HumanE AI Net:

The HumanE AI Network

Grant Agreement Number: 952026
Project Acronym: HumanE AI Net

Project Dates: 2020-09-01 to 2023-08-31 (extended)
Project Duration: 36 months (extended)

D3.1: Third year microproject results on learning, reasoning, perception and interaction (WPs 1, 2, 3) (Papers, tools, datasets)

Author(s): Antti Oulasvirta, John Shawe-Taylor,James Crowley, George Kampis
Contributing partners: Aalto, DFKI, UCL, Grenoble
Date: September 24, 2023
Approved by: Paul Lukowicz
Type: Report (R)
Status: Final
Contact: antti.oulasvirta@aalto.fi

Dissemination Level		
PU	Public	x

DISCLAIMER

DOCUMENT INFO

0.1 Authors

Authors	Institution	e-mail
Antti Oulasvirta (AO)	AALTO	Antti.oulasvirta@aalto.fi
John Shawe-Taylor (JST)	UCL	j.shawe-taylor@ucl.ac.uk
James Crowley (JC)	GRENOBLE	james.crowley@inria.fr
George Kampis (GK)	DFKI	George.Kampis@dfki.de

0.2 Document History

Revision		
Date	**Lead Author(s)**	**Comments**
20.09.2023	AO	Initial draft
22.09.2023	JST, JC, GK	Comments
22.09.2023	AO	Polishing
23.09.2023	AO	Reorg of Sec 3 based on GK's input

HUMANE AI NET

Table of Contents

0.1 Authors 2
0.2 Document History 2
Table of Contents 3
0. Introduction 4
0.1 What is a micro-project? 5
0.2 Microprojects and financing 5
1. WPs1 and 2: Human-in-the-Loop Machine Learning, Reasoning and Planning 6
1.1. Overview of goals 6
1.2. Completed microprojects 6
2021-3-WP123: Memory Aware Conversational AI to aid virtual Team-Meetings 6
2022-9-WP12: Graybox methods for augmenting human-driven narrative analyses 8
2022-9-WP12: Multi-Relational Contextual Reasoning for Complex Scene Generation for Autonomous Vehicle Data 9
1.3. On-going microprojects 11
2022-3-WP12: Knowledge Extraction Through Prompting on Pre-trained Language Models 11
2022-4-WP12: Graybox methods for augmenting human-driven narrative analyses – beta release of the software package 12
2022-2-WP12: Learning from Richer Feedback Through the Integration of Prior Beliefs 13
2022-3-WP123: Intelligent crowdsourcing of geolocation tasks in natural disasters 15
2022-3-WP12: A tool for mitigating algorithmic biases through explanations 16
2022-3-WP1238: Educational module Human-Interactive Robot Learning (HIRL) ... 18
2022-3-WP12: LWMs teaching to teach 19
2022-3-WP123: Proactive announcement based on DEL (Dynamic Epistemic Logic) 21
2. WP2: Multimodal Perception and Modeling 22
2.1. Overview of goals 22
2.2. Completed microprojects 22
2021-5-WP26: Neural Mechanism in Human Brain Activity During Weight Lifting 22
2022-3-WP2: Natural Imitation of Dance Moves and Human Gestures with a Humanoid Robot 24
2022-4-WP246: Polarization with the Friedkin-Johnsen model over a dynamic social network 25

2022-9-WP235:Multilingual and Multimodal conversational agent combined with search engine models. 27

2022-9-WP235: Use of dialog context to boost ASR/NLG/TTS and improve the overall quality of voice dialog systems 30

2.3. On-going microprojects 32

3. WP3: Human AI Collaboration and Interaction 32

3.1. Overview of goals 32

3.2. Overview of WP3 microproject status (September 2023) 32

3.3. Completed microprojects 33

3.1 2021-3-WP3: Exploring the impact of Agency on Human-Computer Partnerships 33

2022-4-WP3: Making for Everyone: Interactive, voice-based support for the design of 3D objects for digital fabrication 35

2022-4-WP35: Human-machine collaboration for content analysis in context of Ukranian war 36

2022-5-WP3: In Sync: Synchronization in interaction between Humans and Embodied AI Systems 37

2022-7-WP3 A Scoping Review on Artificial Intelligence and Extended Reality 39

3.4. On-going microprojects 40

2023-1-WP3: Storyboarder: combining Play Script and Image Generation 40

2022-3-WP123: Gesture-based Interactive Grounding for Mixed-Reality Human-AI Collaboration 40

2022-3-WP3: Exploring the Role of AI in Amplifying Creative Potential: Fostering Ownership In Human-AI Collaborations 42

2022-3-WP3: Exploring Multi-Modal Interaction Concepts to Support Users in AI-aided Co-Creation of 3D Objects in Virtual Environments 43

2022-3-WP3: Grounded dialog models from observation of human-to-human conversation 45

2022-3-WP367: AI Aviation Assistant: An Intelligent Pilot Support Tool 47

0. Introduction

This section introduces the concept of a microproject (MP) and how it has evolved since the original work programme. We will also give some statistics on the development of MPs within the project, concluding with an outline of how we anticipate that the MP activity may develop during the remainder of the network.

Following this introduction, there are three separate sections for the three technical work packages (WPs) focused on learning, representation, and interaction. Note however that

WPs 1 and 2 have been operated closely together. Moreover, many MPs are cross-cutting the WP1/2 and WP3. When an MP contributes to multiple WPs, we have placed it in the first-listed WP in this report. For example, an MP contributing to WPs 1 and 3 is listed under WP1. Following an overview of the MP activity in that WP, descriptions of the individual MPs are included together with deliverables that have been completed for those that are either finished or have been running more than 50% of their anticipated time.

0.1 What is a micro-project?[1]

A micro project is a cooperation of two or more partners over a period of typically 1-6 months aimed at producing a tangible outcome (paper, data set, demo, tutorial etc.) to be made available to the community through the AI4EU platform and appropriately promoted in the community. Key hard requirements are:

1. two or more HumaneAI Net partners working together over a period of 1-6 months
2. cooperation to be documented eg. through joint authorship of the paper.
3. topic clearly tied to one or more tasks as described in the proposal (can be tasks from different WPs)
4. tangible outcome (paper, data set, toolset, demo, etc.)
5. outcome made available through the AI4EU platform or appropriate dissemination
6. a short presentation (5-15 mins) recorded at the end of the project to be made available through appropriate channels, including the project YouTube channel)

Originally there was also a very hard travel requirement so that, for the duration of the micro project, all researchers working on the micro-project are to work together at the same site (thus one organisation is the host, while the other participating researchers travel to the host site for the duration of the micro-project). However, due to the COVID situation the travel requirement were temporarily suspended.

0.2 Microprojects and financing

MPs were the vehicle to promote intensive collaboration between partners, leverage synergies between their groups and “spread the resulting knowledge” to the broader European AI community. Resources were allocated to fund this activity in the sense that most PMs in WPs 1-6 must be spent through micro-projects. We also have additional funds managed by the coordinator for which partners with good micro-project ideas can apply. This includes funds for supporting external (non-HumanE AI Net) as participants in micro-projects.

However, during the first year we have concentrated exclusively on internal MPs that make use of ‘pre-assigned funds’, ensuring that the approval method can be light weight and that activity can begin as quickly as possible. MP submissions were reviewed and approved by WP leaders of the WPs that the MP fell under (this could be typically a principal WP plus one or more additional secondary WPs). Approvals were expected and, in most cases, given within a week, to ensure that delays in initiating the work were minimized.

This deliverable contains reports on the MPs that were initiated and, in many cases, completed during the third year under his procedure.

[1] This and the next subsection are shared with Deliverable 1.1.

Our assessment is that the approach has been very successful in initiating the collaborative activity described in the work programme with many impressive outputs delivered. In the context of the COVID disruption, we believe that this has been a very positive outcome.

There are two outstanding issues that must be addressed during the remainder of the network:

- The need to build opportunities toward larger collaborations (from microprojects to macroprojects).
- The need to find synergies across workpackages.
- The allocation of additional resources beyond the 'pre-assigned' budget including resources to fund the involvement of additional partners needs to be made.

We plan to address these two issues by reviewing the progress of the overall research against the work programme and identify missing elements as well as particularly promising emerging directions. These will then be used to define calls for new MP proposals addressing these topics, but at the same time leaving open the option for proposals addressing different topics. The proposers will also be able to request additional funding either for internal or new external partners. An appropriate reviewing procedure has been developed to ensure that the proposals are objectively and adequately reviewed and decisions about funding can be reached that will ensure gaps in our coverage of the work programme are filled and opportunities for particularly promising work can be supported.

1. WPs1 and 2: Human-in-the-Loop Machine Learning, Reasoning and Planning

1.1. Overview of goals

This WP aims to develop the fundamental Learning, Reasoning and Planing methodologies that allow humans to be interactively involved "in the loop. This goes beyond explainability (which in itself is a challenge) toward methods that allow interactive human input to influence their inner workings. Note that in Y3 this WP has been closely cooperating with WP2.

1.2. Completed microprojects

2021-3-WP123: Memory Aware Conversational AI to aid virtual Team-Meetings

Start Date: 2021-03-02
Proposed Duration: 5 months
End Date: 2023-03-02

Submitted by: Catholijn Jonker, c.r.m.m.oertel@tudelft.nl
Partners:
1. Technical University Delft, Catharine Oertel, c.r.m.m.oertel@tudelft.nl (4 PM)
2. Eotvos Lorand University, Andras Lorincz, lorincz@inf.elte.hu (2 PM)

Contributes to: WP1, WP2, WP3, T3.2

Results:

The MEMO corpus was collected, which contains 45 group discussions around the topic of COVID-19. A total of 15 groups were formed, consisting of 3 to 6 participants who took part in 3 group discussions, with a 3-4 day gap between sessions. A total of 59 individuals with diverse backgrounds took part in the study. Before and after each session participants completed a series of questionnaires to determine which moments they recalled from their conversations, along with their personality traits, values and perceptions.

To capture conversational memory, we collected first-party free-recall reports of the most memorable moments from the discussion immediately after the interaction and again 3-4 days later. For the shorter-term memories, participants also mapped the moments to a particular interval in the video of their discussion, which were used for the ground-truth conversational memory annotations.

For each participant, personality and value profiles were recorded in the pre-screening survey, along with demographic information to identify their social group affected by COVID-19. Pre-session questionnaires also assessed participants' mood before each session. Post-session questionnaire included questions about mutual understanding, personal attitude and perceived social distance. The perception of the discussion and the group as a whole was also monitored in the post-session questionnaire with variables such as Task & Group Cohesion, Entitativity, Perceived Interdependence, Perceived Situation Characteristics, Syncness, and Rapport.

The following automatic annotations were extracted on the corpus:

Transcripts - Transcripts were generated with automatic speech recognition methods and were manually reviewed and corrected where needed. Transcript timestamps are available at the utterance level as well as word-level textgrid files for each recording. Speaker diarization is also available.

Eye gaze and head pose - automatically annotated with EyeWare software, the annotation itself will be provided, but the code uses proprietary API. This includes gaze targets collected through screenshots of participants' screenviews.
Prosody - eGeMAPS feature set was extracted from the default eGeMAPS configuration in OpenSmile

Body pose - Body pose (upper body only) and hand pose when visible were estimated with the models available in the MediaPipe software

Facial action units - Facial action units were estimated for participants using the OpenFace **Software**

A **Paper** describing the corpus and the annotations in more detail is in preparation. Additionally, the collected annotations are to be packaged in an appropriate manner for ease of use for future researchers.

The resulting corpus and corresponding manual and automatic annotations address the goals of work package 3 by providing a new and unique dataset for obtaining a better understanding of conversational memory and the multimodal signals associated with

identifying memorable moments which may be used to provide virtual agents the ability to identify topics and episodes which can enhance the interaction between human and machine.

Publications: :

1. Tsfasman, M., Fenech, K., Tarvirdians, M., Lorincz, A., Jonker, C., & Oertel, C. (2022). Towards creating a conversational memory for long-term meeting support: predicting memorable moments in multi-party conversations through eye-gaze. In ICMI 2022 - Proceedings of the 2022 International Conference on Multimodal Interaction (pp. 94-104). (ACM International Conference Proceeding Series). Association for Computing Machinery (ACM). https://doi.org/10.1145/3536221.3556613

Links to Tangible results:

The corpus is available by request - contact Maria Tsfasman (m.tsfasman@tudelft.nl) and Catharine Oertel (c.r.m.m.oertel@tudelft.nl). It is planned that the corpus will be released by the end of 2024.

