Scientific Research Publishing

Beijing Law Review, 2026, 17(4), 1209-1239
https://www.scirp.org/journal/blr
ISSN Online: 2159-4635
ISSN Print: 2159-4627

Authenticity Is Not Reliability: Algorithmic Facial Identification and Pakistan's Law of Automated Evidence

Saad Khalid Khan[1*], Farkhanda Khaleeq Khattak[2], Umarzish Noor ul Ain[3], Muhammad Afra Siab[4]

[1]School of Law, Beijing Institute of Technology, Beijing, China
[2]Faculty of Political Science and Law, Université de Bordeaux, Bordeaux, France
[3]Islamabad Bar Association, Islamabad, Pakistan
[4]School of Artificial Intelligence, Beijing Institute of Technology, Beijing, China
Email: *saadkhan7919@gmail.com

How to cite this paper: Khan, S. K., Khattak, F. K., ul Ain, U. N., & Siab, M. A. (2026). Authenticity Is Not Reliability: Algorithmic Facial Identification and Pakistan's Law of Automated Evidence. *Beijing Law Review, 17,* 1209-1239.
https://doi.org/10.4236/blr.2026.174059

Received: September 9, 2026
Accepted: November 10, 2026
Published: November 13, 2026

Copyright © 2026 by author(s) and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/

Open Access

Abstract

A court can be satisfied that a recording is genuine and unaltered without being satisfied that a computer has correctly identified the person in it. No reported Pakistani case has yet turned on an algorithmic identification, but Punjab's deployed facial-recognition capability makes it a matter of time. The Qanun-e-Shahadat Order contains a complete chain for machine-generated evidence, inserted in 2002 and read in practice through Article 164: Article 46-A on relevance, Explanation 3 to Article 73 on class and the presumption of working order, Article 78-A on proof, and Article 59 on expert opinion about information systems, to which *Zahir Zakir Jaffar* adds a two-fold test for CCTV that it derives from *Ahmed Omar Sheikh*. Relevance, class, proof, and integrity are provided for; accuracy is not. The presumption of working order should be read as presuming that the machine was running, not that it was correct: on *Millat Tractors*, an Explanation is read by what it was intended to do and does, and this one classifies output and presumes that the system was operating; on *Raja Khurram Ali Khan*, the legal burden cannot be shifted to an accused unless the legislature by express terms commands otherwise. The missing requirement can be supplied from Article 59 with the gatekeeping principle, from the Silent Witness doctrine received in *Zahir Zakir Jaffar* and extended to a system that concludes rather than records, and from identification-parade doctrine. Drawing them together develops doctrine rather than restating a rule already in force. Because Pakistan tries criminal cases without juries, the argument concerns the foundation a court must be given, not a power to exclude. Three institutional routes in Punjab sharpen the problem, which is not wholly antic-

ipatory: in review proceedings in *Zahir Jaffar*, the Supreme Court rejected an objection that forensic reports were inadmissible because their authors were not produced. Indian practice under Section 65B of the Evidence Act answered the same problem with a certificate, and so, in its own field, does Punjab's Electronic Data Certificate; both secure integrity, neither accuracy. China yields lessons about classification and institutional separation, not a model of admissibility. We propose a four-stage framework and apply it to a worked example.

Keywords

Algorithmic Facial Recognition, Article 46-A, Automated Information System, Silent Witness Doctrine, Test Identification Parade, Criminal Evidence, Pakistan

1. Introduction

Pakistani courts ask whether a recording is authentic; they have not yet had to ask whether a machine has correctly identified the person in it. The two questions separate once software is involved. When a camera records an assault, the evidentiary questions are provenance and integrity: where the file came from, who took it out of the recorder, whether anything has been altered. The answers tell the court what it is looking at, not who is on the screen. Facial-recognition software run against the footage does not preserve information but produces a proposition: an inference about identity that no human witness has made.

The distinction is easily lost: a forensic report confirms the footage is unedited, a comparison links the person in it to the accused, and a judgment records both findings in one passage. Nothing in that sequence is wrong, but nothing keeps the propositions apart, and a later court may take the first finding to have carried the second.

No reported Pakistani case has yet turned on an algorithmic identification. The Punjab Information Technology Board and the Punjab Police IT Wing have developed and deployed Face Trace, which PITB describes as an AI-based facial-recognition system providing real-time matching against multiple datasets and allowing police to match CCTV footage (Punjab Information Technology Board, n.d.-a, n.d.-b). The Punjab Safe Cities Authority separately operates analytics identifying persons by attributes including face, height, clothing, gait, and posture (Punjab Safe Cities Authority, n.d.).

1.1. The Literature, and What Remains of It

The measurement science is settled. NIST evaluates verification and identification as distinct tasks and publishes demographic performance data (Grother et al., 2017, 2019a, 2019b). Experimental work establishes that prior identity decisions shift subsequent human face-matching judgments, including cases where the earlier decision is attributed to a computer (Howard et al., 2020), and that perfor-

mance varies across examiners, super-recognisers, and algorithms (Phillips et al., 2018).

In the legal scholarship, Roth's (2017) account of machine output as requiring credibility testing rather than mere authentication is foundational; Grimm, Grossman, and Cormack address artificial intelligence as evidence (Grimm et al., 2021), as do Grossman and Grimm (2025); Garrett and Rudin (2024) argue for interpretable forensic algorithms; Wexler (2018) addresses trade secret claims in criminal discovery; and Gless et al. (2024) take the question comparatively. Facial recognition in the courtroom has its own literature, including Gabel Cino et al. (2022) and Perritt Jr. (2021), who argues that existing criminal-procedure rules are adaptable to it. Courts have also begun to answer the question: *State v. Arteaga* (2023)[1] ordered disclosure of facial-recognition material; *State v. Archambault* (2024)[2] held, after a Frye-Mack hearing, that a facial-recognition tool used by police, although an implementation of generally accepted technology, did not reliably and consistently produce accurate results; and *State v. Miles* (2026)[3] affirmed disclosure while rejecting a rigid checklist and leaving the source-code question open on a particularised-need standard.

Pakistani scholarship has discussed the 2002 architecture and, since *Zahir Jaffar*[4], the Silent Witness doctrine: Hameed et al. (2026) treat the doctrine alongside Articles 164, 73, and 46-A, and reach photogrammetric testing and the establishment of identity. We differ from their account of what the Court decided: they render the test as originality and freedom from tampering together with pertinence to the facts in issue, whereas the guidelines applied in *Ahmed Omar Sheikh*[5] are longer and more demanding. We do not claim that the statutory chain has been overlooked.

Recent Pakistani work directly on AI-generated evidence treats Article 164 as governing evidence produced by modern devices and argues for its reform on that footing (Shahzad et al., 2026), without addressing either the 2023 substitution or the electronic-evidence provisions inserted in 2002. Work on expert evidence identifies the wider condition: such evidence is admissible under Articles 59 and 164, but enacted law and precedent are inadequate to help a trial judge assess its reliability, in a system without national guidelines and with unaccredited laboratories (Rasool & Rasool, 2022).

None of this work asks how a computational identity inference travels through the provisions that govern machine-generated output, or what a Pakistani court must be shown before acting on it.

1.2. Question, Method, and Scope

The question is what Pakistani evidence law requires before a court may rationally

[1] *State v. Arteaga*, 476 N.J. Super. 36 (App. Div. 2023).
[2] *State v. Archambault*, No. 62-CR-20-5866 (Minn. Dist. Ct. Ramsey Cnty. Sept. 13, 2024).
[3] *State v. Miles*, A-41-24 (090275) (N.J. June 24, 2026).
[4] *Zahir Zakir Jaffar v. The State*, 2025 SCP 220 (Sup. Ct. Pak. May 20, 2025).
[5] *The State through P.G. Sindh v. Ahmed Omar Sheikh*, 2021 SCMR 873 (Pak.).

rely on algorithmic facial identification as proof of identity in a criminal case, and what limited lessons China's more developed treatment of the technology offers.

The answer has three parts of different kinds. First, the law as it stands: Articles 46-A and 73 settle relevance and classification and do not establish that an inference is accurate. Second, what existing doctrine makes available: Article 59 with the gatekeeping principle in *Ali Haider*[6], the Silent Witness doctrine, the identification-parade doctrine, and Article 10A give a court the resources to require more than proof that the machine was operating. Third, a proposal: because no Pakistani authority has yet drawn those resources together for an algorithmic identification, the four-stage framework in Part 5 is offered as a doctrinal development, not as a rule already in force. Where a step is analogical or normative, we say so when it is taken.

The method is doctrinal, working from the Order's text and reported superior-court authority; the comparison with China is confined to what it can establish. We claim no functional equivalence between the two systems; Part 4 states its basis. Chinese materials are used where an authoritative text or a published standard number and title are available, and we attribute no clause-level content to standards whose text we have not read. Neither 2025 instrument has an official English text; both are cited from the Chinese texts as promulgated, with unofficial English renderings, and where a proposition turns on a single word, we say so.

Four limits apply. Our statement that no reported superior-court authority prescribes a facial-recognition reliability inquiry is a search finding, not proof of non-existence; the searches are described in the note[7]. Our proposal is anticipatory: no Pakistani appellate case has tendered an algorithmic identification as substantive proof, so every stage is calibrated against constructed facts; Part 6 and the generalisation at the end of Part 3.1 try to reduce that weakness. Institutional descriptions of Punjab's systems establish capability and operational responsibility, not evidentiary practice.

2. What the Technology Does, in the Terms the Law Needs

Software converts a face in an image into a numerical representation, compares it against one or more enrolled representations, and returns a similarity measure. Someone then interprets the output and forms a view about identity.

There are two distinct operations. In one-to-one verification, the system compares a probe against a specific stored image: an identity has been asserted, and the system tests it. In one-to-many identification, it searches the probe against a

[6]*Ali Haider alias Papu v. Jameel Hussain*, PLD 2021 SC 362 (Pak.).

