns defining suitable methodologies for generating social KGs.

Acknowledgments

This work was supported by the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) [grant #Nwa.1332.20.002]. This research was partially funded by the São Paulo Research Foundation (FAPESP) [grant #2022/15816-5].

Declaration on Generative AI

The author(s) have not employed any Generative AI tools.

References

[1] F. Gräßer, H. Malberg, S. Zaunseder, S. Beckert, D. Küster, J. Schmitt, S. A. Klinik, P. f¨r Dermatologie, Application of recommender system methods for therapy decision support, in: 2016 IEEE 18th International Conference on e-Health Networking, Applications and Services (Healthcom), Munich, Germany, 2016, pp. 1–6. doi:https://doi.org/10.1109/HealthCom.2016.7749495.

[2] M. Ryan, V. Blok, Stop re-inventing the wheel: or how ELSA and RRI can align, Journal of Responsible Innovation 10 (2023) 2196151. doi:https://doi.org/10.1080/23299460.2023.2196151.

[3] S. Wang, X. Zhang, Y. Wang, F. Ricci, Trustworthy recommender systems, ACM Trans. Intell. Syst. Technol. 15 (2024) 1–20. doi:https://doi.org/10.1145/3627826.

[4] Y. Ge, S. Liu, Z. Fu, J. Tan, Z. Li, S. Xu, Y. Li, Y. Xian, Y. Zhang, A survey on trustworthy recommender systems, ACM Trans. Recomm. Syst. 3 (2024) 1–68. doi:https://doi.org/10.1145/3652891.

[5] A. Hogan, E. Blomqvist, M. Cochez, C. D'amato, G. D. Melo, C. Gutierrez, S. Kirrane, J. E. L. Gayo, R. Navigli, S. Neumaier, A.-C. N. Ngomo, A. Polleres, S. M. Rashid, A. Rula, L. Schmelzeisen, J. Sequeda, S. Staab, A. Zimmermann, Knowledge graphs, ACM Comput. Surv. 54 (2021) 1–37. doi:https://doi.org/10.1145/3447772.

[6] O. Lassila, R. R. Swick, Resource Description Framework (RDF) model and syntax specification, February 1999, W3C Candidate Recommendation, World Wide Web Consortium (W3C), Massacheusetts, U.S.A, 1999. URL: http://www.w3.org/TR/REC-rdf-syntax.

[7] W. X. Zhao, K. Zhou, J. Li, T. Tang, X. Wang, Y. Hou, Y. Min, B. Zhang, J. Zhang, Z. Dong, et al., A survey of large language models, arXiv preprint (2023). doi:https://doi.org/10.48550/arXiv.2303.18223.

[8] I. Melnyk, P. Dognin, P. Das, Knowledge graph generation from text, in: Y. Goldberg, Z. Kozareva, Y. Zhang (Eds.), Findings of the Association for Computational Linguistics: EMNLP 2022, Association for Computational Linguistics, Abu Dhabi, United Arab Emirates, 2022, pp. 1610–1622. doi:https://doi.org/10.18653/v1/2022.findings-emnlp.116.

[9] B. Zhang, H. Soh, Extract, define, canonicalize: An LLM-based framework for knowledge graph construction, in: Y. Al-Onaizan, M. Bansal, Y.-N. Chen (Eds.), Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, Miami, Florida, USA, 2024, pp. 9820–9836. doi:https://doi.org/10.18653/v1/2024.emnlp-main.548.

[10] Y. Zhang, T. Sadler, M. R. Taesiri, W. Xu, M. Reformat, Fine-tuning language models for triple extraction with data augmentation, in: R. Biswas, L.-A. Kaffee, O. Agarwal, P. Minervini, S. Singh, G. de Melo (Eds.), Proceedings of the 1st Workshop on Knowledge Graphs and Large Language Models (KaLLM 2024), Association for Computational Linguistics, Bangkok, Thailand, 2024, pp. 116–124. doi:https://doi.org/10.18653/v1/2024.kallm-1.12.

