**Report: 4th webinar on Artificial Intelligence in the justice field: Forensic AI in e-Evidence, 27 April 2022**

The topic of the fourth webinar was forensic AI in e-Evidence. The event dealt with various applications of forensic AI, for example with regards to the detection of deep fakes. Furthermore, discussions focused on some of the risks and opportunities related to the use of forensic AI for e-evidence purposes. The webinar featured an introductory part from the European Commission and four panel discussions with external speakers from academia, the Council of Bars and Law Societies of Europe (CCBE) and the private sector.

**The** **introductory part was opened by Mr Richard Sonnenschein, Acting Director in DG JUST (Directorate B, Criminal Justice)**, who welcomed the participants and emphasised the importance of the topic at hand. He stressed the high priority of AI related issues for the European Commission and recalled that this webinar is part of a series of webinars, organised following the Commission Communication on Digitalisation of Justice of 2 December 2020. The aim of these webinars is to bring together experts, implementers, and various stakeholders to exchange views on justice related AI topics. Furthermore, he stressed the significant transformative role AI can play that was already acknowledged in the European Commission’s 2018 Communication “Artificial Intelligence for Europe”. The “Artificial Intelligence Act” from April 2021 is a landmark legislative proposal, which lays down specific rules in order to benefit from safe and trustworthy AI-powered systems put on the European market. Moreover, he referred to several other initiatives, such as the White Paper on AI from February 2020 and the 2030 Digital Compass Communication. He highlighted that the EU also provides corresponding funding, mentioning the Digital Europe Programme and the Recovery and Resilience Facility, both supporting Member States in their digital transition. Furthermore, Mr Sonnenschein gave an overall introduction into the topic of forensic AI in e-evidence, mentioning some of the potential benefits and risks.

**Panel 1: Overview & Assessment, Deployment, Opportunities & Threats**

**A professor from the *Universitat de Barcelona***, gave a general assessment of evidence analysis and the uses of artificial intelligence within that context. The professor stressed that AI-powered tools could be used to check the coherence of witness statements, perform background checks of witnesses to test whether they may have incentives not to tell the truth and remind judges of the need to consider contextual information, such as light conditions and factors related to assessing whether the witness is under the influence of substances. AI could also be used to analyse human behaviour and body language in courtrooms. The speaker noted that body language may fail to reveal deception. The professor mentioned trials, where electronic evidence was used with the aim to increase efficiency, yet the results of such e-evidence naturally focused on the past and are therefore not necessarily reliable: since past decisions may have been judged wrongly, AI databases are biased towards such potential flaws from past decisions. The professor emphasised the drawbacks of AI in such cases and encouraged transparency with regards to the algorithms in place. Moreover, the speaker considered it more important to check the content of the testimonies over the behaviour of the witness while expressing the testimony. Ultimately, the conclusion was that despite potential flaws in AI databases for expert evidence, tools could be useful to evaluate methods, such as scientific acceptance.

**Two representatives from the German Forensic Technology practice, A&M GmbH Wirtschaftsprüfungsgesellschaft**, a company that acts as an independent expert in proceedings, where large amounts of data are analysed, discussed the deployment and opportunities related to the use of AI for electronic evidence. Focusing on two case scenarios, they used (1) supervised machine learning (ML) to find relevant documents in a set of 450 thousand documents and (2) “active learning algo” – training examples selected by machine learning, which aim to differentiate between what is and is not relevant. With iterative human intervention from lawyers, the accuracy of the data set is gradually improved, and scored on a scale of 0-100. Such methods are mostly accepted by courts in the US, yet there is also a trend towards AI use in international arbitration.

**A professor from the Goethe-Universität Frankfurt a. M.** spoke about the data used in AI for e-evidence. The professor emphasised that prosecution mostly still works on digitising paper to electronic files. Furthermore, the academic explained that the search of large and different files is useful, but that there are challenges as well, i.e., with respect to text recognition, for example from scanned files. According to the professor, spelling errors and abbreviations cause information to be misrepresented. There are, however, software products, for example used in economic crime cases in Austria to address these challenges with pattern recognition, classification and content “unification” features. This allows to structure data, build ontologies, make connections between people, places, and dates by machine. Such a tool comes close to AI and is mostly a rule-based recognition system, but there are also ML approaches. The professor stressed that future tools might be used to evaluate the quality of evidence, match information to relevant offences and even generate suggestions. Nevertheless, there is currently a lack of suitable training data.

**Panel 2: Conceptual Framework Proposal & Practical Examples**

**A representative from Amped Software** addressed the question whether we can use AI for image and video evidence analysis. The speaker stressed the need to differentiate between investigative and evidentiary use of processed or so-called “enhanced” media, such as brightened or de-blurred images. In addition to enhancement, there can be analysis, mostly for classification purposes. According to the representative, the methods used in such tools should be accurate, repeatable and the results should be reproducible. Algorithms should equally be explainable, understandable and open to scrutiny, ideally accepted by the scientific community. The speaker emphasised the importance of deterministic algorithms, as opposed to probabilistic ones. Moreover, algorithms should not introduce data external to a case. This relates to image enhancement specifically, where models trained on the faces of other people might extrapolate the little data they have about a low-resolution image at hand. The models generate a higher resolution image, which might however carry noise, reflecting the faces in the training data set rather than being a faithful augmentation of the image at hand. As a conclusion in this context less, but more faithful information is preferred to external information, even if a larger sample exists. The speaker was clear on the fact that image enhancement with AI should not be acceptable for evidentiary use, yet it may be acceptable to generate leads in investigations, with due transparency and safeguards. The representative noted that image analysis does not modify the material and should only be used to *assist* decisions of competent staff and where reliability of a tool was demonstrated for specific type of use. Authentication of images becomes a new challenge since AI is used to “fake” material. Image analysis with AI should therefore be used cautiously, in criminal law contexts as AI, even based on “traditional” algorithms, is not always under our own control. Safeguards such as a decision support system, a known reliability and limited human bias should therefore be in place. Meanwhile, image and video enhancement with AI should in general not be acceptable for evidentiary use due to the influence of data external to the evidence, a processing bias and a lack of *explainability*.