[7]Searches were conducted up to 30 August 2026 in the Supreme Court of Pakistan judgment repository, the Lahore High Court and Sindh High Court case-law databases, the Sindh High Court Law Research Centre collection, and the Pakistan Code, using the terms "facial recognition", "face recognition", "facial comparison", "photogrammetry", "photogrammetric", "automated information system", and "Article 46-A", alone and in combination. Reported decisions were included where an electronic recording or an image-derived identification was at issue; decisions mentioning the terms only in passing were excluded. The Islamabad, Peshawar, and Balochistan High Courts and the Federal Shariat Court were not separately searched, and the finding is stated subject to that limitation.

gallery and returns candidates: nobody has asserted an identity, and the system proposes one.

The second operation carries risks that the first does not. Experimental work on one-to-many search finds that blur and reduced resolution in the probe significantly increase the false positive identification rate, and that operational variables including gallery composition shape search outcomes (Bhatta et al., 2024; Pangelinan et al., 2024). Gallery size and composition affect the probability of a false candidate; threshold choice determines whether a candidate is returned at all; ranking is not proof of correctness; and if the person depicted is not enrolled, every candidate the system still returns is wrong. NIST evaluates the two tasks separately for these reasons and distinguishes rank-based investigative use from threshold-based identification (Grother et al., 2019a). Nor is a similarity score a probability of identity: its meaning depends on the algorithm, the operation, the threshold, the gallery, and the conditions of capture, so presenting it as one gives a court a number without a scale.

Two amici curiae put the same points to the Supreme Court of New Jersey in *Miles*, and the Court set them out in its opinion. A biometrics expert, Dr. Michael King, stated that facial recognition refers not to a single system but to a range of systems, varying widely in reliability; that many of the controls and configuration settings supporting accuracy are relaxed or removed for investigative use, giving as an example that law enforcement may set a low threshold to ensure that a candidate list is fully populated; and that such systems find not a match but similarities, which are not tied to a statistical degree of confidence in the usual sense but depend on the algorithm unique to each system. An eyewitness-identification researcher, Dr. Gary Wells, contended that if the defendant's photograph was not in the database searched, the search was guaranteed to return a false negative (*Miles*, 2026). These were submissions on an interlocutory appeal decided on a limited record, not evidence tested at trial, and we cite them as specialists' statements of the points, not as findings.

Automated recognition must also be distinguished from forensic facial comparison, which FISWG treats as a manual process conducted by a human (Facial Identification Scientific Working Group [FISWG], 2026: § 5.1). When a Pakistani judgment records that facial comparison linked an accused to footage, the phrase does not specify which performed the work. In *Zahir Zakir Jaffar v. The State* (2025), it does not appear to have been an algorithm: the judgment describes a positive forensic facial comparison and identifies no model, software, or database search (paras. 5, 15). That case is not authority for admitting facial-recognition evidence, though we draw on it below for the statutory architecture, the meaning of an automated system, and the authentication doctrine.

3. Pakistan's Evidentiary Architecture

3.1. The Provisions That Govern Machine-Generated Evidence

The Electronic Transactions Ordinance 2002 amended the Qanun-e-Shahadat ra-

ther than creating a parallel regime, producing a chain of provisions to be read together.

Article 46-A makes relevant statements in the form of electronic documents generated, received, or recorded by an automated information system while it is in working order. Explanations 3 and 4 to Article 73 address primary-evidence status: output of an automated information system is not denied that status merely because it is electronic, if the system was in working order at all material times, and absent evidence to the contrary, it is presumed to have been; other electronic documents are primary evidence where a security procedure was applied. Article 78-A provides that where an electronic document is alleged to be signed, or to have been generated wholly or in part by any person through the use of an information system, and that allegation is denied, the application of a security procedure to the signature or the document must be proved. Article 59, as amended, admits expert opinion on the authenticity and integrity of electronic documents made by or through an information system, and on the functioning, specifications, programming, and operations of information systems. The Order takes its definitions from the Ordinance, "automated" meaning without active human intervention (Qanun-e-Shahadat Order, 1984, arts. 46-A, 59, 73 & 78-A; Electronic Transactions Ordinance, 2002, §§ 2(1)(g), 29 & sch.)[8]. The Supreme Court has treated these as a set: in *Mian Khalid Pervaiz v. The State* (2021)[9], it held that Articles 46-A and 78-A with the Ordinance supply the procedure for receiving such evidence and for proving it where it is denied (2021, para. 6).

The architecture is more substantial than the literature's focus on Article 164 suggests: a relevancy rule, a classification rule, a proof rule, and an expert channel, all directed at information systems. None of them addresses accuracy. Article 46-A makes machine-generated information relevant; it does not say the machine was accurate. Explanation 3 concerns the best-evidence principle, not truth. Article 78-A requires proof of a security procedure, which goes to integrity rather than correctness, and may not reach autonomous machine output at all. Article 59 admits expert opinion without saying how to assess scientific method, and Pakistan has adopted no equivalent of the *Daubert* factors; the standard has to come from the gatekeeping principle discussed at Stage II. A computational identity inference can therefore travel the entire chain—relevant, primary, presumed operational, and supported by expert opinion—without anyone showing that it is right.

The objection on which the article turns starts from the fact that Pakistan tries criminal cases without juries. A trial judge is not shielded from weak evidence by rules of admission; she hears it and weighs it, and a judge who cannot evaluate an algorithmic inference will decline to convict on one. On that view, there is no gap, only an American admissibility framing imported into a system with no use for it.

The difficulty, however, is not who decides but what the decider is shown. A

[8]Qanun-e-Shahadat Order, 1984, President's Order No. X of 1984 (Pak.); Electronic Transactions Ordinance, 2002, Ordinance No. LI of 2002 (Pak.).

[9]*Mian Khalid Pervaiz v. The State*, 2021 SCMR 522 (Pak.).

judge can discount an identification she can see through, but weight operates on what is before the court, and Explanation 3 and Section 510 together produce a record that does not show her what she would need: the machine is presumed to have been in working order, the report is admissible without its author, no threshold, gallery, or ranking is on the record, and nothing requires any of it to be. The argument is therefore about the foundation a court must be given, not about a power to exclude that Pakistani law does not need.

The defect may not be peculiar to facial recognition. Nothing in Articles 46-A, 73, 78-A, or 59 is drawn by reference to a particular technology, and none of them requires proof of accuracy for machine output of any kind. The same problem is therefore likely wherever software draws an inference rather than recording an observation: in systems that interpret a DNA profile and report a likelihood ratio, ballistics comparison systems that rank candidate barrels, mobile-device extraction tools that reconstruct deleted material, and automated number-plate recognition. That is a hypothesis, not a finding. Whether reports resting on such systems are in practice tendered in Pakistan without the software being identified, its version recorded, or its error performance stated is an empirical question we have not investigated.

Facial recognition is thus an instance of a general condition rather than a special case, though an extreme one: its output is an assertion about who a person is, made against an accused, drawn from degraded material, with no human perceiver behind it, and no established discipline governing how the assertion is reached. A court that accepts the argument for facial recognition will find it difficult to confine, and the framework in Part 5 is written to be portable for that reason.

3.2. Article 164 Does Not Fill the Gap

Pakistani writing reaches for Article 164, but the provision has never done the work attributed to it.

As originally enacted, it permitted a court, in such cases as it considered appropriate, to allow production of any evidence that had become available because of modern devices or techniques: a permission to receive, addressing neither relevance, proof, class of evidence nor reliability. After 2002, it was largely surplusage for machine-generated output. Its prominence derives from a single much-quoted sentence in *Ali Haider alias Papu v. Jameel Hussain* (2021), where the Supreme Court called it the gateway allowing modern forensic science into Pakistani courtrooms (2021, para. 4). That decision post-dates the 2002 amendments, so the surplusage claim concerns the provision's function, not judicial practice, which has continued to treat Article 164 as central. The description has produced a habit of framing every technological evidence question as an Article 164 question. The Supreme Court has since said as much: in *Zahir Jaffar*, it recorded that digital evidence was formerly admissible only under Article 164 and treated as hearsay, that the Electronic Transactions Ordinance was enacted to change that, and that the

amendments gave such evidence the status of primary evidence, before setting out Section 2(e), Explanations 3 and 4 to Article 73, and Article 46-A in full (2025, para. 10).

Section 12 of the Criminal Laws (Amendment) Act 2023 has since replaced Article 164 in its entirety, with no savings or transitional provision (Criminal Laws (Amendment) Act, 2023, § 12)[10]. The substituted text permits a court to allow any evidence or witnesses to be produced, recorded by the Court through modern devices or techniques, and enumerates video call, viber, skype, imo, whatsapp, facebook messenger, line caller, and video conference. The 2023 Act is a child protection statute, and the substituted Article is the counterpart of its regime for recording victim and witness testimony at a distance (Criminal Laws (Amendment) Act, 2023, §§ 30B(1)(iii), 30C(2)). Nothing in it suggests Parliament turned its mind to forensic evidence.

A proviso added in 2017 states that conviction on the basis of modern devices or techniques may be lawful. The consolidated Order still prints it after the substituted text, and in *Ali Haider*, the Supreme Court described it in the present tense while treating Article 164 as the gateway (2021, para. 4). *Zahir Jaffar* narrates the proviso in the past tense, as having been added in 2017, and then reproduces the substituted Article without it (2025, para. 11), leaving a genuine textual question as to whether the substitution carried the proviso away. The better view is that it did not. Pakistani law, therefore, contains an express provision that conviction on modern devices may be lawful, and none about what must be shown before a court relies on one.

The substituted provision can be parsed as permitting production of "any evidence" as a free-standing limb, in which case a general gateway survives; this reading is weak, since a provision authorising production of any evidence without qualification would displace much of the Order. On either construction, a provision whose marginal note refers to information systems and whose body enumerates messaging applications is not a secure foundation for an evidentiary architecture. Where Pakistani courts have continued to discuss Article 164 in its earlier form, the state of the records is part of the explanation: in *Ali Akbar v. Abdul Jabbar*[11] the order under challenge predated the substitution by several years, and in *Zahir Jaffar* the trial and the first appeal were concluded before 18 July 2023[12]. *Zahir Jaffar* itself did notice the change: the Court observed that the Article "stood substituted in the year 2023" and set out the new text (2025, paras. 10-11). It did not say whether the 2017 proviso survived the substitution, or what the substitution meant for a provision it had just described as the former route for digital

[10]Criminal Laws (Amendment) Act, 2023, Act No. XXXVII of 2023, Gazette of Pakistan, Extraordinary, pt. I, July 24, 2023, at 555-61 (Pak.).

