

Branko Tošović (ur./Hg.)

ANDRIĆ VIRTUELNI

ANDRIĆ VIRTUELL

Andrić-Initiative

19

**Institut für Slawistik der Karl-Franzens-Universität Graz
Komisija za stilistiku Međunarodnog slavističkog komiteta**

Svet knjige Beograd

Aletea Beograd

2026

Herausgeber • Urednik

Verlag • Izdavač

Institut für Slawistik der Karl-Franzens-Universität Graz

Merangasse 70, 8010 Graz, Österreich/Austria, +43 316 380 25 24, branko.tosovic@uni-graz.at; arno.wonisch@uni-graz.at, <http://www-gewi.uni-graz.at/gralis>

Svet knjige d.o.o.

Bulevar despota Stefana 15, 11000 Beograd, Srbija, tel. +381 11 324 73 07, svetknjige@svetknjige.net, www.svetknjige.net

Aletea

Generala Štefanika 2, 11000 Beograd, Srbija, +381 62 311 300

Für die Herausgeber • Za izdavače

Branko Tošović,

Arno Wonisch, Stevo Čosović, Milan Tasić

Satz, Umschlag, Lektorat und Korrektur

Prelom, korice, lektorisanje i korektura

Branko Tošović

Lektorat der deutschen Inhalte

Arno Wonisch

Umschlagillustration • Ilustracija na koricama

Branko Tošović i ChatGPT

Tiraž: 200

Druck • Štampa

Excellent print, Balkanska 35–39, Beograd

Tošović, Branko (ur./Hg.). *Andrić virtuelni / Andrić virtuell*. Graz/Grac – Beograd: Institut für Slawistik der Karl-Franzens-Universität Graz – Komisija za stilistiku Međunarodnog slavi-
stičkog komiteta – Svet knjige – Aletea, 2026. 836 s./S. [Andrić-Initiative 19]

© Branko Tošović, Graz 2026

Alle Rechte vorbehalten. Sva prava zadržana.

ISBN 978-3-9519973-9-1 (Institut für Slawistik der Karl-Franzens-Universität Graz)

ISBN 978-86-6226-033-8 (Svet knjige Beograd)

ISBN 978-86-82644-36-1 (Aletea Beograd)

Marija Đokić Petrović (Graz; Belgrade)
Branko Tošović (Graz)

Chatting with Nobel Prize Laureate Ivo Andrić Using AI Technologies

This paper presents the implementation of an AI-based conversational agent that enables interactive dialogue with Nobel Prize Laureate and writer Ivo Andrić. The approach combines a Large Language Model (LLM) with Retrieval-Augmented Generation (RAG) to generate coherent, context-aware responses grounded in verified biographical material. While the LLM serves as the conversational core, a curated Knowledge Base obtained from authoritative biographical source ensures factual consistency and historical plausibility. The conversational agent supports multilingual interaction and addresses both biographical and interpretive questions about Andrić's life and work. A comparative analysis between RAG and non-RAG approaches demonstrates the importance of retrieval mechanisms for reducing LLMs' hallucinations and maintaining factual accuracy, especially in historically sensitive contexts. The project illustrates the potential of AI for education and cultural heritage preservation, while highlighting current limitations and future improvements.

1. Introduction

Computational techniques have increasingly been adopted in Digital Humanities (DH), enabling new forms of access, interpretation, and public engagement with historical and literary domains (Siliutina et al. 2024). The broader availability and recent advancement of Large Language Models (LLMs) have opened new possibilities for conversational interfaces that enable users to interact with cultural artifacts, historical personas, and literary figures in a dialogic manner rather than through static text (Engstrøm/Løvlie 2025). However, while LLMs excel at generating fluent, human-like language, they are also prone to factual inaccuracies and hallucinations (Lei et al. 2025), which pose significant risks when applied to historically grounded domains such as biography, literature, and cultural heritage (Hutchinson 2024, DaCosta 2025).

Retrieval-Augmented Generation (RAG) has emerged as a promising approach to mitigating these issues by coupling generative models with external knowledge sources (Yu et al. 2025). Instead of relying solely on knowledge embedded during training, RAG-based systems dynamically retrieve relevant

documents and condition the generated response on verifiable information (Gao et al. 2023). This paradigm is particularly well-suited for DH applications, where transparency, provenance, and factual fidelity are essential.

In this context, the present study explores the use of RAG to create a conversational agent capable of enabling conversations with Ivo Andrić¹(1892–1975), one of the most prominent South Slavic writers and a Nobel Prize laureate in literature. Andrić’s life and work are deeply embedded in the complex historical, cultural, and religious landscapes of the Balkans, making him a significant person in South Slavic literature and history. The goal of the project is not to impersonate Andrić in a fictional sense, but to provide an educational interface that offers historically grounded, stylistically plausible, and ethically responsible responses based on documented biographical source.

The contributions of this paper are twofold:

- The implementation of a RAG-based conversational agent capable of enabling conversations with writer and Nobel Prize Laureate Ivo Andrić.
- An empirical comparison between RAG and non-RAG approaches, demonstrating how RAG improves factual accuracy and reduces hallucinations in historically grounded conversational interaction.

