ICCBA WORKING GROUP ON ARTIFICIAL INTELLIGENCE

POSITION PAPER: OTP DRAFT POLICY ON CYBER-ENABLED CRIMES

This Position Paper is submitted by the ICCBA Working Group on Artificial Intelligence (the "WG-AI") according to its mandate to examine legal and ethical issues arising from using AI in international criminal justice. It is submitted for consideration by the ICCBA Executive Council and the wider ICC Community.

The paper welcomes the Office of the Prosecutor's ("OTP") Draft Policy on Cyber-Enabled Crimes ("the Draft Policy"), particularly its inclusion of AI-facilitated conduct within the Rome Statute framework. It further identifies improvement areas and recommendations, drawing on relevant jurisprudence, OTP policy instruments, and comparative standards, through a prism of AI.

1 / Guiding Principles

The ICCBA WG-AI recalls the following foundational principles, which highlight the obligations of the Rome Statute, the rights of the parties and participants, and the evolving role of technology in legal processes:

I. **Rule of Law and Fair Trial Rights:** The use of AI in proceedings must accord with Article 67 of the Rome Statute - the right to a fair and public hearing, equality of arms, and the ability to challenge and test the evidence.
II. **Transparency and Explainability:** Transparency is a cornerstone of justice. Justice must not just be done, it must be seen to be done. This means that any use of AI must be explainable and transparent. As such, "Black Box" algorithms are incompatible with judicial accountability.
III. **Human Oversight and Accountability:** AI must assist human decision-makers, not substitute them. The ultimate decision-making must remain with prosecutors, judges, and counsel. The space for moral perception and ethical deliberation must be preserved.
IV. **Technological Neutrality of the Rome Statute:** The Rome Statute is sufficiently agile to accommodate AI and modern technologies; however, this must be done with an awareness of AI's specific risks and capabilities.

2 / Opportunities and Use Cases for AI

AI presents opportunities in all facets of ICC proceedings. If governed responsibly, it can enhance efficiency, accuracy, and access to justice:

I. **Digital Evidence Review:** Machine learning tools can facilitate organising copious datasets, such as call data records, intercepted communications, and satellite imagery.
II. **OSINT and Early Warning:** Early in investigations, AI can assist in identifying atrocity indicators across open-source platforms.

III. **Pattern Analysis**: Algorithms can assist in identifying coordinated attacks, patterns, or hierarchical command structures.
IV. **Translation and Accessibility**: Neural machine translation models can increase the availability of ICC materials in multiple languages. Furthermore, real-time translation of materials and proceedings advances the broader principle of access to justice.
V. **Legal Drafting and Research:** NLP tools can provide limited support to prosecutors and counsel with legal research and precedent analysis.
VI. **Victim Engagement:** Chatbots and digital assistants (if properly secured and ethically designed) may facilitate accessible victim registration and information-sharing. This can significantly bolster victim/survivor participation in ICC proceedings.

3 / Response to the OTP Draft Policy on Cyber-Enabled Crimes

The WG-AI welcomes the OTP's comprehensive articulation of how Rome Statute crimes can be committed using cyber means, including AI. The WG-AI endorses the OTP's core message that "*As a matter of law, genocide, crimes against humanity, war crimes and aggression, as well as offences against the administration of justice, can all be perpetrated or facilitated by cyber means*" (para. 10). The Draft Policy demonstrates the Prosecutor's intent to be at the forefront of technological developments. However, key areas require further clarity and precision:

A. Distinction Between AI and Cybercrime

Issue: The Draft Policy conflates AI with general cyber tools, without providing a definition or framework for AI-specific analysis (paras. 20–24). It also lacks a comprehensive definition of data (paras. 70–71). The WG-AI submits that, given AI's unique capabilities and potential, whether constructive or destructive, it warrants clear delineation and separate analysis.

Recommendation: Define AI-facilitated conduct separately. For example, the use of:

- Deepfakes for incitement to violence (para. 50);
- Predictive policing tools to target protected populations (para. 59);
- Autonomous systems for prohibited and/or indiscriminate attacks (paras. 65, 68).

Reference frameworks such as the OTP's Policy on Gender-Based Crimes (2023) could serve as models for structured thematic analysis.

B. Attribution, Mental Element and Modes of Liability

Issue: Regarding Article 30 of the Rome Statute (para. 87), the Draft Policy does not sufficiently address attribution issues involving semi-autonomous or opaque AI systems.

Recommendation:

- Elaborate on how modes of liability under Article 25 of the Rome Statute, as well as doctrines such as command responsibility under Article 28, apply in cases where AI

systems mediate or influence human decisions. Particular attention should be given to scenarios involving semi-autonomous or opaque technologies, where attribution of intent or control may be complex. Reference should be made to the **Tallinn Manual 3.0**, as acknowledged in paragraph 11, which provides relevant insights into state and individual responsibility in the context of cyber operations.

- Further, the Draft Policy should expand on the concept of the "ordinary course of events" under Article 30(3) of the Rome Statute, particularly as it relates to foreseeability and the use of AI. The analysis should consider how intent or knowledge might be inferred where AI tools behave in ways consistent with their design or training parameters, even if their outputs are not fully predictable.

C. Evidentiary Safeguards

Issue: The Draft Policy lacks clear guidance on assessing AI-generated or manipulated evidence (paras. 129–132).

Recommendation: Adopt internal protocols on:

- Chain of custody for synthetic content;
- Forensic review of AI outputs;
- Metadata authentication and verification;
- Disclosure of algorithms behind evidence selection and any reporting.

Align with the OTP's Policy on Case Selection and Prioritisation (2016), paras. 42–44 (Degree of responsibility of alleged perpetrators).

