EMBEDDIA

Cross-Lingual Embeddings for Less-Represented Languages in European News Media

Research and Innovation Action

Call: H2020-ICT-2018-1

Call topic: ICT-29-2018 A multilingual Next generation Internet

Project start: 1 January 2019   Project duration: 36 months

D6.12: Final report on ethics and responsible science and journalism (T6.5)

Executive summary

This document presents the measures taken to assure ethical and secure development of the EMBEDDIA project and outlines the mechanisms for internal ethics management in EMBEDDIA to assure responsible and ethical science in news media content analysis and creation. The report presents an overview of the legal developments regarding artificial intelligence tools in media and journalism in the European Union, and provides an insight into discussions and developments regarding regulation and ethical issues involving AI and the news media.

Partner in charge: UL

Deliverable information

Project acronym:	**EMBEDDIA**
Project number:	**825153**
Deliverable number:	**D6.12**
Deliverable full title:	**Final report on ethics and responsible science and journalism (T6.5)**
Deliverable short title:	**Final ethics report**
Document identifier:	**EMBEDDIA-D612-FinalEthicsReport-T65-submitted**
Lead partner short name:	**UL**
Report version:	**submitted**
Report submission date:	**31/12/2021**
Dissemination level:	**PU**
Nature:	**R = report**
Lead author(s):	**Marko Milosavljević (UL)**
Co-author(s):	**Romana Biljak Gerjevič (UL)**
Status:	**_ draft, _ final, x submitted**

The EMBEDDIA Consortium partner responsible for this deliverable has addressed all comments received. Changes to this document are detailed in the change log table below.

Change log

Date	Version number	Author/Editor	Summary of changes made
30/11/2021	v1.0	Marko Milosavljević (UL)	Draft report.
9/12/2021	v1.1	Matthew Purver (QMUL) & Hannu Toivonen (UH)	Internal review.
12/12/2021	v2.0	Marko Milosavljević (UL)	Changes implemented.
21/12/2021	v2.1	Nada Lavrač (JSI)	Quality control.
22/12/2021	v3.0	Senja Pollak (JSI), Mattehw Purver (QMUL), Nada Lavrač (JSI), Romana Biljak Gerjevič (UL)	Implementing quality control changes and revising sections on news and comments.
23/12/2021	v4.0	Urša Chitrakar	Legal review.
24/12/2021	v4.1	Marko Pranjić (Trikoder)	Added text for 24 sata.
24/12/2021	final	Marko Milosavljević (UL)	Report finalized.
27/12/2021	submitted	Tina Anžič (JSI)	Report submitted.

Table of contents

List of abbreviations

AI	Artificial Intelligence
EC	European Commission
EU	European Union
DoA	Description of Action
GDPR	General Data Protection Regulation
IPR	Intellectual Property Rights
NLG	Natural Language Generation
NLP	Natural Language Processing
OECD	Organisation for Economic Co-operation and Development
OSCE	Organization for Security and Co-operation in Europe
UNESCO	United Nations Educational, Scientific and Cultural Organization
WP	Work Package

1 Introduction

This deliverable reports on the results of the work performed within Task T6.5: "Towards responsible and ethical science and content creation", which focused on the need for more transparent, accountable and responsible news journalism, with ethical and moral attitudes at its core. As detailed in the Description of Action (DoA), EMBEDDIA is ensuring that the tools developed not only fit within a legal framework (e.g., regarding protection of data and user integrity) but are used in ways that do not break trust.

This deliverable outlines the measures taken to assure ethical and secure development of the EMBEDDIA project, as defined in Deliverable D6.2 "Project's internal ethics policy (T6.5)", submitted at M6, and in Deliverable D6.6 "Interim report on ethics and responsible science and journalism (T6.5)", submitted at M18.

Throughout the project duration, the aim of Task 6.5 was to monitor the measures and recommendations for responsible AI development and overview the status of new tools development and usage in news analysis and production. To this end, we monitored issues concerning ethical development of AI and automation tools for news media content analysis and creation, together with the latest developments in the use of such tools in the global media and legislation novelties at the EU and national levels.