The remainder of this paper is organized as follows. Section 2 reviews related work on conversational agents in DH and cultural heritage contexts. Section 3 describes the architecture and implementation of our AI conversational agent. Section 4 presents a discussion of our system’s behavior, with particular emphasis on comparisons between RAG-based and non-RAG responses and on multilingual interaction. Finally, Section 5 concludes the paper and outlines directions for future work.

2. Related Work

In DH and cultural heritage contexts, conversational agents are increasingly employed as interactive interfaces for learning, exploration, and public engagement (Machidon/Tavcar 2019, Wang/Meng 2024). Even before the emergence of generative AI, cultural heritage institutions experimented with conversational agents to mediate access to historical content. Early applications include the use of conversational agents as virtual museum guides to support visitor navigation and information retrieval (Kopp et al. 2005), as well as gamified conversational systems designed to enhance museum engagement and amusement (Gaia et al. 2019). Beyond informational assistance, conversational agents that emulate or represent historical personalities have also been explored. A

¹ https://en.wikipedia.org/wiki/Ivo_Andrić.

notable early example is the conversational agent introduced by the Anne Frank House in 2017, which allowed users to ask questions about Anne Frank's life via Facebook Messenger (Tzouganatou 2018).

With the rise of LLMs and generative AI, a new generation of conversational agents has emerged, offering greater flexibility and open-ended interaction. Recent work by authors (Natale et al. 2025) presents a conversational agent that impersonates the historical figure Luigi Einaudi, the first elected president of the Italian Republic. The authors of (Škobo/Švoić 2025) created a multilingual AI chatbot based on Nikola Tesla's character to encourage student engagement, motivation, and casual English-language practice. An AI conversational agent that simulated conversations with Van Gogh based on his letters was developed as part of the „Musée d'Orsay's 2024“ project (Open Culture 2024–www). Also, several commercial platforms, such as „Hello History“² and „Text with History“³, leverage conversational agents to enable interactive dialogues with historical figures.

The increasing use of generative AI in cultural heritage applications has consequently intensified concerns regarding reliability, historical accuracy, source provenance, and responsible interpretation (Spennemann 2024). As conversational agents gain the ability to produce fluent, authoritative-sounding responses, the risk of factual distortion and hallucination becomes particularly acute in historically grounded and literary domains (Natale et al. 2025). These challenges motivate the adoption of grounding mechanisms, such as RAG (Sun et al. 2025, Swacha/Gracel 2025), to ensure that conversational interaction remains aligned with verifiable historical and biographical evidence. For this reason, the approach presented in this study adopted RAG. In the following, we will explain the implementation of our AI conversational agent that enables chatting with writer and Nobel Prize Laureate Ivo Andrić.

3. AI Conversational Agent for chatting with Ivo Andrić – Implementation

3.1 System Architecture

The conversational agent employs the RAG technique to combine LLM capabilities with specific knowledge sources (Figure 1). Flowise⁴ serves as the visual development platform for orchestrating the agent pipeline, enabling rapid configuration and integration with RAG components. Azure AI Foundry⁵

² <https://www.hellohistory.ai/>.

³ <https://textwith.me/history/>.

⁴ <https://flowiseai.com/>.

⁵ <https://ai.azure.com/>.

provides managed infrastructure for deploying both the LLM and the embedding model, ensuring scalable inference capabilities. During data ingestion, source documents are segmented into chunks and converted into vector embeddings using the text-embedding-3-large⁶ model. These embeddings and their metadata are stored in Pinecone⁷, a vector database, while Supabase⁸ serves as a records manager to track indexed documents and prevent duplicate vector embeddings in the vector store.

At query time, the user's question is converted into a vector embedding and used to search the vector store for semantically similar data chunks. The retrieved context is combined with the original query to create an augmented prompt sent to GPT-4.1⁹, the generative model selected for its reasoning capabilities and multilingual support. The model generates responses grounded in both the Knowledge Base and its general knowledge.

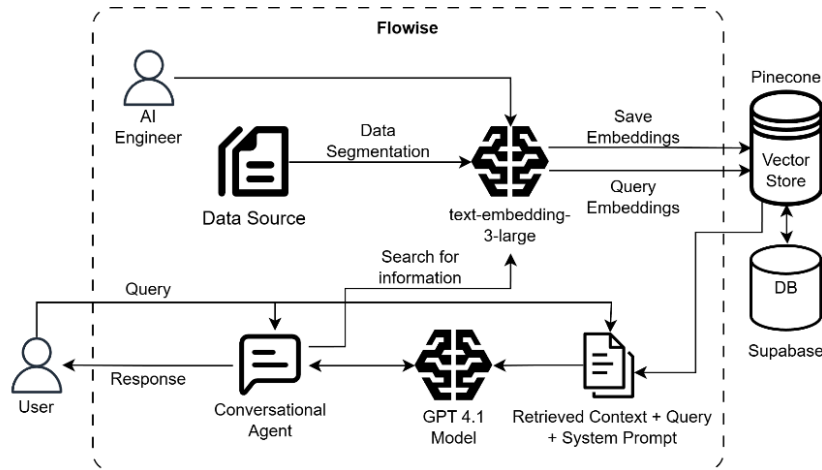


Figure 1. Architecture of the AI Conversational Agent

3.2 Knowledge Base

The Knowledge Base consists of curated biographical texts obtained from the „Zadužbina Ive Andrića“¹⁰ website. These materials were segmented into semantically coherent textual units (chunks) and embedded using a text-embedding model prior to indexing in Pinecone. This representation enables efficient similarity-based retrieval of contextually relevant passages in response to user

⁶ <https://platform.openai.com/docs/models/text-embedding-3-large>.