[11] W. Borst, Construction of Engineering Ontologies for Knowledge Sharing and Reuse, Phd thesis, University of Twente, Netherlands, 1997.

[12] N. Guarino, D. Oberle, S. Staab, What Is an Ontology?, Springer, Berlin, Heidelberg, 2009, pp. 1–17. doi:https://doi.org/10.1007/978-3-540-92673-3_0.

[13] E. de Jong, I. van der Velden, A.-M. C. Smid, J. A. Ida, K. K. Reyher, D. F. Kelton, H. W. Barkema, Dairy farmers' considerations for antimicrobial treatment of clinical mastitis in British Columbia and Alberta, Canada, Frontiers in Veterinary Science 11 (2024). doi:https://doi.org/10.3389/fvets.2024.1417958.

[14] Q. Guo, F. Zhuang, C. Qin, H. Zhu, X. Xie, H. Xiong, Q. He, A survey on knowledge graph-based recommender systems, IEEE Transactions on Knowledge and Data Engineering 34 (2022) 3549–3568. doi:https://doi.org/10.1109/TKDE.2020.3028705.

[15] A. G. Regino, R. O. Caus, V. Hochgreb, J. C. dos Reis, Leveraging knowledge graphs for e-commerce product recommendations, SN Computer Science 4 (2023) 689. doi:https://doi.org/10.1007/s42979-023-02149-6.

[16] N. Tiwary, S. A. Mohd Noah, F. Fauzi, T. S. Yee, A review of explainable recommender systems utilizing knowledge graphs and reinforcement learning, IEEE Access 12 (2024) 91999–92019. doi:https://doi.org/10.1109/ACCESS.2024.3422416.

[17] W. N. Cheng, S. G. Han, Bovine mastitis: risk factors, therapeutic strategies, and alternative treatments - a review, Asian-Australasian Journal of Animal Sciences 33 (2020) 1699–1713. doi:https://doi.org/10.5713/ajas.20.0156.

[18] A. J. Pires, G. Pereira, D. Fangueiro, R. Bexiga, M. Oliveira, When the solution becomes the problem: a review on antimicrobial resistance in dairy cattle, Future Microbiology 19 (2024) 903–929. doi:https://doi.org/10.2217/fmb-2023-0232.

[19] W. Fan, X. Zhao, X. Chen, J. Su, J. Gao, L. Wang, Q. Liu, Y. Wang, H. Xu, L. Chen, Q. Li, A

