

Al4Media Technological Highlights

Al4media

ARTIFICIAL INTELLIGENCE FOR

Discover AI4Media's key research outcomes on Human and Society centred AI



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 951911

Introduction:

The booklet showcases AI4Media's contributions to understanding the impact of artificial intelligence on both society as a whole and on individual users. These contributions are essential due to the widespread deployment and growing significance of AI technologies in everyday applications.

More information on these and other research outcomes of WP6 "Human- and Society-centred AI" can be found in relevant public deliverables available <u>here</u>.

Generative AI enables the creation of realistic multimedia content that media professionals can use to improve their production pipelines. However, the same technologies can be used by malevolent actors to misinform the public, and automatic detection methods are needed. Contributions to automatic content generation and detection are presented. The focus is on video and audio content because their automatic analysis is particularly challenging. Political news available in classical and social media outlets shapes our understanding of socially impactful and fuels democratic debates. End-users and media professionals benefit from a deep understanding of news. The booklet presents contributions addressing sentiment analysis, fallacy detection, bias quantification, and framing. These methods address multiple European languages and national contexts and enable cross-country comparisons.

Social media provides a convenient way to share personal data and interact with other users. Users have a right to understand the risks and opportunities entailed by these interactions. The proposed contributions target the understanding of users' mental states during the consumption of video content and the awareness of the effects of data sharing beyond the initial context. The developed tools are usable by different stakeholders, including media organisations, regulation bodies, end-users, social scientists, and political scientists.

The views and opinions set out in this document are those of the author(s) and do not necessarily reflect the official opinion of the European Union. Neither the European Union institutions and bodies nor any person acting on their behalf.

Synthetic image and video deepfake detection methods

Partner organisations involved CERTH, CNR, UNIPI, UNIFI, UPB \rightarrow Prototype

A few words about this technology

Given the complexity of the challenge of separating between authentic and fabricated content and the rapid rate of advances in the generative AI front, this technology encompasses a number of synthetic image detection and video deepfake detection methods, namely:

- A service for deepfake detection accompanied by a model card (<u>Baxevanakis et al., 2022</u>), a method for improving the cross-domain generalisation of synthetic image detection (<u>Dogoulis et al., 2023</u>) and a method for robust synthetic image detection by leveraging representations from intermediate encoder layers (Koutlis & Papadopoulos, 2024).
- A novel spatio-temporal deepfake detection system capable of effectively working in real-world contexts such as the presence of multiple subjects in the same scene or sudden changes in the distance of the face from the camera (Coccomini et al. 2022).
- Data augmentation and training improvements to improve the generalisation of deepfake detectors in a variety of real-world scenarios (Stanciu et al. 2023).

Who can benefit

Synthetic image detection and deepfake detection methods, once integrated in web services and tools, can support journalists, fact checkers and other kinds of online investigators in their verification tasks. Additionally, the methods themselves can benefit the research community and be used as state-of-the-art solutions against which new solutions can be benchmarked.

Impact and added value for the media industry

The recent emergence and rapid advances in generative models for images (e.g. Midjourney, DALL-E, Stable Diffusion) and videos (Sora, Runway ML) as well as AI-based video manipulation including face swapping,

DeepFake Detection

Use the prompt below to insert the url of an image or video. Our deepfake detection algorithm will process the media and return the probability that this media contains deepfake manipulated faces.

Contact: {olgapapa, gpan, papadop}@iti.gr

https://lh3.googleusercontent.com/8pjOT7bml0u2 Detection



reenactment and lip synching, make it extremely challenging, even for experts to confidently decide on the authenticity of images and videos that appear in digital platforms. Equipping news media and fact-checking organisations with robust tools for supporting the media verification process is therefore an increasingly important need that can guard their reputation and credibility. Given that credible and trustworthy news are essential for a well-informed society, it becomes clear that the value of media verification tools is also crucial for preserving democracy and social cohesion.

Future developments on the technology

Given the rapid advances in generative models, this technology is in need of constant updates, e.g. training new models with up-to-date datasets. Additional foreseen developments pertain to improvements in terms of efficiency, which is of particular importance when a large volume of video content needs to be processed in a short time, or in cases of live deepfake detection, e.g. in teleconferencing applications. Leveraging advances from AI on the edge could further enable the deployment of such models in end users' devices to ensure increased levels of privacy and content confidentiality.