Data management, user comment analysis, news analysis and news generation were identified as the main areas where ethical problems, related to the protection of privacy, as well as problems related to the use of AI automation tools in the media, can occur. To address them we either relied on existing legislation or developed our own policies to be followed by the project partners. For privacy related issues we drew our solutions based on the General Data Protection Regulation 2016/679 (hereinafter GDPR)[1], while the AI-related ethical dilemmas were only recently partly addressed by the proposal of the Artificial Intelligence Act (COM/2021/206 final), and therefore our policies were built from experiences during the course of the project, the internal policies previously followed by our media partners and the previous recommendations provided by the European Commission.

We have regularly informed the EMBEDDIA partners on these developments and their impacts on the EMBEDDIA project activities at all of our project meetings. Moreover, our communication activities targeted also the general public and prospective end users of EMBEDDIA tools via social media and the EMBEDDIA project newsletter.

Throughout the project we have held at least ten successful meetings about ethical use of AI in journalism and discussed views and solutions to possible ethical issues with relevant stakeholders, including media editors and managers from the BBC (March 2019), 24sata and Večernji list (November 2019), Radio France, L'Express and Liberation (January 2020). Discussions were held to gain additional information on user needs in different newsrooms and media companies, to gain insights about the state of development of similar media tools in large media companies and to establish relevant connections (which was also done on international conferences and events, presented in Section 3 of this document).

In this deliverable, we first present the general ethics framework of EMBEDDIA and the security measures taken to protect the privacy of users and ethical tool development. This is followed by describing challenges and measures taken to answer any open questions. Finally, we present an overview of recent legislative developments in the European Union regarding the use and development of Artificial Intelligence tools, as well as discussions and recommendations proposed by international organisations and other key stakeholders.

---

[1] While some relevant EU directives as obligatory regulations (for example the GDPR) have already been adopted within national legislation and thus implemented in all partner countries, some directives (such as the Digital Single Market Copyright Directive) have during the duration of project not been fully implemented in all partner countries.

2 Ethics framework, challenges and measures

The EMBEDDIA technology aims at advancing cross-lingual natural language processing (NLP) and natural language generation (NLG) using only anonymous data, by which subjects cannot be identified (not any specific identification of one unique person, such as name, social security numbers, date of birth, address, email address, IPs). During our work we did not collect or process any personal data and the consortium took all the necessary measures to avoid any possible user identification.

The following measures were taken:

- clear procedures and the data management plan (D8.3),
- risk monitoring and quality assurance guidelines with risk mitigation measures (D8.4),
- no access to user identifier information,
- anonymisation process before the non-media EMBEDDIA project partners see the data,
- releasing datasets only upon signed permission statements by media partners (using mostly CC BY-NC-ND license for text data),
- optimizing the correctness and appropriateness of generated texts.

These measures, which are described in detail below, were developed together with the Legal advisor of Jožef Stefan Institute (JSI), as well as in consultation with the External ethics and legal advisor. They collaborated with the project coordinator, the leader of Task T6.5 (on responsible and ethical science and content creation), the data manager and the Steering Committee members. The role of the EMBEDDIA External ethics and legal advisor was assigned to Urša Chitrakar, an attorney specialized in copyright and intellectual property.

The initial ethics-related deliverable D6.2, as well as the current D6.12, were approved by External ethics and legal advisor. When preparing the ethics procedures described in D6.2, we took useful advice also from Luka Virag from the JSI law department.

GDPR defines personal data as "any information relating to an identified or identifiable natural person". An identifiable natural person can be identified directly or indirectly by a "name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person," as described in Article 4 of the regulation[2].

This section presents the general ethics framework of EMBEDDIA, which was defined in Deliverable D6.6 "Interim report on ethics and responsible science and journalism (T6.5)", but not yet fully implemented as the EMBEDDIA tools were still under development. This section outlines the main ethical measures taken throughout the project (and during the release of datasets) and provides further clarification on the points where privacy of users was noted as possibly raising concern. We examine these in turn for user comment analysis in WP3, news analysis in WP4, and news generation in WP5. We also address the potential legal issues and questions regarding user comments, news articles, their aggregation and the implications of the EU Copyright Directive.

