



Authors	Noémie Krack (KU Leuven Centre for IT & IP Law) Lidia Dutkiewicz (KU Leuven Centre for IT & IP Law) Emine Ozge Yildirim (KU Leuven Centre for IT & IP Law)
----------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

This report is part of the deliverable D2.3 - “*AI technologies and applications in media: State of Play, Foresight, and Research Directions*” of the AI4Media project.

You can cite this report as follows:

F. Tsalakanidou et al., Deliverable 2.3 - AI technologies and applications in media: State of play, foresight, and research directions, AI4Media Project (Grant Agreement No 951911), 4 March 2022

This report was supported by European Union’s Horizon 2020 research and innovation programme under grant number 951911 - AI4Media (A European Excellence Centre for Media, Society and Democracy).

The information and views set out in this report 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.

Copyright

© Copyright 2022 AI4Media Consortium

This document may not be copied, reproduced, or modified in whole or in part for any purpose without written permission from the AI4Media 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.





1 EU policy & regulatory initiatives and their impact on future AI research for the media

AI4Media deliverable D2.1 “*Overview & analysis of the AI policy initiatives at the EU level*” provided an overview of the EU policy initiatives on AI and the forthcoming Commission’s legislative proposals impacting AI. We suggest consulting this document for a wider overview of AI initiatives on different sectors. This section will focus on the initiatives having a clear focus on the media industry.

Two types of initiatives will be presented, namely the policy initiatives (non-binding provisions) and the regulatory initiatives (leading to the adoption of binding legal provisions). They are presented chronologically. We must underline the fact that legislative proposals are only proposals and may be subject to further modifications (for more information on this, see the section on “*Future legislation and regulation*” of this Roadmap).

1.1 The Digital Services Act proposal - DSA (REGULATORY)

Released in December 2020, the proposal on a Single Market for Digital Services (Digital Services Act, DSA) aims to **harmonise rules on the provision of intermediary services** in the internal market.¹ The text sets up **transparency and due diligence obligations** for intermediary services based on an asymmetric approach. Intermediary services, hosting services, online platforms and very large platforms are the providers tackled by the provisions. Intermediary services cover mere conduit², caching³ and hosting services⁴. Such services include wireless local area networks, domain name system (DNS) services, top-level domain name registries, certificate authorities that issue digital certificates, or content delivery networks that enable or improve the functions of other providers of intermediary services, voice over IP services, web-based messaging services and e-mail services. An online platform is a provider of a hosting service which, at the request of a recipient of the service, stores and disseminates to the public information such as social networks, online market places, app stores, online travel and accommodation websites, content-sharing websites. Very large platforms are platforms reaching a number of average monthly active users in the European Union of at least 45 million (10% of the EU population). This would include for instance big social media providers such as Facebook.

¹ European Commission, Proposal for a Regulation on the European Parliament and of the Council on a Single Market for Digital Services (DSA Act) and amending Directive 2000/31/EC, 15 December 2020, COM(2020) 825 final. <https://eur-lex.europa.eu/legal-content/en/TXT/?uri=COM%3A2020%3A825%3AFIN>

² Mere conduit services are services consisting of the transmission of information in, or the provision of access to, a communication network, and include the services of telecommunications operators and internet access providers.

³ Caching services are services that consist of the transmission in a communication network of information provided by a recipient of the service, involving the automatic, intermediate and temporary storage of that information, for the sole purpose of making more efficient the information's onward transmission to other recipients upon their request.

⁴ Hosting services are services consisting of the storage of information provided by, and at the request of, a recipient of the service.





As illustrated in Figure 1, if the provider falls in one or several subcategories (hosting services, online platforms and very large platforms), it will have one, two or three extra layers of obligations applicable to his services.

