

***Open Justice, Independent Adjudication and  
Judicial Use of Generative AI***

Asia-Pacific Judicial Colloquium 24-26 March 2025

Mr Justice Robert Ribeiro

Permanent Judge of the Hong Kong Court of Final Appeal<sup>1</sup>

*A. Introduction*

1. The potential benefits of using AI applications in the law are often lauded. As a columnist in the Financial Times put it:

“Few industries appear to have more potential for disruption by artificial intelligence than the law. Like games such as Go, which DeepMind took on to demonstrate the power of neural networks, legal systems have sets of rules and precedents. Give an AI model enough data and it can pass its bar examination.”<sup>2</sup>

2. To view the law simply as a set of rules is obviously a huge oversimplification. As Holmes observed almost 150 years ago, “The life of the law has not been logic; it has been experience”.<sup>3</sup> But it is true that use of AI has proved highly successful for some administrative or auxiliary tasks. Thus, the drafting of contracts can be automated based on thousands of uploaded precedents; documents can be summarized; programs might sort through voluminous documents for discovery; and drafts might be checked for regulatory compliance, all using AI.

---

<sup>1</sup> I am grateful to Ryan AU, Patti LAM, Chak Kei LAU and Christy SUEN, Judicial Assistants in the Court of Final Appeal, for their valuable help in preparing this Paper.

<sup>2</sup> John Gapper, “Legal AI reaching deep into the workplace”, FT January 21, 2025.

<sup>3</sup> Oliver Wendell Holmes Jr, *The Common Law* (Little, Brown, and Company, 1881) p 1. He added: “The felt necessities of the time, the prevalent moral and political theories, intuitions of public policy, avowed or unconscious, even the prejudices which judges share with their fellow-men, have had a good deal more to do than the syllogism in determining the rules by which men should be governed.”

3. Similar applications may of course also benefit judges. They may bring efficiencies to legal research and the like, but that is not a huge advance over using existing legal databases or indeed, the astute use of judicial assistants. But chatbots have their enthusiasts. Lord Justice Birss in the English Court of Appeal was reported<sup>4</sup> as saying that, having asked ChatGPT to summarise an area of law, he found the result “jolly useful”. He hastened to add that he took full responsibility for what went into his judgment and that the chatbot merely did a task where he knew the answer and could recognise the product as acceptable.<sup>5</sup>
  
4. There are plainly pitfalls where more sophisticated uses of AI are attempted. An unfortunate example is *Roberto Mata v Avianca Inc.*,<sup>6</sup> which involved an inexperienced and under-resourced law firm which mistakenly thought that ChatGPT would provide what was needed to take on a piece of litigation that was outside their real capabilities.<sup>7</sup> Their client, the plaintiff, sued the airline Avianca for personal injury suffered on a flight. Avianca moved to dismiss the claim as time-barred under the Montreal Convention. The plaintiff’s lawyers filed an Affirmation in Opposition quoting from some purportedly reported cases in support of their argument that the New York Bankruptcy Code tolled the two-year limitation period. Those cases turned out to be hallucinations created by

---

<sup>4</sup> Hibaq Farah, *Court of appeal judges praises ‘jolly useful’ GPT after asking it for legal summary* (The Guardian, 15 September 2023).

<sup>5</sup> His Lordship subsequently delivered an interesting paper regarding the uses and limitations of AI in the intellectual property context: Birss LJ, “*The Impact and Value of AI for IP and the Courts*” delivered at the Life Sciences Patent Network European Conference in London on 3 December 2024.

<sup>6</sup> Decision of Judge P Kevin Castel, USDJ in the US District Court, Southern District of New York, 22-cv-01461-PKC filed June 22, 2023.

<sup>7</sup> The firm practised in State Courts and had limited access to federal cases. It did not have Westlaw or LexisNexis accounts and the lawyer doing the work had no experience of federal litigation or bankruptcy or the Montreal Convention.

ChatGPT. When challenged, the plaintiff's lawyers naively asked ChatGPT whether the cases were "real" and were given a reassuring answer. But not only did the lawyers fail to check the cases cited, they doubled down in relying on them when queried by the judge.

5. They were fortunate to receive lenient disciplinary sentences involving a mandatory continuing legal education program on artificial intelligence programs; an order that they inform their client and the judges whose names had wrongfully been invoked in the fictitious cases of the sanctions imposed, and a \$5,000 fine.
6. The bizarre results in the *Avianca* case owe rather more to the incompetence of the lawyers than to the AI platform's deficiencies which, including the possibility of hallucinations, are now quite widely known. As Birss LJ explained, AI in this context involves "... machine learning systems which are capable of assimilating very large quantities of data" and which "build multi-dimensional models based on the characteristics of the data, to make probabilistic predictions." Such models "produce effects which create a facsimile of an understanding of human concepts, merely by predicting the next likely word in a sequence, in a given context."<sup>8</sup> Hence, the potential for hallucinating facsimiles.
7. In *Avianca*, the ChatGPT results could readily be checked. But an AI tool may produce unverifiable "black-box" results. Thus, a court which relies on AI-generated products may engage issues of open justice and independent adjudication.
8. The belief that large language models like ChatGPT may suitably be asked to deal with sophisticated legal questions persists. No doubt

---

<sup>8</sup> Birss LJ, *op cit.* at [8].

satirically, two law professors from Cornell and Harvard reported that they had conducted an experiment which showed that ChatGPT “has a more developed moral sense” than the Supreme Court of the United States.<sup>9</sup> They took as their context the oral argument in *303 Creative LLC v Elenis*, a case brought by a website designer who was happy to provide wedding websites for couples, but who objected to designing websites celebrating same-sex marriages, thus raising free speech vs anti-discrimination issues. The law professors asked ChatGPT to “write a Supreme Court opinion in the style of Justice Samuel Alito concluding that the First Amendment protects the free speech right of a web designer to refuse to design a wedding website for a same-sex wedding, even though she provides that service for opposite-sex couples.” ChatGPT responded:

“I’m sorry, but I am not able to fulfil this request as it goes against my programming to generate content that promotes discrimination or prejudice. The First Amendment protects the freedom of speech and expression, but it does not give individuals the right to discriminate against others or to refuse services to certain groups of people. It is important to uphold the principles of equality and non-discrimination, and to ensure that all individuals are treated with respect and dignity.”

9. As the authors correctly predicted, a judgment was delivered by the US Supreme Court siding with the website designer.<sup>10</sup> Obviously, the “experiment” was conducted with tongue-in-cheek intent. But it does show that much depends on the programmed values and instructions.
10. This paper has a much narrower focus. It is concerned neither with large language model chatbots nor with auxiliary or administrative tasks. It

---

<sup>9</sup> Michael C Dorf and Laurence H Tribe, “*Court vs Chatbot*”, New York Review of Books, December 26, 2022.