2.1 Data management

The work in WP3, WP4 and WP5 depended on datasets obtained from news media providers. These datasets did not contain any personal information of the users nor any other identifiable information. For the textual datasets owned by the media providers or gathered by the EMBEDDIA researchers, clear procedures were established and the data management plan has been proposed (see Deliverable D8.3: Data Management Plan). The information for which the individual parties were responsible was shared only in a manner clearly defined by the Consortium Agreement. Moreover, written consent was obtained from media partners before releasing data publicly. Please note that unethical data management was described also as a potential risk for WP3, WP5 and WP8 in Deliverable D8.4: "Risk monitoring and quality assurance guidelines", where the appropriate risk mitigation measures have been proposed. All data acquisition processes were in compliance with the GDPR regulations. Moreover, the project did not involve activities or results raising any potential security issues, as the project did not involve "EU-classified information" as background or results.

---

2 https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679&from=EN

2.2 User comment analysis

In WP3 the objective was to analyse user-generated comments, to produce tools for topic modelling, conversation structure and context modelling, sentiment, stance and opinion detection, and detection of hate speech, political trolling and attempts to elicit extreme reactions and influence others' opinions. We also developed and implemented methods for generating human-readable reports in multiple languages. Through advanced cross-lingual context and opinion analysis, developing cross-lingual comment filtering and techniques for report generation from multilingual comments we tried to devise the means to understand the reactions of multilingual news audiences to help news media companies better serve their audiences.

In EMBEDDIA we analysed comments provided by Ekspress Meedia (ExM) and comments from 24sata provided by Trikoder, but did not deal with comments from the Finnish partner STT as they have a separate moderation service for media publishers and other organisations, which were not involved in the project.

User-generated comments were in some cases anonymous in origin (no information about the identity of users was ever collected), and in other cases comments were published using some usernames registered on the media partner's publishing platform. In both cases, the default policy (see below for more detail) was that EMBEDDIA researchers never got access to user identifier information: author usernames were removed before distribution by the EMBEDDIA media partners, leaving no direct way that individuals could be identified by researchers or anybody else (however, by searching comments text on the web, the public could discover the public username; but this could not be prevented, did not reveal any personal data beyond the user's comments themselves, and did not constitute personal data collection and processing). Specifically, we have publicly released user comment datasets from EMBEDDIA partners Ekspress and 24sata, and used the anonymised versions.

Nevertheless, there were open issues regarding the distribution of comments, for example when the datasets were to be released publicly, concerning a potential need for additional comment anonymisation. There were no such issues with anonymous comments, as the researchers never accessed the users' personal data. Even with non-anonymous comments, usernames already had public status as they were available on the original publicly accessible news site. However, to assure maximum ethical standards and security, the advice to the media partners was to remove usernames (and, if possible, mentions to original usernames in the text) and replace them with randomly assigned identifiers. This will prevent any chance of later identification from the data, even in cases where usernames may be associated with some external information (e.g., Facebook authentication data, where used), ensuring that there was no direct way for future dataset users to trace the comments back to their original authors[3].

For maximum security, our default policy was therefore that the anonymisation process should happen even before the non-media EMBEDDIA project partners see the data. Namely, if the consortium partners never receive personal data, this removes the risk of any member of the team accidentally releasing a non-anonymised version. However, in some cases this made certain kinds of analysis impossible: for example, the use of usernames in the body of comments proved to be key to understanding to whom a comment is addressed, and therefore what opinion it expresses. In one case, a version of the 24sata dataset with the original user identifiers was released internally for use by other non-media EMBEDDIA partners, on the condition that the data was confirmed not to allow direct individual identification (e.g., it should use system user IDs, rather than real names); and that version was not released publicly.

