Univerza v Mariboru

FRAMEWORK PROGRAMME OF EARLY STAGE RESEARCHER TRAINING[1]

1. BASIC DATA

Mentor's name and surname	**Izidor Mlakar**	Mentor's register number at ARIS (SICRIS)**:**	**50324**
Mentor's e-mail:	Izidor.mlakar@um.si	Mentor's tel. no.:	*+386 2 220 7267*
Research programme (RP) leader's name and surname:	full prof. dr. Zdravko Kačič	RP leader's register number at ARIS (SICRIS):	**06821**
Title of research programme:	Advanced methods of interaction in telecommunication	RP's Register number at ARIS (SICRIS):	P2-0069
Research organisation (RO) of University of Maribor, where training shall be conducted:	**Faculty of Electrical Engineering and Computer Science**	RO Register number at ARIS (SICRIS):	0796
Research field according to ARIS classification:	2.08 Telecommunications	Research field according to Ortelius classification (EURAXESS)	37.3 Communication technology

2. DEFINITION OF RESEARCH PROBLEM AND GOALS OF DOCTORAL RESEARCH[2]

Starting point of research task of the early stage researcher and its position in the research programme, where the mentor is included, work hypothesis, research goals and foreseen result with emphasis on an original contribution to science:

Research Problem Definition

Conversational AI is increasingly deployed in healthcare, education, public services, and intelligent environments, where its ability to adapt to cultural norms, implicit user needs, and multi-topic dialogues is critical. However, traditional conversational AI systems (e.g. generative AI models, such as, ChatGPT-4o, Mistral, Claude, Falcon, LLama, DialogGPT) primarily focus on linguistic adaptation

[1] *Term early stage researcher (ESR) is written in male form and used as neutral for women and men.*

[2] *Research and study programme of training have to harmonise with contents of the research programme, where the mentor is a member.*

but fail to incorporate deeper cultural nuances, dynamic spatial-temporal contexts, and multimodal interaction principles. This gap limits their effectiveness in real-world scenarios such as healthcare, education, and public services, where cultural sensitivity and real-time adaptation are crucial. Moreover, adaptive AI behavior must be formally correct and reliable, especially when applied in high-stakes domains such as healthcare, where errors could have significant consequences. Current AI models lack mechanisms for verifying the correctness, fairness, and robustness of their adaptive responses, which can lead to miscommunication, biases, and potential harm to vulnerable users.

As such we identify the following four key research gaps:

1. The integration of cultural practices, social norms, and environmental contexts into conversational AI. AI system also fail to adjust its responses based on social hierarchies, politeness norms, and regional communication styles.
2. Conversational AI often fails to infer underlying questions users imply rather than explicitly state. This is particularly problematic in high-context domain, such as health, where communication relies on implicit meanings, non-verbal cues, and shared cultural knowledge, rather than direct, explicit statements
3. Formal correctness of adaptive AI behavior to ensure reliability and robustness (e.g. logically consistent and unbiased responses) in critical domains. Namely, lack of formal verification techniques can result in hallucinated, misleading, or harmful AI-generated responses especially in high-stakes scenarios.

Research Goals

The primary goal is to develop a culturally adaptive conversational AI framework that integrates cultural adaptation mechanisms, implicit question understanding, conversational coherence and formal correctness of AI response.

The secondary goal is to verify the conversational AI framework with a series of experiments in critical domain of health.

Specific objectives include:

- To develop a 4D knowledge model representation for cultural, spatial, and temporal adaptation.
- To establish a framework for implicit reasoning and topic continuity in AI dialogues.
- To design meta-learning techniques for rapid adaptation to new cultural contexts.
- To develop formal verification methods for checking AI correctness in high-stakes scenarios.
- To validate the developed models and framework in multiple domains, e.g. patient-AI communication, AI tutoring, and AI-driven assistants.

Work Hypothesis

A context-adaptive conversational AI system integrating cultural-spatial-temporal modeling, implicit reasoning, and formal verification will significantly improve human-AI interaction across healthcare, and education, (and public service) domains, while maintaining system reliability across high-stakes domains through dynamic adaptation to cultural nuances, conversational context, and user requirements.

