

РОЗРОБКА ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ РОЗУМНОГО ВЕДЕННЯ РОДОВОДУ НА ОСНОВІ ШІ-АГЕНТА

Вінницький національний технічний університет

Анотація

У роботі представлено розробку інтелектуального модуля для автоматизованого ведення родоводу з використанням технологій штучного інтелекту. Досліджено сучасні підходи до обробки природної мови та методи векторного пошуку, що застосовуються для інтерпретації користувацьких запитів у діалогових системах. Запропоновано архітектуру інформаційної технології, яка поєднує механізм Retrieval-Augmented Generation (RAG), модуль розпізнавання намірів та підсистему перевірки повноти генеалогічних даних. Проведено аналіз існуючих генеалогічних сервісів, визначено вимоги до системи та здійснено її програмну реалізацію.

Ключові слова: родовід, генеалогія, RAG, діалогова система, обробка природної мови, інтелектуальний агент.

Abstract

The paper presents the development of an intelligent module for automated genealogy management based on artificial-intelligence methods. Modern approaches to natural-language processing and vector search are examined, focusing on their applicability to dialog-based genealogy systems. The proposed architecture integrates a Retrieval-Augmented Generation (RAG) mechanism, an intent-recognition module and a data-validation subsystem. Existing genealogy platforms were analysed and system requirements were formulated. A software prototype of the developed solution was implemented and evaluated.

Keywords: genealogy, family tree, RAG, dialog system, natural language processing, AI agent.

Вступ

Генеалогічні системи давно стали важливими інструментами збереження історії родини. Однак більшості сучасних сервісів притаманна суттєва проблема: введення даних виконується вручну у суворо структурованій формі, що значно ускладнює процес для пересічного користувача [1].

Поява високоточних мовних моделей та механізмів векторного пошуку відкрила можливість створення систем, що здатні інтерпретувати природномовні запити та автоматично здійснювати операції над генеалогічними даними [2]. У таких системах ключову роль відіграє діалоговий інтерфейс, який дозволяє взаємодіяти з даними через розмову, а не через заповнення таблиць чи форм.

Метою роботи є створення інформаційної технології, що забезпечує інтуїтивне поповнення родоводу на основі ШІ-агента, здатного розпізнавати наміри, витягувати ключові дані та автоматично формувати структури для запису у базу.

Постановка задачі дослідження

Для досягнення поставленої мети необхідно:

- провести аналіз методів обробки природної мови, релевантних для генеалогічних систем [3];
- дослідити моделі векторних ембедингів та механізм RAG для визначення релевантних дій системи [4];
- сформулювати вимоги до архітектури інтерфейсу на основі діалогової взаємодії;
- розробити модуль розпізнавання намірів користувача;
- розробити модуль витягування сутностей (імена, дати, типи зв'язків);
- побудувати підсистему перевірки повноти даних і генерування уточнень;
- створити програмний прототип інформаційної технології;
- провести тестування системи в реальних сценаріях та оцінити її ефективність.

Виклад основного матеріалу

Дослідження було зосереджене на створенні інтелектуального модуля, який забезпечує автоматизоване ведення родоводу на основі аналізу природномовних запитів користувача. На першому етапі проведено огляд існуючих рішень у сфері генеалогічних систем, зокрема FamilySearch та MyHeritage, що дозволило визначити ключові проблеми сучасних підходів. Встановлено, що більшість таких систем орієнтовані на ручне введення структурованих даних, не використовують сучасні методи NLP і не підтримують діалогові інтерфейси, що значно ускладнює процес введення великого обсягу інформації [1].

У межах роботи виконано аналіз методів опрацювання природної мови, включаючи класичні моделі NER (Named Entity Recognition), визначення намірів (intent classification) та сучасні підходи на основі трансформерних моделей, що забезпечують точну інтерпретацію текстових описів [3]. Досліджено можливості нейронних ембедингів, які формують векторне представлення тексту й дозволяють використовувати RAG-підхід для пошуку релевантних дій системи [4].

Запропоновано архітектуру інтелектуальної інформаційної технології, що складається з таких основних компонентів.

1. Модуль попередньої обробки тексту, що виконує нормалізацію, токенизацію та очищення введених користувачем даних.
2. Модуль розпізнавання намірів, що визначає тип операції, яку необхідно здійснити: додавання нової особи, встановлення родинного зв'язку, редагування атрибутів, пошук інформації тощо.
3. Модуль витягування сутностей, який автоматично розпізнає імена, дати народження, події, типи зв'язків та інші ключові параметри.
4. Векторна база знань, де зберігаються описи всіх можливих дій системи у вигляді ембедингів. Це дає змогу зіставити запит користувача з найрелевантнішою дією через пошук найближчих сусідів.
5. Механізм уточнення даних, який визначає, чи повні зібрані дані та які параметри необхідно додатково запитати.
6. Підсистема операцій над даними, що перетворює розпізнану дію та набір параметрів у реальні зміни структури родоводу.