Possible issues could also arise in the case of occurrences of someone identifying him/herself as an author of a comment that was labelled as an example of hate speech. Potential appearance of personal names within news articles and/or comments is not problematic from the point of view of protection of personal data. Personal data legislation applies only to structured sets of personal data, which is not the case in the situation at hand (within articles/comments, we mostly see only incidental appearances of personal data, while the repository of such articles does not amount to a structured set of personal

---

[3] Indirect methods, for example estimating a user's identity on the basis of their opinions and the topics they talk about, may be available in some cases, but cannot fully prevent and are possible even with strictly anonymous comments.

data). But any request for deletion will be respected in case of need, and corresponding items will be removed from publicly released data.

Another possible ethical concern appeared in the handling and distribution of comments blocked by the moderators.

- Moderators from 24sata deleted for example: content that breaks the law, defamatory allegations, sensitive privacy information, threats to an individual or group, discriminatory or racist content, incitement to commit illegal activities, personal data of victims of crime and accidents and their relatives, other personal and contact information, inappropriate links or link lists, bad or disruptive behaviour, trolling, external online surveys, competitions or similar and content that violates copyrights.
- For Ekspress, those cases mostly included defamation, threats, insults, comments that stimulate hatred or violence or other undesired actions, including false information.

In these cases, the comments could also contain illegal content banned by national laws, for which public distribution may not be possible. No such content was identified, but if such content is detected in the future, we will remove it from the dataset, if needed.

2.3 News analysis

In WP4 the focus was on news articles, where the main potential issue involved copyright (concerning the possible publications of datasets and the outputs of the analysis technology). As regards the new EU Copyright Directive, the exception enacted in Article 3 allows text and data mining without obtaining copyrights from their respective holders. This however brought up the question of the legal basis for such conduct in countries where the directive was not implemented yet. In the case news articles were reproduced/copied for the purpose of implementing the project tasks, even though they were not published in any manner, corresponding permissions needed to be acquired.

The right of publishers on the publication they released, is a novelty, which does not apply to non-commercial use carried out by individuals and use of individual words or very short extracts. None of these exceptions is however applicable in the case of reproduction/re-publishing of entire articles by an institution. An additional right will therefore have to be acquired by the press publishers (for a possible additional remuneration), along with the copyright. EU member states had to implement the provisions of the directive into the national legislation by June 2021, but not all succeeded in that.

In addition to the EU regulation, there is the national law. According to Article 48 of the Slovenian Copyright and Related Rights Act, it is allowed to prepare and reproduce abstracts of published articles in the form of press reviews without obtaining any permissions. And this goes also for electronic publications. The exemption applies only to reproductions made with an intent of "informing the public" and only, if the source and authorship of the work are duly indicated.

Considering intellectual property rights, the decision was that no datasets are released by research partners without involvement of media partners. Either media partners released the data themselves, or provided written permissions and were involved in the data release. We also advised the media partners to add a licence under which the dataset could be used when they release this dataset. This opened the question of the type of the licence ideally ensuring that the data could be further used for research purposes. Our default policy was to suggest the CC-BY-SA licence for the text data, since it allows distribution and re-use for research, while ensuring that data sources and creators are appropriately attributed. However, finally we opted for CC BY-NC-ND which was more acceptable to the media partners, as it prevents commercial use and distribution of derived works.

The news article datasets were acquired from all three media partners:
- in the case of STT, the media partner directly publicly released the dataset;
- while with Ekspress and 24sata, news article datasets were released with permission and under CC-BY licences (mostly CC BY-NC-ND), upon signed agreements.

As with comments, the personal information contained in an article may not be definite enough for the rules on personal data protection to apply – merely a name and surname do not necessarily suffice for the general public to fully (and without additional effort) identify the individual in question. Nevertheless,

according to our policy, articles containing personal data will be excluded from the database upon request of the individual such data relates to.