Position in the Research Program

This proposed early stage research is aligned with the broader topic of the research program multimodal interfaces and system correctness assurance program Advanced methods of interaction in telecommunication, by contributing to the following topics:

- Multimodal User Interfaces – The research will enhance spoken language technologies for highly inflected languages by improving discourse generation capacity of conversational agents towards more natural and context-aware interactions.

- System Correctness Assurance – The research will investigate into formal verification for AI correctness and fairness, thereby ensuring the reliability of AI-driven interaction systems, even for vulnerable populations.

The research work is also well aligned with the research activities, that the supervisor, dr. Izidor Mlakar, is leading in multiple ongoing Horizon Europe Projects, e.g., HE SMILE (Supporting Mental Health in Young People: Integrated Methodology for cLinical dEcisions and evidence-based interventions), HE AI4HOPE (Artificial Intelligence based health, optimism, purpose and endurance in palliative care for dementia), HE CERTAIN (Certification for Ethical and Regulatory Transparency in Artificial Intelligence).

Expected Results and Original Contribution to Science

Theoretical Contributions.

- A novel framework for culturally adaptive AI with implicit reasoning and conversational coherence.
- A novel dynamic knowledge graph-based representation for cultural, spatial, and temporal awareness in AI adaptation.
- A formal model for validating the correctness of AI-generated responses.

Technical Contributions

- New AI adaptation mechanisms and architectures using meta-learning, few-shot learning, and retrieval-augmented generation.
- Formal verification methodologies for evaluating AI dialogue correctness in high-stakes domains.
- Conversational memory architectures to improve AI's ability to sustain topic coherence.

Social Contributions

- Increased AI inclusivity by accounting for cultural and linguistic diversity.
- Enhanced trust in AI systems through formal verification methods.
- Fair AI development by investigating bias-free, reliable, and context-aware responses.

3. STUDY PROGRAMME

Foreseen study programme, to which early stage researcher shall be enrolled in academic year 2025/2026:

Electrical Engineering

4. DESCRIPTION OF WORK AND TASKS

Year 1 will be dedicated to the formal design and development of the theoretical framework and initial modeling. The tasks will include:

• Literature Review and Gap Analysis: The PhD candidate will carry out in-depth review of existing research on cultural adaptation in conversational AI, implicit question understanding, conversational coherence, and formal correctness verification. Identify key challenges, research gaps, and benchmarking datasets. Define evaluation metrics for assessing the AI system's effectiveness.

• Design and development of Cultural Knowledge Representation Model: Building on the literature review, the PhD candidate will design a knowledge model that encodes cultural, spatial, temporal, and contextual information. This model will be designed to adapt dynamically, integrating social norms, environmental contexts, and real-time changes in cultural settings. A preliminary implementation of the knowledge graph will be developed during this period.
• Design and development of a Framework for Implicit Question Understanding and Conversational: The PhD candidate will develop an initial implicit reasoning model using retrieval-augmented generation (RAG) and neural inference techniques. Additionally, memory-based architecture will be investigated to improve multi-topic conversational tracking and long-term coherence. The output will be a functional prototype of the adaptive reasoning framework.

Year 2 will be dedicated to technical implementations, development and domain-specific adaptation. In Year 2 also protocols for evaluation/validation will be outlined. The tasks will include:

• Implementation of Adaptive Learning Techniques: The PhD student will refine meta-learning and few-shot learning techniques for real-time cultural adaptation, allowing the system to adapt with minimal labeled data. These techniques will be tested on different linguistic and cultural datasets to ensure robustness.
• Extension of the Implicit Question Understanding and Conversational Memory: Further improvements will be made to retrieval-augmented generation for implicit reasoning. Additionally, long-term conversational memory mechanisms will be enhanced to allow the system to track topic continuity across multi-turn dialogues. The outcome will be a more stable and contextually aware conversational AI model.
• Preparation of study protocols: Before conducting validation studies in healthcare and education, the PhD student will develop formal study protocols ensuring compliance with ethical guidelines, data privacy laws, and institutional review board (IRB) requirements. The protocols will include the (i) definition of study objectives, participant inclusion/exclusion criteria, (ii) outlined data collection, storage, and processing methods for conversational AI testing, (iii) and definition of mechanisms ensuring adherence to fairness, transparency, and accountability in AI evaluation.