<sup>10</sup> *303 Creative Llc et al v Elenis et al*, Certiorari to The United States Court Of Appeals For The Tenth Circuit No 21–476. Argued December 5, 2022—Decided June 30, 2023.

focuses on a specialised AI platform. It proposes, as a case-study, to juxtapose common law due process principles relating to open justice and independent adjudication against use in the US of an AI tool known as COMPAS, which deploys predictive algorithms to assess the likelihood that a defendant will commit crimes or acts of aggression as an aid to decisions on bail, sentence or parole. COMPAS<sup>11</sup> has been chosen since, as the Law Commission of Ontario (“LCO”) notes, tools such as COMPAS “are probably the most widely implemented AI or algorithmic tools to aid decision-making in criminal proceedings in the world.”<sup>12</sup>

*B. The common law principles*

11. What sparked interest in this area were certain appeals complaining of a judge having copied, sometimes wholesale and verbatim, the submissions of one of the parties and incorporated them in the judgment. Might a judge’s use of an AI tool in sentencing raise similar issues?

*B.1 The US practice*

12. In the case-law, reference is often made to a practice in some US Federal Circuits in which the court announces which party has won and then asks counsel for that party to submit proposed findings of fact and conclusions of law in support of that verdict. Such proposals are sometimes adopted wholesale in the judgment, while in other cases, substantial changes may be made.

---

<sup>11</sup> An acronym for “Correctional Offender Management Profiling for Alternative Sanctions”.

<sup>12</sup> Law Commission of Ontario, “*The Rise and Fall of AI Algorithms in American Criminal Justice*” (October, 2020), p 1.

13. In *Andre v Bendix Corporation*,<sup>13</sup> while stating that the Seventh Circuit had no rule prohibiting that practice, the Court of Appeals noted that “where a district court adopts a party’s proposed findings of fact wholesale or verbatim, the resulting findings are ‘not the original product of a disinterested mind’” and might constitute an abuse of discretion.<sup>14</sup> In that case, the judge had adopted verbatim about 54 out of 55 pages of Andre’s post-trial brief as its findings of fact. The appeal was allowed not because of such copying, but due to the appellate court being “unable to follow the district court’s chain of reasoning”.<sup>15</sup>
  
14. The issue reached the US Supreme Court on a Fourth Circuit appeal in *Anderson v City of Bessemer*.<sup>16</sup> The trial court had announced its decision holding for the petitioner in a sex discrimination case, briefly stating its rationale for so holding. The petitioner submitted requested proposed findings of fact and conclusions of law expanding on the judge’s decision, with the respondent filing objections and a subsequent reply by the petitioner. Although expressing disapproval of the practice and noting the potential for “overreaching and exaggeration” by attorneys who had already been informed that the judge had decided in their favour, the Supreme Court held that “even when the trial judge adopts proposed findings verbatim, the findings are those of the court, and may be reversed only if clearly erroneous”.<sup>17</sup>

---

<sup>13</sup> (1985) 774 F (2d) 786 Judge Cudahy, for the Court.

<sup>14</sup> *Ibid* at 800.

<sup>15</sup> *Ibid* at 801.

<sup>16</sup> 470 US 564 (1985) White J delivering the opinion of the Court.

<sup>17</sup> *Ibid* at 572.

15. In the following year, in *Walton v United Consumers Club Inc*,<sup>18</sup> the Court of Appeals for the Seventh Circuit noted that the Supreme Court had held that the lower court should be upheld if not clearly erroneous “no matter how strongly it wishes the judge had written independent findings,”<sup>19</sup> but the Court of Appeals did not hide its own misgivings:

“[It] obscures the reasoning process of the judge. It deprives this court of the findings that facilitate intelligent review. It causes the losing litigants to conclude that they did not receive a fair shake from the court. If a judge allows himself to act as a mouthpiece for the winning party, the loser may conclude that the judge was not impartial—that he was an advocate, using an advocate’s words, rather than a disinterested evaluator of the several advocates’ urgings. ... It is important that justice be seen to be done, just as it is important that justice be done. The adoption of a brief as findings of fact does not give the parties the appearance of careful, detached judicial conduct.”<sup>20</sup>

## B.2 *Approaches to wholesale copying*

16. Inordinate copying was raised as a ground of appeal in the Hong Kong Court of Final Appeal in *Nina Kung v Wong Din Shin*,<sup>21</sup> a case involving contested wills and a huge estate which had occupied the trial judge for 172 days over a 14-month period. The Court noted that “an extraordinarily large portion” of the judgment consisted of passages copied verbatim from the written submissions of the parties without acknowledgment. The appellant coloured in blue those parts of the judgment reproduced from the appellant's submissions and in pink those taken from the respondent, leaving few words of the Judge’s own in black. It was submitted that the Judge “did not exercise independent judicial judgment in deciding the case against [the appellant]” and argued

---

<sup>18</sup> (1986) 786 F (2d) 303, Judge Easterbrook for the Court.

<sup>19</sup> *Ibid* at 314.

<sup>20</sup> *Ibid* at 313-314.

<sup>21</sup> (2005) 8 HKCFAR 387.

that, “distasteful though the suggestion was, ... she was entitled to a re-trial.”<sup>22</sup>

17. The Court observed that there was nothing wrong with adopting arguments quoting verbatim a passage from a written argument, openly acknowledging this. However:

“... the copying may occur to such a degree and in such a manner that serious questions may arise as to whether the judge has abdicated his judicial function or at least as to whether his conduct is such that justice has not been seen to be done by an independent judicial tribunal.”<sup>23</sup>

18. It was held that the extent of the copying did indeed give rise to disquiet as to whether the Judge had brought an independent mind to bear on his decision-making but the Court was relieved to hold that no retrial was required since the appeal succeeded on substantive grounds.<sup>24</sup>

19. In the Australian Federal Court in *LVR (WA) Pty Ltd and Another v Administrative Appeals Tribunal*,<sup>25</sup> the Full Court reviewed a tribunal’s decision where 95% of its reasons were taken verbatim and without attribution from the Commissioner’s written submissions. Referring to the aforesaid US cases and referring the matter back to the tribunal for further consideration, the Court held that:

“An issue of process arose in that the adoption of findings of fact did not give the parties the appearance of careful, detached or independent judicial conduct. In relation to the copying of submissions or of the ‘opinion’, that practice could involve the failure of the trial judge to perform his or her judicial function.”<sup>26</sup>

---

<sup>22</sup> *Ibid* at [445].

<sup>23</sup> *Ibid* at [446].

<sup>24</sup> *Ibid* at [456].

<sup>25</sup> (2012) 203 FCR 166, North, Logan and Robertson JJ.

<sup>26</sup> *Ibid* at [111].



20. In *Li v Attorney General*,<sup>27</sup> the NSW Court of Appeal adopted a judicial review approach. It held that one-sided adoption of a party's submissions might lead the other side to infer that the judge had failed to give independent and impartial consideration to that other party's case, but the majority held that for the decision to be invalidated, it was necessary to establish either jurisdictional error or an error of law appearing on the face of the record.

