AT THE EDGE OF REASON: PRESERVING THE SANCTITY OF THE JURY TRIAL IN AN ERA OF ALGORITHMIC ASCENDANCY

Yong Siew Fei (Liam)[1*]

Abstract

The UK jury system, long emblematic of participatory justice and civic responsibility, now stands at a digital inflexion point. As artificial intelligence and wider algorithmic infrastructures become inextricably woven into everyday life, the integrity of jury deliberations is increasingly susceptible to opaque external influences — from algorithmically curated content to the dissolution of traditional informational boundaries. This paper interrogates the evolving tension between juror independence and technological incursion, and asks whether legacy safeguards remain fit for purpose in an era increasingly governed by machine-mediated information flows.

Anchored in UK case law such as ***R v Karakaya*** *and* ***Attorney General v Fraill****, the paper examines the extent to which juror conduct is reshaped in digital domains, where the boundary between passive exposure and active impropriety becomes increasingly indistinct. Statutory instruments like the* ***Criminal Justice and Courts Act 2015*** *are subjected to scrutiny. While punitive measures signal the gravity of misconduct, they do not grapple with the structural dynamics that drive juror transgression in a networked environment.*

Departing from purely retributive paradigms, this research proposes a reimagining of artificial intelligence not as an existential threat to the justice system, but as a carefully bounded instrument of reform. From behavioural-science tools that cultivate bias-awareness to adaptive legal-education platforms that demystify complex doctrines for laypersons, the paper explores judicious integration under strict guardrails. It also engages with the ethical fault lines inherent in any technology that might encroach upon juror privacy, autonomy, and the sanctity of deliberation.

Ultimately, the paper contends that the survival of trial by jury in an AI-mediated legal landscape depends not on resisting technological advancement, but on reconciling innovation

[1*] LLB (Honous) candidate under the UKDTP programme, HELP University.

with the core tenets of procedural fairness. Justice must remain anchored in human discernment, even as its architecture becomes increasingly algorithmic.

Introduction — An Opening at the Edge

As I put it when presenting this paper at the 3rd Tan Sri Dato' Seri Dr. Abdul Malek Law Student Convention at HELP University in June 2025: "*I chose the phrase At the Edge of Reason to mark a methodological threshold*". The jury system has long been a cornerstone of justice, ensuring that verdicts align with community values while acting as a safeguard against judicial overreach. The rise of the internet has made juror impartiality harder to maintain, as easy access to external information threatens the integrity of trials. Despite judicial warnings and legal restrictions, research indicates that jurors continue to seek online material, potentially compromising their decisions. Courts have responded through contempt and, more recently, criminalisation, yet these tools operate ex post and do little to equip jurors to resist contamination *ex ante*. This article argues that trial by jury can be preserved in an algorithmic age by coupling proportionate enforcement under the Criminal Justice and Courts Act ('CJCA') 2015 with primary-prevention measures — staged plain-language directions, pre-service micro-education, and strictly bounded, non-intrusive technology.

Tradition under Pressure — Normative Foundations of the Jury

The UK jury system has been a cornerstone of justice for centuries, dating back to King Henry II's legal reforms in the 12th century.[2] Before its introduction, trials relied on arbitrary methods such as trial by ordeal, where guilt or innocence was determined through physical tests rather than evidence. The ***Assize of Clarendon*** (1166) marked a turning point by replacing these practices with jury trials,[3] establishing adjudication oriented to testimony and fact-finding — an important step toward a fairer legal process.[4]

The ***Magna Carta*** (1215) further reinforced the significance of juries, declaring that no one could be punished without "the lawful judgment of his peers".[5] This principle laid the foundation for public participation in legal decisions. Blackstone later described the jury as

[2] JH Baker, *An Introduction to English Legal History* (5th edn, OUP 2019) p 74.
[3] ***Assize of Clarendon*** (1166), reprinted in David C Douglas and George W Greenaway (eds), English Historical Documents 1042–1189 (Eyre & Spottiswoode 1953) p 220.
[4] JH Langbein, *The Origins of Adversary Criminal Trial* (OUP 2003) p 15.
[5] M Lobban, *The Oxford History of the Laws of England: Volume XI 1820–1914* (OUP 2014) p 35.

the "palladium" of English justice,[6] emphasising its role in preventing tyranny, while Lord Devlin famously called it "the lamp that shows that freedom lives".[7] These perspectives highlight the jury's function in safeguarding individual rights and upholding democracy.