Year 3 will be dedicated to further technological developments and refinements, early testing, execution of research studies, dissemination and wirtting of the doctoral disertation.

• Implementation of a conceptual Prototype Implementation and Initial Evaluation: The conceptual prototype of the culturally adaptive conversational AI system will be delivered, integrating the 4D knowledge graph, implicit reasoning framework, and conversational coherence tracking. Baseline evaluations will be conducted using existing NLP benchmarks.
• Experiments in Healthcare domain (and other targted, high-stakes domains): After obtaining ethics approvals from the relevant ethics committee, initial experiments will be conducted according to the protocols within the healthcare domain, where experiments will be focused on interpreting indirect patient concerns, maintain conversational accuracy, and adapt to AI driven discourse and personalized responses across specific cohorts.
• Refinement of Adaptive AI Mechanisms and Formal Verification: Results from early tests and experiments in healthcare: Feedback from experimental studies and testing will be used to enhance adaptive learning mechanisms, strengthen topic coherence tracking, and improve formal verification processes. AI decision-making transparency will be prioritized.

In addition to the aforementioned tasks, the PhD students will prepare and submit at least at least two journal articles, focusing on:
• Culturally adaptive AI and conversational coherence.
• Formal correctness of AI responses in high-stakes domains.

Findings will also be presented at AI, NLP, and AI-ethics-focused conferences.

5. REQUESTED LEVEL OF EDUCATION

MSCA or Equivalent

6. REQUESTED FIELD OF EDUCATION

Electrical Engineering, Computer Science and Information Technologies, Telecommunications, Informatics and Data Technologies, or Mathematics

7. KLASIUS SRV

Kliknite ali tapnite tukaj, če želite vnesti besedilo.

8. KLASIUS P

Kliknite ali tapnite tukaj, če želite vnesti besedilo.

9. REQUESTED KNOWLEDGE

In depth understanding of Artifical Intelligence and Deep Learning Fundamentals

Experience with Transformer architectures (e.g., BERT, GPT, T5), and their applications in conversational AI.

Experience with Natural Languge Processing (NLP), especially fundamentals of NLP and Multilingual NLP

Basic Understanding of Human-Computer Interaction (HCI) & Multimodal Interfaces

Basic Understanding of Bias, Explainability & Transparency in AI

10. REQUESTED SPECIAL REQUIREMENTS

Basic knowledge of Software Verification Tools for AI Model Checking: e.g. Deepchecks, TruEra, SymGen

Expert understanding and experience with relevant Deep Learning and ML Frameworks e.g.: PyTorch, TensorFlow and Scikit-learn and Transformers.

Expert understanding of and experience with NLP Libraries: spaCy, NLTK, Stanza.

Understanding of Graph Databases & Knowledge Representation: e.g. Neo4j, RDF, OWL, JSON-LD, SPARQL.

11. REQUESTED LANGUAGES

English at least B2, preferably C1

12. REQUESTED WORK EXPERIENCE

Documented work on AI-related projects, including projects relevant to the edcuational settings. Candidates with published work will be favoured

13. FORESEEN POSTDOCTORAL TRAINING

MSCA Postdoctoral Fellowships, Research work in Active European Research Projects, 1-2 month secondments with international partners, e.g. Univeristy of Bari (Italy), URV (Spain), AUTH (Greece), HUA (Greece), etc.

Mentor's signature:

Izidor Mlakar

Digitally signed by Izidor Mlakar
Date: 2025.02.05 14:40:33 +01'00'

Research programme leader's signature:

Digitalno podpisal ZDRAVKO KAČIČ
Datum: 2025.02.11 14:49:39 +01'00'

Name and surname of Dean or authorised person[3]:

Kliknite ali tapnite tukaj, če želite vnesti besedilo.

Signature of dean or authorised person:

Gorazd Štumberger

Digitalno podpisal Gorazd Štumberger
Datum: 2025.02.05 14:58:50 +01'00'

Place and date:

Kliknite ali tapnite tukaj, če želite vnesti besedilo.

Kliknite ali tapnite tukaj, če želite vnesti datum.

Stamp:

[3] The training program is signed by the dean of the member where the ESR's employment and training will take place.