Faculty of Law, Economics and Finance

CRIM_AI

-

Unpacking AI Evidence and (Re-)Defining Procedural Safeguards in Digital Investigations

CRIM_AI Research

FACULTY OF LAW, ECONOMICS AND FINANCE

- Shift from human-centric investigation to data analysis due to
  - expansion of AI systems in investigation and prosecution of crime;
  - omnipresence of AI devices in daily lives of humans.

- CRIM_AI seek to address whether:

  "existing rules on criminal procedure, in particular evidence law and procedural guarantees, are sufficient to address the specific nature and the associated pitfalls of AI evidence?"

- Taking into account
  - the primary role of national courts in building proof and their capacity of judicial interpretation;
  - the potential of data protection principles to fend off the negative effects of AI Evidence's opacity and inaccuracy in criminal proceedings;
  - function creeps;
  - the role of the private sector.

Outline

- Definition and typology of AI Evidence
- Impact of AI Evidence on criminal proceedings
- Responses of national courts to challenges posed by AI Evidence
- New procedural guarantees

1.
AI Evidence:
Definition and Typology

AI Evidence

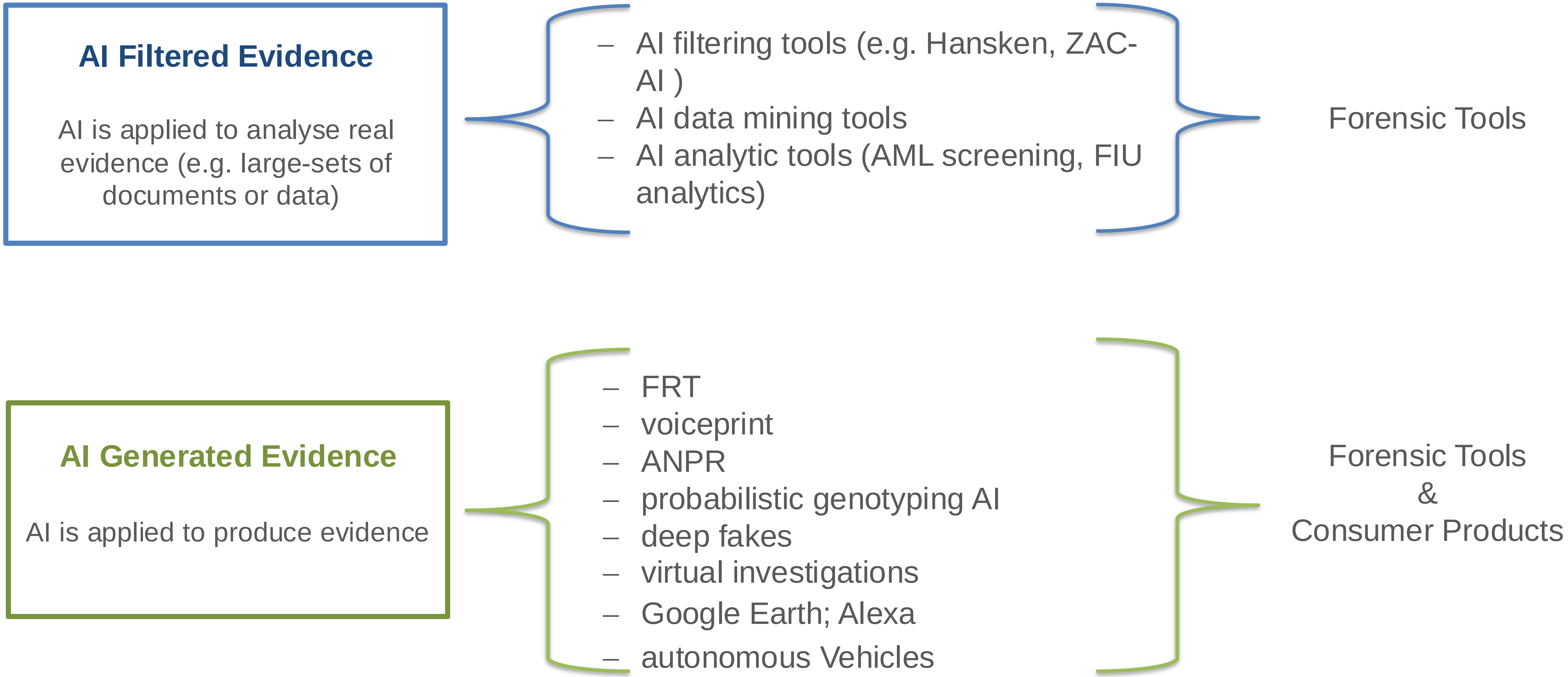
- AI Evidence means the use of AI’s output to establish the guilt or innocence of someone accused of a crime where the AI system generated the output
  - **autonomously**
  - **by using machine learning**.

