Funded by the
European Union

www.gain-twinning.eu

Project Acronym: **GAIN**

Project Full Title: **Georgian Artificial Intelligence Networking and Twinning Initiative**

Grant Agreement Number: 101078950

HORIZON EUROPE - WIDERA-2021-ACCESS-03 (Twinning)

WEBSITE: www.gain-twinning.eu

Project Start Date: 1st October 2022

Project End Date: 30 September 2025

Funded by the
European Union

www.gain-twinning.eu

Deliverable D3.3 GAIN webinar series report

Dissemination level:	PU
Date:	September 30, 2024
Extended date:	October 31, 2024
WP leader:	DFKI
Author:	B. E. Wirth, P. Müller, T. Agrawal,
Contributors:	All
Reviewers:	G. Giorgobiani
Type:	Report
Task responsible:	DFKI

Contents

1. EXECUTIVE SUMMARY

This document describes the **GAIN Webinar Series Report** (D3.3) of the GAIN project.

Document History			
Version	**Date**	**Contributors**	**Description**
1	06/09/2024	B. E. Wirth – DFKI	GAIN Webinar Series Report – First Version
2	15/09/2024	B. E. Wirth, P. Müller – DFKI; T. Agrawal – INRIA; G. Giorgobiani - MICM	GAIN Webinar Series Report – Addition of new subsections
3	30/09/2024	B. E. Wirth, P. Müller – DFKI; T. Agrawal – INRIA; G. Giorgobiani - MICM	Final version

2. VERSIONS AND UPDATES

The initial version of the document was drafted by B.E. Wirth (DFKI). P. Müller (DFKI), T. Agrawal (INRIA) and G. Giorgobiani (MICM) subsequently made some contributions.

3. INTRODUCTION

The main goal of the GAIN project is to support Georgia, one of the widening countries, to integrate into the European research area (ERA) of artificial intelligence (AI). This goal will be achieved by adjusting the research profile of MICM and linking it to the European AI research and innovation community. To this end, the GAIN project initiated joint research projects enabling long-term cooperation and partnership between Georgian and European researchers.

Work Package 3 of the GAIN project – **Scientific Excellence and Networking** – is a crucial contributing factor to achieving this goal. The core objective of this work package is to foster research capacities in MICM, the Georgian partner institute, and to develop and establish careers for young Georgian researchers by education, training, and mentorship. In this regard, the goals of the webinar series are

1. **Transfer of knowledge:** Education of young Georgian researchers about trending topics in AI, their central concepts and theories as well as promising avenues for own research

2. **Transfer of skills:** Training young Georgian researchers on the implementation of ML algorithms, the formulation, and testing of scientific hypotheses, and the presentation of scientific results in scientific articles and conference talks
3. **Transfer of soft skills:** Mentoring of young Georgian researchers regarding project-management, networking, publishing strategies, and career planning

4. INITIAL WEBINAR SESSIONS

In the first three months of the project (October – December 2022), researchers from DFKI and INRIA (i.e., from the established research institutions) gave an initial series of webinar sessions for a large group of Georgian researchers with the aims to

1. Give an overview about current hot topics in AI that would serve as basis for future collaborations between researchers from DFKI and INRIA on the one hand and from MICM on the other hand.
2. Explain the methods and algorithmic basis of these research topics.
3. Assess the individual interests as well as skills and capabilities of the Georgian researchers in order to assign them to fitting research projects.

The initial webinar sessions were designed to give students an overview of the available research topics. These included several topics in multi-modal machine learning, human behaviour analysis, brain-computer interfaces, and applications to psychiatric disorders. In discussions, students articulated their research interests and career goals, and supervisors from DFKI and INRIA gave advice on how these interests and goals can be realized in the specific research projects.

Based on these initial webinar sessions, the following subgroups were formed, each of them with their own scientific focus (see Table 1)

Table 1. Project-related research groups formed at the end of the initial webinar series

Subgroup project	Supervisors	Georgian Researchers
Generalizability of emotion recognition from EEG signals	L.M. Ferrari (INRIA), P. Müller (DFKI), B.E. Wirth (DFKI)	R. Kalandadze, S. Katamadze, N. Kukhilava, T. Tsmindashvili
Emotion recognition from full body movements	T. Agrawal (INRIA), M. Balazia (INRIA), P. Müller (DFKI), B.E. Wirth (DFKI)	I. Kachiashvili, D. Karapetian, L. Tabagari

LLMs for Georgian – a low-resource language	Hali Lindsay (DFKI; until 01/2023); P. Müller (DFKI), Simon Ostermann (DFKI), Joseph van Genabith (DFK), Lonneke van der Plas (Idiap, Switzerland)	B. Mikaberidze, T. Saghinadze, B. Tepnadze
Parameter-efficient transfer learning for CNNs in action recognition	T. Agrawal (INRIA), F. Bremond (INRIA)	T. Saghinadze, D. Datuashvili
Digital phenotyping in psychiatric interactions – a Georgian study protocol based on MePheSTO	Alexandra König (INRIA), Hali Lindsay (DFKI; until 01/2023)	E. Chkonia, E. Gaprindashvili, T. Giorgobiani, K. Sulaberidze, S. Tsagareishvili

5. ONGOING REGULAR WEBINAR SERIES

After the project-related subgroups (see Table 1 above) had been established at the end of the initial webinar series in December 2022, a series of ongoing, regular webinars with the smaller, research-oriented subgroups was started in January 2023. These webinars are ongoing to this date. They will be described in detail below.