ODSS: An Open Dataset of Synthetic Speech detection

Partner organisations involved FHG-IDMT, CERTH Paper & Dataset

A few words about this technology

The Open Dataset of Synthetic Speech (ODSS) is a new resource developed to assist research in synthetic speech detection. It comprises 18,993 audio utterances, evenly split between synthetic and natural speech, spanning approximately 17 hours of audio data. The dataset is characterised by its inclusion of three languages: English, German and Spanish, and it maintains an equal gender ratio across a total of 156 voices.

Furthermore, ODSS is freely available under a licence that encourages its extension and redistribution by the research community. Its development was driven by the twofold objective of promoting progress in synthetic speech generation technology and developing robust detection strategies to guard against potential misuse.

Who can benefit

The primary beneficiaries of this technology are researchers and developers working in the field of synthetic speech detection. By providing a comprehensive dataset that includes a wide range of voices, languages, and speaking styles, the ODSS facilitates the development and benchmarking of algorithms capable of distinguishing between synthetic and natural speech. Additionally, technology companies focusing on security and content authenticity verification can leverage this dataset to improve their solutions against misuse scenarios involving fraudulent or harmful synthetic speech content.

Impact and added value for the media industry

The rapid progression of synthetic speech technologies poses significant challenges in discerning authentic audio from manipulated content, a task vital for preserving the integrity of digital platforms. The ODSS dataset is GDPR-compliant and rich in linguistic diversity and vocal varieties, facilitating the creation of sophisticated detection algorithms that respect user privacy and data protection principles. By adhering to GDPR standards, ODSS underscores the commitment to ensuring that efforts to authenticate audio content do not compromise user privacy. These technologies support new media and fact-checking organisations in the audio verification process, streamlining the curation of information and ensuring the integrity of media content. In doing so, they play a pivotal role in preserving the media's reputation and credibility.

Future developments on the technology

As generative AI and novel synthetic speech algorithms continue to emerge, the ODSS dataset holds substantial potential for significant evolution. This could materialise through the development of a second iteration enriched with these new technologies or its integration into entirely new datasets. Incorporating more data samples that cover an extensive range of voices, dialects, languages, and audio produced by generative AI will significantly enhance the dataset's effectiveness. These enhancements are set to sustain the Open Dataset of Synthetic Speech (ODSS) as a key resource for advancing research and development in synthetic speech detection, continuing to provide an open-source, GDPR-compliant dataset.



Playable Environments: Video Manipulation in Space and Time

Partner organisations involved **UNITN**

ightarrow Paper & Dataset

A few words about this technology

Our tool, called Playable Environments, represents a new approach for interactive video generation and manipulation in space and time. With a single image at inference time, our framework allows the user to move objects in 3D while generating a video by providing a sequence of desired actions. The actions are learnt in an unsupervised manner. The camera can be controlled to get the desired viewpoint. Our method builds an environment state for each frame, which can be manipulated by our proposed action module and decoded back to the image space with volumetric rendering. To support diverse appearances of objects, we extend neural radiance fields with style-based modulation. Our method trains on a collection of various monocular videos requiring only the estimated camera parameters and 2D object locations. To set a challenging benchmark, we introduce two large scale video datasets with significant camera movements. As evidenced by our experiments, playable environments enable several creative applications not attainable by prior video synthesis works, including playable 3D video generation, stylization and manipulation.

Who can benefit

The tool could be useful to all the users (both professionals and amateurs) that are interested in developing technology which uses the domain knowledge (e.g., tennis) and wants to customise it for reusing. The tool enables multiple creative applications, such as 3D- and action-aware video editing, camera trajectory manipulation, changing the action sequence, the agents and their styles, or continuing the video in time, beyond the observed footage. We showcased the playable environment for tennis matches: the user specifies actions to move the players, controls the viewpoint and changes the style of the players and the field. The environment can be played, akin to a video game, but with real objects.



Impact and added value for the media industry

Since we are directly manipulating the existing footage the tool can be used by the media industry to support novel-view synthesis, action-aware video editing, and camera trajectory manipulation. The important aspect we introduce is the fact that the tool is allowing gamification but it uses real footage making the user experience to be closer to the reality.