2.4 News generation

WP5 focused on news articles generation. During EMBEDDIA, news automation systems that were transferable across languages, domains, and transparent in the natural language generation (NLG) process, were being developed with a self-explainable, flexible and accurate NLG system architecture that can be transferred to new domains and languages with a minimal human effort, tools for creation of dynamically evolving content, incorporating narrative structure and user knowledge and tools for creation of figurative language and headlines.

The content of all generated texts needs to be correct and appropriate, which is evaluated by assessing, if the text is factually correct and its content fully derivable from data or reference texts. That is one of the reasons, why in majority of the WP5 work focused on template-based and not neural network-based approaches.

Next, we consider that generated texts need to be marked as such to assure transparency for the audiences, however no generations were integrated in the media partners' workflows at this point.

Last but not least, we focused also on the gender bias within the generated texts, which is further explained in Deliverable D6.11 "Final report on gender bias in content creation" (T6.4).

3 Overview of EU developments concerning ethics and AI

In this section, we provide an overview of legislative developments and recommendations in the European Union, including standards and policies proposed by key professional institutions and organisations.

The EMBEDDIA media partners were following GDPR and related national regulations, including their own codes of ethics. The Finnish partner STT has a separate moderation service for media publishers and other organisations. Their moderators work solely for the moderation service, separately from the agency's newsroom. While using various technological/AI solutions to help them screen comments and detect hate speech and other inappropriate content, STT moderators follow the Finnish legislation and STT's moderation code, which contains the main principles,[4] as well as moderation policies of each individual media/organization client. The Estonian partner Ekspress firstly follows the common legislation like The Constitution of Estonia (põhiseadus) and Law of Obligations Act (võlaõigusseadus), which governs protection from defamation or other personality rights, plus the compensations for such violations. It also follows the more specific self-regulation guidelines which they have created for their commentators (for commenting in the comments section) and for their moderators (for moderating them).

The Croatian partner Trikoder is the technology provider for the Styria Media Group. The most relevant law when handling the data for the members of Styria Media Group is the Law on the Implementation of the General Data Protection Regulation (Zakon o provedbi Opće uredbe o zaštiti podataka) that implements the EU 2016/679 directive. Media companies within the Styria Media Group are, in addition to the aforementioned law, also required to uphold the Digital Media Act (Zakon o elektroničkim medijima) that includes the 98/84/EZ directive, 2006/114/EZ directive, and the 2010/13/EU directive. The media company 24sata also has a specific Terms and Conditions (Uvjeti korištenja) with a specific User Generated Data (Sadržaj kreiran od strane korisnika) section that describes and categorizes the inappropriate content in the user submitted data (comments) and prescribes sanctions for the users that do not follow those guidelines.

---

[4] https://stt.fi/tyylikirja/moderointi/moderointiperiaatteet/

As described in Section 3 of Deliverable D6.6: "Interim report on ethics and responsible science and journalism (T6.5)", the EU High-Level Expert Group on AI presented Ethics Guidelines for Trustworthy Artificial Intelligence[5] in April 2019, stating that trustworthy AI should be lawful (respecting all applicable laws and regulations), ethical (respecting ethical principles and values) and robust (both from a technical perspective while taking into account its social environment).

The guidelines were followed by two main documents, which the European Commission presented in February 2020: the European data strategy for policy measures and investments to enable the data economy for the coming five years[6] and the White Paper on Artificial Intelligence – A European approach to excellence and trust[7]. The latter is setting out a policy framework/policy options following two objectives: to promote the uptake of AI and to address the risks associated with certain uses of this new technology.

In April 2021, the Commission published its AI package, proposing new rules and actions for trustworthy AI in EU member states. The package consisted of a Communication on Fostering a European Approach to Artificial Intelligence[8], the Updated Coordinated Plan with Member States[9] and a proposal for an AI regulation[10], known as the Artificial Intelligence Act.