For any other questions and inquiries feel free to contact Maria Tsfasman (m.tsfasman@tudelft.nl) and/or Catharine Oertel (c.r.m.m.oertel@tudelft.nl).
https://www.youtube.com/watch?v=kTS8XwUVNnI

2022-9-WP12: Graybox methods for augmenting human-driven narrative analyses

Start Date: 2022-09-01
End Date: 2023-01-31

Submitted by: Szymon Talaga, stalaga@uw.edu.pl

Partner:
Univ. Warsaw, Szymon Talaga, stalaga@uw.edu.pl (11 PM)
Institut Polytechnique de Grenoble, James Crowley, james.crowley@univ-grenoble-alpes.fr (0 PM)

Contributes to: WP1: T1.3, T1.3.2, WP2

Description:
This microproject has laid the groundwork for developing a new approach to narrative analysis providing a gray-box (at least partially explainable) NLP model tailored for facilitating work of qualitative text/narrative analysts. We conducted a proof-of-concept study combining existing standard NLP methods (e.g. topic modeling, entity recognition) with qualitative analysis of narratives about smart cities and related technologies and use this experience to conceptualize our approach to narrative analysis, in particular with respect to problems which are not easily solved with the existing tools. This current initial research micro-project will be followed by a microproject dedicated to formalizing our approach to narrative analysis and developing its open-source implementation (for Python)

Results
The aim of the project was to develop a software package (for Python) providing easy to use and understand (also for researchers not trained in computer science or linguistics) tools for extracting narrative information (active and passive actors, the actions they perform as well as descriptions of both actors and actions, which together define events) and organizing them in rich hierarchical data structures (data model is implicitly graphical) from which subsequently different sorts of descriptive statistics can be generated depending on particular research questions. Crucially, for this to be practically possible, a legible and efficient framework for querying the produced data is needed.

The above goal fits into a broader HumanE-AI objective of developing common ground concepts providing better representations shared by humans and machines alike. In particular, the contribution of the project to work on aligning machine analyses with human perspective through the notion of narratives is twofold. Firstly, narrative-oriented tools for automated text analyses can empower human analysts as, arguably, the narrative framework provides a more natural and meaningful context for people without formal training in linguistics and/or computer science for reasoning about textual data. Secondly, the development of the software for narrative analysis is naturally intertwined with conceptual work on the core terms and building blocks of narratives, which can inform subsequent work on more advanced approaches.

Importantly, the software is developed as a graybox model, in which core low-level NLP tasks, such as POS and dependency tagging, are performed by a blackbox statistical model, and then they are transformed to higher order grammar and narrative data based on a set of transparent deterministic rules. This is to ensure high explainability of the approach, which is crucial for systems in which the machine part is supposed to be a helper of a human analyst instead of an implicit leader.

Currently, the core modules of the package responsible for the grammatical analysis are mostly ready (but several improvements are still planned). This includes also a coreference resolution module. Moreover, the core part of the semantic module, which translates grammatical information to more semantic constructs focused on actors, actions and descriptions, is also ready. What is still missing are an interface exposing methods for end users allowing easy access and analysis of rich data produced by the package as well as a principled and convenient query framework on which the interface should be based. This is the main focus of the ongoing and future work. The second missing part is the documentation, but this part is best finished after the interface is ready.
Thus, even though the package in the current state can seem a little rough from the perspective of an end user, its quality and usefulness will increase steadily as new updates are delivered.

Links to Tangible results: https://github.com/sztal/segram

2022-9-WP12: Multi-Relational Contextual Reasoning for Complex Scene Generation for Autonomous Vehicle Data

Start Date 2022-09-01

Proposed Duration: 3 months
End Date 2023-03-31

Submitted by: Loris Bozzato, bozzato@fbk.eu
Partners:
1. Fondazione Bruno Kessler, Loris Bozzato, ghidini@fbk.eu (3 PM)
2. Technical University of Vienna, Thomas Eiter, eiter@kr.tuwien.ac.at (3 PM)
3. BOSCH Deutschland (External Partner), Stepanova Daria

Contributes to: WP1: T1.1, T1.5, WP2: T2.2

Results:
Our work considers the applicability of the MR-CKR framework to the task of generating challenging inputs for a machine learning model. Here, MR-CKR is a symbolic reasoning framework for Multi-Relational Contextual Knowledge Repositories that we previously developed. Contextual means that we can (defeasibly) derive different conclusions in different contexts given the same data. This means that conclusions in one context can be invalidated in a more specific context. Multi-relational means that a context can be "more specific" with respect to different independent aspects, such as regionality or time.

The general idea of generating challenging inputs for a machine learning model is the following: We have limited data on which we can train our model, thus, it is like that the model does not cover all eventualities or does not have enough data in specific contexts to lead to the correct result. Obtaining more data is often very difficult or even infeasible however. We introduce a new approach to solving this problem. Namely, given a set of diagnoses describing contexts in which the model performs poorly we generate new inputs that are (i) in the described contexts and (ii) as similar as possible to a given starting input. (i) allows us to train the network in a targeted manner by feeding it exactly those cases that it struggles with. (ii) ensures that the new input only differs from the old one in those aspects that make the new input problematic for the model. Thus, allowing us to teach the model to recognize aspects relevant for the answer.

This fits very well with the capabilities of MR-CKR: on the one hand, we have different contexts in which the inputs need to be modified to suit a different diagnosis of failure of the model. On the other hand, we can exploit the different relations by having one relation that specifies that inputs are more modifiable in one context than another and another relation that describes whether one diagnosis is a special case of another. Additionally, it allows us to incorporate global knowledge such that we can only modify inputs in such a manner that the result is still "realistic", i.e., satisfies the axioms in the global knowledge.

In this work, we provide a prototype specialized to generating similar and problematic scenes in the domain of Autonomous Driving. This work fits well into Task 1.1 of WP1: "Linking symbolic and subsymbolic learning", since we use a symbolic approach to enable the use of domain knowledge in order to advance the performance of a subsymbolic model. Furthermore, it also loosely fits into Task 1.5 of WP1: "Quantifying model uncertainty", since we can quantify how similar the generated new inputs are to the original ones.

Publications: : (No publications)
Links to Tangible results: Expected results: Prototype implementation, Report on formalization.

1.3. On-going microprojects

2022-3-WP12: Knowledge Extraction Through Prompting on Pre-trained Language Models

Start Date: 2023-03-20
Proposed Duration: 3 months
End Date: 2023-06-16

Submitted by: Mauro Dragoni, dragoni@fbk.eu
Partners:
1. Fondazione Bruno Kessler, Mauri Dragoni, dragoni@fbk.eu (2 PM)
2. External: University of Verona, Marco Rospocher, marco.rospocher@univr.it (2000€)

Contributes to: WP1, WP2,

Tagline: The ambition of the micro-project is to investigate the adoption of a multi-turn dialog strategy and the insertion in prompts of appropriate conceptual knowledge (e.g., definitions of the concepts to extract) or different types of examples (including negative examples) especially for the extraction of tasks and temporal flow relations between tasks

Project description: Procedural documents are a source of temporal procedural knowledge of uttermost importance. These documents are different in format and scope, as they range from the description of administrative procedures to service manuals to medical guidelines and surgical procedures. The extraction of this complex and multidimensional knowledge, which includes a strong temporal dimension usually paired with further static dimensions concerning, for example, resources, tools, objects, costs, and so on, would be of the utmost importance for several tasks ranging from information extraction to validation and verification of the procedures themselves, up to the construction of AI-based systems that have to deal with these procedures (think for instance to an expert surgical system and assistant which may be involved in several different surgery procedures).

Knowledge graphs are a natural and expressive knowledge structure where to represent such multidimensional knowledge, and indeed the insertion of temporal knowledge within knowledge graphs is one of the hot challenges in this area. Nonetheless, the automated construction of knowledge graphs from procedural documents is a challenging research area. Here, the lack of annotated data, as well as raw text repositories describing real-world procedural documents, makes it extremely difficult to adopt deep learning approaches.

Pre-trained language models showed promising results concerning the knowledge extraction tasks from the models themselves. Although several works explored this strategy to build knowledge graphs, the viability of knowledge base construction by using a prompt-based learning strategy from such language models has not yet been investigated

deeply. In this MP we would like to investigate the usage of prompt-based in-context learning strategy to extract, from natural language process descriptions, conceptual information that can be converted into their equivalent knowledge graphs. In particular, we would like to investigate the adoption of a multi-turn dialog strategy and the insertion of prompts of appropriate conceptual knowledge (e.g., definitions of the concepts to extract) or different types of examples (including negative examples), especially for the extraction of tasks and temporal flow relations between tasks. As such the work can contribute to the construction of structured narratives using machine learning models and hopefully enrich with conceptual knowledge in input. Moreover, the adoption of a multi-turn dialog strategy could provide insight into how these models can be used to complement the multi-turn dialog strategy usually adopted by domain experts in traditional knowledge modeling pipelines.

List of tangible Results:

- At least one paper published in a ranked conference or journal.
- An open dataset containing both raw text and their equivalent structured representation providing a ground for future research activities.

2022-4-WP12: Graybox methods for augmenting human-driven narrative analyses – beta release of the software package

Start Date: 2023-04-10
Proposed Duration: 5 months
End Date: 2023-08-11

Submitted by: Szymon Talga, stalaga@uw.edu
Partners:
1. University of Warsaw, Andrzej Nowak, andrzejn232@gmail.com (6 PM)

Contributes to: WP1, WP2,

Tagline: Developing user-friendly software for narrative analysis of text data

Project description: In this project we continue the development of the Segram package for Python. The purpose of the package is to provide tools for automated narrative analysis of text data focused on extracting information on basic building blocks of narratives – agents (both active and passive), actions, events, or relations between agents and actions (e.g. determining subjects and objects of actions), as well as descriptions of actors, actions and events. The development process is also naturally paired with conceptual work on representations of narratives.

The package is designed as a graybox model. It is based on an opaque statistical language model providing linguistic annotations, which are subsequently used by transparent deterministic algorithms for discovering narrative elements. Thus, the final output should be easy to interpret and validate by human users, whenever necessary. Moreover, by lifting the analysis from the purely linguistic level to the arguably more intuitive level of narratives, it is hoped that the provided tools will be significantly easier to

use and understand for end users, including those without training in linguistics and/or computer science.

The proposed framework is aimed at language understanding and information extraction, as opposed to language generation. Namely, the role of the package is to organize narrative information in convenient data structures allowing effective querying and deriving of various statistical descriptions. Crucially, thanks to its semi-transparent nature, the produced output should be easy to validate for human users. This should facilitate development of shared representations (corresponding the WP1 and WP2 motivated goal: „Establishing Common Ground for Collaboration with AI Systems") of narratives, understandable for both humans and machines, that are the same time trustworthy (by being easy to validate for humans), which is arguably a desirable feature, for instance in comparison to increasingly powerful but hard-to-trust large language models. In particular, the package should be useful for facilitating and informing human-driven analyses of text data.

Alpha version of the package implementing core functionalities related to grammatical and narrative analysis is ready. The goal of the present microproject is to improve the package and release a beta version. This will include implementing an easy-to-use interface (operating at the level of narrative concepts) for end users allowing effective querying and analysis of the data produced by Segram as well as developing a comprehensive documentation. Thus, the planned release should be ready for broader adoption to a wide array of use cases and users with different levels of linguistic/computational expertise.

List of tangible Results:

- Segram package for Python published officialy at Python Package Index (PyPI, https://pypi.org/). It may be also published at Conda-forge, but it is not yet guaranteed at this stage.
- Comprehensive package documentation available online at the Read the Docs platform (https://readthedocs.org/).

2022-2-WP12: Learning from Richer Feedback Through the Integration of Prior Beliefs

Start Date: 2023-02-28
Proposed Duration: 6 months
End Date: 2023-08-31

Submitted by: Silvia Tulli , tulli@isir.upmc.fr

Partners:

1. ISIR, Sorbonne University, Silvia Tulli , tulli@isir.upmc.fr (4 PM)
2. External: Colorado State University, Sarath Sreedharan, ssreedh3@asu.edu
3. ISIR, Sorbonne University, Mohamed Chetouani, mohamed.chetouani@sorbonne-universite.fr

Contributes to: WP1, WP2,

Tagline: Extending Inverse Reinforcement Learning to elicit and exploit richer expert feedback by leveraging the learner's beliefs.

Project description: Interactive Machine Learning (IML) has gained significant attention in recent years as a means for intelligent agents to learn from human feedback, demonstration, or instruction. However, many existing IML solutions primarily rely on sparse feedback, placing an unreasonable burden on the expert involved. This project aims to address this limitation by enabling the learner to leverage richer feedback from the expert, thereby accelerating the learning process. Additionally, we seek to incorporate a model of the expert to select more informative queries, further reducing the burden placed on the expert.