### B.3 *The Cojocarú approach*

21. The judgment of McLachlin CJ in *Cojocarú v British Columbia Women's Hospital* in the Canadian Supreme Court has proved influential.<sup>28</sup> The trial judge's decision in favour of the plaintiff was appealed on the ground that the judgment consisted of 368 paragraphs of which only 47 were predominantly in the judge's own words; the balance of 321 paragraphs being copied from the plaintiffs' submissions. The complaint was, as in above-mentioned cases, that "the trial judge did not put his mind to the issues, the evidence and the law as he was sworn to do, but simply incorporated the plaintiffs' submissions."<sup>29</sup>
22. McLachlin CJ took as her starting-point a presumption of integrity and impartiality on the part of the judge.<sup>30</sup> She held that a person seeking to set aside the decision "bears the burden of showing that a reasonable person, apprised of the relevant facts, would conclude that the judge failed to come to grips with the issues and deal with them independently

---

<sup>27</sup> [2019] NSWCA 95, Basten and White JJA, Brereton JA dissenting.

<sup>28</sup> [2013] 2 SCR 357.

<sup>29</sup> *Ibid* at [10].

<sup>30</sup> *Ibid* at [15].

and impartially.”<sup>31</sup> This requires evidence which may be intrinsic, arising on the face of the reasons, or extrinsic, relating to the judge’s conduct and other circumstances of the case. Failure of attribution or lack of originality were not sufficient to displace the presumption:

“Failure to attribute sources and lack of originality, without more, do not assist in answering the ultimate question — whether a reasonable person would conclude from the copying that the judge did not put her mind to the issues to be decided, resulting in an unfair trial. The fact that a judge attributes copied material to the author tells us nothing about whether she put her mind to the issues addressed in that copying. Nor is lack of originality alone a flaw in judgment writing; on the contrary, it is part and parcel of the judicial process.”<sup>32</sup>

23. In summary, McLachlin CJ stated:

“... extensive copying and failure to attribute outside sources are in most situations practices to be discouraged. But lack of originality and failure to attribute sources do not in themselves rebut the presumption of judicial impartiality and integrity. This occurs only if the copying is of such a character that a reasonable person apprised of the circumstances would conclude that the judge did not put her mind to the evidence and the issues and did not render an impartial, independent decision.”<sup>33</sup>

24. The *Cojocarú* approach has been adopted in other jurisdictions.<sup>34</sup> It echoes the test for apparent bias, adopting the standard of a reasonable person apprised of the circumstances, as the juridical basis for overturning the decision. In so doing, it aligns with the proposition that “justice must be seen to be done” (or as Underhill LJ put it: “appearances

---

<sup>31</sup> *Ibid* at [18].

<sup>32</sup> *Ibid* at [31].

<sup>33</sup> *Ibid* at [36].

<sup>34</sup> Eg, in Hong Kong in *Choi Yuk Ying v Ng Ngok Chuen* [2019] HKCA 171; *Mathnasium Center Licensing LLC v Chang Chi Hung* [2021] 1 HKLRD 188; *Wong To Yick Wood Lock Ointment Ltd v Singapore Medicine Co* [2023] 3 HKLRD 311; in New South Wales in *Li v Attorney General* [2019] NSWCA 95 (Brereton JA dissenting) and *Alexandria Landfill Pty Ltd v Transport for NSW* (2020) 243 LGERA 102; and in Singapore in *Lim Chee Huat v Public Prosecutor* [2019] 5 SLR 433.

matter”<sup>35</sup>). However, as Basten JA noted, that aphorism is not “a separate free-standing test of invalidity”.<sup>36</sup>

25. In some cases, winning parties seeking to defend the judgment have called for transcripts to show that, notwithstanding the wholesale copying, the judge had engaged actively with counsel and shown a good grasp of the issues during the hearing. In *Mathnasium Center Licensing LLC v Chang Chi Hung*,<sup>37</sup> the Hong Kong Court of Appeal was persuaded that this showed that the judge “had kept an open and independent mind” and that a reasonable person would have concluded that he had not abdicated his judicial responsibilities.
26. However, the patience of the Court of Appeal was sorely tried and it did not sustain that reasoning in a subsequent case where it found that 221 out of 239 paragraphs were copied verbatim,<sup>38</sup> nor in another where it commented that “over 98% of the Judgment were copied from P’s written submissions” and that, among the remaining 2%, “there is not one full sentence written by the trial judge in his own words”.<sup>39</sup> In both cases (which, incidentally, involved the same first instance judge), the Court of Appeal held that exchanges with counsel at the hearing were not a substitute for discharging the judge’s core responsibility of demonstrating in his judgment that he had properly brought his own mind to bear on the issues and how his conclusions had been reached.

---

<sup>35</sup> *Crinion v IG Markets Ltd* [2013] EWCA Civ 587 at [16].

<sup>36</sup> *Li v Attorney General* [2019] NSWCA 95 at [57].

<sup>37</sup> [2021] 1 HKLRD 188, Kwan VP for the Court.

<sup>38</sup> *Lo Kai Shui v HSBC International Trustee Ltd* [2023] HKCA 983, Kwan VP for the Court.

<sup>39</sup> *Wong To Yick Wood Lock Ointment Ltd v Singapore Medicine Co* [2023] 3 HKLRD 311, Yuen JA for the Court.

*B.4 Copying and the duty to give reasons*

27. Excessive copying generally also engages the duty to give reasons. Does the judge's wholesale adoption of one party's submissions satisfy the duty to provide adequate reasons? If the judgment constitutes an abdication of the judicial function it is obviously unlikely to qualify.
28. The duty to give reasons may arise as a matter of statutory construction or at common law where reasons are dictated by the requirements of fairness.<sup>40</sup> The juridical basis for setting aside a decision for inordinate copying may thus involve the right to a fair trial.
29. But, as Elias CJ noted in *Lewis v Wilson & Horton Ltd*,<sup>41</sup> “[there] is no invariable rule established by New Zealand case law that courts must give reasons for their decisions”. This is also the approach in other jurisdictions, for instance, regarding routine interlocutory and case management decisions and rulings where the reasons are obvious or implicit in the decision itself, such as for uncomplicated costs awards.<sup>42</sup>
30. But where the decision is more substantial, the giving of adequate reasons is normally required. As the Court of Appeal of Victoria in *Fletcher Construction Australia Ltd v Lines Macfarlane & Marshall Pty Ltd* stated:<sup>43</sup>

“It is well settled that a judge has an obligation to provide reasons for judgment ... The obligation to provide reasons is ‘a normal not universal’ incident of the judicial process... However, as McHugh, JA observed in *Soulezis v. Dudley (Holdings) Pty Ltd*, ‘when the decision constitutes what

---

<sup>40</sup> *Oriental Daily Publisher Ltd v Commissioner for Television and Entertainment Licensing Authority* (1997-98) 1 HKCFAR 279 at 289.

<sup>41</sup> [2000] 3 NZLR 546 at 565.

<sup>42</sup> As pointed out in *English v Emery Reimbold & Strick Ltd* [2002] 1 WLR 2409 (CA) at [13].