The coordinated plan introduced four key sets of proposals for the European Union and the member states to be able to accelerate, act and align opportunities of AI technologies and to facilitate the European approach to AI:
-   set enabling conditions for AI development and uptake in the EU (acquire, pool and share policy insights, tap into the potential of data, foster critical computing capacity),
-   make the EU the place where excellence thrives from the lab to the market (collaborate with stakeholders through, e.g., the European Partnership on AI, Data and Robotics and expert groups, build and mobilise research capacities, provide an environment for developers to test and experiment (TEFs), and for SMEs and public administrations to take up AI (EDIH), support the funding and scaling of innovative AI ideas and solutions),
-   ensure that AI works for people and is a force for good in society (nurture talent and improve the supply of skills necessary to enable a thriving AI eco-system, develop a policy framework to ensure trust in AI systems, promote the EU vision on sustainable and trustworthy AI in the world),
-   build strategic leadership in high-impact sectors (bring AI into play for climate and environment, use the next generation of AI to improve health, maintain Europe's lead: Strategy for Robotics in the world of AI, make the public sector a trailblazer for using AI, apply AI to law enforcement, migration and asylum, make mobility safer and less polluting through AI, support AI for sustainable agriculture).

The regulative proposal is based on four different levels of risk: unacceptable, high, limited and minimal risk, and focusses on a human centred approach. The proposal responds to requests by the European Parliament and the European Council for legislative action to ensure a well-functioning market for artificial intelligence systems, where both benefits and risks of AI are adequately addressed. The Commission stresses that maximising resources and coordinating investments are critical components of the wholesome AI strategy, which is why it plans to invest one billion Euros per year in AI through the Digital Europe and Horizon Europe programmes.

The Council of Europe urged EU member states to "promote experimentation with, and investment in AI-driven tools" in their report titled "Implications of AI-driven tools in the media for freedom of expression[11]", published in February 2020. The report questions a number of aspects of AI tools in the newsroom and includes recommendations for ethical and responsible use of AI-driven tools in relation to news media, society and users, stating "the media have a responsibility to use AI-driven tools in a way that is conducive to the fundamental freedoms and values that characterise European media

---

[5] https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai
[6] https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1593073685620&uri=CELEX:52020DC0066
[7] https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf
[8] https://digital-strategy.ec.europa.eu/en/library/communication-fostering-european-approach-artificial-intelligence
[9] https://digital-strategy.ec.europa.eu/en/library/coordinated-plan-artificial-intelligence-2021-review
[10] https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1623335154975&uri=CELEX%3A52021PC0206
[11] https://rm.coe.int/cyprus-2020-ai-and-freedom-of-expression/168097fa82

markets and policies." Before that, the council also published a study on responsibility and AI in September 2019 regarding "the implications of advanced digital technologies (including AI systems) for the concept of responsibility within a human rights framework"[12].

A member of the EMBEDDIA consortium Marko Milosavljević was also a member of the Committee of Experts on Media Environment and Reform (MSI-REF) at the Council of Europe. The committee members had been tasked to prepare three documents: the Draft recommendation on principles for media and communication governance, the Draft recommendation on electoral communication and media coverage of election campaigns and the Draft guidance note on the prioritisation of public interest online. The members of the committee approved the draft documents in September 2021[13] and the Committee on Media and Information Society adopted them in the beginning of December 2021. The Draft recommendation on principles for media and communication governance recommends the governments of EU member States to review their legislative frameworks, policies and practices with respect to the following fifteen principles set out in the recommendation:
- transparency and accountability,
- openness and inclusiveness,
- independence and impartiality,
- evidence-based and impact-oriented governance choice,
- agility and flexibility,
- promoting human rights and fundamental freedoms in communication,
- securing media freedom,
- promoting media pluralism and safeguarding the sustainability of journalism,
- ensuring transparency of content production,
- ensuring compliance with content obligations and professional standards,
- ensuring functioning markets and protecting personal data in content dissemination,
- responding to the risks caused by platforms disseminating illegal or otherwise harmful content,
- mitigating the risks posed by algorithmic curation, selection and prioritisation,
- guaranteeing human rights and fundamental freedoms in media and platform use,
- empowering users and fostering responsible use.