Objectives:
(1) Explore and develop methods for incorporating causal and contrastive feedback, as supported by evidence from psychology literature, into the learning process of IML.
(2) Design and implement a belief-based system that allows the learner to explicitly maintain beliefs about the possible expert objectives, influencing the selection of queries.
(3) Utilize the received feedback to generate a posterior that informs subsequent queries and enhances the learning process within the framework of Inverse Reinforcement Learning (IRL).

The project addresses several key aspects highlighted in the workpackage on Collaboration with AI Systems (W1-2). Firstly, it focuses on AI systems that can communicate and understand descriptions of situations, goals, intentions, or operational plans to establish shared understanding for collaboration. By explicitly maintaining beliefs about the expert's objectives and integrating causal and contrastive feedback, the system aims to establish a common ground and improve collaboration.
Furthermore, the project aligns with the objective of systems that can explain their internal models by providing additional information to justify statements and answer questions. By utilizing the received feedback to generate a posterior and enhance the learning process, the system aims to provide explanations, verify facts, and answer questions, contributing to a deeper understanding and shared representation between the AI system and the human expert.
The project also demonstrates the ambition of enabling two-way interaction between AI systems and humans, constructing shared representations, and allowing for the adaptation of representations in response to information exchange. By providing tangible results, such as user-study evaluations and methods to exploit prior knowledge about the expert, the project aims to make measurable progress toward collaborative AI.

List of tangible Results:
(1) Identification and development of potential informative feedback mechanisms that are more user-friendly, with a focus on determining the appropriate form of queries.
(2) User-study evaluation results that measure the correctness of the information provided by the human and assess the cognitive overhead involved.
(3) Methods to exploit prior knowledge about the expert to improve learning and reduce the burden placed on them, specifically in terms of how to query.
(4) Integration of richer feedback from the expert, including causal knowledge and contrastive information, into the learning process.

(5) Publication of a peer-reviewed paper in a competitive venue, presenting the research findings and contributions to the field.
(6) Creation of a GitHub repository containing all necessary materials to replicate the results and support further research endeavors.

2022-3-WP123: Intelligent crowdsourcing of geolocation tasks in natural disasters

Start Date: 2023-03-15
Proposed Duration: 9.5 months
End Date: 2024-01-31

Submitted by: Jesus Cerquides, cerquide@iiia.csic.es
Partners:
1. CSIC , Jesus Cerquides, cerquide@iiia.csic.es (4.4 PM)
2. External: University of Genova, Jose Luis Fernandez Marquez, joseluis.fernandez@unige.ch

Contributes to: WP1, WP2, WP3

Tagline: Build human-in-the-loop intelligent systems for the geolocation of social media images in natural disasters

Project description: Social media generate large amounts of almost real-time data which can turn out extremely valuable in an emergency situation, specially for providing information within the first 72 hours after a disaster event. Despite there is abundant state-of-the-art machine learning techniques to automatically classify social media images and some work for geolocating them, the operational problem in the event of a new disaster remains unsolved.

urrently the state-of-the-art approach for dealing with these first response mapping is first filtering and then submitting the images to be geolocated to a crowd of volunteers [1], assigning the images randomly to the volunteers.
The project is aimed at leveraging the power of crowdsourcing and artificial intelligence (AI) to assist emergency responders and disaster relief organizations in building a damage map from a zone recently hit by a disaster.

Specifically, the project will involve the development of a platform that can intelligently distribute geolocation tasks to a crowd of volunteers based on their skills. The platform will use machine learning to determine the skills of the volunteers based on previous geolocation experiences.

Thus, the project will concentrate on two different tasks:

- Profile Learning. Based on the previous geolocations of a set of volunteers, learn a profile of each of the volunteers which encodes its geolocation capabilities. This profiles should be unterstood as competency maps of the volunteer, representing the capability of the volunteer to provide an accurate geolocation for an image coming from a specific geographical area.

- Active Task Assigment. Use the volunteer profiles efficiently in order to maximize the geolocation quality while maintaining a fair distribution of geolocation tasks among volunteers.

On a first stage we envision an experimental framework with realistically generated artificial data, which acts as a feasibility study. This will be published as a paper in a major conference or journal. Simultaneously we plan to integrate both the profile learning and the active task assignment with the crowdnalysis library, a software outcome of our previous micro-project. Furthermore, we plan to organize a geolocation workshop to take place in Barcelona with participation from the JRC, University of Geneva, United Nations, and IIIA-CSIC.

In the near future, the system will generate reports and visualizations to help these organizations quickly understand the distribution of damages. The resulting platform could enable more efficient and effective responses to natural disasters, potentially saving lives and reducing the impact of these events on communities.

The microproject will be developed by IIIA-CSIC and the University of Geneva. The micro project is also of interest to the team lead by Valerio Lorini at the Joint Research Center of the European Commission @ Ispra, Italy, who will most likely attend the geolocation workshop which we will be putting forward.

The project is in line with "Establishing Common Ground for Collaboration with AI Systems (WP 1-2)", because it is a microproject that " that seeks to provide practical demonstrations, tools, or new theoretical models for AI systems that can collaborate with and empower individuals or groups of people to attain shared goals" as is specifically mentioned in the Call for Microprojects.

The project is also in line with "Measuring, modeling, predicting the individual and collective effects of different forms of AI influence in socio-technical systems at scale (WP4)" since it ecomprises the design of a human-centered AI architectures that balance individual and collective goals for the task of geolocation.
[1] Fathi, Ramian, Dennis Thom, Steffen Koch, Thomas Ertl, and Frank Fiedrich. “VOST: A Case Study in Voluntary Digital Participation for Collaborative Emergency Management.” Information Processing & Management 57, no. 4 (July 1, 2020): 102174. https://doi.org/10.1016/j.ipm.2019.102174.

List of tangible Results:
1. Open source implementation of the volunteer profiling and consensus geolocation algorithms into the crowdanlysis library
2. - Paper with the evaluation of the different geolocation consensus and active strategies for geolocation
3. - Organization of a one day workshop with United Nations, JRC, University of Geneva, CSIC

2022-3-WP12: A tool for mitigating algorithmic biases through explanations

Start Date: 2023-09-04
Proposed Duration: 5 months
End Date: 2024-01-22

Submitted by: Dino Pedreschi, dino.pedrschi@unipi.it
Partners:
1. External: University of Antwerp, Daphne Lenders, daphne.lenders@uantwerpen.be,
2. University of Pisa - Departement of CS, Dino Pedreschi,
3. External: Scuola Normale Superiore, Roberto Pellungrini roberto.pellungrini@sns.it

Contributes to: WP1, WP2,

Tagline: Presenting users with explanations about AI systems, to let them detect and mitigate discriminatory patterns

Project description: Our project revolves around the topic of fair Artificial Intelligence (AI), a field that explores how decision-making algorithms used in high-stake domains, such as hiring or loan allocations can perpetuate discriminatory patterns in the data they are based on, unfairly affecting people of certain races, genders or other demographics. Early attempts to address bias in AI systems focused on automated solutions, attempting to eliminate discrimination by establishing mathematical definitions of "fairness" and optimizing algorithms accordingly. However, these approaches have faced justified criticism for disregarding the contextual nuances in which algorithms operate and for neglecting the input of domain experts who understand and can tackle discriminatory patterns effectively. Consequently, policymakers have recognized the pitfalls of solely relying on these approaches and are now designing legal regulations, mandating that high-risk AI systems can only be deployed when they allow for oversight and intervention by human experts.

With our project, we investigate how to effectively achieve this human control, by exploring the intersection between fair and explainable AI (xAI), whereas the latter is concerned with explaining the decision processes of otherwise opaque black-box algorithms. We will develop a tool, that provides humans with explanations about an algorithmic decision-making system. Based on the explanations users can give feedback about the system's fairness and choose between different strategies to mitigate its discriminatory patterns. By immediately getting feedback about the effects of their chosen strategy, users can engage in an iterative process further refining and improving the algorithm. Since little prior work has been done on Human-AI collaboration in the context of bias mitigation, we will take on an exploratory approach to evaluate this system. We will set up a think-aloud study where potential end-users can interact with the system and try out different mitigation strategies. We will analyse their responses and thoughts, to identify the tool's strengths and weaknesses as well as users' mental model of the tool. Additionally, we will compare the systems' biases before and after human intervention, to see how biases were mitigated and how successful this mitigation was.

Our work aligns with the goal of the topic "Establishing Common Ground for Collaboration with AI Systems" (motivated by Workpackage 1 and 2). This topic is focused on developing AI systems that work in harmony with human users, empowering them to bring

their expertise and domain knowledge to the table. In particular, our work recognizes humans' ability to make ethical judgements and aims to leverage this capability to make fairer AI systems. By conducting a user study we align with the topics' goal to make this human-AI collaboration desirable from the users' site, ensuring that they understand the inner workings of the AI system and they have full control in adapting it.

List of tangible Results:
- A tool that presents users with explanations about a decision-making system and that can interactively adjust its decision process, based on human feedback about the fairness of its explanations
- A user-centric evaluation of the tool, investigating whether users can effectively detect biases through the tool and how they use the different bias mitigation strategies offered by it
- We aim to present a demo of the tool at a workshop or a conference
- Additionally, we commit to publishing one paper, describing the tool itself and the results of the usability study

2022-3-WP1238: Educational module Human-Interactive Robot Learning (HIRL)

Start Date: 2023-09-01
Proposed Duration: 6 months
End Date: 2024-02-28

Submitted by: Mohamed Chetouani, mohamed.chetouani@sorbonne-universite.fr
Partners:
1. ISIR, Sorbonne University, Mohamed Chetouani, mohamed.chetouani@sorbonne-universite.fr (1 PM)
2. ISIR, Sorbonne University , Silvia Tulli, tulli@isir.upmc.fr (6 PM)
3. Vrije Universiteit Amsterdam, Kim Baraka, k.baraka@vu.nl (6PM)

Contributes to: WP1, WP2, WP3, WP8

Tagline: A graduate level educational module (12 lectures + 5 assignments) covering basic principles and techniques of Human-Interactive Robot Learning.

Project description: Human-Interactive Robot Learning (HIRL) is an area of robotics that focuses on developing robots that can learn from and interact with humans. This educational module aims to cover the basic principles and techniques of Human-Interactive Robot Learning. This interdisciplinary module will encourage graduate students (Master/PhD level) to connect different bodies of knowledge within the broad field of Artificial Intelligence, with insights from Robotics, Machine Learning, Human Modelling, and Design and Ethics. The module is meant for Master's and PhD students in STEM, such as Computer Science, Artificial Intelligence, and Cognitive Science.
This work will extend the tutorial presented in the context of the International Conference on Algorithms, Computing, and Artificial Intelligence (ACAI 2021) and will be shared with the Artificial Intelligence Doctoral Academy (AIDA). Moreover, the proposed lectures and

assignments will be used as teaching material at Sorbonne University, and Vrije Universiteit Amsterdam.

We plan to design a collection of approximately 12 1.5-hour lectures, 5 assignments, and a list of recommended readings, organized along relevant topics surrounding HIRL. Each lecture will include an algorithmic part and a practical example of how to integrate such an algorithm into an interactive system. The assignments will encompass the replication of existing algorithms with the possibility for the student to develop their own alternative solutions.

Proposed module contents (each lecture approx. 1.5 hour):
(1) Interactive Machine Learning vs Machine Learning – 1 lecture
(2) Interactive Machine Learning vs Interactive Robot Learning (Embodied vs non-embodied agents) – 1 lecture
(3) Fundamentals of Reinforcement Learning – 2 lectures
(4) Learning strategies: observation, demonstration, instruction, or feedback
- Imitation Learning, Learning from Demonstration – 2 lectures
- Learning from Human Feedback: evaluative, descriptive, imperative, contrastive examples – 3 lectures
(5) Evaluation metrics and benchmarks – 1 lecture
(6) Application scenarios: hands-on session – 1 lecture
(7) Design and ethical considerations – 1 lecture

List of tangible Results:
Learning objectives along Dublin descriptors:
(1) Knowledge and understanding;
- Be aware of the human interventions in standard machine learning and interactive machine learning.
- Understand human teaching strategies
- Gain knowledge about learning from feedback, demonstrations, and instructions.
- Explore ongoing works on how human teaching biases could be modeled.
- Discover applications of interactive robot learning.
(2) Applying knowledge and understanding;
- Implement HIRL techniques that integrate different types of human input
(3) Making judgments;
- Make informed design choices when building HIRL systems
(4) Communication skills;
- Effectively communicate about own work both verbally and in a written manner
(5) Learning skills;
- Integrate insights from theoretical material presented in the lecture and research papers showcasing state-of-the-art HIRL techniques.