[11]*Ali Akbar v. Abdul Jabbar*, C.P. No. D-965 of 2024 (High Ct. Sindh, Circuit Ct. Mirpurkhas, Feb. 18, 2025) (Pak.).

[12]*Ali Akbar v. Abdul Jabbar* (2025) reviewed an order of the District Judge made in October 2019. In *Zahir Zakir Jaffar* (2025), the judgment under appeal was that of the Islamabad High Court dated 13 March 2023, on appeals and a murder reference instituted in 2022.

evidence; it turned instead to the Silent Witness doctrine.

3.3. What "Automated" Means

Article 46-A's reach depends on a defined word, which has been construed twice: by the Supreme Court in passing and at length by a High Court.

In *Zahir Jaffar*, having set out Article 46-A and Explanations 3 and 4 in full, the Supreme Court recorded that the recording before it was automatically generated without human intervention (2025, para. 15). That applies the Ordinance definition to recorded evidence in a criminal case, establishing that the category is live in criminal practice and not confined to commercial documents. What the Court described was a camera rather than a search, so the application to facial recognition is ours. It does not rest on comparing how much human activity each involves, since an algorithmic search requires more discrete human choices than a recording: the probe must be selected, often processed, a gallery chosen, a threshold set, and the query run. The statute asks which part of the evidentiary proposition was produced without active human intervention, and for a candidate list the answer is the whole of it, because the system alone performs the comparison and the ranking.

In *Mst Sanobar Rizwan v. Mst Naseera Begum* (2025)[13], the Sindh High Court, deciding whether printouts of emails were primary or secondary evidence, had to determine what an automated information system is. The Court distinguished a manual system, in which data is entered by a human being, from an automated one, in which information is entered through an automated process: the illustration given was scanning coupled with a program of algorithms, whereby a machine reads the data itself and processes it, devoid of human intervention (2025, paras. 13-14). It held that an email, because of the active human intervention in its creation, falls outside the category and is governed by Explanation 4 (*Sanobar Rizwan*, 2025, para. 17). The decision is a single-judge order in a civil suit and construes the word for Explanation 3 rather than Article 46-A, so it elaborates what the Supreme Court applied rather than supplying the authority for it.

Its illustration, a system that scans images and applies an algorithm to read and process them without a human entering the content, is close to a description of an algorithmic search. A bare one-to-many search produces a candidate list that nobody has typed, and on this approach that list is the output of an automated information system to which Article 46-A applies.

The question is harder where a human is involved. Prosecutions rarely tender candidate lists: an examiner receives one, looks at the images, and forms a view, which is not automated at all. Article 59 fits the examiner's conclusion; Article 46-A fits the list he looked at, and the two should not collapse. Where a conclusion is formed with knowledge of the machine's ranking or score, the reviewer is not an

[13] *Mst Sanobar Rizwan v. Mst Naseera Begum*, 2025 SHC KHI 495, 2025 SBLR Sindh 568 (Sindh High Ct., Karachi, Jan. 18, 2025) (Pak.).

independent check but a partly dependent one, of unknown degree (Howard et al., 2020). A court has reached the same conclusion on similar facts: in *State v. Ghigliotty*[14], as described in *Arteaga*, the New Jersey Appellate Division held that a hearing into novel ballistics-comparison software was required, although the State argued that its expert's conclusions rested on a standard comparison microscope, because the software had caused him to change his conclusion and had clearly aided and influenced the course of his investigation and informed his ultimate opinion, and he conceded that he could not reach a conclusion by the traditional method until he had used it (2020). A prosecution should not be able to convert machine evidence into expert evidence, and so escape the questions Article 46-A raises, by having a person look at the screen.

Article 78-A raises a related difficulty. Its proof rule is engaged where a document is alleged to have been generated by any person through the use of an information system, but a bare candidate list is generated not by a person but by a system a person has operated. The prosecution can argue that the officer who selected the probe, chose the gallery, and initiated the query generated the output through the use of the system, and that argument will often succeed. But the text does not compel that reading, and where the operator cannot be identified, the argument may fail altogether. The chain is therefore less complete than Part 3.1 suggests: relevance and classification reach machine output squarely, while the proof provision may not reach it at all. That gap in the Order makes the foundation question more pressing.

Where the item tendered is the examiner's conclusion, *Sanobar Rizwan*'s reasoning places it outside the automated category and the presumption discussed next does not arise. However, the candidate list remains the material on which the opinion rests, and a prosecution relying on the opinion relies on the machine result behind it. The presumption problem is therefore acute where a bare machine result is tendered, and operates at one remove where an opinion is built on one.

3.4. The Presumption of Working Order and Its Limits

Article 46-A conditions relevance on the system being in working order, Explanation 3 supplies the presumption that satisfies it, and *Sanobar Rizwan* gives the presumption operative content: it must be rebutted by the person disavowing the output, who must adduce evidence of a defect at a material time (2025, para. 14).

Applied without qualification to facial recognition, this means that the candidate list enters as primary evidence, the system is presumed to have been in working order, and the accused carries the burden of showing otherwise, inverting the condition proposed in Part 7 before it can operate.

One objection must be addressed first. "In working order" is not confined to Explanation 3; it is the condition on which Article 46-A makes machine-generated statements relevant at all, and a court will ordinarily give one phrase one meaning

[14] *State v. Ghigliotty*, 463 N.J. Super. 355 (App. Div. 2020).

in a substantive provision and in the Explanation that presumes it satisfied. Read narrowly, "working order" means the system was running, in which case Article 46-A is satisfied on proof of almost nothing, and the Order contains no accuracy requirement anywhere. Read purposively, it means the system was functioning correctly for the task performed, in which case Article 46-A does demand something close to accuracy, but Explanation 3 then presumes precisely what needs proving. We take the narrow reading, at a cost: the accuracy requirement must then come from outside the presumption, from Article 59 with the gatekeeping principle, from the Silent Witness doctrine, and from Article 10A, which Parts 3.5 and 5 take up.

Four considerations prevent the inversion.

The presumption concerns the machine running, not the inference being right. *Sanobar Rizwan* frames rebuttal in input and output terms: a defect at the point the document is extracted, or at the point the underlying data is entered. Neither route touches the correctness of the algorithm's processing. That is unsurprising for the systems Explanation 3 was drafted around, where a functioning machine produces a correct record and a malfunctioning one does not. Facial recognition breaks the equivalence: a system in perfect working order will return a wrong first-ranked candidate from a degraded probe searched against an over-inclusive gallery, with a number attached.

The presumption also should not be enlarged. In *Commissioner Inland Revenue, Lahore v. Millat Tractors Limited* (2024)[15], the Supreme Court explained that the purpose of an Explanation is ordinarily to explain a concept in the main provision, though the legislature may use one to give that concept an extended or a restricted meaning; that a clarificatory Explanation is to be interpreted according to its own terms and not so as to widen the ambit of the provision; and that, as a general rule, an Explanation is not a substantive provision but a guide to construction (2024, para. 14; discussed in *Sanobar Rizwan*, 2025, para. 12). The rule is general, not absolute: in the case itself the Court looked past an Explanation's declaratory wording to its intention and effect, held that it took away a substantive right, and for that reason refused it retrospective operation (2024, paras. 13, 15). The case therefore cuts both ways; its method is to ask what an Explanation was intended to do and what it does. Explanation 3 was intended to settle whether output generated in electronic form may be classed as primary evidence. Even if the presumption it contains is given substantive effect, its subject is the operation of the system, not the correctness of an inference the system draws; to read it as a presumption that an identification is accurate would widen it beyond both its terms and its purpose.

The obvious retort is that if an Explanation is not ordinarily a substantive provision, it cannot confer primary-evidence status, which the Supreme Court and this article both rely on Explanation 3 to do. But Explanation 3 does not confer

[15] *Commissioner Inland Revenue, Lahore v. Millat Tractors Limited, Lahore*, 2024 SCMR 700 (Pak.).

anything: it provides that output shall not be denied the status of primary evidence solely because it was generated in electronic form. It removes a disqualification that the general law of primary evidence might otherwise impose, which is the clarificatory function an Explanation to Article 73 is fitted for. A presumption operating against an accused on the issue that decides the case is a different kind of instrument, and nothing in the drafting suggests Explanation 3 was meant to be one.

The third consideration is practical. Explanation 3 presumes working order at all material times, but for a facial-recognition search those times are not obvious, since enrolment of the gallery, capture and processing of the probe, and the search itself are distinct events. A presumption directed at a system with one operating state does not transpose cleanly onto a process with several.

The fourth consideration is constitutional, and stronger than the others because it does not depend on how the Order is construed. Article 10A guarantees due process and a fair trial (Constitution of the Islamic Republic of Pakistan, 1973, art. 10A)[16], and the presumption of innocence is among the incidents of that guarantee. Article 117 also carries it: in *Raja Khurram Ali Khan v. Tayyaba Bibi* (2020, para. 54)[17], the Supreme Court held that Article 117 enshrines the foundational principle of the criminal justice system, whereby the accused is presumed innocent unless proved otherwise, that the burden of proving guilt beyond doubt is placed on the prosecution as the legal burden, and that this burden can never be shifted to the accused unless the legislature by express terms commands otherwise. The evidential burden moves afterwards and separately, under a provision such as Article 122, and only once the prosecution has discharged the legal burden by establishing the elements of the offence. A rule of evidence that relieves the prosecution of showing that a machine's inference against an accused was correct, and requires the accused to show that it was not, would partially reverse that presumption on the very issue on which the prosecution's case turned. An Explanation whose function is to classify documents is an implausible vehicle for that result.