**The founder of the private company RightNow** discussed practical risks and opportunities, based on the company’s experiences. RightNow buys small claims from consumers and uses automated tools to process them, for example in the context of flight delay compensation claims. Subsequently, they pool cases for mass lawsuits if needed, i.e., if the airline refuses to pay. Ultimately, they give these claims to law firms or debt collectors for enforcement.

**Panel 3**: **Practical Tools, Assessment, Use & Misuse**

**A representative from Robic LLP** discussed practical tools and the use and misuse of forensic AI in e-evidence, based on company-specific experiences. The speaker delved into the concept of *explainability* in AI and the challenges that weigh on its practical implementation.

**A member of the Surveillance Committee of the CCBE, and member of the IT Law Committee of the CCBE**, illustrated the AI tools available, and discussed the differentiation between civil and criminal law. The representative stressed the need not to assume that systems were robust, and not to trust the marketers that systems do what they say - it is always better to verify. Moreover, the speaker emphasised that even if the system does what it says it does, an additional check if the use is consistent with fundamental rights is highly recommended.

**Panel 4**: **Deep fakes**

**A professor and Research Director with the Information Sciences Institute** elaborated on the current challenges in multimedia forensics and generative adversarial networks (GANs), stressing that seeing ceased to be believing. In the context of deep fakes, the professor emphasised that further checks are needed where stakes are high. It is therefore worth to continue the work on tools to detect deep fakes. Consequently, if the accuracy is good enough, such tools might be accepted as reliable methods, even by courts.

**A professor and Chair of the Forensic Data Science University of Amsterdam and associated with the Nederlands Forensisch Instituut in Den Haag** discussed the topic of forensic deep fake detection tools and methods. The professor specifically mentioned that phonetic fake detection tools can be triggered by a dialect and that deep fakes are potentially used to “swap” the faces of perpetrators in Child Sexual Abuse Material (CSAM) to obfuscate who they are. The speaker also noted that digital camera sensors are unique and leave a particular “fingerprint”.

**A professor and founding co-Director of the Multi-disciplinary Center of Information Integrity at the University at Buffalo, State University of New York** gave an overview of the types and making of deep fake media. The professor stressed that detection algorithms are useful, yet in many cases constitute opaque black boxes that lead to results that could not be used as legal evidence. The professor noted that a tool, that does not provide explanations, may carry out initial detection and humans can look for justifications based on human understanding. The speaker proposed to look for semantic cues rooted in physical or physiological constraints of the real world. As an example, the professor mentioned the unique shapes of ears that can be used to spot deep fakes or to justify an assertion that a video is fake. Moreover, the speaker gave the example that the iris in GAN generated images is sometimes not perfectly round, and reflections are sometimes not similar in both eyes. That said, deep fake generation tools might be improved in view of addressing these elements. Ultimately, the speaker stressed that all deep fake detection happens ex post and that hence, there is a need for pre-emptive approaches as well.

**A PhD researcher at the European University Institute** focused on detecting deep fakes in criminal proceedings and discussed the opportunities and challenges of forensic AI. The speaker started by mentioning that deep fakes can be relevant in criminal proceedings as the object of the proceeding, such as a harassment or defamation crime, or as fake evidence in relation to other crimes, for example a fake video of an armed robbery. In the UK, there was a deep fake audio in a custody case to portray the father as violent and a threat to his children. If not detected, deep fakes can lead to wrong decisions and could be tantamount to a risk of eroding the trust in audio-visual evidence. The speaker flagged that the costs entailed by efforts to detect deep fakes in the interest of a defendant are possibly not covered by legal aid regimes. To ensure fair access to experts and tools (principle of equality of arms), AI Act (AIA) obligations can help with accuracy challenges. However, there are no specific provisions that would allow defendants to gain insights into tools used by prosecutors. There is therefore a need for strong legal safeguards and training on the use of forensic AI in courts. Software providers can help to increase transparency and to reduce cost for defendants.

**Q&A and Discussion**

In the Q&A and discussion round, the PhD Researcher and the representative from the CCBE discussed whether a black box tool can be used for evidence purposes, provided they follow the AIA requirements. The PhD Researcher stressed the need to wait for the final text of the Act, and the evaluation of potential further clarifications regarding the requirement of “interpretability”, whereas the CCBE representative emphasised that AIA compliance does not mean compliance with procedural law of a Member State and concluded that these are different requirements. In conclusion, the AIA should provide safeguards before a system goes to market, but there are also aspects related to the rules on defence in criminal proceedings. That defendants need access to detailed documentation about any AI system used could be clarified, also in the AIA.