2022-3-WP12: LWMs teaching to teach

Start Date: 2023-08-01
Proposed Duration: 8 months
End Date: 2024-03-31

Submitted by: Agnes Grünerbl, agnes.gruenerbl@dfki.de
Partners:
1. DFKI – EI, Agnes Grünerbl, agnes.gruenerbl@dfki.de
2. External: Health Department – University of Southampton, Eloise Monger, E.Monger@soton.ac.uk (5000€)

Contributes to: WP1, WP2,

Tagline: Accelerating nurse training without impacting the quality of education, by leveraging LWM (large whatever models) to provide individual feedback to students and help teachers how to optimize teaching.

Project description: High-quality education and training of nurses are of utmost importance to keep high standards in medical care. Nevertheless, as the covid pandemic has shown quite impressively, there are too few healthcare professionals available. Therefore, education and training of nurse students, or adapting the training of nurses is challenged to accelerate, to have manpower of nurses available when it is required. Still, accelerating training often comes with reduced quality, which can easily lead to bad qualifications and, in the worst case, to a lethal outcome. Thus, in nurse training a pressing question is, how to optimize and with it accelerate training without suffering in quality.

One of the significant questions for teachers in training nurse students is to understand the state of a student's education. Are some students in need of more repetitions? which students can proceed to the next level, who is ready to get in contact with actual patients? In this regard, optimization of training means to individualize, not only individualize the training of students but also individualize the feedback and information a teacher gets about their way of teaching.

We believe this to be a field where Artificial Intelligence (AI) and more specifically the application of foundational models (LLMs large language models, paired with other methods of machine learning) can provide real support.

In the first part of this microproject, together with Nurse-Teachers of the University of Southampton, we want to define and design an LWM that fits the requirements of nurse training. For this, 2-3 nurse teachers from Southampton will visit DFKI in order to get a feeling for systems that are available, and also what applications are feasible. In turn, researchers of DFKI will visit the nurse training facilities in Southampton to get a better picture of how nurse training is conducted. At the end of this first phase of the microproject, an LWM (large whatever model) is defined (existing LLMs combined with additional features and data sources, as required).

In the second phase, this LWM will be implemented and tested against videos of recorded training sessions. Specific focus will be set on:
• How to understand the action of a particular person?
• Actions taken by the trainee, are they correct or false? What would have been the correct action?
• Which teaching efforts work and which do not as much?

• Which useful suggestions and feedback can be provided to the trainees and teachers?

Depending on the outcome of this microproject, in a follow-up project, an online LWM system could be installed at the facilities of the University of Southampton, where the effects of direct feedback on teaching and performance, could be evaluated.

List of tangible Results:
1) Definition and design of the LWM will be documented and if possible published in an adequate scientific journal
2) Developed algorithms and results will be published at a scientific conference (AI and possibly also medical)
3) The developed LWM will be made available to be used in a follow-up project

2022-3-WP123: Proactive announcement based on DEL (Dynamic Epistemic Logic)

Start Date: 2023-09-01
Proposed Duration: 4.5 months
End Date: 2024-01-12

Submitted by: Jasmin Grosinger, jasmin.grosinger@oru.se
Partners:
1. Örebro University- ORU, Jasmin Grosinger, jasmin.grosinger@oru.se (4.5 PM)
2. External: Denmark Technical Unisersity, Thomas Bolander

Contributes to: WP1, WP2, WP3

Tagline: Using Dynamic Epistemic Logic (DEL) so an AI system proactively can make announcements to avoid undesirable future states based on the human's false belief

Project description: Previously we have investigated how an AI system can be proactive, that is, acting anticipatory and on own initiative, by reasoning on current and future states, mental simulation of actions and their effects, and what is desirable. In this micro-project we want to extend our earlier work doing epistemic reasoning. That is, we want to do reasoning on knowledge and belief of the human and by that inform the AI system what kind of proactive announcement to make to the human. As in our previous work, we will consider which states are desirable and which are not, and we too will take into account how the state will evolve into the future, if the AI system does not act. Now we also want to consider the human's false beliefs. It is not necessary and, in fact, not desirable to make announcements to correct each and any false belief that the human may have. For example, if the human is watching the TV, she need not be informed that the salt is in the red container and the sugar is in the blue container, while the human's belief is that it is the other way around. On the other hand, when the human starts cooking and is about to use the content of the blue container believing it is salt, then it is a relevant announcement of the AI system to inform the human what is actually the case to avoid undesirable outcomes. The example shows, that we need to research on not only what to announce but also when to make the announcement.

The methods we will use in this micro-project are knowledge-based, to be precise, we will employ Dynamic Epistemic Logic (DEL). DEL is a modal logic. It is an extension of Epistemic Logic which allows to model change in knowledge and belief of an agent herself and of other agents.

1 week of visit is planned. In total, 7,5 PMs are planned to work on the MP, that is, 1 week we work physically in the same place, the rest of the PMs we work together online.

List of tangible Results:
- Formal model
We expect to develop a formal model based on DEL and based on the findings of J.Grosinger's previous work on proactivity. The model enables an artificial agent to make announcements to the human to correct the human's false belief and false belief about desirability of future states in a proactive way. Being formal we can make general definitions and propositions in the model and provide proofs about its properties, for example, about which proactive announcements are relevant and/or well-timed.
- Conference
We aim for a publication of our work at an international peer-reviewed high-quality conference. Candidate conferences are AAMAS (International Conference on Autonomous Agents and Multiagent Systems), or if this is temporally infeasible, then IJCAI (International Joint Conferences on Artificial Intelligence).

- Further collaboration
The MP can lead to further fruitful collaborations between the applicants (and possibly, some of their colleagues additionally) as the MP's topic is new and under-explored and all cannot be investigated within one MP.

2. WP2: Multimodal Perception and Modeling

2.1. Overview of goals

The ambition of WP2 is to build on recent progress in discriminative and generative networks, to provide integrated multi-modal perception and modeling that combines fast real-time reaction for sensori-motor reflexes, with spatiotemporal and geometric reasoning, prediction of recurrent events and consequences for actions and dynamic processes and linguistic expressions for perceptual concepts to enable communication with and learning from humans. In prticuar we intend to develop systems that can understand complex human actions, motivations and social settings. Note that in Y3, WP2 has been cooperating closely with WP1. The results of WP1 and WP2 should be assessed together.

2.2. Completed microprojects

2021-5-WP26: Neural Mechanism in Human Brain Activity During Weight Lifting

Start Date: 1 May 2022
Proposed Duration: 4 months
End Date: 31 Jan 2023

Submitted by: Sencer Melih Deniz, sencer.deniz@tubitak.gov.tr
Partners:
1. TUBITAK BILGEM, Sencer Melih Deniz, sencer.deniz@tubitak.gov.tr (4 PM)
2. DFKI Kaiserslautern, Hamraz Javaheri, Hamraz.Javaheri@dfki.de (4 PM)

Contributes to: WP2, WP6, T2.2, T2.3, T6.5, T6.6

Description:
In this project, it was investigated whether EEG (electroencephalography) signal can be used for detecting the motion as well as the variable weights a person is lifting. To do this, an experimental paradigm has been designed and EEG data have been acquired during performing biceps flexion-extension motions for different weight categories: lifting with no weight (empty), medium, and heavy lifting. The outcomes of the project can be applied in industrial exoskeleton applications as well as physical rehabilitation of stroke patients.

Results:
Features in EEG data generating difference for each lifted weight of category have been investigated. EEG data via different two EEG headsets have been collected from various participants while they lift different categories of load, namely empty, medium and heavy, in this project. Then, EEG data have been analyzed to realize if different category of weigths result in difference in EEG data by applying different deep learning methods together with different machine learning methods. According to the obtained results, it can be said that that EEG signals can be successfully used as a method to predict different loads during dynamic bicep curl motion. Therefore, this result could result more researches to develop rehabilitation systems robust to dynamic changes in weight. Moreover, information regarding weight change could contribute to a better estimation of fatigue condition to be used in sports and training applications. Finally, it has been evaluated that the approach to predict different categories of lifted weight could be used in further optimizations in industrial applications for which usage of exoskeleton can be given as an example.

Results of micro project in which TUBITAK BILGEM and Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI) Kaiserslautern collaborated was presented at the IEEE-EMBS International Conference on Biomedical and Health Informatics jointly organised with the IEEE-EMBS International conference on Wearable and Implantable Body Sensor Networks organized in Ioannina, Greece between 27-30 September 2022. Also it was published with the title “Prediction of Lifted Weight Category Using EEG Equipped Headgear" in 2022 IEEE-EMBS International Conference on Biomedical and Health Informatics Conference Proceedings.

Publications:
"Prediction of Lifted Weight Category Using EEG Equipped Headgear", published in 2022 IEEE-EMBS International Conference on Biomedical and Health Informatics.

Links to Tangible results:
Paper: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9926744

Dataset: https://www.ai4europe.eu/research/research-bundles/neural-mechanism-human-brain-activity-during-weight-lifting?category=ai_assets

2022-3-WP2: Natural Imitation of Dance Moves and Human Gestures with a Humanoid Robot

Start Date: 1 March 2022
Proposed Duration: 9 months
End Date: 30 Nov 2022

Submitted by: Shivesh Kumar, shivesh.kumar@dfki.de

Partners:
1. DFKI Bremen , Melya Boukheddimi, Shivesh Kumar, shivesh.kumar@dfki.de (5 PM)
2. INRIA Paris, Justin Carpentier, justin.carpentier@inria.fr (4 PM)

Contributes to: WP2

Results:
In order to address the topic, two main contributions were introduced in this project: The first one is the proposition of a generic formalization of robot dancing which allows us to use musical features for choreography generation. Optimal dance trajectories were computed using direct optimal control. From this formalization we derive three different methods of dance generation, that differ in the level of flexibility, human involvement, and automatization. The methods are: imitated, improvised, and automatic choreography generation. The imitated and improvised choreographic methods, are based on beat timing extraction. The automatic choreography generation method,uses the additional music features volume and vocal melody.The results are validated on 4 different music pieces in simulation using the dynamic simulator MuJoCo as well as in experiments on the real robot RH5 Manus. This work was published in the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 2022, and was selected as a finalist for best entertainment and amusement paper.

The second one focuses on the ability of exploiting the full capabilities of a robot through motion generation, with the aim of achieving motions that are more human-like and that can lead to a certain trust and comfort feeling of the human towards the robot acting in its environment. To this purpose, we proposed a first study on resolving all the loop-closure constraints of the series-parallel hybrid robot RH5 Manus within the trajectory optimization process. To this end, we use the open-source software Pinocchio developed by the INRIA Paris team with its recently introduced proximal formulation of the constrained dynamics. This approach allows us to converge to an optimal solution according to the least squares principle, even in the context of singularities. Among the optimization methods available in the literature, the differential dynamics programming (DDP) approach was used to generate optimal trajectories with respect to the constrained dynamics.

Results are presented in simulation as well as experiments on the real robot. This work is significant for humanoid robots based on electric actuation where one must seek to push the robot to its limits to achieve human like agility.

This work has been submitted to the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 2023 and is under review.

Publications :

1. Melya Boukheddimi, Daniel Harnack, Shivesh Kumar, Rohit Kumar, Shubham Vyas, Octavio Arriaga, Frank Kirchner, Robot Dance Generation with Music Based Trajectory Optimization, In IEEE/RSJ *International Conference on Intelligent Robots and Systems (IROS) 2022*, (IROS-2022), IEEE, Nov/2022.
2. Melya Boukheddimi , Rohit Kumar , Shivesh Kumar , Justin Carpentier , and Frank Kirchner, Investigations into Exploiting the Full Capabilities of a Series-Parallel Hybrid Humanoid using Whole Body Trajectory Optimization, In IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 2023, (IROS-2022), IEEE, Nov/2023. (submitted)

Links to Tangible results:

1. Melya Boukheddimi, Daniel Harnack, Shivesh Kumar, Rohit Kumar, Shubham Vyas, Octavio Arriaga, Frank Kirchner, Robot Dance Generation with Music Based Trajectory Optimization, In IEEE/RSJ *International Conference on Intelligent Robots and Systems (IROS) 2022*, (IROS-2022), IEEE, Nov/2022.
2. Melya Boukheddimi , Rohit Kumar , Shivesh Kumar , Justin Carpentier , and Frank Kirchner, Investigations into Exploiting the Full Capabilities of a Series-Parallel Hybrid Humanoid using Whole Body Trajectory Optimization, In IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 2023, (IROS-2022), IEEE, Nov/2023. (submitted)

Videos :

https://www.youtube.com/watch?v=aN_v39p17tg
https://www.youtube.com/watch?v=MA42YUg3e8E

2022-4-WP246: Polarization with the Friedkin-Johnsen model over a dynamic social network

Start Date: 1 April 2022
Proposed Duration: 4
End Date: 31 Dec 2022 (delay due maternity leave).