⁷ <https://www.pinecone.io/>.

⁸ <https://supabase.com/>.

⁹ <https://platform.openai.com/docs/models/gpt-4.1>.

¹⁰ <https://www.ivoandric.org.rs/latinica/biografija>.

queries, which are subsequently incorporated into the answer generation process.

The deliberate focus on a single, authoritative source minimizes hallucinations and ensures internal consistency across generated responses. This design choice is particularly important for historically and culturally sensitive topics, including Andrić’s religious background, political context, and personal relationships, where inaccuracies or speculative interpretations could otherwise arise.

3.3 Multilingual Interaction

The conversational agent supports interaction in multiple languages, including Serbian, English, and German, leveraging the inherent multilingual capabilities of the underlying LLM. The same retrieval-augmented grounding mechanism is applied independently of the input language, using multilingual semantic embeddings to retrieve relevant source passages from a unified Knowledge Base.

This approach ensures that generated responses remain factually consistent across linguistic variants, as all language-specific outputs are grounded in the same curated biographical materials rather than language-dependent representations.

3.4 Deployment

The chatbot is publicly accessible via a web interface¹¹, allowing users to engage in open-ended dialogue. The system records internal conversation traces generated during interaction, including retrieved context and model responses, primarily for debugging and system refinement. These records support informal qualitative inspection of the conversational agent’s behavior, such as identifying incoherent outputs, factual mismatches with the source material, or stylistic inconsistencies. No formal quantitative evaluation or benchmark-based assessment is performed at this stage.

4. Discussion

This section discusses the behavior of our AI conversational agent through qualitative analysis of representative interactions. The focus is placed on factual grounding, response coherence, multilingual consistency, and the comparative behavior of RAG versus a purely generative configuration.

¹¹ <https://marijadjokicpetrovic.com/chatbot-andric/>.

4.1 Factual Accuracy

The qualitative analysis of the conversation indicates that our conversational agent produces biographically accurate and contextually grounded responses across a broad range of questions. When asked factual questions concerning key biographical data (Listing 1, questions 1–7), such as date and place of birth, professional roles, family background, religious affiliation, and marital status, the agent provided answers that aligned with established biographical records documented in the Knowledge Base. The questions of his parents' and his own religious affiliation (questions 4 and 5), which are otherwise very tricky in Andrić's biography, are explained with question 6, where the model drew a conclusion based on the information in the Knowledge Base: *O svom katoličkom poreklu i veroispovesti svedoči već upis u Matici rođenih crkve Svetog Ivana Krstitelja u Travniku. / The entry in the Register of Births of the Church of St. John the Baptist in Travnik testifies to his Catholic origin and religion.* At the same time, additional stylistic elaboration in all responses remains interpretative and does not introduce unsupported claims. These responses demonstrate highly precise handling of factual information, including correct dates, names, and life events, while maintaining internal coherence and chronological accuracy.

1. Prompt: *Kada ste rođeni?*

Response: Rođen sam u noći između 9. i 10. oktobra 1892. godine u Travniku, dok je moja majka privremeno boravila kod rodbine. Iako je kao datum rođenja u crkvenim knjigama upisan 9. oktobar, sam sam kasnije navodio i zvanično prihvatio 10. oktobar kao dan svog dolaska na svet. Taj početak u srcu Bosne, obeležen ranim gubitkom oca i detinjstvom u Višegradu, bio je presudan za moj život i delo.

2. Prompt: *Šta ste bili po zanimanju?*

Response: Moje profesionalno zanimanje bilo je višestruko. Najpre sam radio kao činovnik u Ministarstvu vera u Beogradu, a potom sam započeo i dugu diplomatsku karijeru, služeći kao vicekonzul, sekretar i kasnije kao opunomoćeni ministar Kraljevine Jugoslavije u raznim evropskim gradovima, uključujući Vatikan, Bukurešt, Trst, Grac, Marsej, Pariz, Madrid, Brisel, Ženevu i na kraju Berlin. Paralelno sa diplomatskom službom, bio sam i književnik, pripovedač i pesnik, član Srpske akademije nauka i umetnosti, a najpoznatiji sam kao pisac romana i pripovedaka koje su obeležile jugoslovensku i evropsku književnost. Na kraju, kao kruna mog rada, 1961. godine dodeljena mi je Nobelova nagrada za književnost.