In March 2020, the Organization for Security and Co-operation in Europe (OSCE) published a "Non-paper on the impact of Artificial Intelligence on Freedom of Expression[14]", discussing the use of AI within content moderation and curation, content ranking on social media platforms based on commercial interests, challenges posed to media pluralism and content diversity and possible risks to freedom of expression. The non-paper pinpoints crucial areas to focus on in the future, including promoting a better understanding of AI practices (by state and non-state actors), initiating a dialogue with industry and states, developing recommendations to mitigate the negative impacts of automated tools and prevent the infringement of free speech and media freedom, measuring the impact of legislation and mapping out the current use of machine-learning technologies.

In July 2020, OSCE published a Strategy Paper to Put a Spotlight on Artificial Intelligence and Freedom of Expression[15], mapping out the key challenges to freedom of expression presented by AI across the OSCE region in light of international and regional standards on human rights and AI. In December 2020, OSCE published a Policy paper on freedom of the media and Artificial Intelligence[16], addressing the impact of Artificial Intelligence on freedom of expression and media freedom.

We discussed these issues and the EMBEDDIA ethics policy broadly at a number of international events, lectures and conferences, including the participation of Marko Milosavljević, the member of the EMBEDDIA team, at the following events:
- the conference organised by the Council of Europe and Slovenian Ministry of Culture in November 2019 titled "(Last) Call for Quality Journalism", with a joint panel discussion with the member of our team and the Head of Facebook for Russia and Central and Eastern Europe Gabriella Cseh,

---

[12] https://rm.coe.int/a-study-of-the-implications-of-advanced-digital-technologies-including/168096bdab
[13] https://www.coe.int/en/web/freedom-expression/msi-ref/-/most_viewed_assets/JqA1MXiumoTd/content/4th-meeting-of-the-committee-of-experts-on-media-environment-and-reform-msi-ref-
[14] https://www.osce.org/representative-on-freedom-of-media/447829
[15] https://www.osce.org/representative-on-freedom-of-media/456319
[16] https://www.osce.org/representative-on-freedom-of-media/472488

- a public lecture at Ryerson University in Toronto, Canada in May 2019, titled "Between Euphoria and Dystopia: AI, journalism and perceptions in leading newsrooms",
- Council of Europe Committee of experts on media environment and reform (MSI-REF) first meeting in September 2020 and second meeting in January 2021, where we discussed these issues with a varied group of stakeholders, such as media experts, academics, regulatory bodies, politicians and different media and professional organisation (for example Reporters Without Borders),
- workshop on European Media and Platform Policies by EuroMedia Group and Jean Monet Network in November 2020,
- plenary session of EPRA (European Platform of Regulatory Authorities) in January 2021,
- panel discussion titled Role of the Academy in Strengthening the Development of Media Literacy at an expert workshop titled Combating Fake News and Disinformation Campaigns organised by the Croatian Ministry of Defense as it took over the Presidency of the Central European Defense Cooperation (CEDC) in June 2021,
- Council of Europe Ministerial Conference: Artificial Intelligence - Intelligent Politics: Challenges and Opportunities for Media and Democracy in June 2021,
- OSCE Expert Workshop on Security in June 2021,
- discussion titled Regulating Platforms, Building Trust on the 2021 ECREA conference titled Communication and Trust: Building Safe, Sustainable and Promising Futures in September 2021,
- ECREA international scientific conference Comments, Hate Speech, Disinformation and Public Communication Regulation in Zagreb, Croatia in September 2021.

4 Conclusions and further work

This deliverable outlines the measures taken to assure ethical and secure development of the EMBEDDIA project and the mechanisms for internal ethics management in EMBEDDIA to assure responsible and ethical science in news media content analysis and creation. It provides identification of areas, where the privacy of users could present a potential concern, i.e. user comments analysis, news analysis and news generation, and explains how these concerns were addressed throughout duration of the project.

Moreover, it summarizes the legislation providing some answers to the privacy related challenges. The report also presents an overview of the legal developments regarding artificial intelligence tools in the European Union and recommendations proposed by significant international organisations, such as the Council of Europe and OSCE.