Submitted by: Elisabetta Biondi, elisabetta.biondi@iit.cnr.it
Partners:

1. Consiglio Nazionale delle Ricerche (CNR),Elisabetta Biondi, elisabetta.biondi@iit.cnr.it (2 PM)
2. Central European University (CEU), Janos Kertesz, kerteszj@ceu.edu, Gerardo Iniguez, IniguezG@ceu.edu (4 PM)

Contributes to: WP2-T2.5, WP4, WP6

Description:
The Friedkin-Johnsen model is a very popular model in opinion dynamics, validated on real groups, and well-investigated from the opinion polarization standpoint. Previous research has focused almost exclusively on static networks, where links between nodes do not evolve over time. In this micro-project, we want to fill this gap by designing a variant of the Friedkin-Johnsen model that embeds the dynamicity of social networks. Furthermore, we will design a novel definition of global polarization that combines network features and opinion distribution, to capture the existence of clustered opinions. We have analyzed the polarization effect of the new dynamic model, and identify the impact of the network structure.

Results:
Human social networks are very complex systems and their structure has an essential impact on opinion dynamics. However, since my main goal is to study the impact of the opinion dynamics model per se, we decided to deal with two different social network typologies: a Erdős–Rényi (ER) and a stochastic block model (SBM).

Design of the Friedkin-Johnsen (FJ) dynamic model.
We have implemented a rewiring policy that has been extensively studied in discrete opinion diffusion models. This involves substituting edges that connect nodes with different opinions with other edges. We have adapted this scheme to work with the FJ model's opinions, which are within the range of [-1,1], in both the asynchronous and synchronous versions. According to two parameters θ (the disagreement threshold) and p_rew (the rewiring probability):

- With probability 1-p_rew the FJ is applied
- With probability p_rew, if i and j disagree, i.e. $|x_i-x_j| > \theta$, the edge (i,j) is replaced with an edge (i,k) where k agrees with i, i.e. $|x_i-x_j| <= \theta$.

The above algorithm was specifically designed and implemented for the ER graph. However, in the case of the SBM, I have limited the potential candidates for rewiring to nodes within a maximum of two hops distance. This decision was made to prevent the block structure from becoming entirely irrelevant. The rationale behind this choice is based on the triadic closure mechanism, which suggests that individuals are more inclined to choose new acquaintances among the friends of their friends.

Design of the polarization metric.
The design of the polarization metric involved developing a definition for identifying highly polarized networks. We defined a highly polarized network as one in which there are two distinct opinions that are clustered into two tightly connected communities. To achieve this, we needed to consider both the network structure and the distribution of opinions. Therefore, we decided to use two different metrics to measure these aspects: bimodality for the opinion distribution and homogeneity for its correspondence with the network structure.

Bimodality.

The bimodality coefficient was used to measure the extent to which a distribution is bimodal. It is calculated using the skewness and kurtosis values and represents how similar the distribution is to one with two modes.

Homogeneity
To measure the homogeneity of the opinion distribution with the network structure, we examined the local distribution of nodes' opinions. We looked at whether each node's opinion was similar to those of its neighbors, which would suggest that it was in line with the overall opinion distribution over the network. The final homogeneity value was close to zero if the distribution of opinions was close to linear.

Experimental evaluation.
We have developed a Python simulator that can compute the dynamic FJ (rewiring included), and polarization metrics over time based on the given network and initial opinions. To test the model, we ran simulations on a small network comprising 20 nodes and compared the outcomes of the FJ with rewiring to those without rewiring. For the ER network, we used a vector of uniformly distributed opinions over [-1,1] as the initial opinions. However, for the SBM networks, we employed a different configuration, where the initial opinions were uniformly extracted over the intervals [-0.5,0-0.1] and [0.1,0.5], depending on whether the nodes belonged to one or the other block.

In conclusion, this microproject involves the design of a dynamic version of the FJ model for synchronous and asynchronous cases. Additionally, we have developed a new definition of polarization that considers both the distribution of opinions and the network topology. To assess the model's effectiveness, we conducted simulations on two different network types: an ER network and an SBM network. Our findings indicate that the rewiring process has significant effects on polarization, but these effects are dependent on the initial network.

Publications:
No publications yet. The collaboration is still ongoing.

Links to Tangible results:
Github link of the code of the simulator for the new dynamic model:
https://github.com/elisabettabiondi/FJ_rewiring_basic.git

2022-9-WP235:Multilingual and Multimodal conversational agent combined with search engine models.

Start Date 2022-09-01
Proposed Duration: 6
End Date 2023-04-28 (on-going)

Submitted by: Eric Blaudez, eric.blaudez@thalesgroup.com

Partners:

1. Thales Eric Blaudez, eric.blaudez@thalesgroup.com
2. Unibo Paolo Torrini, p.torroni@unibo.it
3. LISN Christophe Servan c.servan@qwant.com

Contributes to: WP2, WP3, WP5

Description:
The micro-project provides a demonstration of the hierarchical Framework for collaboration decribed in the Humane-AI Net revised strategic workplan, by constructing amultimodal and multilingual conversational agents focused on search.
The framework is based on hierarchical levels of abilities:

- Reactive(sensori-motor) Interaction: Interaction is tightly-coupled perception-action where actions of one agent are immediately sensed and interpreted as actions of the other. Examples include greetings, polite conversation and emotional mirroring
- Situated (Spatio-temporal) Interaction Interactions are mediated by a shared model of objects and relations (states) and shared models for roles and interaction protocols.

On this microproject we focused on the 2 first levels (Reactive and Situational) and designed the global framework architecture. The results are to be demonstrated in a Proof of Concept (PoC).

Results:
Components development:
The development of the components is on going. The mini project deals with three modules: the Human Machine Interface (HMI) to propose an interaction interface, the reactive module to recognize emotions and the situational module to manage the dialog situation. The framework is developed with the strategy proposed by [1]; in this project, we focus on reactive and situational part.

Human Machine interface [Status: DONE]
The HMI is a simple Flask application connected to Framework REST Service.

Reactive module [Status: ON GOING, framework integration almost finished]
The objective of the reactive module was to recognize emotions in conversation, so as to enable the other modules to drive the conversation accordingly. We thus addressed an Emotion Recognition in Conversation (ERC) task. The first step was a literature review and a survey of the available datasets and architectures. We chose the MELD [3] dataset as it is considered a benchmark for ERC and one of the few datasets available for building our ERC module. We considered different options for capturing the context, including the current utterance and its corresponding dialogue history were analyzed. As an architecture, we chose EmoBERTa [1], due to its strong performance and availability of implementation. We trained EmoBERTa using MELD splits for training, validation and test dataset. We carried out several experiments to establish baselines and examine the impact of different context representations. Moreover, since MELD conversations differ from those that may arise in the application domain envisaged for our chatbot, since we do not have any validation datasets available, we explored the generalization capability of the model by transfer learning and few-shot learning. In particular, we considered SetFit [4] as a few-shot learning technique and DailyDialog [2] for transfer learning validation. The experiments conducted on EmoBERTa have yielded promising results, indicating that this

architecture is a suitable starting point for constructing the reactive module. However, the task of generalization and adaptability of the model is challenging, and ongoing experiments are being carried out to address these issues. References

Situational Module [Status: ON GOING]
This module is responsible of the “Interactions mediated by shared models of objects (entities) and relations and shared models of roles and interaction protocols”. Knowledge graph are used to represent the situation, the relations and keep the context of the dialog.

- The document Analysis module has been developed in AI4EU project and is adapted and used as submodule for the miniproject. The module provides a data processing and a knowledge extraction functions by using Natural Language Processing. It builds a structured representation of the data (knowledge graph, named entities, keywords, summaries …) and provide semantic search. This module is in constant evolution; it will provide at the end of year a capability to manage more than 100 languages.
- The LISN part aims to create a spoken language understanding (SLU) part and the dialogue manager part. LISN proposed a first proof of concept based Speech recognition on RASA [1] in order to propose a full dialogue framework. For this part, at LISN, we focus on the SLU part of the module. The SLU module is based on the DIET-Classifier model (DIET stands for Dual Intent and Entity Transformer) [2] for both performing intent detection and concept detection. For the next part, we will move to another framework since we observed some limitation in the use of the RASA framework. We will also use another model to perform both intent and concept detection, by using the JointBERT model [3]. Preliminary results are encouraging.
- Thales & LISN works on Situational context representation based on (Temporal)-Knowledge graph and is based on Endsley model [4].

Bibliography:
[1] Crowley, James & Coutaz, Joëlle & Grosinger, Jasmin & Vázquez-Salceda, Javier & Angulo, Cecilio & Sanfeliu, Alberto & Iocchi, Luca & Cohn, Anthony. (2022). A Hierarchical Framework for Collaborative Artificial Intelligence. 10.48550/arXiv.2212.08659.
Reactive Module (UniBO)
[1] Kim, T., Vossen, P.: Emoberta: Speaker-aware emotion recognition in conversation with roberta. CoRR abs/2108.12009 (2021)
[2] Li, Y., Su, H., Shen, X., Li, W., Cao, Z., Niu, S.: Dailydialog: A manually labelled multi-turn dialogue dataset. In: IJCNLP(1). pp. 986–995. Asian Federation of Natural Language Processing (2017)
[3] Poria, S., Hazarika, D., Majumder, N., Naik, G., Cambria, E., Mihalcea, R.: MELD: A multimodal multi-party dataset for emotion recognition in conversations. In: ACL (1). pp. 527–536. Association for Computational Linguistics (2019)
[4] Tunstall, L., Reimers, N., Jo, U.E.S., Bates, L., Korat, D., Wasserblat, M., Pereg, O.: Efficient few-shot learning without prompts. CoRR abs/2209.11055 (2022)
Situational Module (Thales & LISN)
[1] Bocklisch, T., Faulkner, J., Pawlowski, N., and Nichol, A. (2017). Rasa: Open source language understanding and dialogue man-agement. In the proceedings of the first NIPS workshop on Conversational AI.
[2] Bunk, T., Varshneya, D., Vlasov, V., and Nichol, A. (2020). Diet: Lightweight language understanding for dialogue systems. arXiv preprint arXiv:2004.09936.

[3] Chen, Q., Zhuo, Z., and Wang, W. (2019). Bert for joint intent classification and slot filling. arXiv preprint arXiv:1902.10909.
[4] M. R. Endsley and D. J. Garland (Eds.). Situation Awareness Analysis and Measurement. Mahwah, NJ: Lawrence Erlbaum, 2000.

Links to Tangible results:
Jira & Confluence: https://humane-dialog.atlassian.net/jira/software/projects/HAD/pages
Github [PRIVATE]: project: https://gitlab.com/humane-ai-chatbot/chatbot-fmk
submodules:
T-KEIR: https://github.com/ThalesGroup/t-keir.git
erc-unibo-module: https://github.com/helemanc/erc-unibo-module

2022-9-WP235: Use of dialog context to boost ASR/NLG/TTS and improve the overall quality of voice dialog systems

Start Date: 2022-11-01
Proposed Duration: 4
End Date: 2023-02-28

Submitted by: Petr Schwarz, schwarzp@fit.vutbr.cz
Partners
1. Brno University of Technology, Petr Schwarz, schwarzp@fit.vutbr.cz (4 PM)
2. Charles University, Ondrej Dusek, odusek@ufal.mff.cuni.cz (3 PM)

Contributes to: WP2 T2.7 (T2.2), WP3 T3.6 (T3.4),

Results:
This project brings us data, tools, and baselines that enable us to study and improve context exchange among component and dialog sides (AI agent and human) in voice dialog systems. A better context exchange allows us to build more accurate automatic speech transcription, better dialog flow modeling, more fluent speech synthesis, and more powerful AI agents. The context exchange can be seen as an interactive grounding in two senses - among dialog sides (for example, technologies like example automatic speech transcription rarely use the other dialog side information to adapt itself) and among dialog system components (the speech synthesis rarely uses dialog context to produce more fluent or expressive speech).

The individual project outputs are summarized below:
1) Audio data collection software based on the Twilio platform and WebRTC desktop/mobile device clients. The purpose is to collect audio data of communication between agents (company, service provider, for example, travel info provider) and users. This software enables us to collect very realistic voice dialogs that have high-quality audio (>= 16kHz sampling frequency) on the agent side and low telephone-quality audio on the user side. The code is available here: https://github.com/oplatek/speechwoz
2) We have established a relationship with Paweł Budzianowski (Poly.AI) and Izhak Shafran (Google). Paweł created the MultiWoz database – an excellent dialog corpus

(https://arxiv.org/abs/1810.00278) that we use for the text-based experiment. We decided to collect our audio data similarly. Izhak organized DSTC11 Speech Aware Dialog System Technology Challenge (https://arxiv.org/abs/2212.08704) and created artificial audio data for MultiWOZ through speech synthesis, reading, and paraphrasing. Both provided us with the necessary advice for our data collection.