Secrecy, Safety, and the Architecture of Independence — The Law on Jury Irregularities

Today, juries remain central to criminal trials, serving as independent fact-finders. Composed of randomly selected citizens, they ensure verdicts reflect societal values rather than judicial discretion alone.[8] This acts as a safeguard against judicial overreach and reinforces public trust in the legal system. The ***Juries Act 1974*** and ***Article 6*** of the ***European Convention on Human Rights (ECHR)*** further enshrines the right to a fair trial, underscoring the jury's continued relevance.[9]

Courts operationalise that independence through a simple architecture: deliberation secrecy is protected, while external irregularities are policed. As a matter of principle, internal reasoning errors are inviolable, whereas external contamination is justiciable — hence ***R v Mirza*** and ***R v Connor*** *and another* protect deliberative privacy, while ***R v Young (Stephen)*** and ***R v Karakaya*** warrant intervention when extraneous material intrudes.[10]

However, the jury system is not without its challenges. Darbyshire argues that lay jurors may struggle with complex legal concepts, leading to inconsistencies in verdicts.[11] Biases, emotional reactions, and external influences can also impact decision-making. Studies suggest that factors such as a defendant's appearance or demeanour can unconsciously sway jurors, raising concerns about impartiality.[12]

Another key issue is juror comprehension. Legal terminology and intricate case details can be difficult for those without legal training, and research indicates that jurors often

[6] William Blackstone, *Commentaries on the Laws of England*, Vol 4 (1st edn, 1776) p 349.
[7] Patrick Devlin, *Trial by Jury* (Stevens & Sons 1956) p 164.
[8] Penny Darbyshire, 'The Lamp That Shows That Freedom Lives – Is It Worth the Candle?' (1991) Criminal Law Review 740.
[9] ***Juries Act 1974, s. 1***; ***European Convention on Human Rights*** (adopted 4 November 1950, entered into force 3 September 1953) ***Art 6***.
[10] ***R v Mirza*** [2004] UKHL 2; ***R v Connor and another*** [2004] UKHL 2; ***R v Young (Stephen)*** [1995] QB 324 (CA); ***R v Karakaya*** [2005] EWCA Crim 346; ***Contempt of Court Act 1981 s. 8***.
[11] Darbyshire (n 7) 749-750.
[12] ibid.

misinterpret instructions on legal standards such as the burden of proof. [13] In complex cases, particularly those involving forensic evidence, jurors frequently struggle to understand the subject matter, raising doubts about their ability to make fully informed decisions.[14]

To enhance jury reliability, legal experts propose several reforms. Darbyshire advocates for clearer judicial instructions, written guidelines, and structured deliberation processes.[15] Providing jurors with written summaries of legal directions and allowing note-taking during trials could improve evidence retention and ensure decisions are based on a sound understanding of the law rather than assumptions or personal interpretations. [16]

Despite its shortcomings, the jury system remains an essential pillar of justice. It ensures that legal decisions are not left solely in the hands of judges and that community values shape verdicts. Rather than abolishing the system, the most effective approach is to implement reforms that improve juror education and protect against undue influence. By refining procedures and enhancing juror comprehension, the legal system can strengthen the integrity of trial by jury. While imperfect, the jury remains a vital mechanism, ensuring fairness and accountability in legal decisions.

When the Feed Meets the Forum — Digital Susceptibility in Doctrine and Data

The principle of jury impartiality is a cornerstone of the criminal justice system, ensuring that verdicts are based solely on evidence presented in court. Jurors are expected to remain passive recipients of these materials, deliberating strictly within the framework set by judicial instructions. However, the internet has blurred these boundaries, providing jurors with unrestricted access to external information that directly threatens the integrity of deliberations. This increased accessibility introduces significant risks, as unverified or prejudicial material can undermine due process and compromise the fairness of trials.