comprehensive survey on trustworthy recommender systems, 2022. doi:https://doi.org/10.48550/arXiv.2209.10117.
[20] J. Stray, A. Halevy, P. Assar, D. Hadfield-Menell, C. Boutilier, A. Ashar, C. Bakalar, L. Beattie, M. Ekstrand, C. Leibowicz, C. Moon Sehat, S. Johansen, L. Kerlin, D. Vickrey, S. Singh, S. Vrijenhoek, A. Zhang, M. Andrus, N. Helberger, P. Proutskova, T. Mitra, N. Vasan, Building human values into recommender systems: An interdisciplinary synthesis, ACM Trans. Recomm. Syst. 2 (2024) 1–57. doi:https://doi.org/10.1145/3632297.
[21] A. Bosselut, H. Rashkin, M. Sap, C. Malaviya, A. Celikyilmaz, Y. Choi, COMET: Commonsense transformers for automatic knowledge graph construction, in: A. Korhonen, D. Traum, L. Màrquez (Eds.), Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, Association for Computational Linguistics, Florence, Italy, 2019, pp. 4762–4779. doi:https://doi.org/10.18653/v1/P19-1470.
[22] J. C. dos Reis, A. G. Regino, Generating e-commerce related knowledge graph from text: Open challenges and early results using LLMs, in: Joint proceedings of the 3rd International workshop on knowledge graph generation from text (TEXT2KG) and Data Quality meets Machine Learning and Knowledge Graphs (DQMLKG) co-located with the Extended Semantic Web Conference (ESWC 2024), Hersonissos, Greece, May 26-30, 2024, volume 3747 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2024, pp. 21–39. URL: https://ceur-ws.org/Vol-3747/text2kg_paper11.pdf.
[23] W. Golik, O. Dameron, J. Bugeon, A. Fatet, I. Hue, C. Hurtaud, M. Reichstadt, M.-C. Salaün, J. Vernet, L. Joret, et al., ATOL: the multi-species livestock trait ontology, in: Metadata and Semantics Research: 6th Research Conference, MTSR 2012, Cádiz, Spain, November 28-30, 2012. Proceedings 6, Springer, 2012, pp. 289–300.
[24] C. Khoo, J.-C. Na, V. Wang, S. Chan, Developing an ontology for encoding disease treatment information in medical abstracts, DESIDOC Journal of Library & Information Technology 31 (2011) 103–115. doi:https://doi.org/10.14429/djlit.31.2.862.
[25] B. Varkey, Principles of clinical ethics and their application to practice, Medical Principles and Practice 30 (2020) 17–28. doi:https://doi.org/10.1159/000509119.
[26] T. L. Beauchamp, Principlism in Bioethics, volume 70, Springer International Publishing, Cham, Switzerland, 2016, pp. 1–16. doi:https://doi.org/10.1007/978-3-319-43419-3_1.
[27] P. Guyer, Kant on the theory and practice of autonomy, Social Philosophy and Policy 20 (2003) 70–98. doi:https://doi.org/10.1017/S026505250320203X.
[28] T. L. Beauchamp, J. F. Childress, Principles of biomedical ethics, 4 ed., Oxford University Press, New York, 1994.
[29] B. P. Alcock, W. Huynh, R. Chalil, K. W. Smith, A. Raphenya, M. A. Wlodarski, A. Edalatmand, A. Petkau, S. A. Syed, K. K. Tsang, S. J. C. Baker, M. Dave, M. McCarthy, K. M. Mukiri, J. A. Nasir, B. Golbon, H. Imtiaz, X. Jiang, K. Kaur, M. Kwong, Z. C. Liang, K. C. Niu, P. Shan, J. Y. J. Yang, K. Gray, G. R. Hoad, B. Jia, T. Bhando, L. Carfrae, M. Farha, S. French, R. Gordzevich, K. Rachwalski, M. Tu, E. Bordeleau, D. Dooley, E. Griffiths, H. L. Zubyk, E. D. Brown, F. Maguire, R. Beiko, W. W. L. Hsiao, F. S. L. Brinkman, G. Van Domselaar, A. G. McArthur, Card 2023: expanded curation, support for machine learning, and resistome prediction at the comprehensive antibiotic resistance database, Nucleic Acids Research 51 (2022) D690–D699. doi:https://doi.org/10.1093/nar/gkac920.
[30] D. Kuiper, J. Jansen, R. Renes, C. Leeuwis, H. van der Zwaag, Social factors related to mastitis control practices: the role of dairy farmers' knowledge, attitude, values, behaviour and networks, in: H. Hogeveen (Ed.), Mastitis in Dairy Production, 4th IDF International Mastitis Conference, Wageningen 2005, Wageningen Academic Publishers, Netherlands, 2005, pp. 576–582.
[31] J. Jansen, B. van den Borne, R. Renes, G. van Schaik, T. Lam, C. Leeuwis, Explaining mastitis incidence in dutch dairy farming: The influence of farmers' attitudes and behaviour, Preventive Veterinary Medicine 92 (2009) 210–223. doi:https://doi.org/10.1016/j.prevetmed.2009.08.015, special Issue: SVEPM 2008.
[32] J. Jansen, G. van Schaik, R. Renes, T. Lam, The effect of a national mastitis control program on the attitudes, knowledge, and behavior of farmers in the netherlands, Journal of Dairy Science 93 (2010) 5737–5747. doi:https://doi.org/10.3168/jds.2010-3318.