3) Speech dialog data – the data collection platform preparation and data collection are very time-consuming. The data collection is in progress and will be released before June 26th, 2023.
4) Initial experiments with context exchange between dialog sides (user and agent) were performed. These experiments show a nice improvement in the component of automatic speech recognition side. The results will be re-run with the collected data and published when the collection is finished.
5) Initial experiments with training instance weighting for response generation – which brings context to dialog system response generation, were performed. Experiments were based on the AuGPT system, previously developed at CUNI. The code is available here: https://github.com/knalin55/augpt. Instance weighting increases the re-use of context, compared to normal training, and can go even beyond natural occurrences in data. Simple weighting (threshold) seems better than designing a complex instance weight (in terms of automated metrics, limited manual evaluation is not conclusive). Cross entropy loss works better than unlikelihood loss, where dialogue success may be reduced.
6) There is ongoing work on building a team for JSALT research summer workshop. https://www.clsp.jhu.edu/2023-jelinek-summer-workshop, https://jsalt2023.univ-lemans.fr/en/index.html. This is a prestigious workshop organized by John Hopkins University every year. This year it is supported and co-organized by the University of Le Mans. Our topic is the Automatic design of conversational models from observation of human-to-human conversation (https://jsalt2023.univ-lemans.fr/en/automatic-design-of-conversational-models-from-observation-of-human-to-human-conversation.html). The topic passed a scientific review by more than 40 world-class researchers in AI in Baltimore, USA, in December 2022, and was selected for this workshop out of 15 proposals together with three others. The workshop topic builds on the outcome of this microproject and will reuse the collected data.

Publications:
Not published yet.

Links to Tangible results:
(WIP) Code for audio data collection: https://github.com/oplatek/speechwoz
(WIP) Code for end-to-end response generation: https://github.com/knalin55/augpt
(WIP) Report for end-to-end response generation:
https://docs.google.com/document/d/1iQB1YWr3wMO8aEB08BUYBqiLh0KreYjyO4EHnb395Bo/edit
(WIP = Work in Progress?)

Workshop links:
https://www.clsp.jhu.edu/2023-jelinek-summer-workshop,
https://jsalt2023.univ-lemans.fr/en/index.html

Workshop Proposal:
https://docs.google.com/document/d/19PAOkquQY6wnPx_wUXIx2EaInYchoCRn/edit?usp=sharing&ouid=105764332572733066001&rtpof=true&sd=true
Workshop team (not finalized yet):
https://docs.google.com/spreadsheets/d/1EsHZ-_OREkvf8ODiN7759OYHqSb6MBAX/edit?usp=sharing&ouid=105764332572733066001&rtpof=true&sd=true

2.3. On-going microprojects

As noted, most MPs in this period are carried out together with WP1. See the list of on-going MPs in WP1.

3. WP3: Human AI Collaboration and Interaction

3.1. Overview of goals

This work package aims to establish new methodological and conceptual basis for human-AI collaboration. The goal is to develop methodology for social basis for human-AI partnership, especially group cognition and emotional expression. For AI to understand people, it needs to both be able to infer intentions and emotions from observations as well as make its own intentions understandable to human partners via grounding, emotional expression, and explanation. We believe that these capabilities need to be to some extent be engineered into AI, in order to ensure more natural behavior from first interaction and to reach a desirable level of controllability and transparency. However, they need to be made interactive for users to control and understand. The main objective of this WP is machine-learning methods and suitable interaction techniques based on theoretically grounded models of human-human communication, which can drive the inference and planning of an AI agent in a more human way and with less training data. These models include models of multimodal communication, for grounding, theory of mind, and emotion. They work with interaction histories collected over a longer timespan and over a richer set of sensors than previously.

In Y3, the WP has focused on *interactive grounding* as a joint theme. The key question is how to make AI and human understand each other by establishing what is called common ground. The more recently started MPs in this reporting period reflect that joint topic.

3.2. Overview of WP3 microproject status (September 2023)

Finished and reported within 1st and 2nd reporting period

2021-3-WP123:	Memory Aware Conversational AI to aid virtual Team-Meetings
2021-3-WP3:	Exploring the impact of Agency on Human-Computer Partnerships
2022-4-WP3:	Making for Everyone: Interactive, voice-based support for the design of 3D objects for digital fabrication

2022-4-WP35:	Human-machine collaboration for content analysis in context of Ukrainian war
2022-5-WP3:	In Sync: Synchronization in Interaction between Humans and Embodied AI Systems
2022-7-WP3	A Scoping Review on Artificial Intelligence and Extended Reality
Ongoing at the end of reporting 2nd reporting period, but finished now	
2022-9-WP235:	Multilingual and Multimodal conversational agent combined with search engine models.
2022-9-WP235:	Use of dialog context to boost ASR/NLG/TTS and improve the overall quality of voice dialog systems
2022-11-WP3	A Scoping Review on Artificial Intelligence and Extended Reality
2022-11-WP23:	Use of dialog context to boost ASR/NLG/TTS and improve the overall quality of voice dialog systems
2023-1-WP3:	Storyboarder: combining Play Script and Image Generation
Ongoing	
2023-6-WP3	Grounded dialog models from observation of human-to-human conversation
2023-6-WP367	AI Aviation Assistant: An Intelligent Pilot Support Tool
2023-7-WP3	Exploring Multi-Modal Interaction Concepts to Support Users in AI-aided Co-Creation of 3D Objects in Virtual Environments
2023-7-WP123	Proactive announcement based on DEL (Dynamic Epistemic Logic)
Starting soon	
2023-9-WP1238	Educational module Human-Interactive Robot Learning (HIRL)
2023-9-WP3	Exploring the Role of AI in Amplifying Creative Potential: Fostering Ownership In Human-AI Collaborations
2023-9-WP123	Gesture-based Interactive Grounding for Mixed-Reality Human-AI Collaboration

3.3. Completed microprojects

3.1 2021-3-WP3: Exploring the impact of Agency on Human-Computer Partnerships

Start Date: 2021-03-10 - ????
Proposed Duration: 4
End Date: 2023-09-01

Submitted by: Janin Koch, Janin.Koch@inria.fr

Partners:
1. Ludwig-Maximilians-Universität München, Albrecht Schmidt
2. Københavns Universitet, Kasper Hornbaek
3. Stichting VU, Koen Hindriks
4. Umeå University, Helena Lindgren

Contributes to: WP3: T3.1, T3.3

Results:
The project's objective was to examine the theoretical and practical contributions of agency to successful human-computer partnership. Understanding agency is critical for establishing effective collaboration in human-centered AI research.

The goal of this project was to 1) produce an overview of existing HCI and AI literature on agency, 2) hold a workshop to brainstorm and categorize agency components, define metrics and experimental protocols, and 3) create interactive demos demonstrating various forms of human and system agency.

We conducted individual and collaborative brainstorming sessions with all participants to create an initial overview of current literature in order to establish a common starting point (1). We talked about potential overlaps in our work and how such perspectives influence our current work.
We will hold a workshop at CHI'23 on 'Integrating AI in Human-Human Collaborative Ideation' to examine the role AI can play in such interactive environments in order to identify distinct dimensions and measures of agency within human-ai interaction (2) [Shin et al. , 2023].
Umeå has also investigated how conversations between a human and a socially intelligent robot in a home environment can influence perceptions of agency [Tewari and Lindgren, 2022] and the importance of goal setting in such a scenario [Lindgren and Weck, 2022; Kilic et al., 2023] (3).

While there is currently project underway, we will report our findings in the second half of this year.

Publications:
1) Tewari M and Lindgren H (2022), Expecting, understanding, relating, and interacting - older, middle-aged and younger adults' perspectives on breakdown situations in human–robot dialogues. Front. Robot. AI 9:956709. doi: 10.3389/frobt.2022.956709.
2) Kilic K, Weck S, Kampik T, Lindgren H. Argument-Based Human-AI Collaboration for Supporting Behavior Change to Improve Health. to appear in Front. AI, 2023.
3) Lindgren H and Weck S. 2022. Contextualising Goal Setting for Behaviour Change – from Baby Steps to Value Directions. In 33rd European Conference on Cognitive Ergonomics (ECCE2022), October 4–7, 2022, Kaiserslautern,Germany. ACM, New York, NY, USA. https: //doi.org/10.1145/3552327.3552342
4) Joongi Shin, Janin Koch, Andrés Lucero, Peter Dalsgaard, Wendy E. Mackay. Integrating AI in Human-Human Collaborative Ideation. CHI 2023 - SIGCHI conference

on Human Factors in computing systems, Apr 2023, Hamburg, Germany. pp.1-5. ⟨hal-04023507⟩

Links to Tangible results:
https://www.frontiersin.org/articles/10.3389/frobt.2022.956709/full
https://www.frontiersin.org/articles/10.3389/frai.2023.1069455/full
https://dl.acm.org/doi/abs/10.1145/3552327.3552342
Tb released in the proceedings of CHI

2022-4-WP3: Making for Everyone: Interactive, voice-based support for the design of 3D objects for digital fabrication

Start Date: 1 May 2022
Proposed Duration: 4 Months
End Date: 31 Aug 2022

Submitted by: Florian Müller, florian.mueller@um.ifi.lmu.de
Partner:
1. LMU Munich, Florian Müller, florian.mueller@um.ifi.lmu.de (3 PM)
2. University of Bari (External Partner), Giuseppe Desolda, giuseppe.desolda@uniba.it

Contributes to: WP3, T3.2

Results:
Manufacturing tools like 3D printers have become accessible to the wider society, making the promise of digital fabrication for everyone seemingly reachable. While the actual manufacturing process is largely automated today, users still require knowledge of complex design applications to not only produce ready-designed objects, but also adapt them to their needs or design new objects from scratch. To lower the barrier for the design and customization of personalized 3D models, we imagine an AI-powered system that assists users in creating 3D objects for digital fabrication. Reaching this vision requires a common understanding - a common ground - between the users and the AI system.

As a first step, in this micro project, we explored novices' mental models in voice-based 3D design by conducting a high-fidelity Wizard of Oz study with 22 participants without skills in 3D design. We asked the participants to perform 14 tasks revolving around some basic concepts of 3D design for digital modeling, like the creation of objects, the manipulation of objects (e.g., scaling, rotating, and/or moving objects), and the creation of composite objects. We performed a thematic analysis of the collected data assessing how the mental model of novices translates into voice-based 3D design.

We found that future AI assistants to support novice users in voice-based digital modeling must: manage the correction the users do during and after the commands to fix certain errors; deal with vague and incomplete commands by automatically completing the commands with sensible defaults or by asking the users for clarification; consider the prior novices knowledge, for example, about the use of undo e redo functions; provide only a simplified set of operations for creating simple and composite 3D objects; design a

workflow similar to what novices would do if they were building real objects, for example, providing wizard procedures that guide novices in designing composite 3D models starting from the bottom; provide different commands to select 3D objects; understand and execute chained commands; understand commands that are relative to the users' point of view; grant multiple ways to refer to the axes, for example, by using their names, colors and user direction; favor explicit trigger words to avoid unintentional activation of the voice assistant; embrace diversity in naming approaches since novices often use other words to refer to 3D objects.

Publications:
Paper was not accepted at CHI 23. Currently under submission to INTERACT 23.

Links to Tangible results:
The transcribed and coded data we collected in our study, together with the codebook. We plan to make this data available to the community through a publication after the paper is published.
https://syncandshare.lrz.de/getlink/fiEFHiEQVQYtHDj5ZSWBdp/

2022-4-WP35: Human-machine collaboration for content analysis in context of Ukranian war

Start Date: 1 June 2022
Proposed Duration: 9 months
End Date: 28 Feb 2023

Submitted by: Virginia Dignum
Carmela Comito, carmela.comito@icar.cnr.it
Partners
1. Umeå University (UMU), Nina Khairova, nina.khairova@umu.se (6 PM)
2. Consiglio Nazionale delle Ricerche (CNR), Carmela Comito, carmela.comito@icar.cnr.it
3. Università di Bologna (UNIBO), Andrea Galassi, p.torroni@unibo.it

Contributes to: WP3, WP5, T5.6, T3.6

Description:
In this project, which we work with a Ukranian academic refugee, to combine methods for semantic text similarity with expert human knowledge in a participatory way to develop a training corpus that includes news articles containing information on extremism and terrorism.

Results
1) Collection and curation of two event-based datasets of news about Russian-Ukrainian war.
The datasets support analysis of information alteration among news outlets (agency and media) with a particular focus on Russian, Ukrainian, Western (EU and USA), and international news sources, over the period from February to September 2022. We

manually selected some critical events of the Russian-Ukrainian war. Then, for each event, we created a short list of language-specific keywords. The chosen languages for the keywords are Ukrainian, Russian, and English.