The claim needs qualification: Pakistani law contains presumptions against accused persons that have not been struck down. Section 29 of the Control of Narcotic Substances Act 1997 provides that in trials under the Act it may be presumed, unless and until the contrary is proved, that the accused has committed an offence under the Act in respect of any narcotic drug, psychotropic substance, or controlled substance for the possession of which he fails to account satisfactorily (Control of Narcotic Substances Act, 1997, § 29)[18]. If Parliament may do that, Article 10A does not forbid every presumption bearing on guilt. But Section 29 is what the Supreme Court's formulation contemplates: a legislature commanding otherwise, by express terms, in a substantive provision, for a defined offence, with

[16]Constitution of the Islamic Republic of Pakistan, 1973 (Pak.).
[17]*Raja Khurram Ali Khan v. Tayyaba Bibi*, PLD 2020 SC 146 (Pak.).
[18]Control of Narcotic Substances Act, 1997 (Pak.).

the policy open to debate and to constitutional review. Explanation 3 is not: it commands nothing about guilt, its express terms concern whether electronic output may be classed as primary evidence, and any effect it has on the burden of proving an identification is a by-product of a drafting device about the classification of documents. On the test *Raja Khurram* states, it cannot shift the legal burden at all, because the legislature has not commanded otherwise by express terms.

3.5. The Silent Witness Doctrine, and Where Authentication Stops

A doctrine the technology literature has not connected to this problem bears directly on it. In *Zahir Jaffar*, the prosecution case was circumstantial, and CCTV and DVR recordings were among the evidence on which it rested (2025, para. 5). The Supreme Court held that CCTV footage may also be admitted under the Silent Witness theory, which permits recorded material to serve as substantive proof of what it depicts without a sponsoring witness, after authentication by evidence supporting the reliability of the process or system that produced it (*Zahir Zakir Jaffar*, 2025, paras. 12-13, applied at para. 15). Primary-evidence status comes separately, from the 2002 amendments rather than the doctrine. The Court further held that CCTV satisfying the two-fold test it attributed to *Ahmed Omar Sheikh* requires no corroboration, being primary evidence (*Zahir Zakir Jaffar*, 2025, paras. 9, 16).

Pakistani law therefore contains a reliability inquiry directed at a machine: where there is no human perceiver to cross-examine, the reliability of the recording system takes the place of the witness's credibility. The Court's sources for the doctrine are Black's Law Dictionary and a line of American decisions, among them *Harris*, *Taylor*, *Rembert*, and *Fisher*, rather than the Federal Rules. The Court's phrase is the reliability of the process or system that produced the material, but as those decisions apply it, the doctrine is narrower than the phrase suggests: the federal rule states it as evidence describing a process or system and showing that it produces an accurate result (Fed. R. Evid. 901(b)(9)), and the Court's own account of *Rembert* turns on evidence establishing the reliability of the camera and its functioning. For a camera, the two ideas are interchangeable; for a system that draws conclusions, they come apart. If the doctrine has been received with its rationale and not merely its label, the reliability it requires is reliability at what the system does, which is to produce an identification. We press this no further than the materials allow: the Court was authenticating a camera and had no occasion to distinguish the two ideas.

There is a limit. The process the Court trusted was a camera, which captures rather than concludes; reliability evidence adequate for a camera concerns installation, operation, storage, continuity, and freedom from alteration, none of which bears on the accuracy of an inference.

Whether the doctrine extends to a system that draws conclusions is undecided. In favour: on the reading just given, the evidence required is evidence appropriate

to what the machine does. Against: the doctrine was developed for systems whose failure modes are malfunction, tampering, and gaps in the record, and stretching it invites courts to treat a general finding that the system is reliable as disposing of a particular match. We favour extension with the discipline Part 5 supplies, on condition that courts keep distinct the reliability of a system in general and of a result in a particular case.

Muhammad Ammar Shafi v. The State (2025)[19] points the same way. The Lahore High Court found deficiencies in the source and proof of Safe City material, including the absence of any person who prepared the CD and photographs (para. 9), and separately observed that a photogrammetry test from the Punjab Forensic Science Agency should have been obtained to establish that the appellants were the men visible in the footage (*Ammar Shafi*, 2025, para. 11). Authentication concerned the recording; identification required its own foundation. The judgment names the technique without describing what it would involve or what the institution's method would be, and we do not treat it as authority on facial-comparison methodology.

3.6. Identification Doctrine Already Knows This Problem

Pakistani law has a developed doctrine on the contamination of identity evidence that has not been connected to this technology. Article 22 makes relevant those facts which establish the identity of any thing or person whose identity is relevant (Qanun-e-Shahadat Order, 1984, art. 22). It supplies relevance only; the safeguards come from judicial guidance and executive instructions. In *Kanwar Anwaar Ali* (2019)[20], the Supreme Court consolidated the requirements for a test identification parade and expressly approved a statement of them by a Division Bench of the Lahore High Court in *Muhammad Yaqoob*: that identification evidence is not substantive but only corroborative; that a parade is necessary only where the offender was a complete stranger to the witness; that such proceedings are a test of the witness's memory; and that a test identification in which the possibility of the witness having seen the accused after arrest cannot be ruled out "is worth nothing at all" (2019, para. 3). The Court gave the same reason in its own words in describing identification in the dock as generally unsafe: the witness will have had many opportunities to see the accused at earlier appearances (2019, para. 6). It added that failure to hold a parade is not always fatal to the prosecution (2019, para. 7), but where one is held its safeguards are, in the Court's words, "to be meticulously followed" (2019, para. 3).

The prior-exposure proposition maps onto an algorithmic search wherever a human confirmation follows it. A one-to-many search returns the accused as a ranked candidate, and an officer or a witness then looks at him and confirms the match, having been shown the suspect first by a machine that also expressed its

[19] *Muhammad Ammar Shafi v. The State*, 2025 YLR 1681 (Lahore), 2024 LHC 6125 (Lahore High Ct. Jan. 3, 2025) (Pak.).
[20] *Kanwar Anwaar Ali, Special Judicial Magistrate: In the matter of*, PLD 2019 SC 488 (Pak.).

confidence. The analogy is functional rather than exact, and its force depends on the sequence. It is strongest where a candidate is afterwards shown to an eyewitness, where an examiner reviews a ranked candidate knowing the rank or score, or where a later human identification is tendered as independent although the machine has primed it. Where a bare machine result is tendered with no human confirmation, parade doctrine supplies an analogy about the dangers of suggestion, not a governing rule. The experimental work establishes that a prior identity decision shifts the later human judgment, including where that decision is attributed to a computer (Howard et al., 2020); it does not establish that a number carries more authority than a photograph, and we do not rely on that stronger claim.

The concern has a partial counterpart in *Arteaga*, where the New Jersey Appellate Division held that the reliability of the software that had placed the defendant in a photo array has obvious implications for the accuracy of the identification, because an array constructed around a mistaken potential match would leave the witness with no actual perpetrator to choose (2023). But it also held that the reason a particular person was a suspect, and so in the array, was not relevant to whether the procedure had been impermissibly suggestive, and it located the significance of the software in impeachment of the identification, challenge to the investigation, and third-party guilt rather than in the validity of the procedure. *Arteaga* therefore supports the accuracy concern and disclosure, not the treatment of a machine-selected candidate as itself a suggestive procedure. The *Kanwar Anwaar Ali* concern, that the witness may have been shown the suspect before being asked to find him, reaches further than the New Jersey rule.

Where an algorithmic search precedes a parade and its result has been shown to the witness, the fact and content of that search bear on the parade's worth and must be disclosed for that reason alone. Where no parade is held because the machine has supplied the identification, a court should be slow to let the machine result do the parade's work, since none of the safeguards that give a parade its worth (dummies, confidentiality, a contemporaneous record, the absence of prior exposure) has any counterpart in a candidate list.

3.7. Three Institutional Tracks

Facial-recognition capability is in use in more than one Pakistani jurisdiction, but Punjab is where the institutional arrangements are documented well enough to trace how such evidence would reach a court. There are three routes, depending on which institution produced the evidence.

On the first, an automated one-to-many search is run through Face Trace, developed by PITB with the Punjab Police IT Wing and operated by police. Its output is generated by an automated information system, so Article 46-A and Explanation 3 apply: relevant, primary, presumed in working order. However, the operator is a police officer, not an expert, and we have located no publicly available forensic protocol, validation study, or documented methodology governing the

evidentiary use of the system[21].

On the second, a photogrammetric test or forensic facial comparison is conducted by the Punjab Forensic Science Agency within its Audio-Visual Analysis function. Section 9(3) of the Agency's Act deems its experts to be experts appointed under Section 510 of the Code of Criminal Procedure and persons specially skilled in forensic material under Article 59, and Section 510 was amended in 2022 to add the forensic scientist to the experts whose reports may be used without calling the maker (Punjab Forensic Science Agency Act, 2007, § 9(3); Code of Criminal Procedure, 1898, § 510, as amended by Code of Criminal Procedure (Amendment) Act, 2022, § 2)[22]. The report is therefore admissible without the examiner, subject only to the court's power to summon under the proviso. The Supreme Court confirmed that route in *Ali Haider*, explaining that Section 510 lists certain experts only, so a DNA report is not per se admissible in itself, but that because Section 9 of the Punjab Forensic Science Agency Act deems an Agency expert to be an expert under Section 510, a report prepared by one is (2021, paras. 7-8). The same passage explains that per se admissibility is a procedural facility for tendering evidence and does not bear on whether the document is admissible under the Order at all.

On the third, material captured by the Punjab Safe Cities Authority's own cameras reaches a court under a regime of its own, which the Lahore High Court set out in *Ammar Shafi*. Section 14(2) of the Authority's Act provides that an employee of the Authority working with information technology-related matters shall be deemed to be an expert within the meaning of Article 59 and to have been appointed under Section 510 of the Code of Criminal Procedure (Punjab Safe Cities Authority Act, 2016, § 14(2))[23]. Under the Punjab Safe Cities Authority Electronic Data Regulations 2016, framed under Section 19 and published by gazette notification in December 2016[24], footage is requisitioned on an Electronic Data Request Form and supplied with an Electronic Data Certificate, defined as a document issued by an authorised officer certifying the genuineness of the electronic material recorded at the integrated command centre, with a detailed report of the protocols applied for its collection, extraction, and submission. Regulation 9 makes that certificate the Authority's formal confirmation of the genuineness of the data, requires it to carry the technical details, and provides for clarification and re-examination of the expert opinion by a panel at the court's direction. The High Court concluded that data obtained through that process is deemed genuine

[21]Searches of the Punjab Information Technology Board and Punjab Police published materials, and of the Punjab Forensic Science Agency site, to 30 August 2026. Absence of a published protocol is not evidence that no operational protocol exists.