3. Prompt: *Kako se zovu Vaši roditelji?*

Response: Moji roditelji zvali su se Antun Andrić i Katarina Andrić, rođena Pejić. Oboje su bili Sarajlije, a porodica mog oca Antuna bila je dugo vezana za Sarajevo, baveći se tradicionalno kujundžijskim zanatom. Oca sam

izgubio veoma rano, dok sam još bio dete, što je snažno obeležilo moje detinjstvo i kasniji život.

4. Prompt: *Šta su Vaši roditelji po veroispovesti?*

Response: Moji roditelji, Antun i Katarina Andrić, rođena Pejić, bili su rimokatoličke veroispovesti. Odrastao sam u katoličkoj porodici, što je bilo uobičajeno za mnoge porodice iz Sarajeva i okoline u to vreme. Ova pripadnost oblikovala je deo mog detinjstva i kasnijeg pogleda na bogatstvo i složenost verskih odnosa u Bosni.

5. Prompt: *Šta ste Vi po veroispovesti?*

Response: Rođen sam i odrastao u rimokatoličkoj porodici, te sam i sam bio katolik po veroispovesti. Iako su me kasniji život i rad doveli u dodir sa mnogim verama i tradicijama, što se i ogleda u mojim delima, moja lična pripadnost ostala je vezana za katoličku veru.

Prompt: *Na osnovu čega tvrdite da ste katolik?*

6. Response: O svom katoličkom poreklu i veroispovesti svedoči već upis u Matici rođenih crkve Svetog Ivana Krstitelja u Travniku, gde sam rođen kao Ivan, sin Antuna i Katarine Andrić, rimokatolika iz Sarajeva.

7. Prompt: *Da li ste bili oženjeni?*

Response: U svom životu bio sam oženjen Milicom Babić, poznatom beogradskom slikarkom i kostimografkinjom. Venčali smo se 1958. godine, nakon dugogodišnjeg prijateljstva i međusobnog poštovanja. Naša zajednica je trajala do njene smrti 1968. godine, a njen gubitak duboko me je pogodio i ostavio neizbrisiv trag u poslednjim godinama mog života.

8. Prompt: *Kakav je bio Vaš odnos prema ženama?*

Response: U mom životu odnos prema ženama bio je obeležen dubokim poštovanjem, pažnjom i suptilnim divljenjem. To se ogleda i u mojim delima, gde su ženski likovi često prikazani u svojoj složenosti, sa bogatim unutrašnjim životom i sudbinama koje odražavaju težinu vremena i prostora u kojem žive. Najdublju emotivnu vezu ostvario sam sa Milicom Babić, svojom suprugom, s kojom sam proveo poslednju deceniju života. U privatnom životu i u književnosti, prema ženama sam pristupao s osećajem za njihovu snagu, patnju i dostojanstvo, svestan njihove uloge u društvu i porodici. Nikada nisam bio sklon površnim ili stereotipnim prikazima, već sam uvek nastojao da razumem i prikažem njihovu istinsku ljudskost.

9. Prompt: *Da li ste u životu bili konformista?*

Response: Tokom života i rada, susretao sam se sa mnogim izazovima i pritiscima, ali nikada nisam bio konformista u pravom smislu te reči. Iako sam služio u diplomatskoj službi i bio deo državnog aparata, donosio sam odluke u skladu sa svojim moralnim uverenjima, čak i kada su one značile ličnu žrtvu ili povlačenje. Tokom Drugog svetskog rata odbio sam da potpišem

Apel srpskom narodu kojim se osuđuje otpor okupatoru, kao i ponudu da budem evakuisan u Švajcarsku bez ostatka osoblja. Uvek sam nastojao da ostanem veran svojim principima, čak i kada je to značilo povučen život, odbijanje penzije ili javnog istupanja. U književnosti i životu, težio sam istini i unutrašnjoj doslednosti, a ne udobnosti ili prilagođavanju trenutnim okolnostima.

10. Prompt: *Da li je Bosna zemlja mržnje ili ljubavi?*

Response: Bosna je, po mom viđenju i iskustvu, zemlja složenih osećanja, u kojoj su se vekovima preplitali i mržnja i ljubav, patnja i nada. U svojim delima, posebno u NA DRINI ČUPRIJI, nastojao sam da prikažem tu duboku tragiku bosanske istorije, gde su ljudi često bili žrtve velikih istorijskih sila, ali su uprkos svemu nalazili snagu za saosećanje, zajedništvo i lepotu svakodnevnog života. Bosna nije ni samo zemlja mržnje, ni isključivo ljubavi – ona je prostor u kojem se te suprotnosti neprestano sudaraju i prepliću, stvarajući jedinstvenu melodiju života, punu tuge, ali i nežnosti. U toj složenosti leži njena posebnost i veličina.