Finally, besides the scraping operation over the selected sources, we also gather articles using an external news intelligence platform, named Event Registry which keeps track of world events and analyzes media in real-time. Using this platform we were able to collect more articles from a larger number of news outlets and expand the dataset with two distinct article sets. The final version
of the RUWA Dataset is thus composed of two distinct partitions.

2) Development of an unsupervised methodology to establish whether news from the various parties are similar enough to say they reflect each other or, instead, they are completely divergent and therefore one is likely not trustworthy. We focused on textual and semantic similarity (sentence embeddings methods such as Sentence-BERT), comparing the news and assess if they have a similar meaning. Another contribution of the proposed methodology is a comparative analysis of the different media sources in terms of sentiments and emotions, extracting subjective points of view as they are reported in texts, combining a variety of NLP-based AI techniques and sentence embeddings techniques. Finally, we applied NLP techniques to detect propaganda in news article, relying on self-supervised NLP systems such as RoBERTa and existing adequate propaganda datasets.

Preliminary Qualitative results:
When the events concern civilians all sources are very dissimilar. But Ukraine and Western are more similar. When the event is military targets, Russian and Ukraine sources are very dissimilar from other sources, there is more propaganda in Ukraine and Russian ones.

Contribution to the objectives of HumaneAI-net WPs:

The micro-project has been realized together with a refugee Ukranian academic, addressing, thus, WP5 goals by ensuring an AI system operating within a moral and social framework, in verifiable and justified ways. We focused on methods ensuring responsible design of AI Systems and compliance to ethical, trust, fairness, public perception and societal principles.

Publications:
Working on a conference and a journal papers.

Links to Tangible results:
Github repository of datasets and software:
https://github.com/fablos/ruwa-dataset

2022-5-WP3: In Sync: Synchronization in interaction between Humans and Embodied AI Systems

Start Date: 2022-05-01
Proposed Duration: 4 months
End Date: 2022-08-31

Submitted by: Florian Müller, florian.mueller@um.ifi.lmu.de
Partners
1. LMU Munich, Florian Müller, florian.mueller@um.ifi.lmu.de Institution (4 PM)
2. University Warsaw, Andrzej Nowak, andrzejn232@gmail.com (4 PM)

Contributes to: WP3: T3.2

Description:
When we go for a walk with friends, we can observe an interesting effect: From step lengths to arm movements - our movements unconsciously align; they synchronize. Prior research in social psychology found that this synchronization is a crucial aspect of human relations that strengthens social cohesion and trust. In this micro project, we explored if and how this effect generalizes beyond human-human relationships. We hypothesized that synchronization can enhance the relationship between humans and AI systems by increasing the sense of connectedness in the formation of techno-social teams working together on a task.

Results
To evaluate the feasibility of this approach, we built a prototype of a simple non-humanoid robot as an embodied representation of an AI system. The robot tracks the upper body movements of people in its vicinity and can bend to follow human movements and vary the movement synchronization patterns. Using this prototype, we conducted a controlled experiment with 51 participants exploring our concept in a between-subjects design. We found significantly higher ratings on trust between people and automation in an established questionnaire for synchronized movements. However, we could not find an influence on the willingness to spend money in a trust game inspired by behavioral economics. Taken together, our results strongly suggest a positive effect of synchronized movement on the participants' feeling of trust toward embodied AI representations.

List of pubications,
1. Wieslaw Bartkowski, Andrzej Nowak, Filip Ignacy Czajkowski, Albrecht Schmidt, and Florian Müller. 2023. In Sync: Exploring Synchronization to Increase Trust Between Humans and Non-humanoid Robots. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23), April 23–28, 2023, Hamburg, Germany. ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3544548.3581193

Links to Tangible results:
Paper:
https://syncandshare.lrz.de/getlink/fiAUnSaqgQJyEdcw5XJqpN/in_sync_final.pdf
Video:
Short:
https://syncandshare.lrz.de/getlink/fiRjwbk1AoYxKujaEaZ5ax/in_sync_video_short.mp4

Full:
https://syncandshare.lrz.de/getlink/fiGEX3bGahhbzrChUiXqvL/in_sync_video_full.mp4
Repository: https://github.com/wbartkowski/In-Sync-Robot-Prototype

2022-7-WP3 A Scoping Review on Artificial Intelligence and Extended Reality

Start Date: 2022-07-11
Proposed Duration: ??
End Date: 2023-03-01

Submitted by: Teresa Hirzle, tehi@di.ku.dk
Partners
1. UCPH, Teresa Hirzle, kash@di.ku.dk
2. LMU, Florian Müller, albrecht.schmidt@ifi.lmu.de
3. Saarland University (External partner), Martin Schmitz
4. Universität Innsbruck (External partner), Pascal Knierim

Results:
We conducted a scoping review covering 311 papers published between 2017 and 2021. First, we screened 2619 publications from 203 venues to cover the broad spectrum of XR and AI research. For the search, we I inductively built a set of XR and AI terms. The venues include research from XR, AI, Human-Computer Interaction, Computer Graphics, Computer Vision, and others. After a two-phase screening process, we reviewed and extracted data from 311 full papers based on a code book with 26 codes about the research direction, contribution, and topics of the papers, as well as the algorithms, tools, datasets, models, and data types the researchers used to address research questions on XR and AI. The extracted data for these codes form the basis for our predominantly narrative synthesis. As a result, we found fve main topics at the intersection of XR and AI: (1) Using AI to create XR worlds (28.6%), (2) Using AI to understand users (19.3%), (3) Using AI to support interaction (15.4%), (4) Investigating interaction with intelligent virtual agents (IVAs) (8.0%), and (5) Using XR to Support AI Research (2.3%). The remaining 23.8% of the papers apply XR and AI to an external problem, such as for medical training applications (3.5%) or for simulation purposes (3.0%). Finally, we summarise our findings in 13 research opportunities and present ideas and recommendations for how to address them in future work. Some of the most pressing issues are a lack of generative use of AI to create worlds, understand users, and enhance interaction, a lack of generalisability and robustness, and a lack of discussion about ethical and societal implications.

In terms of the call topics, we analysed whether XR can serve as a tool to establish and enhance interactive grounding in human-AI interaction. Here, we found that there is a lack of understanding user experience during human-AI interaction using XR technology. Typically, AI is used for content creation and to enhance interaction techniques. We did, however, not find a lot of papers that use XR to support human-AI interaction. There are some works that look into artificial agents and how an interaction with them can be realised through XR. However, most of these works do not yet work in real=time and are mostly based on mock-up scenes.

List of pubications

1. Teresa Hirzle, Florian Müller, Fiona Draxler, Martin Schmitz, Pascal Knierim, and Kasper Hornbæk. 2023. When XR and AI Meet - A Scoping Review on Extended Reality and Artifcial Intelligence. *In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (CHI '23), April 23–28, 2023, Hamburg, Germany. ACM, New York, NY, USA, 45 pages. https://doi.org/10.1145/3544548.3581072

Links to Tangible results:
Paper: https://thirzle.com/pdf/chi23_xrai_scoping_review_hirzle.pdf
Reviewed Papers and Coding Spreadsheet:
https://thirzle.com/supplement/chi23_xrai_scoping_review_hirzle.zip
Videos: There will be talk videos at a later stage.

3.4. On-going microprojects

2023-1-WP3: Storyboarder: combining Play Script and Image Generation

Start Date: 2023-01-02
Proposed Duration: 6
End Date: 2023-07-03 (on-going)

Submitted by: Victor Chitolina Schetinger, victor.schetinger@tuwien.ac.at
Partners:
1. Charles University, Rudolf Rosa, rosa@ufal.mff.cuni.cz
2. IADE – Faculdade de Design, Tecnologia e Comunicação da Universidade Europeia
3. (External Partner), Edirlei Lima, edirlei.lima@universidadeeuropeia.pt

Contributes to: WP3

Results:
In the short time since our start we were able to achieve the results we proposed, namely the generation of storyboards through the use of text and image generative models combined. Due to the fast development of these fields in the past months, however, the quality of the results is not up to the state of the art. We plan to use the remainder time of the project to improve them, explore adjacent research directions and work on a publication plan.

Links to Tangible results: (Here is the link to the current prototype*
https://ufallab.ms.mff.cuni.cz/cgi-bin/rosa/theaitre-demo/demo_images.py

2022-3-WP123: Gesture-based Interactive Grounding for Mixed-Reality Human-AI Collaboration

Start Date: 2023-09-01
Proposed Duration: 6 months
End Date: 2024-02-29

Submitted by: Rui Prada, rui.prada@tecnico.ulisboa.pt
Partners:
1. IST, Rui Prada, rui.prada@tecnico.ulisboa.pt (3 PM)
2. Eotvos Lorand University, Andras Lorincz, lorincz@inf.elte.hu
3. DFKI Lower Saxony, Daniel Sonntag, Daniel.Sonntag@dfki.de
4. External: CMU, László Jeni, laszlojeni@cmu.edu

Contributes to: WP1, WP2, WP3

Tagline: Develop AI interactive grounding capabilities in collaborative tasks using a game-based mixed reality scenario that require physical actions.

Project description: The project addresses research on interactive grounding. It consists of the development of an Augmented Reality (AR) game, using HoloLens, that supports the interaction of a human player with an AI character in a mixed reality setting using gestures as the main communicative act. The game will integrate technology to perceive human gestures and poses. The game will bring about collaborative tasks that need coordination at the level of mutual understanding of the several elements of the required task. Players (human and AI) will have different information about the tasks to advance in the game and need to communicate that information to their partners through gestures. The main grounding challenge will be based on learning the mapping between gestures to the meaning of actions to perform in the game. There will be two levels of gestures to ground, some are task-independent while others are task-dependent. In other words, besides the gestures that communicate explicit information about the game task, the players need to agree on the gestures used to coordinate the communication itself, for example, to signal agreement or doubt, to ask for more information, or close the communication. These latter gesture types can be transferred from task to task within the game, and probably to other contexts as well.

It will be possible to play the game with two humans and study their gesture communication in order to gather the gestures that emerge: a human-inspired gesture set will be collected and serve the creation of a gesture dictionary in the AI repertoire.
The game will provide different tasks of increasing difficulty. The first ones will ask the players to perform gestures or poses as mechanisms to open a door to progress to the next level. But later, in a more advanced version of the game, specific and constrained body poses, interaction with objects, and the need to communicate more abstract concepts (e.g., next to, under, to the right, the biggest one, …) will be introduced.
The game will be built as a platform to perform studies. It will support studying diverse questions about the interactive grounding of gestures. For example, we can study the way people adapt to and ascribe meaning to the gestures performed by the AI agent, we can study how different gesture profiles influence the people's interpretation, facilitate grounding, and have an impact on the performance of the tasks, or we can study different mechanisms on the AI to learn its gesture repertoire from humans (e.g., by imitation grounded on the context).

We see this project as a relevant contribution to the upcoming Macro Project on Interactive Grounding, and we would like the opportunity to join the MP later. Our focus is on the

grounding based on gestures being critical in certain scenarios. The setting can include language if vocalization is allowed and can be heard. Our game scenarios are simple and abstract and can be the basis for realistic ones.

List of tangible Results:
A game that serves as a platform for studying grounding in the context of collaborative tasks using gestures.

A repertoire of gestures to be used in the communication between humans and AI in a collaborative task that relies on the execution of physical actions. We will emphasize the gestures that can be task-independent.

The basis for an AI algorithm to ground gestures to meaning adapted to a particular user. One or two papers, describing the platform and a study with people.

2022-3-WP3: Exploring the Role of AI in Amplifying Creative Potential: Fostering Ownership In Human-AI Collaborations

Start Date: 2023-09-15
Proposed Duration: 6months
End Date: 2024-03-15

Submitted by: Steeven Villa, steeven.villa@um.ifi.lmu.de
Partners:
1. LMU München, Sebastian Feeger, sebastian.feger@ifi.lmu.de (4 PM)
2. External: Sheffield Hallam University Enterprises Limited, Daniela Petrelli, D.Petrelli@shu.ac.uk

Contributes to: WP3

Tagline: Exploring the balance between ownership and AI assistance in creative collaboration through an interactive exhibition.
Project description: Novel AI systems enable individuals to maximize their creative potential by rapidly prototyping ideas based on initial sketches or idea descriptions. A generative AI system is the bridge between an individual's thought and its physical manifestation. Traditional approaches, on the other hand, require a greater investment of effort, involvement, and time, which was historically associated with a sense of ownership over the creation and agency (or control) over the creation process. As the paradigm shift caused by AI significantly reduces the amount of work required to achieve a desired result, individuals consistently report low agency and ownership over their creations, and such boundaries are unclear even in the legal sphere. Therefore, it is essential to understand how these variables can be balanced to foster a strong sense of ownership while allowing users to fully exploit the potential of AI systems.

