The latest by AIDA Project

Fourth Edition of AIDA's Bi-Annual Newsletter

August 2022

Welcome to the fourth edition of AIDA's newsletter!

The AIDA project is now in its 24th month of implementation and already went through its first review and conducted its first Pilot. In these months, the consortium has achieved several milestones, as it has:

- Developed and tailored of AI tools and solutions;
- Analysed the legal environment relevant to the project;
- Conducted the analysis of citizens' security perceptions and impact of AI capabilities by LEAs on societal resilience;
- Developed pilot scenarios in line with the use cases (cybercrime and terrorism);
- Conducted its first round of testing through its first pilot;
- Defined the architecture of the AIDA system;
- Presented the project in numerous events and workshops.

AIDA will equip Law Enforcement Authorities in Europe with AI-enabled tools to prevent, detect, analyse and combat criminal activities.

In the upcoming months, partners will focus on:

- The final development of tools for data acquisition and for AI actionable intelligence;
- Conducting two additional pilots to test and train AIDA solutions;
- Prepare the close and report on the project, due to end in February 2023.

Update from AIDA's Coordinator

Dear readers,

After two years of intense and exciting work, AIDA is getting into its final phase. Crucial milestones have been accomplished and important events are being prepared at the Law Enforcement Agencies and EUROPOL premises and in synergy with different sister's projects.
The first round of pilots was successfully completed, and a refinement of the AIDA solutions is going to be performed before the next pilots' iteration already

planned for October, immediately after the training activities that will be kindly hosted by Guardia Civil. Two additional iterations will be performed in November and in January, with more details communicated on this topic in the next newsletter.

Outcomes of the project are actively being disseminated thanks to several invitations the consortium is receiving to participate workshops and scientific conferences. On the 7th of April 2022 AIDA had the pleasure of participating in the panel discussion on the "Protection of Public Spaces and Privacy" that took place during the “CERIS FCT Workshop on protection of public spaces” oraganised by DG-HOME in Brussels.

AIDA is also glad to announce its participation in the upcoming **EU Innovation Hub for Internal Security Annual Event, organised in cooperation with the CERIS community, on “Shaping responsible solutions for internal security”**. The event will take place on Tuesday 13 September 2022 in Brussels. The project will present some of its technical solutions in the “**project market place**” and will contribute to the panel discussion on “**Fundamental rights compliant use of data for internal security**”.

AIDA partner developing a multilingual tool using Neural Machine Translation (NMT)

Multilingual NMT for Slavic languages

AIDA partner Lingea attended the 2022 New Trends in Translation and Technology Conference, held in Rhodes in July 2022. The aim of the conference was to bring together academics in linguistics, translation studies, machine translation and natural language processing, as well as developers, practitioners, language service providers (LSPs) and vendors who work on or are interested in various aspects of technology for translation. The conference was set up as a distinctive and interdisciplinary event for discussing the latest developments and practices in translation technology. There, Lingea presented their continuous work on improving translation of Slavic languages using multilingual models. Traditionally, neural machine translation model is trained to translate languages in one direction (e.g. English->Czech). When translating tens or hundreds of languages between each other, this approach becomes computationally infeasible due to a large number of models that would have to be trained. Another issue is that most of the combinations actually have very little training data available. This challenge is usually solved using the so called "pivoted translation". This means that only systems for combinations of all languages with high-resource, pivot language – usually English – are trained and other combinations are translated through this intermediate language, via a two-step translation process. The disadvantage of this approach is the accumulation of errors and loss of information. Firstly, the accumulation of errors is caused by the imperfections of the translation systems, errors thus happen, and accumulate, in both the translations into and from pivot language. Secondly, the loss of information arises even with hypothetical perfect translation system. Indeed, it is possible that both source and target languages possess linguistic features (e.g. gender in adjectives) that do not exist in the pivot language and these are consequently lost in the pivot translation. This is especially true when translating between related languages, for example for Slavic ones.

Lingea's solution to counter this challenge is to employ multilingual models. These models are trained to translate from and into multiple languages jointly, so that they can share knowledge obtained from training data for one language with all the other languages. Lingea presented at the New Trends in Translation and Technology Conference, their experience with development and deployment of such models and with adaptation of these models to particular

domain, for example to the counter-terrorism domain, one of the AIDA's use cases.

Working together to combat criminal activities: AIDA collaborate with APPRAISE

APPRAISE is a Horizon 2020 EU funded project that started in September 2021 and will run until August 2023. The project's main focus is Cybercrime and Terrorism, by approaching specific issues and challenges related to soft target protection with a particular focus on active, audited, and well-defined information and intelligence exchange among private and public sector security practitioners to enable an effective collaboration, at the information and the operational levels.

Whilst AIDA's overarching goals differ, to develop a Big Data and Analytics framework equipped with effective, efficient and automated data mining and analytics solutions to facilitate Cybercrime and Terrorism investigations; the commonalities and challenges both projects are facing are very clear and provide an exciting opportunity to collaborate and gain new insights into potential challenges and threats in the Cybercrime and Cyber Terrorism spheres alongside the sharing of methods, best practices and technologies which may be used to tackle such issues.

Such co-dissemination activities will involve the reciprocal posting of events, training, and project information, the marketing of newsletters, attendance at workshops, conferences, etc. And the exchange of best practices, knowledge and information.