Despite explicit judicial warnings, empirical research suggests that jurors frequently disregard these restrictions. A study by Cheryl Thomas found that in high-profile cases, 12% of jurors admitted to conducting independent online research during the trial, compared to 5%

[13] Blake M. McKimmie, Emma Antrobus, Chantelle Baguley, 'Objective and Subjective Comprehension of Jury Instructions in Criminal Trials' (2014) 17(2) New Criminal Law Review 163, 73-75.
[14] Laurence J. Severance and Elizabeth F. Loftus, 'Improving the Ability of Jurors to Comprehend and Apply Criminal Jury Instructions' (1982) Law & Society Review 17(1), 153, 65-72 (see Table 1. Sources of Juror Confusion about Criminal Jury Instructions: Questions Juries Ask).
[15] Darbyshire, 'What Can We Learn from Published Jury Research?' (2001) Criminal Law Review 970, 76-77.
[16] 'Note-taking Jurors Influence Verdicts, Study Finds' (University of Liverpool News, 12 March 2019) <https://news.liverpool.ac.uk/2019/03/12/note-taking-jurors-influence-verdicts-study-finds/> accessed 15 February 2025.

in standard cases.[17] This highlights a clear correlation between media attention and juror misconduct. The availability of extraneous information — whether accurate or misleading — poses a serious risk to fair trials, as such material is neither tested under evidentiary rules nor subject to judicial scrutiny.

Consider a juror in a multi-week fraud trial: over lunch, his/her phone surfaces a push-notification linking a blog speculating about the defendant's "pattern". He/she does not search for the case; the feed finds him/her. The headline lingers in deliberations — not as evidence, but as salience. The harm is external and untested, yet cognitively persistent.

The courts have acknowledged these risks in key cases. In ***R v Karakaya***,[18] a juror conducted an online search about the case and introduced printed materials into the deliberation room, contaminating the trial with inadmissible information. The Court's ruling underscored growing judicial concerns over digital interference, affirming that reliance on external material undermines the integrity of verdicts. Similarly, in ***Attorney General v Fraill***,[19] a juror was convicted of contempt of court after engaging in Facebook discussions with a defendant. This marked the first instance where social media misuse resulted in a custodial sentence for juror misconduct, reinforcing the need for stricter enforcement of jury impartiality.

The influence of digital media extends beyond instances of active misconduct to the subconscious shaping of jurors' perceptions. The prevalence of algorithm-driven content on social media, means that jurors may be exposed to biased news articles, speculative commentary, or misleading information that reinforces their pre-existing beliefs. In high-profile cases, even jurors who comply with judicial instructions may inadvertently encounter prejudicial material, further compromising the impartiality of the trial process.[20] Empirical patterns are summarised in Table 1, which motivates the prevention-first reforms developed below.[21]

Indicator	Finding (summary)	Implication	Source

[17] Cheryl Thomas, 'Avoiding the Perfect Storm of Juror Contempt' [2013] Criminal Law Review 483, 484.
[18] ***R v Karakaya*** [2005] EWCA Crim 346.
[19] ***Attorney General v Fraill*** [2011] EWHC 1629 (Admin).
[20] John Jackson, 'Juror Decision-Making and the Internet' (2013) 14(3) Legal Studies 393, 402.
[21] Cheryl Thomas, *Are Juries Fair?* (Ministry of Justice, Research Series 1/10, 2010) 52; John Jackson, 'Juror Decision-Making and the Internet' (2013) 14(3) *Legal Studies* 393, 402; 'Note-taking Jurors Influence Verdicts, Study Finds' (University of Liverpool News, 12 March 2019) <https://news.liverpool.ac.uk/2019/03/12/note-taking-jurors-influence-verdicts-study-finds/> accessed 15 February 2025.

Extraneous online searching in high-profile cases	Non-trivial rates of case-adjacent searching reported among jurors exposed to publicity.	Strengthen staged directions; emphasise why extra-curial material is unreliable.	Cheryl Thomas, Are Juries Fair? (MoJ 2010) 52.
Effect of written directions and note-taking	Written aids and permission to take notes are associated with improved comprehension/recall.	Default to written directions and note-taking permissions.	University of Liverpool News (12 March 2019).
Algorithmic curation and passive exposure	Feeds can shape perceptions even absent deliberate research.	Explain algorithmic personalisation in micro-education modules.	John Jackson, Legal Studies 14(3) (2013) 393, 402.

