AUTHENTICATING DIGITAL EVIDENCE IN THE AGE OF AI

CARISSA PRETORIUS

EDUVOS 101 Hendrik Van Eck Boulevard, Vanderbijlpark S.E. 6, 1911 South Africa (VAT: 4680293406)
E-mail: Carissa.Pretorius@eduvos.com

Digital evidence has become routine in criminal trials. Screenshots. Voice notes. CCTV clips. Metadata. Courts increasingly treat these artefacts as neutral reflections of fact. That assumption is no longer safe. Generative artificial intelligence can now fabricate text, audio, images and video that are difficult to distinguish from authentic material. It can also manipulate timestamps, simulate conversations and produce synthetic confessions. In African criminal justice systems, where digital forensic capacity and judicial training remain uneven, the risk is acute. The evidentiary framework has not kept pace with the technology. This paper identifies a central research problem: how should criminal courts assess authenticity and reliability when digital evidence may be AI generated or AI manipulated? It asks three questions. First, are existing doctrines on admissibility, authentication and chain of custody sufficient to address synthetic media risks? Second, what institutional vulnerabilities make African jurisdictions particularly exposed to evidentiary manipulation? Third, what regulatory and procedural reforms are necessary to safeguard trial integrity without undermining legitimate digital innovation? The paper finds that current evidentiary rules are formally flexible but practically under-equipped. They rely heavily on presumptions of authenticity and adversarial challenge, both of which weaken in the face of highly convincing synthetic artefacts and unequal technical expertise between parties. The result is a credibility vacuum. It argues for a calibrated reform model: heightened authentication thresholds for high-risk digital evidence; mandatory disclosure of AI tools used in evidence preparation; judicial education on synthetic media; and minimum accreditation standards for digital forensic experts. Criminal justice depends on the integrity of proof. If reality itself becomes contestable, the presumption of innocence is destabilised. Evidentiary resilience must become a priority of African cyber law.

Keywords - Artificial Intelligence; Digital Evidence; Authentication Standards; African cyber law Criminal Trials