- Autonomous working of AI is key element of the definition
  - **lack of human control** in processing or generating such evidence;
  - some form of machine learning is required; i.e. rule based systems are excluded (limitation on the technology considered), but foundational models of Generative AI are considered.

AI Evidence

FACULTY OF LAW, ECONOMICS AND FINANCE

UNIVERSITÉ DU LUXEMBOURG



AI Filtered Evidence

AI is applied to analyse real evidence (e.g. large-sets of documents or data)

- AI filtering tools (e.g. Hansken, ZAC-AI)
- AI data mining tools
- AI analytic tools (AML screening, FIU analytics)

Forensic Tools

AI Generated Evidence

AI is applied to produce evidence

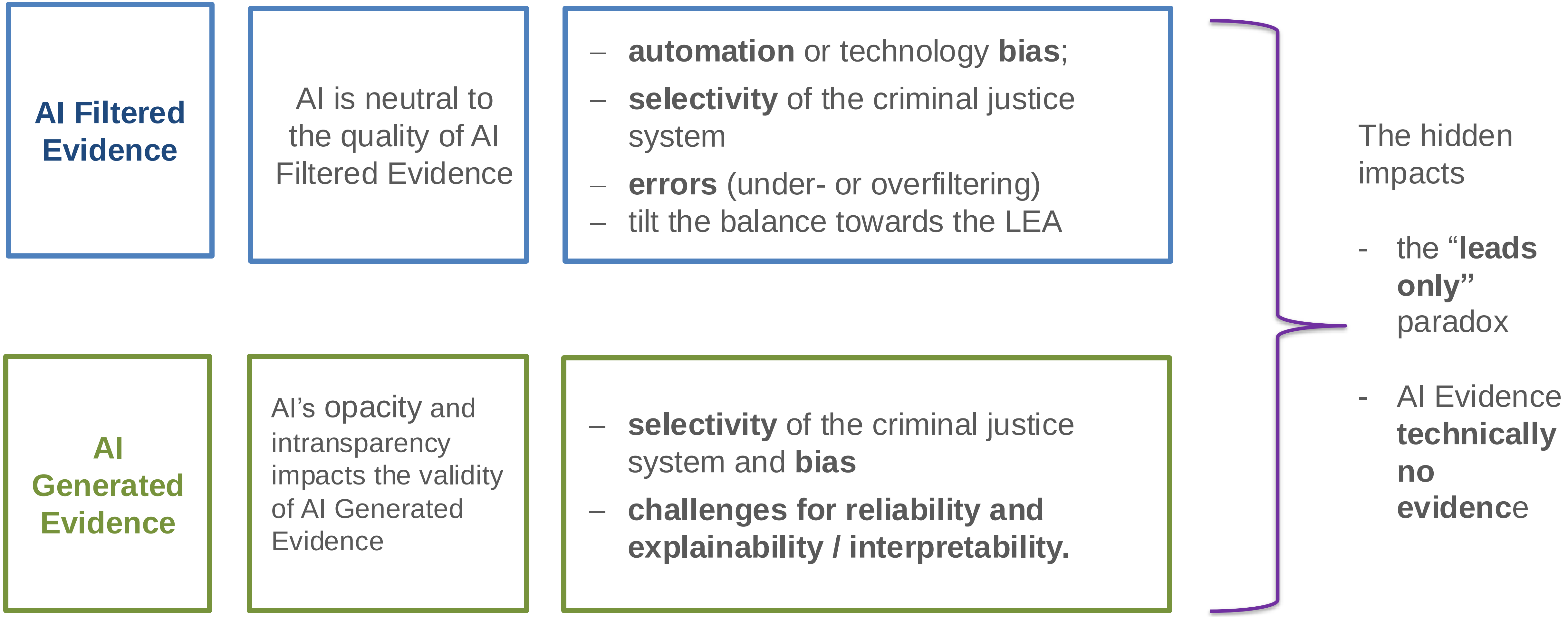
- FRT
- voiceprint
- ANPR
- probabilistic genotyping AI
- deep fakes
- virtual investigations
- Google Earth; Alexa
- autonomous Vehicles