Table 1: Selected empirical indicators relevant to jury directions, digital exposure, and comprehension

Recognising these threats, the UK enacted the ***Criminal Justice and Courts Act ('CJCA') 2015***, shifting internet-related juror misconduct from contempt of court to a distinct criminal offence. Under Sections 71–74, jurors found guilty of conducting independent research, sharing improperly obtained information, or engaging in online discussions about a trial now face up to two years' imprisonment and a ten-year disqualification from jury service. This legislative shift reflects a more punitive, deterrence-focused approach, underscoring the judiciary's recognition of the internet's potential to undermine fair trials.

Ultimately, ensuring jury impartiality in the digital age remains a formidable challenge. The internet has introduced unprecedented risks to the criminal justice system, and without stringent safeguards, digital interference could erode public confidence in the jury system. To uphold the fundamental principle of fair and unbiased trials, the judiciary must remain vigilant in enforcing existing laws and adapting to emerging threats.

Criminalisation without Capability? — The *CJCA 2015* and Its Limits

While legislative reforms have sought to curb juror misconduct, enforcing contempt of court laws in the digital age remains a significant challenge, particularly given the difficulty in monitoring online activity. The ***Contempt of Court Act 1981***, designed to prevent prejudicial influences on trials, was enacted in a pre-digital era and fails to account for the complexities of modern internet use. Unlike traditional media, which can be regulated through reporting restrictions, personal internet access remains largely unmonitored, allowing jurors to discreetly seek for extraneous information without detection.[22] This poses a substantial risk to the integrity of the trial process.

In response to these challenges, the ***CJCA 2015*** (Sections 71–74) introduced stricter penalties for juror misconduct, criminalising independent research and online discussions about trials. However, enforcement remains largely reactive rather than preventive. Misconduct is typically identified only after it has occurred, often through self-disclosure or external reports, making intervention difficult before prejudicial harm is done.

The judicial system currently lacks effective mechanisms to detect and prevent jurors from accessing external information, rendering existing laws insufficient to address the realities of the digital age.

Legal scholars continue to debate whether punitive measures alone are an effective deterrent against juror misconduct. Crosby argues that harsh criminal sanctions risk discouraging civic participation in jury duty and may alienate jurors by treating them as potential offenders rather than trusted participants in the legal process.[23] Furthermore, research suggests that the existence of punitive laws does not necessarily eliminate the temptation to seek outside information, particularly in high-profile or complex cases where jurors may feel compelled to fill gaps in their legal understanding.[24] This suggests that relying solely on criminal penalties is unlikely to resolve the issue effectively. Sentencing practice under the ***CJCA 2015*** should track culpability and risk — distinguishing inadvertent exposure, negligent curiosity, and intentional research/dissemination — in line with proportionality.

[22] Cheryl Thomas, *Are Juries Fair?* (Ministry of Justice, 2010) p 52.
[23] Kevin Crosby, 'Juror Punishment, Juror Guidance and the Criminal Justice and Courts Act 2015' [2015] Criminal Law Review 578, 591-592.
[24] Kevin Crosby, 'Controlling Devlin's Jury: What the Jury Thinks, and What the Jury Sees Online' (2012) Criminal Law Review 15, 26.

This raises a critical question: should the legal system prioritise punishment, or should it focus on preventative measures to uphold jury integrity? While punitive measures reinforce the seriousness of juror misconduct, they fail to address its root causes — curiosity, a misunderstanding of legal restrictions, or the perceived need for additional context. A more sustainable solution requires a proactive approach, including enhanced juror education, clearer judicial instructions, and stricter controls over media exposure during trials. Implementing structured digital policies — such as systems-level monitoring of publicly available chatter about active cases (without tracking identifiable jurors) — could help judges time and reinforce directions.[25]

Ultimately, ensuring jury impartiality in the digital age demands a balanced strategy. The internet has introduced unprecedented risks to the justice system, and without stringent safeguards, digital interference may continue to undermine public confidence in jury trials. While criminal sanctions play a crucial role in deterring misconduct, they must be complemented by preventative measures that address the evolving nature of digital information consumption. Without ongoing reform, the increasing accessibility of online content will continue to challenge the fundamental principle of fair and unbiased trials.