[22]Punjab Forensic Science Agency Act, 2007, Punjab Act No. XIII of 2007 (Pak.).; Code of Criminal Procedure, 1898, Act No. V of 1898 (Pak.); Code of Criminal Procedure (Amendment) Act, 2022, Act No. XXXIII of 2022, Gazette of Pakistan, Extraordinary, pt. I, Nov. 18, 2022, at 1181-82 (Pak.).

[23]Punjab Safe Cities Authority Act, 2016, Punjab Act No. XXXVII of 2016 (Pak.).

[24]Punjab Safe Cities Authority Electronic Data Regulations, 2016, Notification No. 6027/PSCA/2016, Gazette Notification No. 178/2016 (Dec. 2, 2016) (Pak.).

and admissible without being sent to the Punjab Forensic Science Agency at all (*Ammar Shafi*, 2025, paras. 12-13).

Each route defeats scrutiny at a different point. Face Trace yields a machine result presumed sound with no expert behind it; the Agency, an opinion with no expert in the box; the Authority, material certified genuine by its own officer on a certificate whose stated subject is the protocols of collection and extraction and nothing else. None is required to produce evidence that an identification is correct, and the third makes the gap explicit, because a document whose express subject is genuineness cannot be read as speaking to accuracy.

Having set out the regime, the High Court held that such electronic data "shall be read in evidence in conjunction with other explanation like photogrammetry test etc. of accused visible therein" (*Ammar Shafi*, 2025, para. 13). That is this article's central distinction, stated by a Pakistani court about a Pakistani certificate: material may be genuine, admissible, and unimpeachable as a record and still establish nothing about who is in it. The Court reached it for a recording rather than for an inference, and the extension to algorithmic identification is ours; the premise itself is domestic and expressly held.

The difficulty with the Agency route is compounded by the absence of any public account of its method. The Agency's published description of Audio-Visual Analysis lists clarification of audio and video evidence and analysis of images extracted from video, mentioning neither facial comparison nor photogrammetry (Punjab Forensic Science Agency, n.d.). It nevertheless performs such work: reporting on the investigation into the events of 9 May 2023 quotes a supplementary report stating that a photogrammetric test or forensic facial comparison could be conducted if investigators supplied reference video ("Punjab Forensic Science Agency Report," 2023)[25], and in *Ammar Shafi* the Lahore High Court proceeded on the footing that such a test was available. The work is not thereby improper, but we have located no published methodology, validation study, or quantified error rate for a test a High Court has directed a party to obtain and whose report is admissible without its author. The defence may seek laboratory documentation through the routes discussed at Stage IV, and Part 5 argues that it should; but nothing in the present position requires any of it to be produced unprompted.

In *Honeysett v. The Queen* (2014)[26], the High Court of Australia quashed a conviction and ordered a new trial, where the prosecution had relied on an anatomist's comparison of a robber in security footage with the accused. Although the witness was a qualified anatomist, his opinion was held not to be based on specialised knowledge: he had looked carefully at the images and described his impressions in anatomical terms, his expertise did not extend to the comparison he had actually performed, and his evidence gave an unwarranted appearance of science to matters of ordinary observation (2014, paras. 23-24, 42-46). Australian

[25]Punjab Forensic Science Agency Report Shows the Sequence of Events on May 9. (2023, July 3). Dawn. Punjab Forensic Science Agency Report.

[26]*Honeysett v. The Queen*, [2014] HCA 29, (2014) 253 C.L.R. 122 (Austl.).

scholarship on image-comparison evidence has made the same point at greater length (Edmond et al., 2009). The lesson for Pakistan is that an institution's standing, and the statutory admissibility of its reports, say nothing about whether the method used on the day was one the institution's expertise supports; where the method is unpublished, nobody is in a position to ask.

3.8. The Comparison Pakistani Courts Would Reach First

Indian authority on the cognate provisions is the natural comparator and is routinely persuasive in Pakistani courts, the two statutes descending from the same 1872 parent. Section 65B of the Indian Evidence Act 1872[27] admitted electronic records on a certificate from a person occupying a responsible position, deposing to the conditions of production. In *Anvar P. V. v. P. K. Basheer* (2014)[28], the Supreme Court of India held the certificate mandatory for secondary electronic evidence, and in *Arjun Panditrao Khotkar v. Kailash Kushanrao Gorantyal* (2020)[29], a bench of the same strength reaffirmed it, overruling the contrary line. The Court rejected as wholly incorrect the reasoning that a party not in possession of the device cannot be required to produce the certificate, holding that it may apply to the court to compel the certificate from the person who can give it, and clarified that the certificate is unnecessary where the original is itself produced, as where the owner of the device steps into the witness box and proves it; on the facts, a party that had done everything possible to obtain a certificate and been refused was relieved of the obligation (2020, paras. 39, 43, 49, 72). The scheme has since been carried into sections 62 and 63 of the Bharatiya Sakshya Adhiniyam (2023)[30].

India addressed the problem of machine output entering a trial without a human sponsor by requiring someone to attest to the conditions of production. So, within its own field, did Punjab: the Electronic Data Certificate examined in Part 3.7 is issued by an officer of the Authority and attests to the protocols by which the material was collected and extracted. The Order's general regime presumes what the Indian statute required, but a foundation condition is already at work within this statutory family, in the province whose systems this article analyzes.

Both certificates also show the limits of that answer. Section 65B required a deposition about the computer's operation and the regularity of its use; Regulation 9 requires confirmation of the genuineness of the data and of the protocols of its extraction. Each attests to integrity, not accuracy: neither would say whether a facial-recognition identification was correct. Transplanting either would give Pakistan a more rigorous version of the wrong requirement. The foundation must instead be specified by reference to what the machine does, as Part 5 argues.

[27]Indian Evidence Act, 1872, No. 1 of 1872 (India).
[28]*Anvar P.V. v. P.K. Basheer*, (2014) 10 S.C.C. 473, [2014] 11 S.C.R. 399 (India).
[29]*Arjun Panditrao Khotkar v. Kailash Kushanrao Gorantyal*, (2020) 7 S.C.C. 1 (India).
[30]Bharatiya Sakshya Adhiniyam, No. 47 of 2023 (India).

4. China as a Limited Comparator

China has moved further than Pakistan toward express legal treatment of facial recognition as a distinct object. The claim is narrow: China has not solved the evidentiary problem, and Chinese scholarship says so. The comparator is chosen for institutional rather than functional similarity: in both places, facial recognition has been deployed by police at scale under state direction, ahead of any evidentiary rule addressed to it.

Its usefulness lies in classification. The 2025 Measures (Cyberspace Administration of China [CAC] & Ministry of Public Security [MPS], 2025) separate verification of personal identity, meaning one-to-one comparison against a specific stored face, from identification of a specific individual, meaning one-to-many comparison against a defined gallery (art. 19(3), (5) & (6))[31]. They are a personal-information instrument imposing notice, consent, assessment, and filing duties, excluding from their scope processing for research and algorithm training (CAC & MPS, 2025: art. 2), and are silent on criminal evidence and accuracy. The lesson is that once a system must name the operation performed, the difference between confirming an asserted identity and generating one becomes visible. China also maintains national standards for facial-recognition systems, cited here for their existence and institutional role rather than their content[32], and accreditation has improved the credibility of appraisal results without displacing judicial assessment (Guo & Hou, 2018).

Chinese criminal procedure supplies less. The 2025 Identity Review Opinion (Supreme People's Court et al., 2025) requires identity to be verified by information system comparison, archival comparison, or comparison of biometric information (art. 2(3))[33], allows identity to be confirmed where the biometric information is determined to be consistent with the matched information (art. 3(2)), and makes documentary proof primary with biometric comparison as a fallback (art. 4(3)). Facial recognition is named only in Article 5, which defines identity that genuinely cannot be ascertained as identity still unverified after household-registration enquiry, on-site investigation, facial recognition, fingerprint comparison, and DNA appraisal. It appears as a technique whose exhaustion licenses proceeding without an established identity, so accounts reporting that Chinese criminal procedure incorporates it into identity verification overstate the text. Article 3 is the useful provision: somebody must determine that the biometric information is consistent, though the instrument does not say who, on what material, or by what standard. The Opinion governs a person already within the criminal

[31]English rendering checked against the Chinese text at https://www.cac.gov.cn/2025-03/21/c_1744174262156096.htm; the unofficial translation at China Law Translate numbers both definitions in Article 19 as (5).

[32]National standards issued by the Standardization Administration of China: GB/T 41772-2022; GB/T 42981-2023; GB/T 44248-2024; GB/T 38671-2020; GB/T 41819-2022. Public security industry standards issued by the Ministry of Public Security: GA/T 1756-2020; GA/T 1470-2018; GA/T 1324-2017; GA/T 1325-2017. Cited by number, title, and issuing body only; the authors have not read the clause-level text of these standards.

[33]Chinese text as released by the Supreme People's Court; English renderings are unofficial.

process, a closed-gallery problem, and the 2016 electronic data provisions address collection and examination, not accuracy (Fa Fa [2016] No. 22)[34].

Wang's (2026) study of 396 criminal judgments concludes that AI evidence is absent from Chinese law as a category, though it operates as one in practice (Wang, 2026), so China's position resembles Pakistan's more than it differs. The comparison yields three things: name and classify the technology, build testing capacity outside the courts, and keep certification separate from judicial judgment. It yields no admissibility test, no disclosure rules, and no access rights.

5. A Framework for Pakistani Courts

The framework is normative; no Pakistani court has adopted it, and we identify the source of each component as we proceed. Two questions come first.

What role is the evidence playing? The technology may generate a lead that drops out once identity is proved independently, corroborate an identification established elsewhere, or itself be the evidence of identity. FISWG's taxonomy tracks the distinction, treating investigative and operational lead comparisons as generally not intended for presentation in court and forensic comparisons as assisting a trier of fact (FISWG, 2026: §§ 5.1.2.3-5.1.2.4). Where the search produced a lead and identity was later proved independently, a detailed reliability contest is usually pointless. Where the output is offered as corroboration or substantive proof, the more the prosecution's case depends on the algorithm, the more of the algorithm must be shown. Disclosure does not track this distinction: material about a search used only for lead generation may still bear on the thoroughness of the investigation, the suggestiveness of a later identification, and third-party guilt. That is the position the New Jersey Supreme Court reached in *Miles*, where the obligation applied although the prosecution did not intend to tender the evidence at trial (2026).