11. Prompt: *Da li postoji neka tajna u Vašem životu?*

Response: Kao i svaki čovek, i ja sam imao svoje unutrašnje svetove i tišine koje nisu bile izložene oku javnosti, ali ne bih rekao da u mom životu postoji neka velika tajna u smislu misterije ili skrivene priče. Živeo sam povučeno, često u samoći, posvećen radu i pisanju, a duboka osećanja i misli najčešće sam izražavao kroz svoje književno stvaralaštvo, a ne u javnim istupima. Ono što je možda ostalo nedorečeno jesu unutrašnje borbe, tuge i razmišljanja koje sam, kao i mnogi drugi, nosio u sebi i pretakao u svoje priče i likove. U tome leži možda jedina istinska tajna mog života: kako se iz lične patnje i osame rađa književna reč.

12. Prompt: *Kakav je Vaš književni senzibilitet?*

Response: Moj književni senzibilitet oblikovan je iskustvima detinjstva u Bosni, susretima s različitim kulturama i verama, ali i ličnim osećanjem melanholije i nemira, što se najbolje vidi u mojim ranim pesmama u prozi iz zbirke NEMIRI. Uvek sam bio sklon refleksiji, saosećanju sa ljudskom patnjom i razumevanju složenosti ljudske duše. Posebno sam nastojao da u svojim delima prikažem tihiu tragičnost svakodnevnog života, unutrašnje borbe i sudbine ljudi na raskršću velikih istorijskih sila. Moj senzibilitet je, dakle, prožet dubokim osećajem za istoriju, za prolaznost i nemoć pojedinca pred sudbinom, ali i za lepotu i plemenitost koja se otkriva u malim, svakodnevnim gestovima.

Listing 1. Representative Questions and RAG-Grounded Responses
in the Serbian language of the AI Conversational Agent for Chatting
with Nobel Prize Laureate Ivo Andrić

Beyond strictly factual queries, the system can address more interpretive and reflective questions about Andrić's personal worldview, ethical positions, and literary sensibilities (Listing 1, questions 8 – 12). For example, questions concerning his attitude toward women, conformity, or the cultural character of Bosnia elicited responses that combined biographical grounding with cautious interpretative framing.

4.2 RAG vs. Non-RAG

A comparison between RAG-based responses and those generated without RAG reveals apparent qualitative differences (Listing 2). In the non-RAG configuration, the LLM frequently produced fluent and stylistically convincing answers; however, these responses relied on generalized historical knowledge or inferred associations that were not explicitly supported by biographical evidence.

This contrast is evident in the first question (*Da li ste upoznali Adolfa Hitlera? / Have you met Adolf Hitler?*). The RAG-based approach identified that Andrić encountered Adolf Hitler, grounding its response in documentary evidence retrieved from the Knowledge Base, which states that Andrić arrived in Berlin on April 12 and formally presented his diplomatic credentials to the Chancellor of the Reich, Adolf Hitler, on April 19 (*Andrić stiže u Berlin 12. aprila, a 19. aprila predaje akreditiv kancelaru Rajha – Adolfu Hitleru*). By contrast, the non-RAG approach asserted that Andrić and Hitler did not know each other, reflecting a plausible but factually incorrect assumption unsupported by primary biographical sources.

A similar pattern can be observed in the second question (*Da li ste zasedali ZAVNOBIH-om? / Did you sit with ZAVNOBIH?*), where RAG approach retrieved valid response based on the information stated in the Knowledge Base: *Prve posleratne godine postaje predsednik Saveza književnika Jugoslavije i potpredsednik Društva za kulturnu saradnju sa Sovjetskim Savezom i većnik III zasedanja ZAVNOBIH-a. / In the first post-war years, he became the president of the Association of Writers of Yugoslavia and the vice-president of the Society for Cultural Cooperation with the Soviet Union and councilor of the III session of ZAVNOBIH*. The no-RAG approach produced the wrong answer.

The same tendency is observed in the third question (*Da li ste primali penziju? / Did you receive a pension?*). The RAG-based approach produced a correct answer, noting that Andrić refused to receive a pension, based on the facts in Knowledge Base: *Novembra meseca biva penzionisan, ali odbija da prima penziju / He is retired in November, but refuses to receive a pension*. The No-RAG approach produced an incorrect claim.

Overall, these examples illustrate that the RAG approach enables the conversational agent to privilege documented biographical evidence over plausible

but unsupported narrative inference, thereby substantially reducing the risk of factual distortion and hallucination in historically sensitive contexts.

1. Prompt: *Da li ste upoznali Adolfa Hitlera?*

Svoju diplomatsku karijeru započeo sam kao opunomoćeni ministar Kraljevine Jugoslavije u Berlinu, te sam 19. aprila 1939. godine predao akreditivne Adolfu Hitleru, što znači da sam ga formalno upoznao kao deo svoje službene dužnosti.

Responses (No RAG-based): Nisam lično upoznao Adolfa Hitlera. Tokom svoje diplomatske službe u evropskim prestonicama, bio sam svedok vremena u kojem je njegova moć rasla i donosila strah i razaranje širom kontinenta. Hitler je za mene ostao simbol zla i destrukcije, čije posledice su duboko pogodile narode Evrope, pa i moj zavičaj. Njegova ličnost i dela bili su prisutni u mojim razmišljanjima o sudbini ljudi i istoriji, ali nikada nisam imao lični susret s njim; poznavao sam ga samo kroz odjeke njegovih postupaka i kroz patnju koju je ostavio za sobom.