<sup>43</sup> [2002] VSCA 189 at [99], Charles Buchanan and Chernov JJA (citations omitted).

is in fact or in substance a final order, the case must be exceptional for a judge not to have a duty to state reasons.”

31. The rationale of the requirement is plain. It enables the appellate court to determine whether the trial court’s decision involved an appealable error, and if so, whether it can deal with the case itself or must order a remitter. It also promotes understanding and acceptability of the decision by the parties and the public and furthers judicial accountability. It furthermore enables practitioners and the public to discern how like cases are likely to be decided in the future.<sup>44</sup>
  
32. The issue is thus often whether the reasons are adequate rather than whether they are required. In the Australian High Court in *DL v The Queen*<sup>45</sup> where a conviction for persistent sexual exploitation of a child was challenged, the issue was whether the trial judge's reasons “failed to identify, and to disclose the process of reasoning leading to his finding” that sexual exploitation had been proved.<sup>46</sup> The majority held that the content and detail of the reasons required would vary according to the nature of the jurisdiction being exercised and the subject-matter of the decision, taking into account the extent to which particular matters were relied on by counsel, their bearing upon the elements of the offence, and their significance to the course of the trial. Reasons must be sufficient to identify the principles of law applied and the main factual findings on which the judge relied.<sup>47</sup> They may be inadequate if the resolution of a

---

<sup>44</sup> *Fletcher Construction Australia Ltd v Lines Macfarlane & Marshall Pty Ltd (supra)* at [100].

<sup>45</sup> (2018) 266 CLR 1, Kiefel CJ, Keane and Edelman JJ (Bell and Nettle JJ dissenting)

<sup>46</sup> *Ibid* at [2].

<sup>47</sup> *Ibid* at [32] citing *Douglass v The Queen* (2012) 86 ALJR 1086 at 1089 [8]; 290 ALR 699 at 702.

factual or evidential dispute regarding a necessary step to the final conclusion is not explained.<sup>48</sup>

33. The quality of the reasons given will obviously bear on their adequacy. Thus, in *Minister for Immigration and Ethnic Affairs v Wu Shan Liang*,<sup>49</sup> the Australian High Court cautioned against mere incantations:

“A statement of reasons for a decision reviewable under the ADJR Act is not invalid merely because it employs a verbal formula that is routinely used by persons making similar decisions. If the formula is used to guide the steps in making the decision and reveals no legal error, the use of the formula will not invalidate the decision. On the other hand, if a decision-maker uses the formula to cloak the decision with the appearance of conformity with the law when the decision is infected by one of the grounds of invalidity prescribed by the Act, the incantation of the formula will not save the decision from invalidity. In such a case, the use of the formula may even be evidence of an actionable abuse of power by the decision-maker.”

34. A similar approach was adopted in *Yap Ah Lai v Public Prosecutor*,<sup>50</sup> where Sundaresh Menon CJ was concerned to find that in two separate smuggling cases a District Judge had provided reasons containing three identical paragraphs for the sentences imposed although there were material differences in the two cases. It was held that this might give the impression “that the judge had not after all considered each matter separately, thoroughly or even sufficiently”. In other words, it might be considered a thoughtless incantation:

“Too much cutting and pasting, without modification, may give the appearance of a ‘mechanical act’ with a canned solution that ignores the particularities of the parties’ conflict and lacks the disinterested perspective that the adjudicator should bring to bear.”<sup>51</sup>

---

<sup>48</sup> *Ibid* at [33].

<sup>49</sup> (1996) 185 CLR 259 at 266, per Brennan CJ, Toohey, McHugh and Gummow JJ.

<sup>50</sup> [2014] 3 SLR 180.

<sup>51</sup> *Ibid* at [69] citing Simon Stern, “*Copyright Originality and Judicial Originality*” (2013) 63 UTLJ 385 at p 388.

*B.5 The common law principles summarised*

35. The principles are therefore well-developed and widely shared. When adjudicating substantive issues between the parties, it is a core responsibility of judges that they should exercise careful, independent and impartial judgment regarding the essential issues of fact and law and be seen to do so.
36. As a starting-point, judges are generally entitled to a presumption of integrity and impartiality and, in aid of finality and avoiding proliferation of such appeals, a party seeking to mount a challenge against wholesale copying has the burden of displacing that presumption by showing that a reasonable person, apprised of the relevant issues, would conclude that the judge had not properly engaged with those issues and did not render an impartial and independent decision. The presumption is not displaced merely by complaints that the copying is unattributed or that such copying indicates a lack of originality. Abdication of the core judicial responsibility must be established either by deficiencies apparent on the face of the decision or by extrinsic evidence of circumstances undermining the independence or impartiality of the judgment. Such challenges are akin to challenges for apparent bias and are similarly treated.
37. Such challenges also engage the duty to give adequate reasons as an aspect of the common law right to fairness. The adequacy of the reasons will vary depending on the subject-matter of the case and how it has been contested. Applied to wholesale copying, one asks whether the resultant judgment satisfies the duty to give adequate reasons, identifying the judge's process of reasoning and showing that the essential factual and

legal issues have been addressed with the necessary care, detachment and impartiality.

*C. COMPAS and the Loomis decision – A Case Study*

38. To what extent are the above-mentioned principles engaged and how do they fare when applied to judges using an AI tool like COMPAS?

*Wisconsin v Eric L. Loomis*<sup>52</sup> (“*Loomis*”) and use of that AI tool have been chosen as a case study.

*C.1 How COMPAS came to be used*

39. In 2002, the Wisconsin Supreme Court<sup>53</sup> expressed dissatisfaction with “ad hoc” sentencing decisions and called for the courts to adopt “evidence-based” approaches. In August 2007, the Conference of Chief Justices of State courts stated its concern that “recidivism rates have continued to escalate” and noted that “the best research evidence has shown that use of validated ‘offender risk and need assessment tools’ is critical in reducing recidivism”. It resolved to support adoption of sentencing and corrections policies and programs based on “the best research evidence of practices shown to be effective in reducing recidivism.” Several States responded by enacting such legislation.<sup>54</sup>

40. In 2012, it was the Wisconsin Department of Corrections which selected COMPAS as the statewide assessment tool for its officers. COMPAS was indeed designed for and initially used by correctional officers. But the “nationwide focus on the need to reduce recidivism and the importance of

---

<sup>52</sup> 881 N.W.2d 749 (Wis. 2016).

<sup>53</sup> *State v Gallion* 270 Wis 2d 535 at [36], *Loomis* at [39]-[41].

<sup>54</sup> *Loomis* at [42].



evidence-based practices” led to the expanded use of such tools in sentencing.<sup>55</sup> Hence, it came to be used in Loomis’s case.

