

Nordic AI Meet 2026

P2: AI in Health, Medicine, Care & Human Well-being

9. A Guideline-Schema-Guided LLM Framework for Auditing CAUTI Documentation in Finnish Inpatient Records

Authors: Linli Zhang¹; Tiina Kurvinen², Hans Moen³, Sanna Salanterä², and Pekka Marttinen¹

Affiliations: 1. Department of Computer Science, Aalto University, Finland 2. Department of Department of Nursing Science, University of Turku, Finland 3. Department of Clinical Medicine, Aalborg University, Denmark

Abstract: *Objective.* Catheter-associated urinary tract infection (CAUTI) is among the most common healthcare-associated infections, yet evidence of preventive practice remains buried in unstructured clinical notes. We develop and evaluate a guideline-schema-guided large language model (LLM) framework for extracting and summarizing CAUTI-related information from heterogeneous patient visit records. Using the resulting automated audit, we examine the association between documented CAUTI-prevention practice and CAUTI incidence among Finnish inpatients, and quantify changes in incidence rates. The data involves inpatients who were admitted with a primary diagnosis other than CAUTI and acquired the infection during their stay. Throughout, we prioritize completeness, relevance, and fidelity to the source data.

Materials and methods. We first designed a structured schema derived from clinical CAUTI guidelines, defining key elements across the stages of urinary catheterization, including indication and insertion, maintenance, symptoms, diagnostic findings, and treatment and care plans. The schema guides rule-based and hybrid retrieval to identify candidate evidence, which is then summarized by an LLM. The framework was applied to 2,544 patient visit records of heterogeneous structure and content. To evaluate summary quality, we developed a query-based assessment framework: schema-derived clinical queries were posed to both the original notes and the LLM-generated summaries, and the resulting answers were compared using semantic agreement metrics as a measure of completeness. Hallucination was defined as summary content unsupported by the source notes, and irrelevance as the proportion of non-UTI-related content in the generated summaries. A subset of cases underwent expert review to assess clinical validity.

Results. Relative to manual extraction and summarization, the framework substantially reduced processing time while maintaining accuracy against a gold standard authored by clinical experts, with low rates of hallucinated and irrelevant content.

Discussion. By coupling a guideline-derived schema with LLM-based summarization, this work addresses persistent challenges in clinical text processing: fragmentation, linguistic variability, and lack of structure. The framework offers a scalable route to faithful, complete summarization without exhaustive manual reading and annotation, with applications in automated guideline-adherence assessment, clinical decision support, and large-scale clinical data analysis. It also raises questions of how documentation practices themselves shape what can be measured in care quality.

Conclusion. We present a guideline-schema-guided LLM framework that produces structured, clinically meaningful summaries of CAUTI-related information from unstructured clinical notes, maintaining fidelity to the source while reducing irrelevant and hallucinated content.