The Media, the Message, and the Margin of Prejudice — Contempt in a Networked Age

The strict-liability rule in the 1981 Act seeks to prevent publications that create a "substantial risk that the course of justice…will be seriously impeded or prejudiced".[26] That doctrine worked tolerably in a broadcast or print era, aided by postponement orders and professional media compliance. The social web dissolves many of those assumptions. Content is decentralised, persistent, and algorithmically amplified; its audience cannot be straightforwardly cabined by geography or timing; and its authors include the public itself. Contempt remains essential — particularly against professional publishers — but it is a blunt instrument against a participatory ecosystem that multiplies untraceable speakers. The result is an enforcement gap: we cannot meaningfully control the information environment around a trial, and jurors cannot be perfectly insulated from it without intolerable burdens.

[25] Daniel William Bell, 'Juror Misconduct and the Internet' (2010) American Journal of Criminal Law, Vol 38, 81, 86.
[26] ***Contempt of Court Act 1981 s. 2(2)***.

Jury Integrity in the Digital Age — Balancing Transparency, Education, and Impartiality

The internet poses significant challenges to jury integrity. However, when properly regulated, it can also serve as a valuable tool for enhancing juror education and strengthening public confidence in the legal system. Greater access to legal information allows jurors to develop a clearer understanding of legal principles and courtroom procedures, addressing one of the key criticisms of the jury system — jurors' difficulty in grasping complex legal concepts. Research suggests that structured legal education can significantly improve juror comprehension, enabling more informed deliberations.[27] Gobert challenges the notion that jurors fundamentally lack legal understanding, arguing that with access to appropriate legal materials, they can engage more critically with evidence and judicial directions, ultimately leading to more reasoned verdicts.[28]

Beyond juror education, digital media also plays a crucial role in fostering public scrutiny, promoting transparency, and reinforcing trust in the justice system.[29] However, Jackson warns that unrestricted internet access exposes jurors to prejudicial media coverage, which can lead to unconscious bias and compromise impartial decision-making.[30] As a result, judicial oversight is essential to ensure that jurors do not rely on extraneous information that falls outside the evidentiary framework. Similarly, Smartt highlights that legal databases such as BAILII, provide direct access to case law, potentially aiding jurors in interpreting complex legal principles.[31] However, the risk lies in jurors misinterpreting legal doctrines without judicial guidance, which could inadvertently distort trial outcomes.

Despite the potential benefits of digital access, strict regulation remains necessary to preserve trial integrity. The ***CJCA 2015*** reflects this necessity by criminalising juror misconduct related to independent research and online discussions. The key challenge lies in striking a balance between leveraging the internet's educational potential and enforcing safeguards to prevent external influences from compromising jury impartiality. While digital resources can enhance jurors' legal understanding, judicial oversight remains critical in ensuring that public confidence in the jury system is maintained.

[27] Thomas (n 20) p 63.
[28] James Gobert, 'The Myth of the Jury's Ignorance' (1997) 62 MLR 733, 742.
[29] David Eady, 'Trial by Media: The Betrayal of the First Principle' (2003) 11(2) Tort L Rev 71, 75.
[30] John Jackson, 'Juror Decision-Making and the Internet' (2013) 14(3) Legal Studies 393, 402.
[31] Ursula Smartt, 'Juries and the Impact of the Internet' (2011) 75(6) J Crim L 446, 750.

Trial by Jury in the Digital Age — Addressing the Internet's Challenge to Justice

In November 2010, Lord Judge delivered an address in Belfast, stating that: "If the jury system is to survive as the system of a fair trial…the misuse of the internet by jurors must stop".[32] His assertion underscores legitimate concerns about the integrity of jury trials in the digital age.

Unrestricted online access allows jurors to encounter information that has not undergone judicial scrutiny, threatening the fundamental principle that verdicts should be based solely on courtroom evidence.[33] However, completely eliminating internet influence is neither feasible nor necessary. Instead, reforms should prioritise enhanced regulation, juror education, and procedural safeguards to preserve the integrity and viability of the jury system.

Can Internet Influence on Juries Be Eliminated?

Given the internet's deep integration into modern life, expecting jurors to remain entirely insulated from online content is unrealistic. In this context, enforcing a strict ban on internet use is impractical. A more effective approach involves regulating jurors' online behaviour through clearer judicial instructions, digital literacy education, and stricter penalties for misconduct.[34]

Abolition or Reform — Is the Jury System Still Viable?