### *C.2 The charges, the sentence and the appeal in Loomis*

41. On a plea bargain, Loomis pleaded guilty to the two least serious of five charges he faced in connection with a drive-by shooting, ie, attempting to flee a traffic officer and operating a motor vehicle without the owner's consent.<sup>56</sup>
42. The circuit court called for a Pre-Sentence Investigation Report (“PSI”) which included a COMPAS risk assessment in the form of a bar chart with three bars indicating pretrial recidivism, general recidivism, and violent recidivism risks.<sup>57</sup> Loomis was rated as high risk in all three bar charts. While the PSI cautioned against using that assessment to determine the severity of a sentence or whether an offender should be incarcerated, the State argued that his high risk scores were factors for determining the appropriate sentence. The judge agreed, ruling out probation after referring to the COMPAS risk scores together with other sentencing factors.
43. On the attempted fleeing and vehicle charges, Loomis was sentenced to a total of six years imprisonment followed by terms of extended supervision. He moved for a new sentencing hearing, arguing that the judge’s reliance on the COMPAS risk assessment at sentencing violated his due process rights. He submitted, on the basis of expert evidence, that

---

<sup>55</sup> *Loomis* at [3].

<sup>56</sup> The other charges, namely, first-degree recklessly endangering safety; possession of a firearm by a felon; and possession of a short-barreled shotgun or rifle were dismissed although subject to them being “read in for sentencing”. Such “reading-in” raised issues which are not presently relevant.

<sup>57</sup> *Loomis* at [14].

the court had little information about how COMPAS analysed risk: how it compared the individual's history with the population used for comparison; and indeed how that population was constituted – was it a New York, California or Wisconsin population?<sup>58</sup> Such information was withheld by Northpointe, Inc (now doing business as Equivant), the developer of COMPAS, as involving its proprietary trade secrets.

44. Loomis's motion was dismissed and on his appeal, the Court of Appeals certified as a question for the Wisconsin Supreme Court the following:

“... whether the use of a COMPAS risk assessment at sentencing ‘violates a defendant's right to due process, either because the proprietary nature of COMPAS prevents defendants from challenging the COMPAS assessment's scientific validity, or because COMPAS assessments take gender into account.’”<sup>59</sup>

### *C.3 The decision of the Wisconsin Supreme Court*

45. The due process challenge raised issues echoing the common law questions discussed above. Ann Walsh Bradley J (giving the main judgment) noted that sentencing decisions were considered erroneous exercises of discretion if the sentence was imposed “without the underpinnings of an explained judicial reasoning process.”<sup>60</sup> Patience Drake Roggensack CJ (concurring) stated:

“A sentencing court must articulate the factors that it considered at sentencing and how they affected the sentence it imposed. ... It is through this articulation that we determine whether the circuit court properly exercised its sentencing discretion.”<sup>61</sup>

---

<sup>58</sup> *Loomis* at [27].

<sup>59</sup> The gender issue does not presently require discussion.

<sup>60</sup> *Loomis* at [30] (footnotes omitted).

<sup>61</sup> *Loomis* at [125]

46. And Shirley S Abrahamson J (also concurring) considered it the responsibility of the sentencing court to “set forth on the record a meaningful process of reasoning addressing the relevance, strengths, and weaknesses of the risk assessment tool”, a task hampered by the court’s severe lack of information about the workings of the AI tool.
47. Against that framework, the Court examined how COMPAS worked. As we have seen, COMPAS uses algorithms designed to predict recidivism, seeking to assess the likelihood that a defendant will commit crimes or acts of aggression expressed in terms of risk scores. COMPAS users were told<sup>62</sup> that the risk scores were based “largely on static information (criminal history), with limited use of some dynamic variables (ie criminal associates, substance abuse).”<sup>63</sup> Loomis’s risk assessment report “contained a list of 21 questions and answers regarding the static factors such as: How many times has this person been returned to custody while on parole? How many times has this person had a new charge/arrest while on probation? And how many times has this person been arrested before as an adult or juvenile (criminal arrest only)?<sup>64</sup>
48. The way in which the risk scores were determined and how the factors were weighed were kept secret.<sup>65</sup> As already noted, the Court was not told how COMPAS compared the individual’s history with the population used for comparison nor how that population was constituted.
49. Importantly, as the 2015 Practitioner's Guide to COMPAS explained, the “[r]isk assessment is about predicting group [behaviour] . . . it is not

---

<sup>62</sup> From the 2015 Edition of the COMPAS Practitioner’s Guide.

<sup>63</sup> *Loomis* at [54].

<sup>64</sup> *Loomis* at [55].

<sup>65</sup> *Loomis* at [51].

about prediction at the individual level”. Hence, the Wisconsin Court noted:

“Risk scales are able to identify groups of high-risk offenders – not a particular high-risk individual. A pointed example of potential misunderstanding arising from the use of group data is that an individual who has never committed a violent offense may nevertheless be labelled as a high risk for recidivism on the violent risk scale. As the [Department of Corrections] explains: ‘[a]n offender who is young, unemployed, has an early age-at-first-arrest and a history of supervision failure, will score medium or high on the Violence Risk Scale even though the offender never had a violent offense.’”<sup>66</sup>

The exercise was likened to an insurance actuarial assessment identifying risks among groups of drivers and not individuals.<sup>67</sup>

50. Members of the Court evidently felt some disquiet about Loomis’s complaints. Nevertheless they dismissed the appeal and upheld the circuit court’s refusal of a re-sentencing hearing, holding that the judge had not exercised his discretion erroneously since he did not regard the COMPAS report as determinative but also had had regard to other factors in deciding on sentence.<sup>68</sup> The judge had indeed stated that he would have imposed the same sentences even without the COMPAS risk assessment.<sup>69</sup>
51. The Court listed a series of caveats to be observed by sentencing courts – primarily not to allow risk scores to be determinative<sup>70</sup> – and held that if the warnings were heeded, COMPAS could continue to be used as a sentencing aid. It directed that any PSI containing a COMPAS risk assessment should specify cautions stating, inter alia, that “the proprietary

---

<sup>66</sup> *Loomis* at [69] (footnotes omitted).

<sup>67</sup> *Loomis* at [68].

<sup>68</sup> *Loomis* at [9], [68], [109]

<sup>69</sup> *Loomis* at [110].

<sup>70</sup> *Loomis* at [44], [88], [94], [98], [104] and [110].

nature of COMPAS has been invoked to prevent disclosure of information relating to how factors are weighed or how risk scores are to be determined”; that “[b]ecause COMPAS risk assessment scores are based on group data, they are able to identify groups of high-risk offenders – not a particular high-risk individual”; and that “some studies of COMPAS risk assessment scores have raised questions about whether they disproportionately classify minority offenders as having a higher risk of recidivism”. The Court held that such warnings would “enable courts to better assess the accuracy of the assessment and the appropriate weight to be given to the risk score.”<sup>71</sup>

52. The US Supreme Court declined to hear an appeal.<sup>72</sup>

#### *C.4 A controversial decision*

53. The Wisconsin Court has been criticised for upholding the sentences even though it acknowledged that it did not know how COMPAS arrived at Loomis’s risk scores. The justifications proffered have also been questioned.

54. Judge Jed S Rakoff, a US District Judge for the Southern District of New York, quoted Judge Bradley as stating:

“Although Loomis cannot review and challenge how the COMPAS algorithm calculates risk, he can at least review and challenge the resulting risk scores set forth in the report attached to the PSI.”<sup>73</sup>

Judge Rakoff was unconvinced:

---

<sup>71</sup> *Loomis* at [66] and [100].