Future developments on the technology

Representing games as the evolution of an environment's state driven by the actions of its agents is a powerful concept. However, while such a paradigm enables users to play a game action-by-action, its rigidity precludes more semantic forms of control. To overcome this limitation, we plan to augment game models with prompts specified as a set of natural language actions and desired states. The result-a Promptable Game Model (PGM)- will make it possible for a user to play the game by prompting it with high- and low-level action sequences. Most captivatingly, the PGM can enable the director's mode, where the game is played by specifying goals for the agents in the form of a prompt.

GreekPolitics: Sentiment Analysis on Greek Politically Charged Tweets

Partner organisations involved **AUTH**

Paper & Dataset

A few words about this technology

Automatic social media text content analysis is a rather underexplored topic, due to the scarcity of publicly available annotated datasets. Thus, we present and release "GreekPolitics", i.e., a dataset of Greek tweets with politically charged content. The tweets were independently annotated across four different sentiment dimensions: polarity, figurativeness, aggressiveness and bias. GreekPolitics has been evaluated comprehensively in a classification setting, separately for each sentiment, using state-of-the-art Deep Neural Networks (DNNs) and data augmentation methods. Based on the evaluation, best practices are also identified for achieving the highest classification accuracy at the DNN deployment stage.

Who can benefit

Scholars in the fields of natural language processing (NLP), sentiment analysis, and political science could leverage this dataset for studying sentiment analysis in the context of political discourse. They could explore techniques for sentiment analysis, figurativeness detection, aggressiveness identification, and bias detection in tweets.

Impact and added value for the media industry

Media companies could use sentiment analysis on Greek political tweets to gain insights into the attitudes, opinions, and sentiments of their audience regarding various political topics and figures efficiently. This understanding can inform content creation strategies, allowing media outlets to produce more engaging and relevant content tailored to their audience's preferences. By analysing sentiments expressed in Greek political tweets, especially with Neural techniques, media organisations can curate news articles, opinion pieces, and other content that resonates with the prevailing sentiment of the public automatically. This can help improve audience engagement and increase reader satisfaction by delivering content that aligns with their interests and viewpoints.

Future developments on the technology

Interesting future work could explore more sophisticated data augmentation for natural language text, which could ideally be applied without any human supervision.



Political Tweet Sentiment Analysis for Public Opinion Polling

Partner organisations involved **AUTH**

ightarrow Paper

A few words about this technology

Public opinion measurement through polling is a classical political analysis task, e.g. for predicting national and local election results. However, polls are expensive to run and their results may be biassed primarily due to improper population sampling. With this technology we offer two innovative methods for employing tweet sentiment analysis results for public opinion polling. The first method utilises merely the tweet sentiment analysis results outperforming multiple well-recognised methods. In addition, we introduce a novel hybrid way to estimate electorally results from both public opinion polls and tweets. This latter, enables more accurate, frequent, and inexpensive public opinion estimation and is used for estimating the result of the 2023 Greek national election. Our algorithm demonstrated lower deviation from the actual election's results than the conventional public opinion polls, introducing new possibilities for public opinion estimation using social media platforms.

Who can benefit

Scholars and researchers studying political behaviour and election dynamics could use this software to automatically analyse large datasets of tweets and identify patterns and correlations between sentiment and election outcomes. It could help advance our understanding of how social media influences political processes. Furthermore, Campaign managers and strategists could use such software to gauge public sentiment in real time, allowing them to adjust their messaging and campaign strategies accordingly. It could help them target undecided voters more effectively and tailor their messaging to resonate with specific demographics.

Impact and added value for the media industry

Media outlets can provide up-to-the-minute updates on public sentiment regarding political parties and candidates. This real-time reporting adds value by keeping audiences informed about evolving trends and dynamics in the political landscape. Additionally, journalists can use the insights generated by the software to enrich their reporting with data-driven analysis. This allows for more nuanced and informed coverage of political campaigns and elections, enhancing the credibility and depth of media content.

Furthermore, Coverage based on sentiment analysis of tweets can be more engaging for audiences, particularly in the digital realm where social media plays a significant role. Media organisations can leverage this content to spark discussions and interactions with their audience, increasing reader engagement and loyalty. Also by incorporating sentiment analysis into their reporting, media outlets can offer predictive analysis on potential election outcomes. This type of content can attract a wider audience interested in understanding the potential impact of sentiment trends on election results. Finally, media organisations that adopt this technology early can gain a competitive advantage by offering unique insights and analysis that set them apart from competitors. This can lead to increased audience trust and loyalty, as well as potentially attracting new audiences seeking innovative and data-driven reporting.