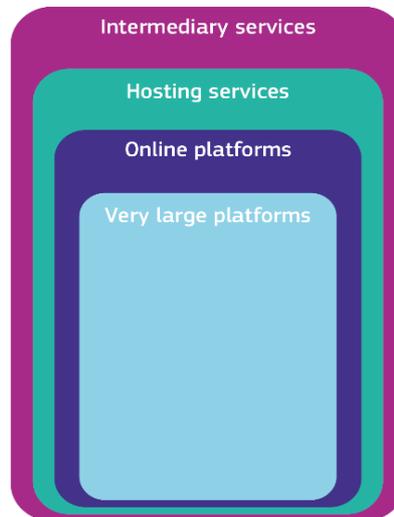


Figure 1: Categories of intermediary services according to the AI Act. Each category corresponds to extra layers of obligations.

The DSA proposal considers the impact of the use of AI based tools used in online media. The preamble of the proposal underlines how algorithmic systems shape information flows online (e.g. via content prioritisation, advertisement display and targeting or content moderation). They also create or may reinforce existing discrimination in content moderation. Below, we discuss how the DSA proposal regulates the use of AI tools in media by different online service providers.

AI and Content moderation

Providers of intermediary services must include in their **terms and conditions** (article 12 DSA proposal), in a clear and unambiguous language, information on any policies, procedures, measures and tools used for the purpose of content moderation, including algorithmic decision-making and human review.

Providers of hosting services shall put mechanisms in place to allow any individual or entity to notify them of the presence on their service of specific items that they consider to be illegal content. When confirming receipt of the **notice**, they must provide information on the use of automated means for the processing or decision-making of the notice submitted (see article 14 DSA proposal). When communicating their decision to remove or disable access to specific items of information provided by the uploader, hosting providers must include information on the use of automated means in taking the decision. For instance, if the content got detected or identified by the use of automated means (see article 15 DSA proposal).





Online platforms must now put in place an **internal complaint-handling system** for managing the complaints against a decision taken against an information provided/uploaded by a recipient of their services. The decision on the complaint must not be solely taken on the basis of automated means (article 17§5 DSA proposal). The online platforms also get additional obligation when it comes to **transparency reporting** as they must include in their yearly report on content moderation, any use made of automatic means for the purpose of content moderation, including a specification of the precise purposes, indicators of the accuracy of the automated means in fulfilling those purposes and any safeguards applied (see article 23 DSA proposal).

AI and advertising

Online platforms displaying **advertising** on their online interfaces must also ensure that information on the ad is provided in a clear and unambiguous manner and in real time: such as on whose behalf the ad is displayed and meaningful information on the main parameters used to determine the recipient to whom the advertisement is displayed, in other words what are the criteria used for targeting (article 24 DSA proposal).

VLOPS (Very Large Online Platforms) must additionally hold a **publicly available online advertisement repository**. This repository must indicate whether the advertisement was intended to be displayed specifically to one or more particular groups of recipients of the service and if so, the main parameters used for that purpose (see article 30 DSA proposal).

AI and risks assessment

VLOPS must conduct at least once a year a **risks assessment** of the functioning and use of their service which must take into account how their content moderation systems, recommender systems and systems for selecting and displaying advertisement influence any of the systemic risks (see article 26 DSA proposal). Once risks are identified, mitigation measures must be put in place and can lead to adapting content moderation or recommender systems, their decision-making processes, and the like.

AI and recommender systems

The **VLOPS** shall, in addition to the intermediary providers, set out in their **terms and conditions**, in a clear, accessible and easily comprehensible manner, the main parameters used in their **recommender systems**, as well as any options for the recipients of the service to modify or influence those main parameters that they may have made available, including at least one option which is not based on profiling, within the meaning of the GDPR. Not only providing transparency, VLOPS must also empower recipients to parameter their recommender systems.

AI access for research and enforcement purpose

VLOPS are subject to more stringent enforcement measures when it comes to transparency of automated decision-making systems. They shall **provide access** (see article 31 DSA proposal), on the one hand, to the data that are necessary to monitor and assess compliance with this



Regulation to the European Commission or the Digital Services Coordinator. On the other hand, to the data for vetted researchers focusing on identification and understanding of systemic risks created by VLOPS, including their content moderation systems, recommender systems and their systems for selecting and displaying advertisements. This data access regime will be further clarified in delegated acts. However, the provision mentions that only the system for researchers should be in compliance with the GDPR and take into account the interest of VLOPS such as the protection of confidential information, in particular trade secrets, and maintaining the security of their service. Questions remain about the implementation of this regime and about the limitation attached to the restrictive vetted researcher scheme.