In this project, we seek to achieve this understanding by creating an interactive exhibition where visitors to a science museum will interact with a generative AI system to create

illustrations for a children's book based on rough sketches and prompts. The participants will be instructed to collaborate with an image-generating AI system to illustrate a children's storybook with a simple plot. Participants will start with their own sketch or by selecting one from a set. When an illustration is finished, participants will be asked if they want to sign the illustration with their name, the name of the AI model, or both. Participants will have the option to display their illustrations on the exhibition's billboard. We will conclude by asking them three brief questions about self-efficacy.

The interaction will be logged to record the degree of intervention (iterating over the illustrations, using a starting sketch instead of drawing their own sketch, signature ownership). We plan to carry out a limited quantitative study with observations of visitors' behavior paired with interviews. With the collected data, we will be able to analyze the correlations between time, effort, ownership, and self-efficacy in the AI-assisted creative process, and ultimately gain insights into how to design such systems to promote a sense of ownership in the user.

This project falls under WP3. It examines the Pragmatic aspects of communication and collaboration between humans and AI by exploring how participants collaborate with an AI system to translate their initial sketches or prompts into meaningful illustrations for visual narratives for Storytelling. Everything from the lens of influence of the participants' sense of ownership and agency and how it impacts the outcome and design process.

List of tangible Results:
A manuscript reporting the intervention and the results of the field-study.

A video explanation of the intervention and the insights gained from it.

An open-source repository of the materials used in the intervention.

2022-3-WP3: Exploring Multi-Modal Interaction Concepts to Support Users in AI-aided Co-Creation of 3D Objects in Virtual Environments

Start Date: 2023-07-01
Proposed Duration: 4 months
End Date: 2023-11-01

Submitted by: Teresa Hirzle, tehi@di.ku.dk
Partners:
1. UCPH, Kasper Hornbaek, Teresa Hirzle, tehi@di.ku.dk (4 PM)
2. LMU, Albrecht Schmidt, Florian Müer, Julian Rasch, florian.mueller@um.ifi.lmu.de; julian.rasch@um.ifi.lmu.de (3 PM)
3. External: Saarland University, Martin Schmitz, mschmitz@cs.uni-saarland.de

Contributes to: WP3,

The project aims to explore multi-modal interaction concepts for collaborative creation of 3D objects in virtual reality with generative AI assistance.

Project description: The use of generative AI in the creation of 3D objects has the potential to greatly reduce the time and effort required for designers and developers, resulting in a more efficient and effective creation of virtual 3D objects. Yet, research still lacks an understanding of suitable interaction modalities and common grounding in this field.

Objective: The objective of this research project is to explore and compare interaction modalities that are suited to collaboratively create virtual 3D objects together with a generative AI. To this end, the project aims to investigate how different input modalities, such as voice, touch and gesture recognition, can be used to generate and alter a virtual 3D object and how we can create methods for establishing common ground between the AI and the users.

Methodology: The project is split into two working packages. (1) We investigate and evaluate the use of multi-modal input modalities to alter the shape and appearance of 3D objects in virtual reality (VR). (2) Based on our insights on promising multi-modal interaction concepts, we then develop a prototypical multi-modal VR interface that allows users to collaborate on the creation of 3D objects with a generative AI. This might include, but is not limited to the AI assistant generating 3D models (e.g. using https://threedle.github.io/text2mesh or Shap-E) or providing suggestions based on the users' queries.

The project will use a combination of experimental and observational methods to evaluate the effectiveness and efficiency of the concepts. This will involve conducting controlled experiments to test the effects of different modalities and AI assistance on the collaborative creation process, as well as observing and analyzing the users' behavior.

Expected outcomes: The research project is expected to produce several outcomes, including a software package to prototype multi-modal VR interfaces that enables collaborative creation of 3D objects, insights into the effectiveness and efficiency of different modalities and AI assistance in enhancing the collaborative process, and guidelines for the design of multi-modal interfaces and AI assistance for collaborative creation of 3D objects. The project's outcomes may have potential applications in fields such as architecture, engineering, and entertainment.

Relation to call: This research project is directly related to the call for proposals as it addresses the
challenge of coordination and collaboration between AI and human partners in the context of creating 3D objects. The project involves the use of multi-modal interfaces and AI assistance to enhance the collaborative process, which aligns with the call's focus on speech-based and multimodal interaction with AI. Additionally, the project's investigation of co-adaptive processes in collaborative creation aligns with the call's focus on co-adaptive processes in grounding. The project's outcomes, such as the development of guidelines for the design of multi-modal interfaces and AI assistance for collaborative creation, may also contribute to the broader theme of interactive grounding. Finally, the project's potential applications in architecture, engineering, and entertainment also align with the call's focus on special application areas.

List of tangible Results:
1. VR co-creation software package: The project aims to develop a publicly-available open-source software package to quickly prototype Multi-Modal VR interfaces for co-creating virtual 3D objects. It enables practitioners and VR application developers to more easily create virtual 3D objects without requiring expert knowledge in computer-aided design.
2. Recorded dataset and derived guidelines for the design of multi-modal interfaces with AI assistance: The project aims to publish all recorded datasets and further provides a set of guidelines for the design of efficient and effective multi-modal interfaces for generating and altering 3D objects with an AI assistant
3. We aim for publishing the results of this research as a paper in a leading XR or HCI venue, such as CHI, UIST, or ISMAR.

2022-3-WP3: Grounded dialog models from observation of human-to-human conversation

Start Date: 2023-06-12
Proposed Duration: 3.5 months
End Date: 2023-08-31

Submitted by: Ondrej Dusek, odusek@ufal.mff.cuni.cz
1. BUT, Ondrej Dusek, odusek@ufal.mff.cuni.cz (3 PM)
2. Eotvos Lorand University, Andras Lorincz, lorincz@inf.elte.hu (3 PM)

Contributes to: WP3

Tagline: This project targets the design of grounded dialogue models from observation of human-to-human conversation, typically from a set of recordings. It will bring trustable conversation models as well as a tool for the analysis of dialogue behavior.

Project description: This microproject aims to design grounded dialogue models based on observation of human-to-human dialogue examples, i.e., distilling dialogue patterns automatically and aligning them to external knowledge bases. The current state-of-the-art conversation models based on finetuned large language models are not grounded and mimic their training data, or their grounding is external and needs to be hand-designed. Furthermore, most commercially deployed dialogue models are entirely handcrafted. Our goal is to produce grounding for these models (semi-)automatically, using dialogue context embedded in vector spaces via large language models trained specifically on conversational data. If we represent dialogue states as vectors, the whole conversation can be seen as a trajectory in the vector space. By merging, pruning, and modeling the trajectories, we can get dialog skeleton models in the form of finite-state graphs or similar structures. These models could be used for data exploration and analysis, content visualization, topic detection, or clustering. This can bring faster and cheaper design of fully trustable conversation models. The approach will serve both to provide external model grounding and to analyze the progress in human-to-human dialogues, including negotiation around the participants' common ground.

The microproject will investigate the optimal format of the dialogue context embeddings (such as temporal resolution) as well as the optimal ways of merging dialogue trajectories and distilling models. Here, Variational Recurrent Neural Networks with discrete embeddings (Shi et al., NAACL 2019) are a promising architecture, but alternatives will also be considered.

We plan to experiment with both textual and voice-based dialogues. We will use the MultiWOZ corpus (Budzianowski et al., EMNLP 2018) as well as the DIASER extension developed in a Humane AI microproject by CUNI+LIMSI (Hudecek et al., LREC 2022) for text-based experiments. For voice-based experiments, we will use MultiWOZ spoken data released for the DSTC11 Challenge and dialogue data currently developed in a Humane AI microproject by BUT+CUNI.

The work will be done as a part of the JSALT workshop hosted by the University of Le Mans, France, and co-organized by Johns Hopkins University (JHU) and the Brno University of Technology.
https://jsalt2023.univ-lemans.fr/en/index.html

The JSALT workshop topic leader is Petr Schwarz from BUT (MP partner). The topic passed a scientific review by about 40 researchers in Baltimore, USA, in December 2022 and was selected among four workshop topics.
https://jsalt2023.univ-lemans.fr/en/automatic-design-of-conversational-models-from-observation-of-human-to-human-conversation.html

workshop Topic Proposal:
https://docs.google.com/document/d/19PAOkquQY6wnPx_wUXIx2EaInYchoCRn/edit?usp=sharing&ouid=105764332572733066001&rtpof=true&sd=true
Workshop Topic Presentation:
https://docs.google.com/presentation/d/1rt7OFvIu34c3OCXtAkoGjMhVXrcxCSXz/edit?usp=sharing&ouid=105764332572733066001&rtpof=true&sd=true

Workshop Team:
https://docs.google.com/spreadsheets/d/1EsHZ-_OREkvf8ODiN7759OYHqSb6MBAX/edit?usp=sharing&ouid=105764332572733066001&rtpof=true&sd=true

The workshop and attendants will be supported by several sources – JHU sponsors, European Esperanto project, private companies, and the HumanE AI project. Ondrej Dusek from CUNI is responsible for the HumanE AI participants (as MP PI). The aim is to cover mainly travel, accommodation, per diem to move participants to Le Mans, and some preparation.

A joint place for four workshop topics having teams with world-top researchers and initial summer school gives participants an excellent opportunity for networking and personal growth, with high visibility and high impact of the work results. We expect that this effort can start new long-term collaborations among the participating institutions.

List of tangible Results:
- Software – code for dialogue embeddings & trajectory merging
- Trained embedding models
- Paper describing the dialogue skeleton models

2022-3-WP367: AI Aviation Assistant: An Intelligent Pilot Support Tool

Start Date: 2023-06-01
Proposed Duration: 5 months
End Date: 2023-10-31

Submitted by: Florian Müller, florian.mueller@um.ifi.lmu.de
Partners:
1. LMU München, Florian Müller, florian.mueller@um.ifi.lmu.de (5 PM)
2. External, ENAC, Anke Brock, anke.brock@enac.fr

Contributes to: WP3, WP6&7

Tagline: We are going to build and evaluate a novel AI aviation assistant for supporting (general) aviation pilots with key flight information that facilitate decision making, placing particular emphasis on their efficient and effective visualization in 3D space
Project description: Pilots frequently need to react to unforeseen in-flight events. Taking adequate decisions in such situations requires to consider all available information and demands strong situational awareness. Modern on-board computers and technologies like GPS radically improved the pilots' abilities to take appropriate actions and lowered their required workload in recent years. Yet, current technologies used in aviation cockpits generally still fail to adequately map and represent 3D airspace. In response, we aim to create an AI aviation assistant that considers all relevant aircraft operation data, focuses on providing tangible action recommendations, and on visualizing them for efficient and effective interpretation in 3D space. In particular, we note that extended reality (XR) applications provide an opportunity to augment pilots' perception through live 3D visualizations of key flight information, including airspace structure, traffic information, airport highlighting, and traffic patterns. While XR applications have been tested in aviation in the past, applications are mostly limited to military aviation and latest commercial aircrafts. This ignores the majority of pilots in general aviation, in particular, where such support could drastically increase situational awareness and lower the workload of pilots. General aviation is characterized as the non-commerical branch of aviation, often relating to single-engine and single-pilot operations.
To develop applications usable across aviation domains, we plan to create a Unity project for XR glasses. Based on this, we plan to, in the first step, systematically and iteratively explore suitable AI-based support on pilot feedback in a virtual reality study in a flight simulator. Based on our findings, we refine the Unity application and investigate opportunites to conduct a real test flight with our external partner ENAC, the French National School of Civil Aviation, who own a plane. Such a test flight would most likely use latest Augmented Reality headsets like the HoloLense 2. Considering the immense safety requirements for such a real test flight, this part of the project is considered optional at this stage and depends on the findings from the previous virtual reality evaluation.

The system development will particularly focus on the use XR techniques to create more effective AI-supported traffic advisories and visualizations. With this, we want to advance the coordination and collaboration of AI with human partners, establishing a common ground as a basis for multimodal interaction with AI (WP3 motivated). Further, the MP relates closely to “Innovation projects (WP6&7 motivated)”, calling for solutions that address “real-world challenges and opportunities in various domains such as (...) transportation [...]”.

List of tangible Results:
- Requirements and a prototype implementation for an AI-based assistant that provides recommendations and shows selected flight information based on pilot workload and current flight parameters
- A Unity project that implements an extended reality support tool for (general) aviation and that is used for evaluation in simulators (Virtual Reality) and possibly for a real test flight at ENAC (Augmented Reality)
- Findings from the simulator study and design recommandations
- (Optional) Impressions from a real test flight at ENAC
- A research paper detailing the system and the findings