Is lawfulness in issue, and separately? Lawfulness is not a component of scientific reliability; an unlawfully obtained image can produce an accurate match, and a lawfully obtained one a false match. Pakistan has no single acquisition code for facial imagery, so the rules are source-specific: the Investigation for Fair Trial Act 2013 for warranted surveillance, Chapter III of the Prevention of Electronic Crimes Act 2016 as amended for preservation, search and seizure, and real-time collection, and the Punjab Safe Cities Authority Act 2016 for the provincial camera infrastructure (Investigation for Fair Trial Act, 2013, §§ 9-11; Prevention of Electronic Crimes Act, 2016, ch. III, as amended by Prevention of Electronic Crimes (Amendment) Act, 2025; Punjab Safe Cities Authority Act, 2016)[35]. Where a concrete legality challenge is raised, the question is which rule governed the con-

[34] *Provisions on Several Issues Concerning the Collection, Extraction, Examination and Judgment of Electronic Data in Criminal Cases*, Fa Fa [2016] No. 22 (China).

[35] Investigation for Fair Trial Act, 2013, Act No. I of 2013 (Pak.).; Prevention of Electronic Crimes Act, 2016, Act No. XL of 2016 (Pak.), as Amended by the Prevention of Electronic Crimes (Amendment) Act, 2025, Act No. II of 2025 (Pak.).

duct and what consequence it attaches to breach. We do not propose a general exclusionary rule, and illegality should not be relabelled as unreliability.

Two terms are used in a fixed sense. *System reliability* means the validated performance of the system as a class of instrument, established by testing independent of the case: what the model does, on what kind of material, at what threshold, and with what error rates. *Reliability of this result* means the correctness of the particular identification produced on the particular probe, gallery, and conditions before the court. A system may be reliable in the first sense and its result wrong in the second, and the difficulty this article addresses lies in that gap. Stage II asks the first question and Stage III the second; neither answers the other.

Stage I—Provenance and integrity of the input

This stage is almost entirely domestic and asks whether the material fed to the algorithm can be reconstructed: where the image came from, who extracted it, whether the original was preserved, and whether the image analysed traces back to that source and remained intact between capture and analysis. Existing authentication doctrine supplies this, and more than the summaries suggest. In *Ishtiaq Ahmed Mirza v. Federation of Pakistan* (2019)[36], the Supreme Court set out the requirements for proving an audio tape or video, and in *Ahmed Omar Sheikh* (2021, para. 15), it reproduced and applied them. *Zahir Jaffar* (2025, para. 9) compresses that list into a two-fold test, an account of source and a forensic report of non-alteration, which it attributes to *Ahmed Omar Sheikh*. The two limbs do not appear in that form in the majority judgment; the nearest statement is in Yahya Afridi J's dissenting opinion, which read *Ishtiaq Ahmed Mirza* as holding that failure to disclose a recording's source or to confirm forensically that it was not tampered with "would render its admissibility at naught" (2021, para. 68). The source-and-proof requirements were in any event applied in *Ammar Shafi* (2025, para. 9), with Article 78-A supplying the route to proof where the allegation is denied.

Two of the full guidelines decide this article's question. The first is that the accuracy of the recording must be proved, with satisfactory evidence produced so as to rule out any possibility of tampering. Paired with tampering, accuracy means fidelity of the record to the event, not correctness of any conclusion drawn from it: the distinction this article is built on. The Court did not draw it in terms, and we do not claim that it did; the guideline's pairing carries it, in the one place the word occurs.

The second requires that the person who recorded the event be produced and identify the person seen, or that the person seen be identified by some other person who recognizes him, and that the persons shown be properly identified. Pakistani law already requires that identity be established, and every route it contemplates runs through a human being who recognizes the person. An algorithmic

[36] *Ishtiaq Ahmed Mirza v. Federation of Pakistan*, PLD 2019 SC 675 (Pak.) (Guidelines Reproduced and Applied in *The State through P.G. Sindh v. Ahmed Omar Sheikh*, 2021 SCMR 873, para. 15).

identification fits none of them. The gap is a mismatch rather than an absence of doctrine: the requirement exists, and the technology satisfies it in no way the guidelines anticipate.

Probe images are frequently processed before searching, whether cropped, resized, sharpened, or denoised. Processing is often necessary and not improper, but if the court is not told what was done, neither it nor the defence knows what the algorithm analysed; FISWG maintains guidance on the point (FISWG, 2025c). The record should identify the original, the processed material, and the steps between them.

Stage II—System reliability

This stage asks whether the system has a scientific foundation adequate to the task attributed to it. Its domestic anchor is Article 59, which makes expert opinion on the functioning, specifications, programming, and operations of information systems a relevant fact, with the gatekeeping principle in *Ali Haider*: a judge, and more so a trial judge, is the gatekeeper of scientific evidence, and courts should embrace new techniques provided the technique or device is well established and widely accepted in the scientific community as credible and reliable (2021, para. 4). That formulation is closer to general acceptance than to a multi-factor test, and on its face exhorts courts to receive new techniques rather than scrutinize them. We read it as a condition of admission, a use the Court did not have in mind but which the proviso makes available. We do not resolve whether one-to-many facial recognition, as deployed in policing, meets that condition; *Archambault* suggests that the condition is the wrong one to rely on alone. The Minnesota Frye-Mack standard has two prongs: whether the underlying science is generally accepted in the relevant scientific community, and whether the particular evidence has foundational reliability (2024, slip op. at 11). The court found facial recognition generally accepted, and the process in issue of using the tool to generate a candidate list for human review generally accepted too, but excluded the evidence on the second prong, holding that the tool did not consistently and reliably produce accurate results (2024, slip op. at 11, 13-14). Its reasoning went to the tool's design as well as its use: the system reported a confidence score but applied no threshold and returned low-scoring results anyway; it was designed to produce possibilities rather than accurate results, and so produced false positives by design; a large dataset searched with a low threshold or none will always return false positives; and an analyst's review of the results did not make the process more reliable (2024, slip op. at 5, 7, 14-18, 21). A court applying *Ali Haider*'s general-acceptance element at the level of the technique would therefore risk admitting what *Archambault* excluded, which is the clearest reason for giving this stage content beyond general acceptance.

The division between this stage and the next is PCAST's distinction between foundational validity and validity as applied for forensic feature-comparison methods (President's Council of Advisors on Science and Technology, 2016: pp. 4-6, 47-56); we offer only its localization into the Order. The stage's content comes

from facial-recognition science, not from *Ali Haider*, and it has an external check. In *Miles*, the Supreme Court of New Jersey held that, on the facts before it, the State must produce information identifying the facial-recognition tools and materials it had used—the name and manufacturer of the software, the version numbers, the years of development, key metrics of performance including publicly available documentation of error rates, and information as to the database and the operator—and observed that such basic information, together with information on how the tools were used, will in most cases be the minimum necessary to safeguard the right to a fair trial, while declining to lay down a general checklist (2026). That list was reached independently, in another legal system and on a discovery question rather than one of foundation, yet it corresponds closely with what follows, which is some evidence that these are the things anybody must know before an algorithmic identification can be assessed. The proponent should identify the system, provider, and model or version, since "facial recognition software" describes a category rather than a technique, and the record should state whether the operation was verification or identification. Validation should be independent of the vendor and relevant to the conditions of the case: performance on cooperative frontal photographs does not transfer to degraded surveillance video, which is why NIST evaluates video separately (Grother et al., 2017). The court should receive error performance appropriate to the operation and be told the operative threshold where one determined whether a candidate was returned or treated as a match (FISWG, 2022). Whatever demographic performance data exists should be produced, with materiality argued rather than assumed (Grother et al., 2019b); FISWG's deployment-testing and implementation guidance indicates what operational assurance for a fielded system looks like (FISWG, 2025a, 2025b).

Validation will rarely match the case: published testing uses the material the testers had, and a Pakistani prosecution will involve a probe, gallery, and capture conditions that no evaluation reproduces. The mismatch is a reason neither to disregard the validation nor to treat it as though it transferred. A court should ask three things of it. First, in which direction does the difference run? Testing on cooperative frontal images against a surveillance probe overstates performance, and a court can say so without knowing by how much. Second, does the evidence identify the variables that drive the difference? Probe blur, reduced resolution, and gallery size are known to raise the false positive identification rate in one-to-many search (Bhatta et al., 2024), and a validation silent on them is silent on the conditions that matter most. Third, has the proponent offered anything closer, such as internal operational testing, a quality assessment of the probe, or a gallery description? Where the mismatch is wide and nothing bridges it, the stage has not been attempted rather than failed, and the weight the result can bear at Stage III falls accordingly.

Stage III—Reliability of this result

The record should allow assessment of whether the probe was suitable at all: resolution, pixels across the face, pose, illumination, blur, and occlusion. For a

one-to-many search, it should show enough about the gallery, meaning its size, the population represented, and enrolment conditions, for the court to understand what a returned candidate means. Where multiple candidates were returned, it should show the accused's rank, identify plausible alternatives, and state what investigators did about them. That last step is the most likely to be omitted and the hardest to detect later, and it is why rank must be disclosed rather than characterised: a statement that the system matched the accused is equally true of a first-ranked and an eighth-ranked candidate, and tells a court nothing. In *Miles*, the investigating officer's report recorded that a search had "revealed a positive match" for the defendant; the Court's account adds that the search also revealed multiple other possible matches, and the candidate list disclosed afterwards placed the defendant eighth of ten. The opinion describes a second search inconsistently, once as returning five images of the defendant in the first five places and once as showing five other people matched to the probe, and records that the State Police had no record of any request relating to the case (2026). Whether or not the report mentioned the other candidates, the words "positive match" said nothing about where the accused stood among them; seven ranked above him. *Miles* treats this material as the second of the two categories the State had to produce—how the tools were used, including the original probe photograph, any edited copies of it, and the photographs returned as matches—holding it potentially relevant to impeaching the identifications, challenging the investigation, and demonstrating third-party guilt.