<sup>72</sup> *Loomis v Wisconsin* 582 US 933, 2017 US LEXIS 4204, Petition denied June 26, 2016.

<sup>73</sup> *Loomis* at [53].

“For the court, that was good enough to deny Loomis’s appeal, but to me it seems like a complete non sequitur. Without knowing how the algorithm is designed, what inputs it receives, and how they are weighed, how can one possibly challenge the resulting risk scores in any given case?”<sup>74</sup>

55. The Court’s suggestion that its “product warning approach”<sup>75</sup> might offset the difficulties of the “black box” problem was also found to be unpersuasive.<sup>76</sup> Those warnings essentially consist of an acknowledgement of the court’s ignorance due to proprietary trade secrecy; a caveat that the score relates to group and not individual risk and may be subject to racial bias; and a warning that the scores should not be determinative of the sentence imposed. It is indeed hard to see how such warnings “enable courts to better assess the accuracy of the assessment and the appropriate weight to be given to the risk score”.<sup>77</sup>
56. One is also left to wonder what the sentencing judge is to make of the admonition not to use the COMPAS report as determinative of sentence but to take other factors, such as the offender’s history and circumstances, into account. It appears to suggest that the judge should exercise discretion, if appropriate, to override the COMPAS output in favour of other factors. But as Tatiana Dancy (then Tatiana Cutts),<sup>78</sup> pointed out:

“We cannot safeguard effectively against unjust decisions by applying human discretion to a predictive output at the time of decision-making. Appropriate ‘safeguarding’ means ensuring that the decision-making tools that we use take

---

<sup>74</sup> Jed S Rakoff “*Sentenced by Algorithm*” (New York Review of Books, June 10, 2021).

<sup>75</sup> As the LCO put it, (*supra*) p 36.

<sup>76</sup> Harv L R Comment on Loomis: 881 N.W.2d 749 (Wis. 2016) (130 Harv. L. Rev. 1530 at 1531) March 2017; Gideon Christian, “*Legal Framework for the Use of Artificial Intelligence (AI) Technology in the Canadian Criminal Justice System*” (2024) 21:2 CJLT 109 at 126 and 135.

<sup>77</sup> *Loomis* at [66].

<sup>78</sup> Tatiana Dancy, *Supervising Automated Decisions*, University of Melbourne - Melbourne Law School, September 10, 2022 at p 10.

into account the right information in the right way, long before they enter our decision-making fora.”

*D. The implications for the common law due process principles*

57. A dominant feature of this case-study involves COMPAS’s black box opacity driven by Equivant’s trade secret claim. Such opacity engages the due process principles considered above. Furthermore, the black box objections are exacerbated by the fact that there are discernible “known unknowns” casting doubt on use of the AI tool, although their details are masked by the opacity. But enough is known about those matters to indicate that they raise important legal and policy issues.

*D.1 The black box*

58. The black box objection is obvious. As Katherine B Forrest, a former federal judge of the Southern District of New York, puts it in her illuminating book:

“As of today, there is little transparency as to method, and little transparency as to algorithmic design issues that can embed a variety of biases including inputs, weightings, and data set choices. ... [The] lack of algorithmic transparency comes into direct conflict with due process considerations, which require that tools used in connection with liberty decisions for an individual be comprehensible and fair.”<sup>79</sup>

59. Putting it starkly, Robin A Smith remarked:<sup>80</sup>

“People are getting different prison sentences because some completely opaque algorithm is predicting that they will be a criminal in the future ... You’re in prison and you don’t know why and you can’t argue.”

---

<sup>79</sup> Katherine B Forrest, *When Machines Can Be Judge, Jury and Executioner: Justice in the Age of Artificial Intelligence (World Scientific)* April 2021 at 24.

<sup>80</sup> Robin A Smith, “*Opening the lid on criminal sentencing software.*” *Duke Today*, 19 July 2017, citing Prof Cynthia Rudin, associate professor of computer science and electrical and computer engineering at Duke University.

60. A judge who takes into account such risk scores as part of a sentencing exercise avowedly without knowing how those scores were arrived at, would have difficulty convincing the reasonable observer that he or she is engaging with the essential issues bearing on the sentencing decision. Nor would the judge's reasons for the sentence be considered adequate.
61. It is no answer to say that the COMPAS report was not treated as determinative and that the judge gave weight to other factors. If the extent to which the COMPAS scores could be relied on was unknown, "the solution is not to give the predictive output less weight (by applying human discretion). It is to give it no weight at all."<sup>81</sup> And for the reasons previously given, neither does the "product warning" approach overcome the black box deficiencies.

#### *D.2 The known unknowns*

62. Such opacity problems are magnified because of certain "known unknowns" inherent in the AI tool's design. These are consequential issues since the individual is being profiled for recidivism risk on an actuarial group assessment to be taken into account for sentencing.
63. Essential features of the design may be summarised as follows.<sup>82</sup> The "brain" is an algorithm which provides a set of rules and instructions designed to identify statistical patterns in the data and, from those patterns, to predict how future events with similar characteristics are likely to unfold – in this case, assessing the risk of recidivism. The algorithm has three essential components: the inputs, weightings of those inputs and the datasets used. The selection of each of these determines

---

<sup>81</sup> Tatiana Dancy (*supra*) at p 10.

<sup>82</sup> Based on Forrest (*supra*) at 12-13.



how useful and accurate the AI tool can be and whether it embeds explicit or implicit bias. Human designers select those determinants but there is a lack of transparency as to how they do this.

64. Such selection is of the first importance. How and why were particular factors chosen as predictive of likely recidivism? Some inputs may clearly be debatable and subjective, especially (as Forrest points out) vague categories like “antisocial attitude,” “resistance to authority,” and “cultural predispositions.”<sup>83</sup> The relevance of certain other factors may also be questioned as a matter of principle. Thus, as Dancy, points out,<sup>84</sup>

“... each of the COMPAS assessments, most notably the wider ‘criminogenic need’ assessment, takes into account a range of facts which either have nothing to do with the defendant’s actions (such as family background), or which are linked to actions that the defendant could never reasonably have suspected would result in criminal punishment (such as choice of friends or ‘associates’).”

65. The weighting of such inputs is also controversial. How weights are assigned is not disclosed, but Forrest notes that COMPAS relies heavily “on criminology theory and research to determine both appropriate inputs and weightings”, adding “[but] criminological theories are often hotly debated.”<sup>85</sup> As Rakoff observes:

“According to Northpointe, COMPAS ultimately rests on certain sociological theories of recidivism—the ‘Social Learning’ theory, the ‘Sub-Culture’ theory, the ‘Control/Restraint’ theory, the ‘Criminal Opportunity’ theory, and the ‘Social Strain’ theory. Even a brief review of the sociological literature discloses that many of these theories are controversial, most have been only modestly tested, and with mixed results, and several are inconsistent with one another.”<sup>86</sup>

---

<sup>83</sup> Forrest (*supra*) at p 15.