Some legal scholars advocate for replacing juries with judge-only trials in complex cases, particularly those involving financial fraud or organised crime. This approach has already been implemented in cases such as ***R v Twomey***,[35] where the risk of jury tampering necessitated a non-jury trial. While judge-only trials can reduce the risk of juror misconduct, they also raise concerns about democratic legitimacy and public confidence in the justice system. Rather than abolishing jury trials, reforms such as enhanced jury vetting, clearer

[32] Igor Judge LCJ, 'Jury Trials' (Judicial Studies Board Lecture, Belfast, 16 November 2010) <https://www.judiciary.uk/wp-content/uploads/JCO/Documents/Speeches/speech-lcj-jury-trials-jsb-lecture-belfast.pdf> accessed 31 August 2025.
[33] Thomas (n 20) p 37.
[34] Lyndon Harris, 'Has the Internet Destroyed Trial by Jury?' (2013) 177 Criminal Law and Justice Weekly 561, 562.
[35] ***R v Twomey*** [2009] EWCA Crim 1035.

judicial directions, and digital monitoring tools could help safeguard their integrity while preserving public participation in the legal process.[36]

Sequestration as Solution or Symptom? — Practicality, Psychology, and Comparative Retreat

Sequestration, which isolates jurors from external influences for the duration of a trial, has been proposed as a solution to internet-related juror misconduct. While historically employed in high-profile cases to prevent prejudicial exposure, its effectiveness in the digital age has become increasingly uncertain. Research indicates that sequestration is often imposed only during deliberations, leaving jurors free to conduct independent research during earlier stages of the trial.[37] Moreover, the widespread availability of smartphones and internet access makes enforcing complete isolation significantly difficult.

Beyond practical concerns, sequestration imposes significant financial and psychological burdens. It strains court resources, contributes to trial delays, and increases costs, making it an impractical long-term solution.[38] In Australia, for instance, jury sequestration has largely been abandoned due to these challenges.[39] Given these limitations, a more viable alternative is the implementation of AI-driven systems-level signals of publicity spikes to detect potential juror misconduct. Such measures would allow courts to intervene proactively, preserving trial integrity without resorting to extreme restrictions.

Comparative Vistas — Calibrating the UK Response by Looking Outward

Comparative law offers cautionary tales and heuristics. In the United States, ***Remmer v United States*** held that private communications with jurors about a pending matter are presumptively prejudicial, shifting the burden to the Government to show harmlessness;[40] ***Sheppard v Maxwell*** emphasised trial-management duties amid prejudicial publicity;[41] and ***Pena-Rodriguez v Colorado*** carved a narrow exception to the no-impeachment rule for

[36] Harris (n 33) 562.
[37] Bell (n 24) 87.
[38] Nicola Monaghan, 'The Problem of Jury Misbehaviour in an Internet Age: Recent Cases and the Law Commission's Consultation', Coventry Law Journal (2013) 18(1) 69, 80.
[39] Roxanne Burd and Jacqueline Horan, 'Protecting the Right to a Fair Trial in the 21st Century – Has Trial by Jury Been Caught in the World Wide Web?' (2012) Criminal Law Journal 36(2) 103, 117.
[40] ***Remmer v United States*** 347 US 227 (1954).
[41] ***Sheppard v Maxwell*** 384 US 333 (1966).

evidence of racial animus in deliberations.[42] The American path is not transplantable wholesale, but it underscores that sanctity of the jury room yields where fundamental fairness is credibly impugned.

Australia's experience points to a retreat from routine sequestration coupled with stronger directions and statutory support for written aids, exemplified by the ***Jury Directions Act 2015 (Vic)***.[43] These strands converge on a common insight: where the media environment overwhelms traditional prophylaxis, legitimacy is restored by explanation — clearer rules, clearer reasons, and procedures that respect juror capacities while protecting the accused.

A jurisdictional comparison is set out in Table 2;[44] it underscores why explanation, not isolation, is the sustainable path.

Issue	UK (current)	US (selected)	Australia (selected)	Strengths	Risks
Extra-curial influence	Contempt; ***CJCA 2015*** offences; judicial directions.	***Remmer*** presumption; ***Sheppard*** management duties.	Retreat from sequestration; stronger written directions (e.g., ***Victoria***).	Clear signalling; doctrinal clarity.	Reactive; under-inclusive for passive exposure.
Jury secrecy vs fairness	***s. 8*** secrecy; external-irregularity doctrine.	***Pena-Rodriguez*** exception for racial animus.	Generally maintains secrecy with pragmatic	Protects candour; tailored exceptions possible.	Diagnosis difficulties; risk of over- or under-intervention.