Forensic Tools
&
Consumer Products

2.
Impact of AI Evidence

Inmpact of AI Evidence on the Proceedings

FACULTY OF LAW, ECONOMICS AND FINANCE



AI Filtered Evidence

AI is neutral to the quality of AI Filtered Evidence

- **automation** or technology **bias**;
- **selectivity** of the criminal justice system
- **errors** (under- or overfiltering)
- tilt the balance towards the LEA

AI Generated Evidence

AI's opacity and intransparency impacts the validity of AI Generated Evidence

- **selectivity** of the criminal justice system and **bias**
- **challenges for reliability and explainability / interpretability.**

The hidden impacts

- the "**leads only**" paradox
- AI Evidence **technically no evidence**

3.

National Courts Responding to the Challenges of AI Evidence

Admissibility of AI Evidence

- **Divergent national rules on admissibility** and exlcusion of evidence:
  - NL, FR, DE follow the inquisitorial tradition and place a lot of importance on how the evidence was obtained (i.e. regulating investigative measures) and contain less detailed rules on admission, presentation and evaluation of evidence;
  - UK & US follow the adversarial tradition and have detailed rules on admissibility; the judge has a gate-keeper role ensuring that the trier of the fact sees only admissible evidence.

- **General tendecy to admit** AI Evidence without too-detailed scrutiny as to validity, reliability, or credibility
  - criminal justice systems lack standardized tests for forensic evidence;
  - determinations on reliability and authenticity require quite a bit of specialized fact-finding in the case of AI Evidence.

Disclosure of AI Evidence

- The prosecutor is obliged to disclose both inculpatory and exculpatory evidence to the defense including if the evidence contains forensic reports.

- **Fair trial requires that the prosecutor, judge, jury, and affected parties know that AI Evidence** is part of the evidence.

- National approaches vary whether information on AI Evidence is provided in the case file
  - UK requires indication if parts of the evidence were computer-generated or assisted;
  - NL reports introducing complex forensic evidence must indicate whether the evidence contains original or processed data;
  - US no requirement regarding disclosure of the use of AI Evidence;
  - the file may not contain information on the use of AI because it generated only leads.

Scrutinizing AI Evidence by the Defense

- To challenge the admissibility the defense needs to demonstrate that the AI output is
  - either not valid and/or not reliable;
  - and/or it has not been correctly applied in the case of the defendant.
- To verify the reliability of AI Evidence, the defense needs both **opportunity** and **means** to do so.
- **Courts often deny defense requests** to access the information required for an independent validation on grounds
  - of trade secrets of the proprietary AI (US);
  - that such discovery is not necessary for disposing fairly of the action (NL, UK) or
  - that it will incur unnecessary costs (UK).
- **Defense often lacks means** to pay for forensic experts

From “Blind” Trust to Real Scrutiny

- (European) Courts in increasingly take a protective stance:
  - AI Evidence needs to be supported by other proof
  - dominance of human judgement
  - novel approaches to recognise (new) defence rights

- European frameworks offer new guarantees
  - transparency
  - right to explanation

New Procedural Guarantees?

- Right to information on the AI tool
- Right to access to the full collection of data
- Right to access the AI tool
- Right to explanation of forensic methods and results
- Right to have digital forensic assistance

Conclusion

Thank you very much for your attention!