The power to conduct **on-site inspections** is being granted to the Commission and it can request **VLOPS** to provide *explanations on the algorithms* used (see article 54§3 DSA proposal). The Commission can take actions to monitor the effective implementation and compliance with the DSA by **ordering access** and explanations to the **VLOPS databases and algorithms**.

1.2 Proposal for a Regulation laying down harmonised rules on artificial intelligence - AI ACT (REGULATORY)

Released by the EC in April 2021, the AI Act proposal aims to ensure the proper functioning of the internal market by setting harmonised rules in particular on the development, placing on the Union market, and use of products and services making use of AI technologies or provided as stand-alone AI systems.⁵ A **risk-based approach** is pursued by the proposal, which means that specific obligations and requirements are foreseen for distinct categories of AI systems. It includes AI systems creating **Unacceptable risk** (they are hence prohibited), **High-risk** (requirement specific ex ante, ex post requirements and a supervision and enforcement structure and regime), **limited risks** (only requiring transparency obligations) and **minimal risks** (free use) (see Figure 2).

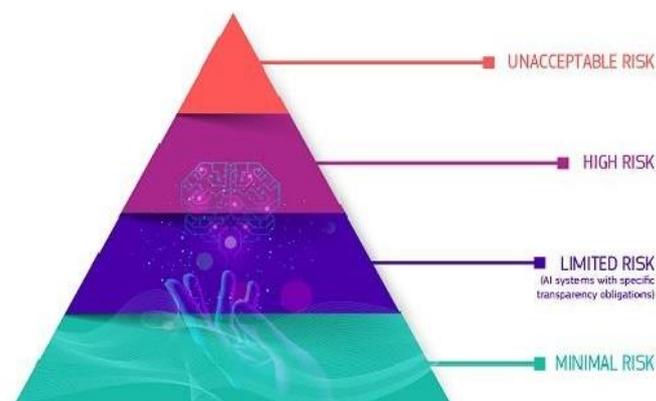


Figure 2: The risk-based approach proposed by the AI Act: AI systems are classified in four categories based on the risk they create.

⁵ European Commission, Proposal for a Regulation of the European Parliament and of the Council laying down harmonized rules on artificial intelligence (AI Act) and amending certain Union legislative acts, 21 April 2021, COM(2021) 206 final, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206>

AI media applications are not specifically tackled by the regulation, except chatbots and deepfakes. This has raised concerns from the media community since from the text description, it is not clear whether AI media applications could fall or not in a certain category.

Prohibited Practices. Article 5 of the AI Act proposal prohibits certain practices, such as the use of subliminal techniques or the exploitation of a specific vulnerability of a specific group of persons. Could recommender systems and targeted advertising fall within the subliminal techniques being considered as manipulative systems?⁶ It is unclear whether that would be the case, therefore additional guidelines or clarifications would be welcomed.

High risk. ‘Media’ is not considered as a ‘high-risk’ area and is not listed in the high-risk list in Annex III. However, research is being conducted on the risks associated to the use of AI in media sector: discrimination, bias, threats to fundamental freedom such as freedom of expression when AI systems are used for content organisation, moderation. One can wonder whether recommender systems could fall under this category. It is likely not for e-commerce, online dating, entertainment content or search engines.