Future developments on the technology

Interesting future work could explore video and audio tweets, which are becoming increasingly popular but present unique analytical challenges. #6 Twitter COVID19 Discussions Topics dataset

Partner organisations involved **BSC**

 $ightarrow rac{ extsf{Paper}}{ o extsf{Dataset}}$

A few words about this technology

We collected tweets from the early COVID-19 pandemic period. Within this collection, we gather tweets belonging to a given thematic conversation (a topic) by defining a set of keywords specific to such a theme. We characterise these topics by looking solely at the volume and time of their activity. Under the hypothesis that the ephemerality of conversations is related with their quality, we formalise different measures and compute their values for the gathered topics. Our results indicate that ephemeral topics represent a big family of short-lived and burst-like discussions (unlikely to be informative), while non-ephemeral topics correspond to sustained, persistent and argued discussions that last on time (with potential to be informative and healthy).

The dataset contains tweet IDs and timestamps in CSV format. Contains a total of 3,423,260 entries between 10 topics and covers 2 time periods: March to August 2020 and August 2020 to March 2021. General information about the topics can be found in info.csv. All timestamps are in UNIX seconds.

Who can benefit

While the tweet contents cannot be shared due to Twitter's restrictive licence, their aggregated volume and topic mapping are made available. This can be useful for researchers and practitioners trying to understand the dynamics of social network discussions, enabling the definition of conversation quality indicators and healthy discourse recommenders.



(a) Burst topics cluster

(b) "Uniform" topics cluster

Illustration of the volume of activity over time for topics identified as ephemeral (left) and topics found non-ephemeral (right).

Impact and added value for the media industry

The developed ephemerality indicators of quality can help guide users and journalists towards discussions that have well built arguments, references and citations, producing more reliable news. At the same time, it can inform users journalists of the volatility of certain discussions, and the lack of quality in certain content, reducing their exposure to polarised and unreliable conversations.

Future developments on the technology

Improved indicators can be built on top of the proposed ones. Particularly, focusing on early indicators that can produce warnings in the early stages of ephemeral discussions.



#7 <u>News</u> framing method

 \rightarrow Paper

Partner organisations involved **IDIAP**

A few words about this technology

The identification of frames (perspectives that are emphasised) in journalistic texts has been a field widely studied in journalism. For example, given two news items on migration, one of them can emphasise the economic impact of migration on a country (economic frame), while the other shows the case of a migrant who almost lost his life arriving in Europe (human interest frame). What we propose is the use of LLMs, as assistants to help identify frames (human interest, economic, morality, attribution of responsibility, conflict or no frame), so that given a text and a definition of the different types of frames, the LLM is able to choose the one it considers most predominant in that text.

Who can benefit

The main beneficiaries are journalists, who have to spend a great deal of time analysing the frames with which news is presented. Publishers may also be interested, as they have an interest in understanding how their news is perceived by society, and the LLM can play the role of a reader. Finally, society at large may benefit if an assistant is able to tell you which frame is being emphasised and why.



Impact and added value for the media industry

The use of LLMs is revolutionising several sectors, including the media sector. This solution is in its very early stages, exploring the potential use with different types of texts. We started with headlines, and as a last step we have been working with TV transcripts. We have spoken with experts in the field of framing, and they all agree on the potential of the idea and the interest in moving this research forward.

Future developments on the technology

The next step would be to study different types of prompts, such as following a few shot learning approach, where some examples are given to help the model better understand the differences between the different types of frames.



#8 News analysis through an NLP framework

Partner organisations involved **IDIAP**

 \rightarrow Paper

A few words about this technology

Much of the technology developed in the media sector has focused on the issue of misinformation. The technology we propose aims to use different NLP techniques to give a holistic view of how information is presented in our case in the European written press on the topic of the anti-vaccine movement against the covid vaccination. But this can be extrapolated to other media or topics. Through techniques such as Topic modelling to identify topics, sentiment analysis to see if a more positive or negative tone is spread on a topic, Named ENtity Recognition, to identify the actors / organisations / places that are part of a topic and their relationship between them, as well as a qualitative analysis of the differences in the presentation of a news item on the same topic by different media with a different political orientation.