Human review must be examined rather than assumed. Where a confirmation contributes materially, the record should show who the reviewer was, what materials they had, whether they knew the ranking or score, whether competing candidates were considered, and what method supported it. FISWG's categories indicate what to expect: assessment is often undocumented and the least rigorous, while review and examination carry documentation requirements and an independent technical review (FISWG, 2026: § 5.2).

Independent confirmation is not proposed as a rule of admissibility, and no fixed quantity of other evidence is required. The proposition is narrower: as the prosecution's dependence on the algorithm increases, the presence or absence of genuinely independent evidence matters more to the weight the identification can safely bear. That is consistent with *Zahir Jaffar*, whose holding that CCTV passing the two-fold test needs no corroboration, being primary evidence, concerns the recording (2025, para. 16); an inference drawn from that recording is a different proposition.

Stage IV—Contestability

Stage IV rests on Article 10A and on Section 94 of the Code of Criminal Procedure, which the Supreme Court has held contains no stage limitation, is available to either party where production is necessary or desirable, and serves the fairness and cross-examination interests of the defence (*The State v. Chaudhry Muhammad Usman*, 2023, paras. 6-10)[37]. That case does not concern facial recognition

[37] *The State v. Chaudhry Muhammad Usman*, 2023 SCMR 1676 (Pak.).

and is used analogically.

Where facial recognition materially contributes to proof, the defence should receive enough to investigate the reliability issue in dispute: as the case requires, the probe image, the processing record, the system and version, the operation, the gallery, the score and ranking, the threshold, validation material, error data, examiner documentation, and audit records. The governing principle is materiality, not a fixed list; *Miles* (2026) rejected mechanical application of the thirteen items requested in *Arteaga*, eleven of which that court had ordered produced on that ground.

The entitlement does not depend on the identification evidence being excluded. In *Arteaga*, the trial court had held a hearing on the eyewitness identifications and found them admissible, and the Appellate Division agreed that the procedure, while flawed, was not so irreparably flawed as to be inadmissible. It ordered the discovery all the same, holding that the discovery dispute stands separate from that hearing and that the material went to the defendant's ability to impeach the identification, to challenge the investigation, and to establish reasonable doubt at trial (2023). A Pakistani court has no power of exclusion to withhold here and no jury to protect, which makes the point easier rather than harder: the question is what the defence can put, not what the court may hear.

Source code is separate and should not be disclosed automatically. Where validation reports, operating records, input data, and expert access allow the issue to be examined, code will usually add little; where they do not, protected technical inspection may be required. We state that ordering tentatively, because the recorded instances run the other way. *Pickett* collects them: a federal court's unsealing of the source code of a probabilistic genotyping tool already used in thousands of prosecutions showed that it was unreliable, did not work as intended, and had to be abandoned; disclosure of the code of a product competing with the one in issue in *Pickett* revealed coding errors that led to misleading results; and in *State v. Chun* (2008)[38], the Supreme Court of New Jersey ordered a breathalyser manufacturer to disclose its source code for independent review, which found significant errors (*State v. Pickett*, 2021)[39]. In each instance, on *Pickett*'s account, the defect came to light through review of the code. That is a reason to treat the proportionate position as a starting point rather than a rule, and to ask in a given case whether the other material would expose errors of that kind. Proprietary assertions cannot end the inquiry (Wexler, 2018; *Pickett*, 2021), but nor do they yield automatically: *Miles* left the source-code question open on a particularised-need standard, the record below being insufficiently developed (2026).

The three institutional routes give this stage added weight: where a Face Trace result is tendered through a police officer, there is no expert to cross-examine on method; where a forensic report is admissible without its author, there is no examiner in the box unless the court exercises the Section 510 proviso; and where

[38] *State v. Chun*, 194 N.J. 54 (2008).

[39] *State v. Pickett*, 466 N.J. Super. 270 (App. Div. 2021) (A-4207-19T4, decided Feb. 3, 2021).

Safe Cities material arrives under an Electronic Data Certificate, the certifying officer is deemed an expert by statute and need not be called either.

In review proceedings in *Zahir Jaffar*, the Court rejected the argument that the DNA and video forensic reports were inadmissible because their authors were not produced, holding the 2022 amendment procedural and retrospective, and observing that the petitioner had applied to neither court to summon the expert under the proviso and that neither had thought it necessary of its own motion (*Zahir Zakir Jaffar*, 2026, para. 9)[40]. That disposal turned on the stage at which the point was taken rather than its merits, but its effect is that the initiative for producing the examiner rests with the party who did not commission the report, and failing to take it early forecloses the question. Where the author of a report or certificate is not called, production under Section 94 is the principal remaining mechanism by which the reliability question can be reached, and declining it forecloses the inquiry. Section 10 of the Punjab Forensic Science Agency Act supplies a further route that has attracted no attention: a court may itself send forensic material to the Agency for examination and expert opinion, which the Agency must authenticate and return to the court (Punjab Forensic Science Agency Act, 2007, § 10). Technical evidence must also remain technically challengeable, and funding expert assistance where most accused cannot pay requires legislative rather than judicial attention.

A reasoned judgment should keep the propositions separate on the record: the integrity of the input, the general reliability of the system, the reliability of this result, the adequacy of the opportunity to challenge, and the weight given.

The consequence of failure differs by stage. Failure at Stage I is ordinarily fatal to substantive use, not because a rule excludes the evidence but because the input cannot be established and nothing downstream can be assessed; that is what the High Court found in *Ammar Shafi*. Failure at Stage II, where nothing is offered about the system at all, leaves no rational basis for treating the output as an identification; the court should decline substantive reliance while leaving the material to do whatever investigative work it did. Failure at Stage III goes to weight in the ordinary way: a court assesses how much the identification can bear alongside the rest of the case. Failure at Stage IV is procedural, and its first remedy is production under Section 94 rather than any consequence for the evidence; only where production is refused or impossible, and a genuine reliability issue is thereby rendered untestable, does the question arise whether the identification should carry substantive weight at all. Part 7 takes that hierarchy further.

6. The Framework Applied

We run the framework through facts drawn from a decided judgment, with one

[40] *Zahir Zakir Jaffar v. The State*, Crim. Rev. Pet. No. 89 of 2025 (Sup. Ct. Pak. June 4, 2026). In its operative passage, the Court correctly identified the amending statute as the Code of Criminal Procedure (Amendment) Act 2022, which inserted the forensic scientist, with others, after "fire-arm expert" in Section 510. A later sentence refers instead to the Criminal Laws (Amendment) Act 2022, a different enactment. The Act numbers given in the reference list are supplied by the authors.

alteration. In *Ammar Shafi*, the Lahore High Court found deficiencies in the source and proof of Safe City material and observed that further evidence would be needed to establish that the accused were the men visible in the footage. Take those two features, but suppose that the prosecution extracts a still frame, runs it through Face Trace against a police-linked gallery, and tenders the result as proof of identity. Everything else is hypothetical, and we assume the algorithm is the identification evidence rather than a lead that dropped out.

Stage I produces the first difficulty, the one the High Court actually found: with source and proof deficient, nothing downstream can be assessed. Even where provenance is cured, Stage I still requires the processing record, since a still pulled from nighttime footage will probably have been enhanced before searching.

Stage II asks which system and version, and whether the operation was verification or identification. Against a police gallery, it is one-to-many, which brings gallery composition and ranking into play; the stage also asks for validation relevant to degraded surveillance stills, the operative threshold, and error performance. Because the operator is a police officer rather than an Agency expert, that material must come from the institution that deployed the system.

Stage III is where the case would probably be decided. A single frame from an elevated camera at an oblique angle is a poor probe, and the questions that matter are how many candidates were returned, where the accused ranked, whether anyone else resembling him was returned, and whether the person depicted could have been enrolled at all. If an officer confirmed the match after seeing the ranking, the confirmation is not an independent check. Stage IV asks whether the defence had enough to test any of this; on the record as it stood, no.

The stages are not a checklist for every case, and most will succeed or fail at Stage I or Stage III. The questions are those a competent cross-examiner would put if he knew they existed, which suggests the deficiency is as much one of practice as of doctrine.

7. Admissibility, Weight, and Burdens

Where nobody can say what system produced the result, what image it analysed, or what the output means, there is nothing to evaluate. We do not claim that Pakistani law presently prescribes exclusion in those circumstances, only that a result with no intelligible foundation should not perform substantive identification work merely because a machine produced it. Other deficiencies go to weight, reducing confidence without destroying the basis for evaluation. Keeping foundation and weight apart prevents "admissible" from being read as "reliable". Procedural failures have their own hierarchy. Where a reliability challenge cannot be investigated because information has been withheld, production under Section 94 is the first and best-grounded response; more severe consequences, such as restricting the use made of the evidence or declining to rely on an untestable identification, are proposals rather than established remedies, and depend on the seriousness of the failure, the centrality of the evidence, and whether a lesser measure

would restore challenge.

The word "burden" does three jobs here. The ultimate legal burden is the prosecution's obligation to prove guilt to the criminal standard, which Article 117 of the Qanun-e-Shahadat Order (1984) states and which *Raja Khurram Ali Khan v. Tayyaba Bibi* (2020, para. 54) holds can never be shifted to the accused unless the legislature by express terms commands otherwise. The evidential burden concerns producing enough to put a particular issue properly before the court and moves only once the legal burden has been discharged; we do not suggest that every technical criticism shifts anything to the prosecution. What we propose is neither, but a condition of foundation. If the extension argued in Part 3.5 is accepted, the condition has a domestic home: the Silent Witness doctrine receives machine-produced material on evidence about the machine, and where the machine concludes rather than records, evidence that says nothing about the correctness of its conclusions is not evidence about the machine. If the extension is not accepted, Article 59 with the gatekeeping principle and Article 10A are enough to carry the condition.

The party seeking substantive reliance on an algorithmic identification, ordinarily the prosecution, should therefore establish enough about the system and its application for the court to evaluate the inference it is asked to accept. That is reconciled with Explanation 3 as set out in Part 3.4: the Explanation presumes operation, not accuracy; its rebuttal routes concern data entry and extraction, not processing; on the approach in *Millat Tractors* (2024) it is read by what it was intended to do and does, namely classify output and presume the system's operation; and a presumption in an Explanation cannot displace the incidents of Article 10A in a criminal trial or shift a legal burden the legislature has not by express terms commanded be shifted. A defendant cannot sensibly be required to prove a system unreliable before anyone has shown what the system is.