D. State-sponsored actors and Hacktivists

Issue: The Draft Policy provides limited elaboration on the types of actors commonly involved in cyber-enabled crimes (e.g. "hackers" at para. 53). A clearer taxonomy would aid understanding of the range of potential perpetrators, while remaining non-exhaustive.

Recommendation: Offer more structured definitions of key actor categories typically associated with cyber-dependent or cyber-enabled conduct:

- **State-sponsored actors**: Typically government-funded, technically proficient, and politically motivated. These actors often target critical infrastructure or high-value systems that increasingly rely on digital support. Recent geopolitical conflicts illustrate their persistent activity across multiple jurisdictions.
- **Hackers and hacktivists**: Non-state actors often driven by political or social objectives. Their methods may include denial-of-service attacks, website defacement, or data exfiltration intended to highlight perceived injustices or promote specific causes.
- **Cyber-terrorists**: Actors who employ cyber means to advance political or ideological agendas through disruption, fear, or destabilisation. Attribution is often complex due to overlaps with state sponsorship, armed groups, or loosely affiliated entities. The blurred

boundaries between cybercrime, terrorism, and state activity pose evidentiary and definitional challenges.

Source: Understanding Cybercrime, EPRS, 2024.

E. Defence and Victims' Capacity

Issue: Paragraph 107 omits reference to training and resources for defence teams and legal representatives.

Recommendation: Ensure that all parties, including defence and victims' representatives, receive adequate technical support and digital literacy training. The Registry should facilitate equal access to relevant tools and forensic expertise to uphold the principle of equality of arms and ensure compliance with Article 67(1)(e) of the Rome Statute.

F. Standalone AI-Facilitated Crimes

Issue: Paras. 94–97 (Gravity, impact, and practicality) suggest cyber-enabled acts are only prosecutable when part of larger criminal campaigns.

Recommendation: Affirm that AI-facilitated acts (e.g. autonomous targeting, algorithmic persecution) may independently satisfy Rome Statute thresholds. Reference the OTP's Policy on Situation Completion (2021), para. 21, on the expressive value of prosecutions.

G. Oversight of Private Partnerships

Issue: The OTP notes partnerships with Microsoft and other tech firms (para. 17) without transparency or accountability measures.

Recommendation: Introduce a protocol or advisory board for vetting such partnerships.

H. Clarity on *Proprio Motu* Investigations

Issue: While complementarity is affirmed (paras. 140–141), the Draft Policy does not articulate when the OTP may initiate proceedings *proprio motu* in cyber/AI cases.

Recommendation: Provide clearer criteria for such interventions, especially where digital repression is state-sponsored or domestic investigations are absent.

I. Intersectional Harm Analysis

Issue: Victim impact is referenced (para. 129) but lacks disaggregated analysis.

Recommendation: Adopt impact assessments recognising the heightened vulnerabilities of children, women, LGBTQ+ persons, and ethnic minorities in AI-enabled harms, following the Policy on the Crime of Gender Persecution (2022), paras. 55–58.

5 / Conclusion

The OTP's Draft Policy is a pivotal step in enabling the ICC to address the challenges posed by cyber and AI-enabled criminality. The ICCBA WG-AI fully supports its development and encourages the adoption of a final version that reflects the doctrinal, evidentiary, and operational complexities associated with AI in conflict settings.

We offer this position paper as a constructive contribution to that goal and stand ready to assist through expert dialogue, drafting support, and technical consultations.

Submitted on behalf of the ICCBA Working Group on Artificial Intelligence

ICCBA Working Group on Artificial Intelligence
International Criminal Court Bar Association (ICCBA)

Date: 21 May 2025

Sources

Primary Legal Texts and ICC Policy Instruments

1. Rome Statute of the International Criminal Court (adopted 17 July 1998, entered into force 1 July 2002) https://www.icc-cpi.int/sites/default/files/RS-Eng.pdf

2. Office of the Prosecutor, *Policy Paper on Gender-Based Crimes* (ICC, December 2023) https://www.icc-cpi.int/sites/default/files/2023-12/2023-policy-gender-en-web.pdf

3. Office of the Prosecutor, *Policy Paper on Case Selection and Prioritisation* (ICC, 15 September 2016) https://www.icc-cpi.int/sites/default/files/itemsDocuments/20160915_OTP-Policy_Case-Selection_Eng.pdf

4. Office of the Prosecutor, *Policy on the Crime of Gender Persecution* (ICC, December 2022) https://www.icc-cpi.int/sites/default/files/2022-12/2022-12-07-Policy-on-the-Crime-of-Gender-Persecution.pdf

5. Office of the Prosecutor, *Policy on Situation Completion* (ICC, 15 December 2021) https://www.icc-cpi.int/news/policy-situation-completion

Cybercrime and AI Legal Frameworks

6. Michael N Schmitt (ed), *Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations* (2nd edn, CUP 2017) https://www.cambridge.org/gb/universitypress/subjects/law/humanitarian-law/tallinn-manual-20-international-law-applicable-cyber-operations-2nd-edition?format=PB

7. European Parliamentary Research Service, *Understanding Cybercrime* (Briefing 760356, February 2024) https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI(2024)760356

Academic and Theoretical Foundations

8. Sylvie Delacroix, *Moral Perception and Uncertainty Expression in LLM-Augmented Judicial Practice* (2024) SSRN https://ssrn.com/abstract=4787044

9. Henning Lahmann, ‘Self-Determination in the Age of Algorithmic Warfare’ (2025) LT Special Issue *European Journal of Legal Studies* 161, DOI: 10.2924/EJLS.2025.LT.005