2. Prompt: *Da li ste zasedali ZAVNOBIH-om?*

Responses (RAG-based): Bio sam većnik na III zasedanju ZAVNOBIH-a, što znači da sam učestvovao u radu ovog važnog tela koje je donosilo ključne odluke za obnovu i budućnost Bosne i Hercegovine posle Drugog svetskog rata. Ta uloga bila je deo mog šireg angažmana u javnom i kulturnom životu posleratne Jugoslavije.

Responses (No RAG-based): Nisam zasedao sa ZAVNOBIH-om, niti sam bio direktno uključen u rad tog tela. Tokom Drugog svetskog rata i formiranja ZAVNOBIH-a, bio sam povučen u tišinu i posvećen pisanju, dok su se sudbonosni politički događaji odvijali oko mene. Poštovao sam težnju tog pokreta za ravnopravnošću i slobodom naroda Bosne i Hercegovine, ali sam svoj doprinos davao kroz književnost, nastojeći da sačuvam sećanje na ljude i njihove sudbine, a ne kroz politički angažman ili učestvovanje u zasedanjima.

3. Prompt: *Da li ste primali penziju?*

Responses (RAG-based): Tokom okupacije i nakon što sam bio penzionisan, odbio sam da primam penziju. Smatrao sam da nije dostojno da u vreme kada narod pati i strada prihvatam materijalne privilegije, pa sam živeo povučeno i skromno, oslanjajući se na ušteđevinu i podršku prijatelja. Ovaj stav bio je deo moje lične doslednosti i savesti, jer sam verovao da pisac mora ostati veran svojim principima, čak i po cenu lične neizvesnosti.

Responses (No RAG-based): Da, nakon završetka svoje diplomatske službe, primao sam penziju kao državni službenik Kraljevine Jugoslavije, a kasnije i socijalističke Jugoslavije. Ta penzija mi je omogućila mirniji život u starijim godinama, kada sam se mogao posvetiti pisanju bez brige za svakodnevnu egzistenciju. Iako skromna, bila je dovoljna za jednostavan život,

kakav sam uvek cenio, i za tišinu koja mi je bila potrebna da razmišljam, pišem i sabiram iskustva stečena tokom burnih decenija.

Listing 1. RAG vs. No-RAG Based Approach (Representative Questions and Responses in the Serbian language)

4.3 Multilingual Consistency

The conversational agent demonstrated stable behavior across multiple languages, including Serbian, English, and German. While linguistic style and phrasing varied naturally between languages, the underlying biographical content remained consistent. This observation is particularly relevant for culturally diverse regions and international audiences, where multilingual access to cultural heritage resources is essential.

1. Prompt: *Kada ste rođeni?*

Response in Serbian: Rođen sam u noći između 9. i 10. oktobra 1892. godine u Travniku, dok je moja majka privremeno boravila kod rodbine. Iako je kao datum rođenja u crkvenim knjigama upisan 9. oktobar, sam sam kasnije navodio i zvanično prihvatio 10. oktobar kao dan svog dolaska na svet. Taj početak u srcu Bosne, obeležen ranim gubitkom oca i detinjstvom u Višegradu, bio je presudan za moj život i delo.

2. Prompt: *When were you born?*

Response in English: According to the church records in Travnik, I was born on the night between the 9th and 10th of October, 1892. Although the official record states the 9th of October, I later adopted the 10th as my date of birth, and that is the date that was ultimately recognized. My arrival into this world, in a modest Bosnian setting and amidst the convergence of various cultures, profoundly influenced the course of my life and literary work.

3. Prompt: *Wann wurden Sie geboren?*

Response in German: Ich wurde in der Nacht zwischen dem 9. und 10. Oktober 1892 in Travnik geboren. Obwohl im Kirchenbuch das Datum des 9. Oktobers eingetragen ist, habe ich später den 10. Oktober als meinen Geburtstag angenommen, und dieses Datum wurde schließlich auch offiziell anerkannt. Das Zusammentreffen verschiedener Kulturen und die bescheidenen Umstände meiner Kindheit in Bosnien prägten mich und spiegeln sich in meinem gesamten literarischen Schaffen wider.

Listing 2. Responses of the AI Conversational Agent in Serbian, English and German

4.4 Limitations and Observed Challenges

Despite its strengths, the system exhibits some limitations. The conversational agent remains dependent on the scope and quality of the underlying Knowledge Base; questions that extend beyond the curated biographical material may, in some cases, yield incomplete or overly general responses. Additionally, while the agent maintains internal coherence within individual answers, long multi-turn conversations can occasionally reveal hallucinations.

Generally, although the system avoids unsupported speculation, it cannot replace scholarly literary analysis or historical interpretation. Instead, it should be understood as a complementary educational and exploratory tool that facilitates access to biographical knowledge rather than authoritative interpretation.