5.1. GENERALIZABILITY OF EMOTION RECOGNITON FROM EEG SIGNALS

The first subgroup decided to work in the field of emotion recognition from EEG signals. Within this field, participants are asked to watch emotional films while EEG is recorded. At the end of each film, participants are asked to rate their own emotional state. The goal of this research field is to predict participants' emotional state based on the EEG signals. However, studies in this field suffer from low generalizability because of the use of different datasets, different train-test splits of the data, different data pre-processing, and inconsistent evaluation measures. Thus, in their research project, the subgroup aims to

1. Systematically address the existing inconsistencies that hamper comparability, generalizability, and – ultimately – progress in the field.
2. Provide guidelines and gold-standard research practices for future research.

3. Devise a software framework that enables other researcher to adhere to these gold-standard practices.

In order to meet these goals, researchers from INRIA (Laura M. Ferrari) and DFKI (P. Müller and B.E. Wirth) gave weekly webinar sessions to researchers from MCMI (R. Kalandadze, S. Katamadze, N. Kukhilava, T. Tsmindashvili). During crucial times of the project, these sessions took even place twice a week.
The following topics were covered in these meetings.

5.1.1 EEG: Theoretical foundations and technical issues

Electroencephalography (EEG) is a technique that aims to measure the electrocortical signals of the brain during specific processes (here: during the emotional processing of a film). In the beginning of the webinar, the Georgian researchers were taught about the fundamental principles of this technique, for example:

- What roles do the ground and the reference electrode play?
- Why are the impedances between electrodes and participant scalp critical?
- How does the international 10-20 system for naming the electrodes work?
- How does the sampling rate relate to aliasing?

However, the EEG signal that is obtained directly after the recording is rather noisy. Therefore, a number of pre-processing steps are applied to the signal before it is analyzed in the second half of this thematic block, different preprocessing techniques were explained, for example:

- Bandpass and notch filters
- Downsampling.
- Windowing/epoching.
- Artifact removal.
- Normalization.

5.1.2 Procedural basics of literature reviews

The subgroup decided to write a literature review with the aim of making the research community of EEG-based emotion recognition aware of the problem that inconsistencies in the use of (a) different datasets, (b) different pre-processing strategies (c) different train-test splits (d) different ground-truth definitions, and (e) different evaluation metrics impedes comparability, generalizability, and thus progress in the field. Therefore, the second part of the webinar covered theoretical and technical aspects of writing literature reviews, such as:

- Formulating a search strategy to find and identify relevant publications.

- Defining inclusion and exclusion criteria.
- Specifying a multi-dimensional classification scheme for included articles.
- Applying the classification scheme to concrete publications.
- Reporting literature reviews in scientific publications (e.g., PRISMA guidelines, flow chart, etc.)

5.1.3 Framework evaluation

As part of the research project, the subgroup created a software framework that allows researchers to directly import different datasets used in the field of EEG-based emotion recognition. In order to ascertain that the framework is working as intended, the third part of the webinar focused on the evaluation of the framework. More specifically, the researchers of INRIA and DFKI taught the Georgian researchers how to perform validation experiments, which evaluation metrics to use, and how to report these metrics.

5.1.4 Writing of scientific articles

The subgroup is currently writing a paper reporting the literature review they conducted and the software framework they created. Thus, during the fourth part of the webinar, the young Georgian researchers were taught how to present their work in a scientific article. The topics of this part included:

- Scientific language: How to write concisely and precisely
- Basic article structure
- Scientific crediting: When and how to cite (in-text citations, bibliography, etc.)
- Reporting literature reviews (e.g., flowcharts, PRISMA guidelines, etc.)
- How to clearly present scientific results
- Creation of illustrative figures.

5. 2. EMOTION RECOGNITION FROM FULL BODY MOVEMENTS

The second subgroup decided to pursue a project in the field of emotion recognition from body movements. To this end, the group worked on the MPII-Emo dataset, a collection of film clips depicting people that are engaged in dyadic (i.e., two-person) interactions with emotional content in a kitchen setting. The dataset is characterized by high ecological validity because the content of the interactions is only minimally scripted. The group aims to develop emotion recognition approaches that are able to classify the emotions of the depicted persons based on the body movements (i.e., gestures, postures) they make. In order to support the subgroup in attaining their goal, researchers from INRIA (Michal Balazia and Tanay Agrawal) as well as DFKI (P. Müller and B.E. Wirth) gave regular webinar sessions

to researchers from MICM (I. Kachiashvili, D. Karapetian, L. Tabagari). The frequency of these sessions was contingent upon the current workload at hand and ranged from bi-weekly to twice a week. The following topics were covered in these sessions.