Who can benefit

The benefits of this technology are two:

- On the one hand the newspapers themselves, who can have a detailed analysis of how their newspaper presents the news regarding a certain topic.
- On the other hand, society in general, if it has access to this type of tools, would be a better informed society, and more critical with the content that it consumes.

Impact and added value for the media industry

With the latest developments in AI, the industry is becoming more open and less disbelieving of AI's potential as an assistant and task facilitator. We believe that these types of NLP techniques could be implemented relatively easily and would shed a lot of light for journalists themselves as to what topics they are covering, what is the predominant tone of their articles, or what people or institutions they name most often. In terms of society, we are talking about tools that would allow access to external and independent analysis, helping to achieve a more informed society.

Future developments on the technology

As a future work we could think about the creation of an interactive tool that would allow both newspapers and society to have access to all this type of analysis, without the need for technical knowledge that would limit their access as it currently happens, if not there would be a website or application that would run this type of analysis in the back-end and the user could directly access the analysis of the results.

Learning from Label Relationships in Human Affect

Partner organisations involved

 $\rightarrow \underline{Paper}$ $\rightarrow Code$

A few words about this technology

Human affect and mental state estimation in an automated manner face a number of difficulties, including learning from labels with poor or no temporal resolution, learning from few datasets with little data (often due to confidentiality constraints) and, (very) long, in-the-wild videos. For these reasons, deep learning methodologies tend to overfit, that is, arrive at latent representations with poor generalisation performance on the final regression task. To overcome this, we introduce two complementary contributions in this work. First, we introduce a novel relational loss for multilabel regression and ordinal problems that regularises learning and leads to better generalisation. The proposed loss uses label vector inter-relational information to learn better latent representations by aligning batch label distances to the distances in the latent feature space. Second, we utilise a two-stage attention architecture that estimates a target for each clip by using features from the neighbouring clips as temporal context.

Who can benefit

Industries and applications that involve human interaction, such as healthcare, education, customer service, and

market research, can benefit from this method. In healthcare, it could aid in the automated monitoring of patient well-being and mental health, which is one of the experimental setups of the method. In education, it could enhance personalised learning experiences by detecting student engagement and emotional states.

Impact and added value for the media industry

As more media companies aim to deliver personalised content that resonates with the user and is not intrusive, using such methods can better understand viewers' emotional responses and preferences. Furthermore, in film-making, directors and producers can gauge audience reactions during the production process. By analysing emotional responses, they can adjust storytelling techniques, pacing, and visual elements to optimise engagement. Finally, as media content becomes increasingly immersive and emotionally impactful, ethical considerations around content creation and consumption become more pronounced. The proposed method can help media companies identify potentially distressing or harmful content and implement safeguards to protect viewers, particularly vulnerable audiences such as children or individuals with mental health concerns.



Method Overview

Method, dataset, and prototype for raising user awareness about the effects of online photo sharing

Partner organisations involved **CEA, UPB**

 $ightarrow rac{\mathsf{Paper}}{\mathsf{Pataset}}$

A few words about this technology

Online social networks use AI techniques to automatically infer profiles from users' shared data. However, these inferences and their effects remain, to a large extent, opaque to the users themselves. We propose a method and a dataset that raise user awareness about the potential use of their profiles in impactful situations, such as searching for a job or an accommodation. These situations illustrate usage contexts that users might not have anticipated when deciding to share their data. User photographic profiles are described by automatic object detections in profile photos, and associated object ratings in situations. Human ratings of the profiles per situation are also available for training. These data are represented as graph structures which are fed into graph neural networks in order to learn how to automatically rate them. An adaptation of the learning procedure per situation is proposed since the same profile is likely to be interpreted differently, depending on the context. Automatic profile ratings are compared to one another in order to inform individual users of their standing with respect to others.

Who can benefit

The intended public of the app prototype is composed of social media users who share their photographic content online. They are made aware about the effects of sharing via tangible examples of how their data can be analysed using artificial intelligence technologies, and can affect



impactful situations associated with their daily lives. The dataset is shared with the research community, and particularly researchers working on data privacy, to facilitate the proposal of new methods assessing the impact of online data sharing.