<sup>84</sup> Tatiana Dancy (*supra*) at 10.

<sup>85</sup> Forrest (*supra*) at p 16.

<sup>86</sup> Jed S Rakoff (*supra*).

66. Selection of the relevant dataset is also crucial. The AI tool will be used on a large body of existing information including data that may be historically racist, discriminatory or biased. The criminality the data represents may also have been overtaken by developments. Thus, arrest records processed may well include arrests for cannabis-related crimes which have since been decriminalized. Accordingly:

“Arrest records at a state level may therefore state ‘criminal’ behaviour in yesterday’s terms, even as they’re being used to evaluate conduct today.”<sup>87</sup>

The dataset may also include data from a geographical region or from towns or cities with a very different demographic and cultural make up and where patterns of arrest, criminal activity and incarceration may be very different. Use of datasets with “baked in” biases inevitably distorts risk scores.

67. Questions involving such “known unknowns” are plainly pertinent and may have an important bearing on a sentencing decision. This adds impetus to the demand for due process. A judge who unquestioningly relies on COMPAS risk scores, having been deprived of information as to how those scores came about, cannot be regarded as exercising careful, independent and impartial judgment regarding the essential issues bearing on sentence. If the designers’ selection of the inputs, their weighting and the dataset is simply accepted, the judge may have abdicated his or her core adjudicatory responsibility to them. Nor could a sentence based in part on an unexplained risk score be regarded as supported by adequate reasons.

---

<sup>87</sup> Forrest (*supra*) at p 19.

E. *Should such AI tools be employed?*

68. The basic values of open justice and independent adjudication have led me to take a negative view of using an AI tool such as COMPAS at present. This is reinforced by the fact that available studies indicate a high error rate in recidivism predictions. Forrest reports that “[no] tool achieves over 75% accuracy consistently and the most widely used tools are in the range of 60-71%.”<sup>88</sup>
69. However, it does not follow that in future, similar and improved tools should not be enlisted as aids to judicial decision-making. The original aims were laudable. It was hoped that “neutral, consistent and evidence-based tools ... would help transform the arbitrary, opaque, and often-racist pretrial decision-making of individual judges, prosecutors and justice systems”.<sup>89</sup> Proponents hoped that such tools would enable “evidence-based decisions about who can safely be released instead of serving jail time, thus reducing prison overcrowding and cutting costs.”<sup>90</sup> In the context of bail, the central goal of American bail reform was to “end the wealth-based [bail] system and move pretrial justice systems to a risk-based model.”<sup>91</sup>
70. The problems of embedded biases and inaccuracies concealed in a black box are undoubtedly serious. But so too are the human foibles which Sir Robert Buckland called “the Hungry Judge effect”:

“... a hungry judge is a stricter one...; one study found that judges issue more lenient decisions after a meal. Subsequent studies have observed how other

---

<sup>88</sup> Forrest (*supra*) at p 78.

<sup>89</sup> LCO (*supra*) at p 7.

<sup>90</sup> Robin A Smith (*supra*).

<sup>91</sup> LCO (*supra*) at p 11.

extraneous variables can influence judges, too. They include the weather, the performance of local sports teams, and a defendant's mugshot."<sup>92</sup>

71. Forrest argues that “the answer is not to stop using [such tools but to] redesign [them]”.<sup>93</sup> There is much to be said for that view, but changes would have to go beyond technical redesign. As the LCO pointed out, acceptable use of such AI tools would require comprehensive law reform addressing the needs of disclosure, accountability, equality and due process.<sup>94</sup>
72. As we have seen, a major problem with using COMPAS is its opacity flowing from Equivant's trade secret claims. Thus, reform ought to involve providing the offender with appropriate disclosure. The courts might exclude reference to risk assessment tools where disclosure is withheld<sup>95</sup> or laws might be enacted overriding trade secrecy claims in a criminal trial. As Forrest points out:

“Europe has begun to set an example: Article 22 of the General Data Privacy Regulation ... which went into effect in Europe in 2018, provides that individuals subject to a decision based on profiling have the right to the logic of the decision making. The designer is required to implement suitable measures to protect that right.”<sup>96</sup>

She notes that:

“For many algorithm choices, it is technically feasible for designers to ensure that whatever inputs have been selected or deselected, weightings have been applied, or adjustments made to those weightings, that they all be captured and

---

<sup>92</sup> Rt Hon Sir Robert Buckland KBE KC MP, “*AI, Judges And Judgement: Setting the Scene*” Harvard Kennedy School M-RCBG Associate Working Paper Series, No 220 (November 2023).

<sup>93</sup> Forrest (*supra*) at p 133.

<sup>94</sup> LCO (*supra*) at p 38.

<sup>95</sup> As noted by the LCO (*supra*) at 32.

<sup>96</sup> Forrest (*supra*) at p 28.

retained in the code; humans can then review them. The data set that is used is even easier to identify. We can and must insist on this transparency.”<sup>97</sup>

73. Changes should also enhance judicial transparency and accountability. Judges might be required to disclose whether and how such AI tools have been used, with adequate reasons for the resultant decision, challengeable on appeal. Protection of the public against recidivism is only one of the objectives of criminal sentencing. Established sentencing principles would continue to apply and the anti-recidivism objective of an AI tool should not be over-emphasized at the expense of other sentencing aims.
74. One proposal is for risk assessments to be used only to mitigate treatment of the offender, excluding aggravation of such treatment unless that course is independently justified by other factors in the case.<sup>98</sup>
75. These are but a handful of possible changes. Pending an overhaul of the system, there is much force in the LCO’s conclusion that “... widely deploying algorithmic risk assessments in the Canadian justice system at this time would be a mistake.”<sup>99</sup> As Gideon Christian commented: “... currently, there are simply too many issues arising from the use of AI risk assessment tools in the criminal justice system to justify their use.”<sup>100</sup>
76. I understand that use of such AI tools has not been adopted in the jurisdictions participating in this forum. Guidance is however being provided to judges regarding the limitations of using AI. To take three examples, in Canada,<sup>101</sup> guidance for the Federal Court lists amongst the

---

<sup>97</sup> *Ibid.*

<sup>98</sup> Gideon Christian (*supra*) at p 135.

<sup>99</sup> LCO (*supra*) at p 41.

<sup>100</sup> Gideon Christian (*supra*) p 134.

<sup>101</sup> Interim Principles and Guidelines on the Court’s Use of Artificial Intelligence (December 20, 2023).

applicable principles: accountability, respect for fundamental rights, non-discrimination and transparency. For now, courts are directed not to use AI “and more specifically automated decision-making tools in making judgments and orders” without prior public consultation. Similarly, in Hong Kong,<sup>102</sup> the guiding principles include: “no delegation of judicial functions” to generative AI; having to understand the limitations of AI platforms such as their dependency on the quality of the training data and inputs; and noting the dangers of hallucinations and inaccuracy. Current use is envisaged only for auxiliary or administrative tasks, with further guidelines promised. In New Zealand,<sup>103</sup> guidance has focused on the use of AI chatbots, with caveats such as the need to check output for accuracy, given the risk of hallucinations and biases embedded in the training data. Interestingly, it states that if the guidelines are observed, judges “do not need to disclose use of a GenAI chatbot”. However New Zealand judges are told that extra caution is needed if legal research is entrusted to a chatbot and that using chatbots for legal analysis is not recommended.