Limited risks. The current wording of article 52 establishing transparency obligations lacks clarity, which could lead to legal uncertainty. Article 52§1 deals with AI systems intended to interact with natural persons. They shall be designed and developed in such a way that natural persons are informed that they are interacting with an AI system. Could this apply not only to chatbots but to recommender systems, automated journalism, as well? Article 52§3 on deepfakes provides that users of an AI system that generates or manipulates image, audio or video content that appreciably resembles existing persons, objects, places or other entities or events and would falsely appear to a person to be authentic or truthful (‘deepfake’) shall disclose that the content has been artificially generated or manipulated. However, this obligation comes with a few exceptions. Firstly, the provision only applies to users. According to Article 3§4 of the proposal, users are defined as follows: “*Any natural or legal person, public authority, agency or other body using an AI system in the course of a personal non-professional activity shall not be considered as a user*”. Secondly, the provision shall not apply where “*necessary for the exercise of the right to freedom of expression and the right to freedom of arts and sciences and for law enforcement purposes*” (article 52§3,2). These limitations are considerably restricting the added value of the transparency requirements for deepfakes as many of the malicious ones out there could be considered covered by one or several exemptions. For researchers, this labelling obligation could be optional when the absence of labelling would be necessary for the exercise of freedom of sciences. It remains to be seen in which situation the obligation could be applicable. In any case, labelling is always a good practice for ensuring trustworthy use of AI in media applications.

1.3 Resolution on AI in education, culture and the audiovisual sector (POLICY)

⁶ M. MacCarty, K. Propp, Machines learn that Brussels writes the rules: the new EU’s new AI Regulation, Lawfare Blog, April 2021, <https://www.lawfareblog.com/machines-learn-brussels-writes-rules-eus-new-ai-regulation> .



On 19th May 2021, the European Parliament (hereafter EP) adopted a resolution on AI in education, culture and the audio-visual sector.⁷ The EP acknowledged the considerable influence of education, cultural programmes and audio-visual content in shaping people's beliefs and values. This influence requires that any development, deployment and use of AI and related technologies in these sensitive sectors must fully respect the fundamental rights, freedoms and values enshrined in the EU treaties.

AI media applications have already entered the entire creation value chain and are present in users' daily lives. This is well reflected by the EP resolution, which makes a non-exhaustive list of AI media applications, their opportunities and challenges (for more information, see AI4Media deliverable D2.1 "*Overview & Analysis of the AI Policy Initiatives in EU level*").

Additionally, the resolution asked the EC to present a **general regulatory framework**, which applies to all applications of AI, and to complement it with sector-specific rules, for example for audio-visual media services.

MEPs invited the EC to introduce **strict limitations on targeted advertising** based on the collection of personal data, starting with a **ban on cross-platform behavioural advertising**. Incorporating the **gender and the diversity dimension in datasets, training, education, composition of developers' team, research** was also requested. Having a framework in place to help determine which content is protected under the IPR legislation would be necessary to solve ambiguities. They called for the establishment of a clear ethical framework for the use of AI technologies in media to prevent all forms of discrimination and ensure access to culturally and linguistically diverse content at Union level, based on accountable, transparent, and inclusive algorithms, while respecting individuals' choices and preferences. Transparency would also be a key tool for all the stakeholders: the media users, the media organisation providing content, the researchers, and the intermediaries. The following questions came into prominence: "*How is the content ranked?; What are the parameters used?; How can the users be empowered to influence to set their preference?*" They pointed that this is only partially achieved in the DSA and that more could be done in this regard (see below).

Focusing on the audio-visual sector, further recommendations were made. Algorithms used by media service providers, video sharing platforms (VSPs) and music streaming services should ensure that **personalised suggestions do not put forward the most popular works, for targeted advertising, commercial purposes or to maximise profit**. Indicators on content promoting diversity, European work should be included. The EP called both for a recommendation on algorithms and personalised marketing striving for **explainability, transparency and non-discriminatory outputs** and on **user control and empowerment over algorithms used for content recommendation** with an option to opt-out from recommendation and personalised services. They also put forward that algorithms should only be used as a flagging mechanism in content moderation, subject to human intervention. Furthermore, access to data held by intermediary providers was underlined.

⁷ European Parliament, Resolution on artificial intelligence in education, culture and the audio-visual sector, 19 May 2021, 2020/2017(INI), <https://oeil.secure.europarl.europa.eu/oeil/popups/summary.do?id=1663438&t=d&l=en>



On AI media applications, disinformation and deepfakes, the EP asked for more visibility, education on the matter. MEPs also requested an **appropriate legal framework to govern deepfake** creation, production, or distribution, including automatic tagging, stating that the content was manipulated and strict limitation when used for electoral purposes. Interestingly, EP's requests strive to be way more ambitious than the provision contained in the AI Act (see section 1.2 above). They also urged for more user control over the content they are suggested.