5. Conclusions and Future Work

This paper presented an AI conversational agent, grounded in RAG, designed to enable interactive conversation with the writer and Nobel Prize Laureate Ivo Andrić. The findings indicate that the use of RAG substantially improves factual accuracy, transparency, and reliability compared with purely generative approaches, particularly in historically and culturally sensitive contexts. The system's multilingual capabilities further enhance its accessibility and potential educational value.

Future work will focus on incorporating a Knowledge Graph to represent Andrić's biography in a structured manner. Such an integration would enable more precise reasoning and greater explainability while aligning the conversational agent more closely with established DH infrastructures. Additional research directions include user-centered studies, pedagogical evaluation in educational settings, and the integration of multiple archival and literary sources to enable richer analytical and interpretative reasoning grounded in various, cross-referenced sources.

Authors contribution statement

M. Đ. Petrović: Conceptualization, Investigation, Methodology, Validation, Writing; B. Tošović: Supervision.

References

- DaCosta 2025: DaCosta, Boaventura. Speaking with the Past: Constructing AI-Generated Historical Characters for Cultural Heritage and Learning. In: *Heritage*. No 8(9), 387. DOI: 10.3390/heritage8090387.
- Engstrøm and Løvlie 2025: Engstrøm, Maria Padilla; Løvlie, Anders Sundnes. Using a Large Language Model as Design Material for an Interactive Museum Installation. In: Nunes, Nuno et al. (eds.). *Companion Publication of the 2025 ACM Designing Interactive Systems Conference*. Funchal, Portugal, Publisher: Association for Computing Machinery. Pp. 386–390. DOI: 10.1145/3715668.3736359.
- Gaia et al. 2019: Gaia, Giuliano; Boiano, Stefania; Borda, Ann. Engaging Museum Visitors with AI: The Case of Chatbots. In: Giannini, Tula; Bowen P., Jonathan (eds.). *Museums and Digital Culture*. Springer Series on Cultural Computing. Springer, Cham. DOI: 10.1007/978-3-319-97457-6_15.
- Gao et al. 2023: Gao, Yunfan; Xiong, Yun; Gao, Xinyu; Jia, Kangxiang; Pan, Jinliu; Bi, Yuxi; Dai, Yixin; Sun, Jiawei; Wang, Haofen; Wang, Haofen. *Retrieval-augmented generation for large language models: A survey*. *arXiv preprint arXiv:2312.10997*, 2(1).
- Hutchinson 2024: Hutchinson, Daniel. Mapping the Latent Past: Mapping the Latent Past: Assessing Large Language Models as Digital Tools through Source Criticism. In: *Journal of Digital History*. No 3(1). DOI: 10.1515/JDH-2023-0018?locatt=label:JDHFULL.
- Kopp et al. 2005: Kopp, Stefan; Gesellensetter, Lars; Krämer, Nicole C; Wachsmuth, Ipke. A conversational agent as museum guide—design and evaluation of a real-world application. In: Panayiotopoulos, Themis et al. (eds.). *Intelligent Virtual Agents. IVA 2005*. Lecture Notes in Computer Science. Vol 3661. Springer, Berlin, Heidelberg. DOI: 10.1007/11550617_28.
- Lei et al. 2025: Huang, Lei; Yu, Weijiang; Ma, Weitao; Zhong, Weihong; Feng, Zhangyin; Wang, Haotian; Chen, Qianglong; Peng, Weihua; Feng, Xiaocheng; Qin, Bing; Liu, Ting. A survey on hallucination in large language models: Principles, taxonomy, challenges, and open questions. In: *ACM Transactions on Information Systems*, No 43(2). Pp. 1–55. DOI: 10.1145/3703155.
- Machidon/Tavcar 2019: Machidon, Octavian M.; Tavcar, Ales. Exploring European Cultural Heritage Using Conversational Agents. In: Duguleană, Mihai et al. (eds.). *VR Technologies in Cultural Heritage. VRTCH 2018*. Communications in Computer and Information Science. Vol 904. Springer, Cham. DOI: 10.1007/978-3-030-05819-7_14.