We very much look forward to working with APPRAISE!

AIDA is publishing!

AIDA aims to provide LEAs with innovative tools and technologies towards fighting cybercrime and terrorism, focusing on big data analysis, knowledge mining and fusion, graph analytics, visual analytics, legal, ethical, societal and data privacy practices. Within the year 2022, the research efforts of AIDA partners have started to bear fruits. Several research pieces have been published in peer-reviewed international conferences and journals.

The AIDA project ensures open access to peer-reviewed scientific publications relating to its results. Towards this direction, an AIDA Project Community was created in ZENODO. This community collection gathers open-access publications, as well as open datasets, related to the results of the project.

So far, the following works have been published in the AIDA community of ZENODO:

- "Leveraging Transformer Self Attention Encoder for Crisis Event Detection in Short Texts", conference paper by P. Kyriakidis et al. (CERTH);

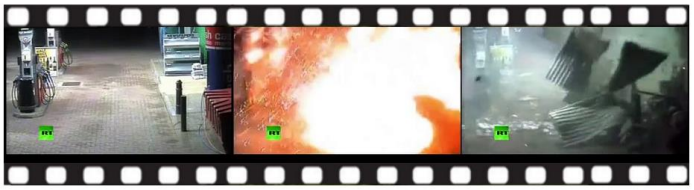
- "A Faster Converging Negative Sampling for the Graph Embedding Process in Community Detection and Link Prediction Tasks", conference paper by K. Loumponias et al. (CERTH);

- "PolygloNet: Multilingual Approach for Scene Text Recognition Without Language Constraints", conference paper by S. Gómez et al. (VICOM);

- "UCF-CAP dataset", dataset by C. Chatzikonstantinou et al. (CERTH); A video captioning dataset, extracted from the UCF-Crime dataset videos and described at the ICIP 2022 paper "UCF-CAP, video captioning in the wild".

More publications are expected within 2022; please check the ZENODO community for any updates!

An open-access video captioning dataset with crime-related data



Caption: an explosion is happening at a gas station

Caption: someone is shooting a house

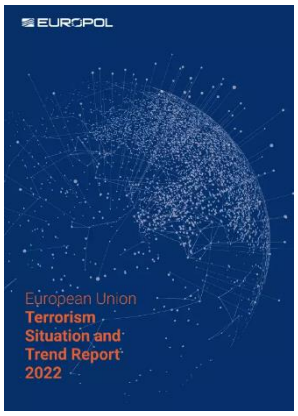
Caption: people are fighting at a room while three security guards interfere

Recent technological advances in Deep Learning (DL) have significantly improved the accuracy of automated tools for analysis of multi-source data. In the crime detection domain, it is of great importance to develop robust tools, based on Artificial Intelligence (AI), capable of detecting, preventing, and combating crime. To benefit from the DL paradigm for creating such tools, large volumes of annotated data are required, while the open-research approach must be followed to attract as many researchers as possible. Within the context of AIDA project, a new publicly available crime-related video captioning dataset has been introduced, aiming to fill the gap of appropriate video captioning data in the crime domain.

This dataset is based on content from the well-known UCF-crime dataset, which provides a set of normal videos and one of videos with anomalies. The UCF-CAP dataset provided by CERTH in the context of AIDA enhances a subset of UCF-crime videos with natural language descriptions, making them appropriate for the video captioning paradigm. The complete dataset contains captions for clips of 862 different videos, 2012 sentences and a vocabulary of 888 words. UCF-CAP

dataset is available for download in the ZENODO community of AIDA project via this link.

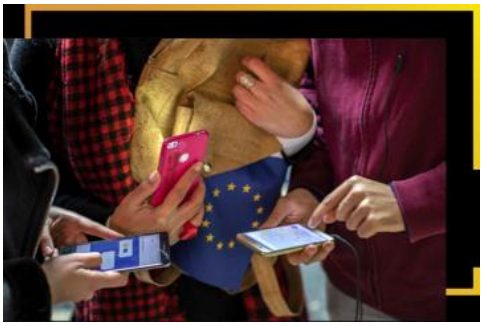
News and opportunities:



EU Terrorism Situation & Trend Report (TE-SAT) 2022 is out!

Europol's annual EU Terrorism Situation and Trend Report (TE-SAT) provides an overview of the terrorism phenomenon in the EU in a given year.

Read more



A Europe fit for digital age!

The European Commission is determined to make this Europe's "Digital Decade". Read more about the EU's digital strategy which aims to make this transformation work for people and businesses, while helping to achieve its target of a climate-neutral Europe by 2050.

Read more

IOCTA 2021 is out!

The 2021 Internet Organised Crime Threat assessment (IOCTA) outlines the features, facts, figures and trends concerning cybercrime in the EU in the past year.

Read more

Debunking common myths on the criminal use of cryptocurrencies

Europol's new report Cryptocurrencies: tracing the evolution of criminal finances offers insights into how criminals are using cryptocurrencies today.

Read more

This project has received funding from the EU's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 883596. The content of the publication herein is the sole responsibility of the publishers and it does not necessarily represent the views expressed by the European Commission or its services.

AIDA is coordinated by

Mr. Ernesto La Mattina

Engineering Ingegneria Informatica S.p.A.

Contact us at:

info-aida-project@eng.it

© 2020 AIDA project