[42] *Pena-Rodriguez v Colorado* 580 US 206 (2017).
[43] Jury Directions Act 2015 (Vic).
[44] Criminal Justice and Courts Act 2015 ss 71–74; Contempt of Court Act 1981; *Remmer v United States* 347 US 227 (1954); *Sheppard v Maxwell* 384 US 333 (1966); *Pena-Rodriguez v Colorado* 580 US 206 (2017); Jury Directions Act 2015 (Vic).

			case management.		
Jury education	Non-standardised; judge-by-judge practice.	Varies; some pattern instructions; juror handbooks.	Statutory support for simpler directions ***(Vic)***.	Improves comprehension; public confidence.	Cost, consistency, evaluation challenges.

Table 2: Comparative safeguards and risks — UK, US, Australia

From Policing to Empowerment — Primary Prevention for a Digital Forum

Treating jurors as problems to be managed is both normatively unattractive and instrumentally weak. The system should equip jurors to succeed. First, plain-language, staged directions about digital conduct should be given at empanelment, after openings, and pre-retirement, in oral and written form. A draft, court-ready text appears in Appendix A.[45] Experimental literature indicates that structured guidance and permission to take notes, improve comprehension and recall.[46]

Second, pre-service micro-education — neutral modules of five to seven minutes — should explain burdens and standards of proof, why extra-curial materials are unreliable, and how algorithms personalise content. Far from "coaching," such modules level the informational playing field without entering the merits.

Third, supportive court architecture should supply neutral glossaries for technical terms, short written summaries of directions, and permission to take notes.[47] These

[45] *See Appendix A ('Model Digital-Conduct Directions'), paras A.1–A.4.*
[46] Blake M McKimmie, Emma Antrobus and Chantelle Baguley, 'Objective and Subjective Comprehension of Jury Instructions in Criminal Trials' (2014) 17(2) New Crim L Rev 163; Laurence J Severance and Elizabeth F Loftus, 'Improving the Ability of Jurors to Comprehend and Apply Criminal Jury Instructions' (1982) 17(1) Law & Society Rev 153; University of Liverpool News (n 15).
[47] Penny Darbyshire, 'What Can We Learn from Published Jury Research?' (2001) *Crim LR* 970.

interventions protect deliberative independence, not by surveillance, but by increasing juror capability and confidence.

Augmentation, Not Automation — A Narrow Role for AI with Hard Guardrails

AI should neither sit in the jury room nor peer over its shoulder; it should support the conditions under which human judgment can flourish. Three roles are defensible.

First, pre-service bias-awareness tools can help prospective jurors recognise cognitive pitfalls through interactive hypotheticals and reflective prompts.[48]

Second, adaptive legal-education modules can translate burdens, standards, and inferential reasoning into accessible formats, calibrated for lay comprehension, and audited for neutrality.

Third, systems-level risk mapping of open-source public chatter about cases can inform judicial timing and the reinforcement of directions without tracking identifiable jurors. These uses are contingent on non-negotiable guardrails: strict necessity and proportionality; no surveillance of identifiable jurors; no analytics on deliberative content; human judicial oversight; data minimisation and default deletion; and compatibility with the dignity of the jury room, and due process.[49]

This is augmentation, not automation: technology in service of reason, never its substitute — and any systems-level tool should be preceded by a documented Data Protection Impact Assessment (DPIA), a necessity/proportionality analysis, and strict minimisation under the ***Data Protection Act 2018*** and ***UK General Data Protection Regulation (GDPR) Article 35***, with default deletion and judicial oversight.[50]

Neutral pre-service micro-education can explain burdens and standards of proof, why extra-curial material is unreliable, and how algorithms personalise content without coaching verdicts. A concise outline is provided in Appendix B.[51]

From Principle to Practice — Doctrinal and Policy Proposals in Full

[48] James Gobert, 'The Myth of the Jury's Ignorance' (1997) 62 *MLR* 733, 742.
[49] Ursula Smartt, 'Juries and the Impact of the Internet' (2011) 75(6) *J Crim L* 446, 750; Law Commission, *Contempt of Court (1): Juror Misconduct and Internet Publications* (Law Com No 340, 2013).
[50] *Data Protection Act 2018*; *UK GDPR Art 35* (Data Protection Impact Assessment).
[51] See *Appendix B ('Pre-Service Micro-Education Outline')*, paras B.1–B.5.