Not all of this belongs to courts. Judicial development can handle the minimum foundation, case-specific scrutiny, and weight; standing obligations about audit logs, documentation of searches, and disclosure practice suit procedural legislation; and testing methodology and validation protocols belong to technical bodies. Courts should not invent testing standards, and a technical body's certificate does not settle an evidentiary question.

8. Conclusion

Pakistani evidence law is not as unprepared on this subject as the reform literature suggests. Article 46-A makes machine-generated information relevant, Explanation 3 to Article 73 makes it primary and presumes the system to have been operating, Article 78-A governs proof on denial, and Article 59 admits expert opinion on the functioning, specifications, programming, and operations of information systems. In *Zahir Jaffar*, the Supreme Court set out that architecture, stated a two-fold test for receiving CCTV that it derived from *Ahmed Omar Sheikh*, recorded that the footage was automatically generated without human intervention, and

authenticated it through a doctrine that turns on the reliability of the producing system.

What is missing is narrower than a statutory gap. Every one of those provisions concerns relevance, class, or proof, as do both limbs of that test; none concerns accuracy. The one source that speaks to accuracy is the Silent Witness doctrine, and only if extended as Part 3.5 argues. Article 78-A may not reach autonomous machine output at all. A computational identity inference can travel the entire chain without anyone being required to show it is correct, and the presumption of working order will do real damage if read as a presumption that the machine was right. The three institutional routes in Punjab make the problem concrete. China's contribution is real but modest: an instrument that names the technology and classifies its operations, and a standards infrastructure that tests systems, neither of which is a rule of evidence.

The framework draws its components from different places. Provenance and integrity come from the *Ishtiaq Ahmed Mirza* guidelines as applied in *Ahmed Omar Sheikh*; the division between system and case-specific reliability from PCAST, localized into Article 59 with the gatekeeping principle in *Ali Haider*; contestability from Article 10A and Section 94; and the demonstration that a foundation condition is at home in this statutory family from Indian practice under Section 65B and from Punjab's own Electronic Data Certificate. The proportionality principle and the reconciliation with Explanation 3 are ours; the foundational condition in Part 7 applies *Zahir Jaffar*, and the connection to test-identification doctrine in Part 3.6 draws on authority that was already there.

The framework's purpose is to prevent one finding from doing the work of another. That a recording is genuine is a finding about the recording; that a system performs well on benchmark data is a finding about the system. Neither, alone or together, establishes that the man in the dock is the man on the screen. Pakistan can work this out before the first case rather than after it, and the Supreme Court has already accepted the premise on which the answer depends.

AI Usage Statement

The authors used AI assistants for review and editing of text subsequently revised and adopted by the authors, for verification of sources, and for assistance in reading Chinese-language materials. The research question, the doctrinal argument, and the framework proposed are the authors' own. No AI tool generated any data, figures, or empirical results. The authors have reviewed and verified the entire manuscript and take full responsibility for its originality, accuracy, and integrity.

Author Contributions

All authors contributed equally to this work.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

Bhatta, A., Pangelinan, G., King, M. C., & Bowyer, K. W. (2024). Impact of Blur and Resolution on Demographic Disparities in 1-to-Many Facial Identification. In *2024 IEEE/CVF Winter Conference on Applications of Computer Vision Workshops (WACVW)* (pp. 422-430). IEEE. https://doi.org/10.1109/wacvw60836.2024.00051

Cyberspace Administration of China [CAC] & Ministry of Public Security [MPS] (2025). *Measures for the Security Management of the Application of Facial Recognition Technology*. Order No. 19 (Promulgated Mar. 13, 2025, Effective June 1, 2025). (In Chinese) https://www.cac.gov.cn/2025-03/21/c_1744174262156096.htm

Edmond, G., Biber, K., Kemp, R., & Porter, G. (2009). Law's Looking Glass: Expert Identification Evidence Derived from Photographic and Video Images. *Current Issues in Criminal Justice, 20,* 337-377. https://doi.org/10.1080/10345329.2009.12035817

Facial Identification Scientific Working Group [FISWG] (2026). *Facial Comparison Overview and Methodology Guidelines (Version 2.1)*. https://fiswg.org/documents

Facial Identification Scientific Working Group [FISWG] (2022). *Facial Recognition Systems Operation Assurance*: *Scoring Thresholds (Version 1.1)*. https://fiswg.org/documents

Facial Identification Scientific Working Group [FISWG] (2025a). *Face Recognition Systems Operational Assurance: Deployment Testing (Version 1.0)*. https://fiswg.org/documents

Facial Identification Scientific Working Group [FISWG] (2025b). *Facial Recognition Technology Implementation Guidelines (Version 1.0)*. https://fiswg.org/documents

Facial Identification Scientific Working Group [FISWG] (2025c). *Standard Guide for Image Processing to Improve Facial Recognition Search Performance (Version 2.1)*. https://fiswg.org/documents

Gabel Cino, J., Kleider-Offutt, H., Stevens, B., Albrecht, K., Evans, R., & Riedley, E. (2022). The Oracle Testifies: Facial Recognition Technology as Evidence in Criminal Courtrooms. *University of Louisville Law Review, 61,* 137-173.

Garrett, B. L., & Rudin, C. (2024). The Right to a Glass Box: Rethinking the Use of Artificial Intelligence in Criminal Justice. *Cornell Law Review, 109,* 561-627.

Gless, S., Lederer, F., & Weigend, T. (2024). AI-Based Evidence in Criminal Trials? *Tulsa Law Review, 59,* 1-37.

Grimm, P. W., Grossman, M. R., & Cormack, G. V. (2021). Artificial Intelligence as Evidence. *Northwestern Journal of Technology and Intellectual Property, 19,* 9-106.

Grossman, M. R., & Grimm, P. W. (2025). Judicial Approaches to Acknowledged and Unacknowledged AI-Generated Evidence. *Science and Technology Law Review, 26,* 110-155. https://doi.org/10.52214/stlr.v26i2.13890

Grother, P., Ngan, M., & Hanaoka, K. (2019a). *Face Recognition Vendor Test Part 2: Identification (NISTIR 8271)*. National Institute of Standards and Technology.

Grother, P., Ngan, M., & Hanaoka, K. (2019b). *Face Recognition Vendor Test Part 3: Demographic Effects (NISTIR 8280)*. National Institute of Standards and Technology.

Grother, P., Ngan, M., & Quinn, G. (2017). *Face in Video Evaluation (FIVE): Face Recognition of Non-Cooperative Subjects (NISTIR 8173)*. National Institute of Standards and Technology.

Guo, H., & Hou, J. (2018). Review of the Accreditation of Digital Forensics in China. *Forensic Sciences Research, 3,* 194-201. https://doi.org/10.1080/20961790.2018.1503526

Hameed, U., Shahid, A., & Noor, S. (2026). CCTV Evidence and the Law: The Emergence

of the Silent Witness Doctrine in State v Zahir Jaffer. *Policy Journal of Social Science Review, 4,* 310-317.

Howard, J. J., Rabbitt, L. R., & Sirotin, Y. B. (2020). Human-Algorithm Teaming in Face Recognition: How Algorithm Outcomes Cognitively Bias Human Decision-Making. *PLOS ONE, 15,* e0237855. https://doi.org/10.1371/journal.pone.0237855

Pangelinan, G., Bhatta, A., Wu, H., King, M. C., & Bowyer, K. W. (2024). Analyzing the Impact of Demographic and Operational Variables on 1-to-Many Face ID Search. *IEEE Transactions on Technology and Society, 5,* 217-230. https://doi.org/10.1109/tts.2024.3416344

Perritt Jr., H. H. (2021). Defending Face-Recognition Technology (and Defending against It). *Journal of Technology Law and Policy*, *25*, 41-139.

Phillips, P. J., Yates, A. N., Hu, Y., Hahn, C. A., Noyes, E., Jackson, K. et al. (2018). Face Recognition Accuracy of Forensic Examiners, Superrecognizers, and Face Recognition Algorithms. *Proceedings of the National Academy of Sciences, 115,* 6171-6176. https://doi.org/10.1073/pnas.1721355115

President's Council of Advisors on Science and Technology (2016). *Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods*. Executive Office of the President.

Punjab Forensic Science Agency (n.d.). *Audio Video Analysis*.

Punjab Information Technology Board (n.d.-a). *PITB Launches First AI-Based "Face Trace" System for the Punjab Police*.

Punjab Information Technology Board (n.d.-b). *PITB Has Developed the AI-Powered "Face Trace System (FTS)" for Punjab Police*.

Punjab Safe Cities Authority (n.d.). *CBOR*.

Rasool, N., & Rasool, M. (2022). Challenges for Expert Evidence in the Justice System of Pakistan. *Journal of Forensic Science and Medicine, 8,* 62-67. https://doi.org/10.4103/jfsm.jfsm_16_21

Roth, A. (2017). Machine Testimony. *Yale Law Journal*, *126*, 1972-2053.

Shahzad, M. K., Imdad, A., & Sultana, N. (2026). Rethinking the Authentication of AI-Generated Evidence under Article 164 of the Qanun-E-Shahadat Order, 1984: Reforming Pakistan's Evidentiary Framework. *Journal for Social Science Archives, 4,* 2553-2570. https://doi.org/10.59075/jssa.v4i1.684

Supreme People's Court, Supreme People's Procuratorate & Ministry of Public Security (2025). *Opinion on Improving the Mechanism for Reviewing the Identity of Criminal Suspects and Defendants*. Fa [2025] No. 144 (Sept. 29, 2025). (In Chinese)

Wang, C. (2026). AI Evidence in Criminal Proceedings in China: Challenges and Solutions. *Asia Pacific Law Review, 34,* 572-599. https://doi.org/10.1080/10192557.2025.2583310

Wexler, R. (2018). Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System. *Stanford Law Review, 70,* 1343-1429. https://doi.org/10.2139/ssrn.2920883