Окрему увагу приділено проектуванню моделі даних. На основі аналізу існуючих підходів сформовано гнучку структуру родоводу, яка включає сутності Person, Relation, Attribute, SubAttribute, а також асоціативні таблиці для збереження складних багатозначних характеристик. Така модель забезпечує масштабованість системи та можливість розширення переліку атрибутів без змін базової схеми.

Програмна реалізація була виконана з використанням серверних технологій, що дозволяють інтегрувати модуль штучного інтелекту зі структурованою базою даних. Було створено прототип діалогового інтерфейсу, який у режимі реального часу обробляє введені користувачем повідомлення та виконує відповідні операції. Задля забезпечення високої точності виявлення намірів використано векторні ембеддинги та алгоритми семантичного пошуку.

Результати дослідження

Розроблена система була протестована шляхом моделювання реальних сценаріїв взаємодії користувача з інтелектуальним модулем. Перевірялися такі функціональні можливості:

- визначення намірів у коротких та складних текстових повідомленнях;
- автоматичне виділення сутностей з неструктурованих описів осіб;
- правильність вибору операції за допомогою векторного пошуку;
- коректність формування уточнень при нестачі даних;
- можливість роботи з різними типами родинних зв'язків;
- стійкість до неоднозначних формулювань та граматичних помилок у текстах.

Під час тестування було встановлено, що система самостійно та коректно визначає більшість намірів користувача, навіть коли запит сформульований неформально або містить зайві деталі. Модуль витягування сутностей продемонстрував високу точність у визначенні імен, родинних ролей та дат, що дозволяє виконувати складні операції над базою даних без ручного введення параметрів.

RAG-механізм показав здатність правильно обирати найбільш релевантну дію за умов, коли декілька можливих операцій є схожими за змістом. Це стало можливим завдяки використанню семантичних ембеддингів, які забезпечують правильне зіставлення текстів за змістом, а не тільки за ключовими словами.

У тих випадках, коли введених даних не вистачало, модуль уточнення формував запит додаткової інформації, що дозволило значно знизити кількість помилкових операцій. Таким чином, система продемонструвала здатність до адаптивної взаємодії з користувачем, що є критично важливим для генеалогічних застосунків.

Отримані результати свідчать про те, що розроблена інформаційна технологія дозволяє істотно зменшити обсяг ручної роботи під час ведення родоводу та підвищує зручність користувача за рахунок діалогового режиму роботи. Успішність тестування підтверджує доцільність подальшого розвитку систем на основі інтелектуальних агентів.

Висновки

У роботі створено інформаційну технологію ведення родоводу на основі інтелектуального чат-бота, що використовує методи NLP та RAG для інтерпретації текстових запитів. Запропонований підхід є новим для генеалогічних систем і забезпечує автоматизацію операцій, які раніше виконувалися вручну. Подальші дослідження можуть бути спрямовані на інтеграцію механізмів перевірки правдоподібності даних, розширення моделі зв'язків та підтримку імпорту з форматів на кшталт GEDCOM.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. FamilySearch — Official Website. URL: <https://www.familysearch.org/>
2. Goodfellow I., Bengio Y., Courville A. Deep Learning. MIT Press, 2016.
3. Згуровський М.З., Зайченко Ю.П. Обробка природної мови. URL: <https://ela.kpi.ua/server/api/core/bitstreams/9813cd05-1aee-48b2-8096-b1952986b7ff/content>
4. OpenAI Platform – Retrieval-augmented generation (RAG). URL: <https://platform.openai.com/docs/guides/optimizing-llm-accuracy#retrieval-augmented-generation-rag>
5. KLU – Що таке LLMops? URL: <https://klu.ai/glossary/llm-ops-ua>

Бойко Андрій Вікторович – студент кафедри комп'ютерних наук, факультет інтелектуальних інформаційних технологій та автоматизації, Вінницький національний технічний університет, м. Вінниця, e-mail: boiko.andrii.y@gmail.com

Денисюк Валерій Олександрович – канд. техн. наук, доцент, доцент кафедри комп'ютерних наук, Вінницький національний технічний університет, м.Вінниця, e-mail: vad64@i.ua.

Boiko Andrii Viktorovich - student of Computer Science Department, Faculty of Intelligent Information Technologies and Automation, Vinnytsia National Technical University, Vinnytsia, e-mail: boiko.andrii.y@gmail.com;

Denysiuk Valerii Olexandrovich – Ph.D., Assistant Professor, Assistant Professor of the Chair of Computer Science, Vinnytsia National Technical University, Vinnytsia, e-mail: vad64@i.ua