Concrete reforms follow from the foregoing analysis. These measures are not efficiency tweaks but fair-trial guarantees: they operationalise the right to an impartial tribunal under ***Article 6 ECHR*** by reducing exogenous risk, while preserving deliberative independence.[52]

Model digital-conduct directions should be annexed to the ***Criminal Practice Directions***, with standard text in plain language and mandatory delivery at set stages. Written directions and juror aids should be the default, not the exception, given the association between written materials, note-taking, and improved comprehension.[53]

Pre-service micro-education should be audited for neutrality, accessibility, and equality impacts, and kept content-agnostic to avoid case-specific influence. Proportionate sentencing guidance under the 2015 Act should distinguish inadvertent exposure from deliberate research and harmful dissemination, favouring education or conditional discharge for low culpability, while escalating to custody for intentional, prejudicial misconduct.

Narrow, non-intrusive technology should operate at the systems-level to assist judges in calibrating directions and scheduling; individualised monitoring of jurors should remain off the table. Finally, judge-only trials (as in ***Twomey***) should remain a last resort for proven tampering or intimidation, accompanied by clear public reasons to safeguard legitimacy.[54]

Any tools must operate at systems-level — not juror level — and comply with necessity, proportionality, non-surveillance of identifiable jurors, data minimisation and default deletion, with human judicial oversight. Where intimidation or tampering credibly threatens trial integrity, judge-only trials — as in ***R v Twomey*** — remain a narrowly cabined last resort, accompanied by public reasons to preserve legitimacy. A practical checklist is set out in Appendix C.[55]

Objections and Replies — Privacy, Cost, Equality, and Creep

Privacy concerns are addressed by rejecting individualised monitoring and forbidding analytics on deliberative content. Cost objections are mitigated by economies of scale and the high expense of aborted trials, and appeals. Equality concerns are met by standardisation, audit, and universal access to pre-service micro-education, which raises baseline

[52] Convention for the Protection of Human Rights and Fundamental Freedoms (*European Convention on Human Rights*, 1950) *Art 6*.
[53] University of Liverpool News (n 15).
[54] n 34.
[55] See *Appendix C ('AI Guardrails Checklist')*, paras C.1–C.7.

comprehension. Fears of ‘function creep’ are constrained by constitutional principles of necessity and proportionality, judicial oversight, public guidance, and sunset review mechanisms to prevent drift from augmentation to automation.

Conclusion — Shielding the Lamp

If the jury is a lamp, the digital ecosystem is a crosswind. Extinguishing public participation is no remedy; nor is criminalising confusion. The task is to shield the flame: clearer directions, modest education, supportive trial architecture, and carefully bounded technology that serves human judgment. Reimagined this way, trial by jury can continue to express the community’s reason at the edge — not beyond it.

Appendix A — Model Digital-Conduct Directions

A1. Members of the jury, you must decide this case only on the evidence you hear and see in this courtroom.

A2. You must not look for information about the case or the people involved. That includes internet searches, social media, podcasts, or news. Online content is often personalised to your interests; it may appear relevant but it has not been tested in this court.

A3. If you are exposed to information outside court, tell the usher immediately. Do not discuss the case with anyone outside your fellow jurors. Do not post about the case online.

A4. These directions apply now and throughout deliberations.

Appendix B — Outline of Pre-Service Micro-Education Module

A five-to-seven-minute neutral module delivered before empanelment:

B1. The role of evidence and the reasons we exclude extra-curial material;

B2. Burdens and standards of proof explained in plain language;

B3. How algorithms personalise online content and why that can mislead;

B4. Practical examples of improper research and how to respond;

B5. Reassurance about support available if exposure occurs.

Appendix C — AI Guardrails Checklist (For Court Administrators)

C1. Purpose limitation (education or systems-level risk mapping only);

C2. Necessity and proportionality documented;

C3. No monitoring of identifiable jurors;

C4. No analytics on deliberative content;

C5. Human judicial oversight;

C6. Data minimisation and default deletion;

C7. Public guidance;

C8. Periodic audit and sunset review.