5.2.1 Video processing and person tracking

The first challenge encountered by the group was to track the acting persons of the videos and to assign them a unique identifier in order to subsequently be able to train a classifier. Therefore, the first part of the webinar series covered the topic of person tracking. The researchers from DFKI and IRNIA taught two general possible approaches to the young Georgian researchers: skeleton-based approaches and bounding-box based approaches. Skeleton-based approaches (e.g., open pose) detect the joints (elbows, knees, wrists, etc.) of a person to create a virtual representation of the skeleton of a person (like a stick figure). Throughout the frames of a video, the coordinates of the different joints of a person are saved. Bounding-box based approaches, in contrast, do not use the joints of a person to estimate their pose, but create a rectangular frame that follows the person around the video. Thus, throughout the video, the coordinates, the height and width of the bounding box per frame are saved.

The Georgian researchers decided to use the bounding-box based approach because popular skeleton-based approaches like open pose do not allow to assign a permanent identifier to the persons depicted in a video. Thus, the Georgian researchers used the bounding-box based ByteTrack tool in order to track people and subsequently create a cropped video only showing one person. Unfortunately, however, the quality of the cropped videos was initially poor (stretching/squeezing, lagging, framedrops, etc.). Therefore, the Georgian researchers were taught video processing techniques that would solve this problem, for example percentile-based outlier techniques and window-based median smoothing.

5.2.2 Feature extractors for videos

The next step towards a system for emotion recognition based on full body movements is the extraction of features from the cropped videos. Therefore, in the second part of the webinar series, Georgian researchers were taught about several different feature extractors, for example:

- OpenTad.
- Swin Transformer.
- Slow-Fast.
- DINOv2.

5. 3. LLMs FOR GEORGIAN – A LOW-RESOURCE LANGUAGE

The scientific focus of this subproject is to research and develop natural language processing (NLP) approaches for Georgian, a language with only few available resources. The young Georgian researchers were taught the entire pipeline of a scientific project: starting with literature review, formulation of hypotheses, operationalization and implementation of experiments, evaluation and data analysis, as well as paper writing.

To cover the entire NLP pipeline, the project started with an investigation of different tokenization methods. In particular, we conducted the first systematic evaluation of tokenization methods for Georgian. The students learned about- and compared standard subword tokenizers, such as WordPiece, Byte Pair Encoding, SentencePiece with Unigram, and a recently proposed token-free approach. They defined and pre-processed language corpora for pre-training as well as for evaluation on downstream tasks. This involved the curation of existing and newly created corpora. Downstream tasks included sentiment analysis, part-of-speech tagging, named entity recognition, and toxicity detection. The students got hands-on experience in training of BERT-style language models from scratch and on evaluating such models on the downstream tasks. These training processes also included familiarization with complete compute architecture and associated tools.

Based on the results achieved in the group, we submitted a paper to the 7th International Conference on Natural Language and Speech Processing which was accepted. In this process, the students learned how to write scientific papers, including crafting a story line, presenting results in the proper way, etc.

At present, we continue working on the next step: Training of a full-scale BERT-style language model on an increased Georgian corpus. We aim to create the best available BERT-style language model for Georgian. In this effort, the students will learn how to best utilise compute resources for optimal training.

5. 4. PARAMETER-EFFICIENT TRANSFER LEARNING IN ACTION RECOGNITION AND DETECTION

The motivation behind this subproject is the growing disparity between model and hardware sizes. Parameter efficient transfer learning (PETL) deals with this problem by making only parts of the model trainable while adding very few additional parameters. Transformers have shown impressive results with various PETL techniques, but there seemed to be something missing. We identify the missing piece of the puzzle, no PETL methods for the patch embedding layer in transformer based architectures. So, the aim here is to develop PETL methods for CNNs so that they can be optimized too, thus improving performance.

Secondly, using methods like adapters can be used to make bigger inputs possible to the model which improves detection performance. Gradients computed for the whole model generally take up a lot of compute memory and this can be reduced by making only certain parameters trainable. This applied to the task of emotion classification will allow a more holistic processing of datasets, than just a few seconds of input at a time, as is usual so far.

Based on these two, there is work going on to complete papers for each of the ideas.

5. 5. DIGITAL PHENOTYPING IN PSYCHIATRIC INTERACTIONS – A GEORGIAN STUDY PROTOCOL BASED ON MEPHESTO

The webinar sessions for the last subgroup have been less regular than for the other groups because this group was busy with the recording of patient-therapist interactions throughout a large period of the project. Nevertheless, whenever necessary, sessions were held with this group as well. This included, for example, an in-person meeting in Nice in January 2023 in order to demonstrate the typical MePheSTO recording setup to the Georgian researchers. Moreover, recently several sessions have been held with this group to discuss potential annotation schemes for the recordings.

At present, 9 videos of doctor/patient interviews are recorded and stored on the MICM server. Researchers of the subgroup and 3 volunteer medical students are working on annotations of the recordings.