[33] J. Swinkels, A. Hilkens, V. Zoche-Golob, V. Krömker, M. Buddiger, J. Jansen, T. Lam, Social influences on the duration of antibiotic treatment of clinical mastitis in dairy cows, Journal of Dairy Science 98 (2015) 2369–2380. doi:https://doi.org/10.3168/jds.2014-8488.

[34] N. Valeeva, T. Lam, H. Hogeveen, Motivation of dairy farmers to improve mastitis management, Journal of Dairy Science 90 (2007) 4466–4477. doi:https://doi.org/10.3168/jds.2007-0095.

[35] K. Huijps, T. J. Lam, H. Hogeveen, Costs of mastitis: facts and perception, Journal of Dairy Research 75 (2008) 113–120. doi:https://doi.org/10.1017/S0022029907002932.

[36] T. Brown, B. Mann, N. Ryder, M. Subbiah, J. D. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, D. Amodei, Language models are few-shot learners, in: H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, H. Lin (Eds.), Advances in Neural Information Processing Systems, volume 33 of *NIPS '20*, Curran Associates, Inc., Vancouver, BC, Canada, 2020, pp. 1877–1901. doi:https://doi.org/10.48550/arXiv.2005.14165.

[37] Z. Wu, R. H. Bai, A. Zhang, J. Gu, V. Vydiswaran, N. Jaitly, Y. Zhang, Divide-or-conquer? which part should you distill your LLM?, in: Y. Al-Onaizan, M. Bansal, Y.-N. Chen (Eds.), Findings of the Association for Computational Linguistics: EMNLP 2024, Association for Computational Linguistics, Miami, Florida, USA, 2024, pp. 2572–2585. doi:https://doi.org/10.18653/v1/2024.findings-emnlp.145.

[38] Y. Yuan, J. Shi, Z. Zhang, K. Chen, J. Zhang, V. Stoico, I. Malavolta, The impact of knowledge distillation on the energy consumption and runtime efficiency of nlp models, in: Proceedings of the IEEE/ACM 3rd International Conference on AI Engineering - Software Engineering for AI, CAIN '24, Association for Computing Machinery, New York, NY, USA, 2024, p. 129–133. doi:https://doi.org/10.1145/3644815.3644966.

[39] M. Honnibal, I. Montani, spaCy: Industrial-strength natural language processing in python, 2020. doi:https://doi.org/10.5281/zenodo.1212303, python library.

[40] P. R. Adkins, J. R. Middleton, Methods for diagnosing mastitis, Veterinary Clinics: Food Animal Practice 34 (2018) 479–491. doi:https://doi.org/10.1016/j.cvfa.2018.07.003.

[41] O. Baumann, D. Nandini, A. Rossanez, M. Schoenfeld, J. C. Dos Reis, How to surprisingly consider recommendations? a knowledge-graph-based approach relying on complex network metrics, in: Proceedings of the 16th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2024) - Volume 2: KEOD, INSTICC, SciTePress, Porto, Portugal, 2024, pp. 27–38. doi:https://doi.org/10.5220/0012936100003838.

[42] J. Liu, C. Zhang, J. Guo, Y. Zhang, H. Que, K. Deng, Z. Bai, J. Liu, G. Zhang, J. Wang, Y. Wu, C. Liu, J. Wang, L. Qu, W. Su, B. Zheng, Ddk: Distilling domain knowledge for efficient large language models, in: A. Globerson, L. Mackey, D. Belgrave, A. Fan, U. Paquet, J. Tomczak, C. Zhang (Eds.), Advances in Neural Information Processing Systems, volume 37, Curran Associates, Inc., 2024, pp. 98297–98319. URL: https://proceedings.neurips.cc/paper_files/paper/2024/file/b206d54ffbb803b5c51d85f405d422e4-Paper-Conference.pdf.