1.4 Code of Practice on disinformation (SELF-REGULATION) & EC Guidance to strengthen the Code (EU POLICY)

On 26th May 2021, the European Commission released its **guidance to strengthen the Code of Practice on Disinformation**.⁸ For information, at the EU level, disinformation is currently left to a self-regulation instrument namely **the Code of Practice on disinformation**⁹ (hereinafter: the Code). The Code has been signed by major online platforms and members of the advertising industry. **Regarding AI media applications**, the Code indicates that the signatories acknowledge the importance to cooperate by providing relevant data on the functioning of their services, including data and general information on algorithms. They also indicate committing to put in place clear policies regarding identity and the misuse of automated bots. They also commit to indicate in their policies what constitutes impermissible use of automated systems and to make this policy publicly available on the platform and accessible to EU users.

In September 2020, after a year of existence, the Code received an unsatisfactory assessment¹⁰ by the EC. Following these events, the EU increased its regulatory efforts to fight disinformation with several EU initiatives. This includes the DSA proposal, which suggests the introduction of a **subtle co-regulatory mechanisms for regulating disinformation**. This also includes the EU Democracy Action Plan¹¹, which announced the publication of the European Commission Guidance to strengthen the Code of Practice on disinformation, inviting signatories of the Code to submit a revised version of the Code.

The Guidance aims to reduce the financial incentives for disinformation actors, empower users and encourage flagging harmful content. It pushes providers of AI-enabled online systems for a **safe design commitment**, especially in light of risks for users' behaviours using their services. They must also take into consideration the relevant provision of the AI Act (p.12-13). The

⁸ European Commission, Communication to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, Guidance on strengthening the Code of Practice on Disinformation, 26 May 2021, COM(2021) 262 final. <https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A52021DC0262>

⁹ Code of practice on disinformation, April 2018, <https://digital-strategy.ec.europa.eu/en/policies/code-practice-disinformation>

¹⁰ European Commission, Staff Working Document – Assessment of the Code of Practice on disinformation – achievements and areas for further improvement, 10 September 2020, SWD(2020) 180 final, <https://digital-strategy.ec.europa.eu/en/library/assessment-code-practice-disinformation-achievements-and-areas-further-improvement>

¹¹ European Commission, Communication to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions on the European Democracy Action Plan, 3 December 2020, COM(2020) 790 final, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2020:790:FIN>





guidance also specifically calls for **recommender system accountability** through transparency measures about the criteria used for content ranking and prioritisation and also the possibility for users to customise the ranking algorithms.

The signatories are now expected to deliver the strengthened Code of Practice by the end of **March 2022**.¹²

1.5 Proposal on transparency and targeting of political advertising (REGULATORY)

In November 2021, the European Commission released a new proposal for a regulation focusing on political advertising, transparency and targeting.¹³ It will establish harmonised transparency measures to bring uniformity to the diverse regulations adopted in the Member States, enhance trust in the political debate and the integrity of the political debate. The text will bring legal certainty to a booming sector and will prevent the replication the Cambridge Analytica events and the interference and influence on democratic events. Rules for political targeting based on personal data and sensitive personal data will be established. The text will apply to both offline and online political advertising and targeting.

The proposal complements the regime set by the DSA proposal for commercial advertising and applies in compliance with the rules sets in the General Data Protection Regulation (GDPR).¹⁴ The harmonious co-existence of the two instruments is a core aspect of the proposal. The text sets a uniform set of rules for growing practices which come with considerable risks for democracy and fundamental rights protection. The European Commission sets a broad scope for political advertising. The material scope is quite wide. So is the political advertising definition, which includes any message for or on behalf of a political actor which is liable to influence the outcome of an election or referendum, a legislative or regulatory process or voting behaviour (Article 2§2). The personal scope targets a wide range of stakeholders involved in political advertising: political parties, political advertising services, publishers, sponsors.