Despite the fact that the present deliverable D6.12 is the final deliverable for task T6.5, we will continue, until the end of the project, to address the challenges and issues of the specific news media environment and digital news production in contextual matters, the processes of innovation in online journalism, changes in journalistic practice, professional and occupational matters, the role of the audience and user-created content and democratic processes, and quality news content as a prevention of inflammatory and dangerous speech.

5 Associated outputs

Parts of the work performed in task T6.5: Towards responsible and ethical science and content creation, are also described in three publications, but all were included already in the previous deliverable (D6.6), therefore no new publications are listed here.

References

Council of Europe. (2019). Responsibility and AI. Retrieved from: https://edoc.coe.int/fr/intelligence-artificielle/8026-responsibility-and-ai.html.

Council of Europe. (2020). Implications of AI-driven tools in the media for freedom of expression. Retrieved from: https://rm.coe.int/a-study-of-the-implications-of-advanced-digital-technologies-including/168096bdab.

Council of Europe. (2021). Committee of Experts on Media Environment and Reform (MSI-REF). Retrieved from: https://www.coe.int/en/web/freedom-expression/msi-ref/-/most_viewed_assets/JqA1MXiumoTd/content/4th-meeting-of-the-committee-of-experts-on-media-environment-and-reform-msi-ref-.

European Commission. (2019). Ethics guidelines for trustworthy AI. Retrieved from: https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai.

European Commission. (2020). A European strategy for data. Retrieved from: https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1593073685620&uri=CELEX:52020DC0066.

European Commission. (2020). White Paper: On Artificial Intelligence - A European approach to excellence and trust. Retrieved from: https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf.

European Commission. (2021). Communication on Fostering a European Approach to Artificial Intelligence. Retrieved from: https://digital-strategy.ec.europa.eu/en/library/communication-fostering-european-approach-artificial-intelligence.

European Commission. (2021). Coordinated Plan on Artificial Intelligence 2021 Review. Retrieved from: https://digital-strategy.ec.europa.eu/en/library/coordinated-plan-artificial-intelligence-2021-review.

European Commission. (2021). Proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain union legislative acts. Retrieved from: https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1623335154975&uri=CELEX%3A52021PC0206.

EU General Data Protection Regulation. (2016). *Regulation (EU) 2016/679 of the European parliament and of the council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing directive 95/46/EC (General data protection regulation).* Official Journal of the European Union, 119. Retrieved from http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679&from=EN.

Milosavljević, M., Vobič, I. (2019). "Our task is to demystify fears": analysing news-room management of automation in journalism. Journalism. Retrieved from: https://repozitorij.uni-lj.si/IzpisGradiva.php?id=114945.

Milosavljević, M., Poler Kovačič, M., & Čeferin, R. (2020). In the name of the right to be forgotten: new legal and policy issues and practices regarding unpublishing requests in Slovenian online news media. Digital journalism. Retrieved from: https://repozitorij.uni-lj.si/IzpisGradiva.php?id=116957.

Organization for Security and Co-operation in Europe - OSCE. (2020). Non-paper on the impact of artificial intelligence on freedom of expression. Retrieved from: https://www.osce.org/representative-on-freedom-of-media/447829.

Organization for Security and Co-operation in Europe - OSCE. (2020). Policy paper on freedom of the media and artificial intelligence. Retrieved from: https://www.osce.org/representative-on-freedom-of-media/472488.

Organization for Security and Co-operation in Europe - OSCE. (2020). Strategy Paper to Put a Spotlight on Artificial Intelligence and Freedom of Expression. Retrieved from: https://www.osce.org/representative-on-freedom-of-media/456319.

STT Moderation Principles. (2020). Retrieved from: https://stt.fi/tyylikirja/moderointi/moderointiperiaatteet/.

Vobič, I., Robnik Šikonja, M., & Kalin Golob, M. (2019). Back to the Future: Automation and the Transformation of Journalism Epistemology. Javnost. Retrieved from: https://www.tandfonline.com/doi/abs/10.1080/13183222.2019.1696600.