Impact and added value for the media industry

User trust is needed to ensure a healthy long-term use of online social networks (OSNs). The proposed work contributes to increasing trust by unveiling the way AI technologies can be leveraged to analyse users' personal data. It can be seen as a practical implementation of GDPR requirements such as data minimization since users can better control what they share. The need for such feedback and control tools became more stringent with the advent of large multimodal models (LMMs), which facilitates the automatic analysis of users' personal data with unprecedented accuracy. Ideally, feedback tools should be integrated by OSNs to improve user awareness. However, the adoption of such tools by OSNs is improbable since they promote more responsible sharing practices. This would negatively affect existing business models that rely on massive data sharing by users. In practice, we implemented a standalone mobile app prototype that acts before they are actually shared. The app works locally, with photos being analysed on the user device and feedback about the effect of individual content and aggregated photo sets. This functioning is important to maximise the impact of the proposed feedback because once photos are shared, OSNs will process them and store the associated inferences even if content is subsequently removed.

Future developments on the technology

The future developments will be driven by generative AI technologies. The current results were obtained using a CNN-based object detection method that was trained with a limited set of objects. The wide availability of LMMs enables an automatic analysis of a much larger set of visual concepts with high accuracy. They can be leveraged to broaden the user profile representation. The existing method encodes four real-life situations to provide feedback. LMMs can be fine-tuned to enrich the set of modelled situations and make the awareness feedback more comprehensive.

Argument-based Detection and Classification of Fallacies in Political Debates

Partner organisations involved Université Côte d'Azur - 31A Côte d'Azur $ightarrow rac{ extsf{Paper}}{ o extsf{Paper}}$ Dataset & Code

A few words about this technology

Fallacies are arguments that employ faulty reasoning. Given their persuasive and seemingly valid nature, fallacious arguments are often used in political debates. Employing these misleading arguments in politics can have detrimental consequences for society, since they can lead to inaccurate conclusions and invalid inferences from the public opinion and the policymakers. In this work, we address the task of automatic detection and classification of fallacious arguments. First, we introduce a new annotated corpus, ElecDeb60To20, as an extension of theElecDeb60To16 dataset of U.S. presidential debates by incorporating the most recent Trump-Biden presidential debates. This resource includes 3 levels of token-level annotations: (i) argumentative components (i.e., claims and premises), (ii) the relations between these components (i.e., support and attack), and (iii) six categories of fallacious arguments (i.e., Ad Hominem, Appeal to Authority, Appeal to Emotion, False Cause, Slippery Slope, and Slogans). Second, we perform fallacy detection and classification by defining neural network architectures based on Transformers models, combining text, argumentative features, and engineered features. That is, the model processes raw text and identifies the specific sequences of words containing each type of fallacy.

Who can benefit

Various agents can benefit from this technology. Media organisations and journalists can use it to enhance the accuracy and credibility of their content. Social media platforms can integrate these methods to detect misleading political content. Policy makers and regulators can use these techniques to evaluate the integrity of political communications. Ultimately, these tools enable the general public to better discern between the arguments expressed by politicians, contributing to a more informed and critical electorate.

Impact and added value for the media industry

Automatically detecting and classifying fallacious arguments represents a crucial challenge to limit the spread of misleading or manipulative claims and promote a more informed and healthier political discourse.

By incorporating neural network architectures to identify fallacies in political debates, media organisations can enhance the accuracy and credibility of their reporting. For instance, this technology could help to identify fallacies in real-time during political debates, so that journalists and reporters can deliver more insightful and analytical coverage. This not only improves the quality of news content but also aids in educating the public about fallacious arguments, which leads to a more critical and informed audience. For social media platforms, integrating this technology helps in unveiling disinformation, manipulative content and harmful narratives online.

Overall, the adoption of such Al-driven tools in the media industry can help to create more accurate reporting, educate the audience, support the fight against disinformation, and contribute to the health of democratic processes.

Future developments on the technology

In future research, we intend to delve deeper into fallacious argumentation by integrating knowledge in order to address more challenging fallacy categories like causal ones, where reasoning and knowledge-based features are required to identify the fallacy. Our further objective is to generate valid arguments from identified fallacious ones and their context. Additionally, a challenge we aim to tackle is to explore ways to counter the formal invalidity of fallacious arguments through the generation of new arguments.



© Copyright 2024 Al4Media Consortium

This document may not be copied, reproduced, or modified in whole or in part for any purpose without written permission from the Al4Media Consortium. In addition to such written permission to copy, reproduce, or modify this document in whole or part, an acknowledgement of the authors of the document and all applicable portions of the copyright notice must be clearly referenced. All rights reserved.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 951911