Firstly, it sets up **transparency obligations for political advertising** services. In short, political advertisements shall be clearly labelled as such and include strict transparency information (who paid for it, how much, for which democratic event, etc.). Periodic reporting on political advertising services would become mandatory. Notification mechanisms for illegal political advertisement should be put in place by advertising publishers. Cooperation is also foreseen

¹² European Commission, Revision of the Code of Practice: the strengthened Code expected by March 2022, December 2021, <https://digital-strategy.ec.europa.eu/en/news/revision-code-practice-strengthened-code-expected-march-2022>

¹³ European Commission, Proposal for a Regulation of the European Parliament and of the Council on the transparency and targeting of political advertising, 25 November 2021, COM(2021) 731 final, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0731>

¹⁴ European Commission, 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), 27 April 2016, OJ L 119, <https://eur-lex.europa.eu/eli/reg/2016/679/oj>.





between the providers of political advertising services and competent authorities and interested parties such as vetted researchers, civil society organisations, electoral observers and so forth.

Secondly, the proposal introduces **specific requirements for targeting and amplification techniques related to political advertising**. Political targeting and amplification techniques would now need to be explained publicly in unprecedented detail (transparency notice, user friendly format and information content, record keeping on techniques and criteria used). The proposal establishes a **ban on political advertising** based on sensitive personal data (article 12§1). Sensitive personal data are data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation (article 9§1 GDPR). This prohibition is however nuanced by some **exceptions** (article 12§2) including data subject's explicit consent. Some already point out that given the dominant position of giant technological companies, users risk to have little choice regarding their data if they want to use their services, which goes against the concept of free and informed consent.¹⁵

Additionally, specific schemes should be put in place to ensure that data subjects can effectively exercise their GDPR rights. Recital 51 clarifies that *"The tools made available to the individuals to support the exercise of their rights should be effective to prevent an individual from being targeted with political advertisements, as well as to prevent targeting on the basis of specific criteria and by one or several specific controllers"*. The AI systems employed need therefore to be explained in a transparency manner.

Data access provisions are also in place for competent authorities exercising their supervisory tasks (article 15§3). The proposal will now be discussed within the European Parliament and the Council of the European Union. The goal of the European Commission is to have the regulation ready before the EU elections of 2024.

1.6 OSCE Policy Recommendations on artificial intelligence and freedom of expression (POLICY)

The Organization for Security and Co-operation in Europe (OSCE) has released on 20 January 2022 a manual containing policy recommendations on the most effective ways to safeguard freedom of expression and media pluralism, when deploying advanced machine-learning technologies such as AI.¹⁶ The publication is part of the project "Spotlight on Artificial Intelligence and Freedom of Expression" (#SAIFE).

¹⁵ Euractiv, How new, binding EU transparency standards for political advertising could be even higher, 30 November 2021, <https://www.euractiv.com/section/digital/opinion/how-new-binding-eu-transparency-standards-for-political-advertising-could-be-even-higher/>

¹⁶ Organization for Security and Co-operation in Europe (OSCE), The Representative on Freedom of Media, Spotlight on artificial intelligence and freedom of expression, a Policy Manual, January 2022, https://www.osce.org/files/f/documents/8/f/510332_0.pdf





14 recommendations are addressed to the OSCE Participating States. The manual contains the findings of expert workshops organised in 2021 by the OSCE and Access Now, which unpacked and analysed the main challenges that AI tools pose to human rights, in particular, the right to freedom of expression and opinion, and media freedom and pluralism. It is organised around AI in content moderation and curation. The analysis focuses on challenges relating to the use of AI in these domains such as security threats, hate speech, media pluralism, and surveillance-based advertising.

1.7 Conclusion

In conclusion of these policy and regulatory initiatives, we see that AI media applications are on the verge of being specifically regulated in legal instruments. The common approach chosen to deal with them is principally transparency requirements and then, at a level, users' empowerment, especially by providing information on the use of automated means for the processing or decision-making in content moderation.





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

info@ai4media.eu

www.ai4media.eu