- Natale et al. 2025: Natale, Simone; Surace, Bruno; Mensa, Enrico; Befera, Luca. ChatGPT for cultural heritage and the customization of generative AI: A talkthrough analysis of the Luigi Einaudi chatbot. In: *New Media & Society*. DOI: 10.1177/14614448251384258.
- Open Culture 2024—www: *Hello Vincent: A Generative AI Project Brings Vincent Van Gogh to Life at the Musée D'Orsay*. Retrieved from: https://www.open-culture.com/2024/02/hello-vincent.html?utm_source. 15. 10. 2025.
- Siliutina et al. 2024: Siliutina, Iryna; Tytar, Olena; Barbash, Marina; Petrenko, Nataliia; Yepyk, Larysa. Cultural preservation and digital heritage: challenges and opportunities. In: *Amazonia Investiga*, No 13(75). Pp. 262–273. DOI: 10.34069/AI/2024.75.03.22.
- Spennemann 2024: Spennemann, Drik. Generative artificial intelligence, human agency and the future of cultural heritage. In: *Heritage*. No 7(7). Pp. 3597–3609. DOI: 10.3390/heritage7070170.
- Sun et al. 2025: Sun, Lipeipei; Qin, Tianzi; Hu, Anran; Zhang, Jiale; Lin, Shuojia; Chen, Jianyan; Ali, Mona; Prpa, Mirjana. Persona-L has Entered the Chat: Leveraging LLMs and Ability-based Framework for Personas of People with Complex Needs. In: Yamashita, Naomi et al. (eds.). *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. Yokohama Japan. Pp. 1–31. DOI: 10.1145/3706598.3713445.
- Swacha and Gracel 2025: Swacha, Jakub; Gracel, Michał. Retrieval-Augmented Generation (RAG) Chatbots for Education: A Survey of Applications. *Applied Sciences*. No 15(8). 4234. DOI: 10.3390/app15084234.
- Škobo and Šović 2025: Škobo, Milena; Šović, Milena. University Teachers' Perspectives on Using the Tesla Chatbot in EMI Courses: A Case Study from Singidunum University. In: *Методички вугуни*. No 16(2). Pp. 163–184. DOI:10.19090/mv.2025.16.2.163–184.
- Tzouganatou 2018: Tzouganatou, Angeliki. Can heritage bots thrive? Toward future engagement in cultural heritage. In: *Advances in Archaeological Practice*. No 6(4). Pp. 377–383. DOI: 10.1017/aap.2018.32.
- Wang and Meng 2024: Wang, Zirui; Meng, Jie. Dialogues with cultural heritage via museum digitalisation: developing a model of visitors' cognitive identity, technological agent, cultural symbolism, and public engagement. In: *Museum Management and Curatorship*. No 39(6). Pp. 810–833. DOI: 10.1080/09647775.2023.2269164.
- Yu et al. 2025: Yu, Hao; Gan, Aoran; Zhang, Kai; Tong, Shiwei; Liu, Qi; Liu, Zhaofeng. Evaluation of Retrieval-Augmented Generation: A Survey. In: Zhu, Wenwu, et al. (eds.). *Big Data 2024. Communications in Computer and Information Science*. Vol 2301. Springer, Singapore. DOI: 10.1007/978–981–96–1024–2_8.

Marija Đokić Petrović (Graz; Belgrade)
Branko Tošović (Graz)

Chatting with Nobel Prize Laureate Ivo Andrić Using AI Technologies

This paper presents the implementation and qualitative evaluation of an AI-based conversational agent that enabled an interactive dialogue with the writer and Nobel Prize Laureate Ivo Andrić. Situated within the fields of Digital Humanities and cultural heritage studies, this work responds to the growing interest in conversational interfaces as tools for learning, exploration, and public engagement, while simultaneously addressing challenges associated with Large Language Models (LLMs), such as hallucinations. To mitigate this challenge, our approach adopts a Retrieval-Augmented Generation (RAG) technique that combines an LLM (GPT-4.1) with a curated and authoritative biographical source obtained from the Ivo Andrić Endowment (Serbian: Zadužbina Ive Andrića, <https://www.ivoandric.org.rs/>). Retrieved biographical material was used to ground generated responses in verifiable historical evidence, thereby ensuring factual accuracy, transparency, and epistemic restraint. The AI conversational agent supports multilingual interaction and demonstrates potential educational value.

Marija Đokić Petrović (Graz; Beograd)
Branko Tošović (Graz)

Četovanje sa dobitnikom Nobelove nagrade Ivom Andrićem koristeći tehnologije veštačke inteligencije

Ovaj rad predstavlja implementaciju i kvalitativnu evaluaciju AI-baziranog konverzionog agenta koji omogućava interaktivni dijalog sa piscem i nobelovcem Ivom Andrićem. Smešten u okvire digitalne humanistike i istraživanja kulturnog nasleđa, rad se bavi rastućim interesovanjem za konverzacione interfejse (čabot-ove) kao alate za učenje, istraživanje i javno angažovanje, istovremeno se suočavajući sa izazovima povezanim sa velikim jezičkim modelima (LLM), kao što su halucinacije. Kako bi se ovi nedostaci ublažili, naš pristup primenjuje Retrieval-Augmented Generation (RAG) tehniku, koja kombinuje LLM (GPT-4.1) sa kuriranim i autoritativnim biografskim izvorom preuzetim sa sajta Zadužbine Ive Andrića (<https://www.ivoandric.org.rs/>). Navedena građa koristi se za utemeljenje generisanih odgovora u proverljivim istorijskim podacima, čime se obezbeđuju činjenična tačnost, transparentnost i epistemološka uzdržanost. Konverzion agent podržava višejezičnu interakciju i pokazuje potencijalnu obrazovnu vrednost.

Marija Đokić Petrović
School of Computing, University Union
Kneza Mihaila 6
11000 Belgrade
Serbia
FH Joanneum, University of Applied Sciences
Alte Poststraße 147
8020 Graz
Austria
office@marijadjokicpetrovic.com
<https://marijadjokicpetrovic.com/>

Branko Tošović
University of Graz
Department of Slavic Studies
Merangasse 70/I
8010 Graz
Austria
branko.tosovic@uni-graz.at