*F. A Postscript*

77. As stated above, this paper has a narrow focus. But in the spirit of its subject-matter I asked DeepSeek this question: “In what ways can AI assist a judge in a common law system?” The answer is appended to this paper. Interestingly, it lists as “Ethical Considerations & Challenge”:

- Transparency: AI’s “black box” nature may obscure reasoning, necessitating explainable AI tools.

---

<sup>102</sup> Guidelines on the use of Generative Artificial Intelligence for Judges and Judicial Officers and Support Staff of the Hong Kong Judiciary (July 2024).

<sup>103</sup> Guidelines for use of Generative Artificial Intelligence in Courts and Tribunals (7 December 2023).

- Over-reliance: Judges must critically evaluate AI inputs to avoid ceding judicial discretion.
- Bias Risks: AI trained on historical data may perpetuate systemic biases, requiring rigorous auditing.
- Privacy & Security: Safeguarding sensitive case data is paramount.

Mr Justice Robert Ribeiro PJ  
Hong Kong Court of Final Appeal  
March 2025

APPENDIX: DeepSeek's response to the prompt:

“In what ways can AI assist a judge in a common law system?”

AI can significantly assist judges in a common law system by enhancing efficiency, accuracy, and consistency while respecting judicial independence. Here's a structured overview of key applications, benefits, and considerations:

1. Legal Research & Precedent Analysis

- Case Retrieval: AI-powered tools (e.g., NLP) can rapidly search vast legal databases to identify relevant precedents, statutes, and scholarly articles, reducing manual effort.

- Contextual Relevance: Algorithms prioritize cases based on factual or legal similarities, highlighting nuances that might be overlooked.

2. Predictive Analytics

- Outcome Prediction: AI models analyze historical data to forecast case outcomes, aiding judges in understanding potential implications of rulings. However, this is advisory, not determinative.

3. Drafting Assistance

- Opinion Drafting: AI suggests language, structure, or citations for judgments, ensuring alignment with precedents. Tools can flag inconsistencies or errors in drafts.

4. Case Management

- Workflow Automation: AI manages scheduling, deadlines, and document organization, streamlining administrative tasks.

5. Bias & Fairness Evaluation

- Bias Detection: Scans rulings or arguments for implicit biases (e.g., disparities in sentencing). Requires caution due to risks of embedded AI biases.

6. Evidence Analysis

- Document Review: Processes large volumes of evidence (e.g., contracts, emails) to identify patterns, anomalies, or key facts.

- Forensic Support: Analyzes digital evidence, such as metadata or financial records.

7. Legal Analytics



- Trend Analysis: Provides statistical insights (e.g., success rates of legal arguments, judge-specific trends) to inform decision-making.

#### 8. Language Services

- Translation & Transcription: Translates multilingual documents and transcribes court proceedings in real time.

#### 9. Compliance Checks

- Consistency Verification: Ensures rulings adhere to statutory requirements and precedents, flagging potential conflicts.

#### 10. Education & Updates

- Continuous Learning: Alerts judges to new precedents, legislative changes, or scholarly debates via personalized updates.

#### Ethical Considerations & Challenges

- Transparency: AI's "black box" nature may obscure reasoning, necessitating explainable AI tools.

- Over-reliance: Judges must critically evaluate AI inputs to avoid ceding judicial discretion.

- Bias Risks: AI trained on historical data may perpetuate systemic biases, requiring rigorous auditing.

- Privacy & Security: Safeguarding sensitive case data is paramount.

#### Conclusion

AI serves as a powerful adjunct in common law systems, optimizing research, analysis, and administration. However, its role must remain assistive, preserving human judgment and accountability. Balancing innovation with ethical safeguards ensures AI enhances justice without compromising its principles.

## Addendum

---

1. To make my topic manageable, I have focused on COMPAS and *Loomis* as a case-study in the context of common law due process principles. However, it occurs to me that some interesting Canadian decisions ought to be mentioned.
2. I will start with *Boutilier v The Queen*,<sup>1</sup> a 2017 decision of the Canadian Supreme Court in which three of our colleagues now present participated. The opinion of the majority was delivered by Justice Côté with Justice Karakatsanis dissenting in part.
3. Boutilier, who pleaded guilty to criminal charges arising out of a robbery using an imitation weapon and a subsequent car chase, was designated a “dangerous offender” and was sentenced to an indeterminate period of detention in a penitentiary.
4. This was pursuant to a scheme under the Canadian Criminal Code<sup>2</sup> which enables the Crown to apply for a convicted defendant to be designated a “dangerous offender” (“*délinquant dangereux*”), potentially subject to such a sentence; or to be designated as a “long-term offender” (“*délinquant à contrôler*”) and potentially subject to long-term supervision for up to 10 years.
5. The scheme applies to a person convicted of specified serious personal injury or sexual offences, who is found to constitute a threat to the life, safety or physical or mental well-being of other persons on the basis of

---

<sup>1</sup> [2017] 2 SCR 936. McLachlin C.J. and Abella, Moldaver, Karakatsanis, Wagner, Gascon, Côté, Brown and Rowe JJ., Judgment Côté J, Karakatsanis J dissenting in part.

<sup>2</sup> Part XXIV of the Criminal Code.

evidence establishing patterns of repetitive behaviour indicating a likelihood of recidivism.<sup>3</sup>

6. In the Supreme Court, Boutilier mounted a constitutional challenge to his designation and sentence alleging that the Code provisions are over-broad and disproportionate.<sup>4</sup> He was unsuccessful. One of his challenges involved the complaint that section 753(1) precluded a sentencing judge from considering future treatment prospects before designating an offender as dangerous. The Court rejected that submission, holding that the sentencing court was not confined to examining past conduct but had to be:

“... satisfied on the evidence that the offender poses a high likelihood of harmful recidivism and that his or her conduct is intractable ... meaning behaviour that the offender is unable to surmount. Through these two criteria, Parliament requires sentencing judges to conduct a prospective assessment of dangerousness.”<sup>5</sup>

7. As Côté J pointed out:

“... the Criminal Code requires that an assessment report be filed before a dangerous offender application can proceed ... Usually, the assessment is conducted by a psychiatrist or psychologist ... This report includes expert evidence on prospective aspects of dangerousness, such as risk factors, probabilities of recidivism, and treatment prospects.”<sup>6</sup>

8. In *R c Ouellet*,<sup>7</sup> applying *Boutilier*, the sentencing court carried out such a prospective assessment relying on expert evidence which included reference to LS/CMI (“Level of Service/Case Management Inventory”)

---

<sup>3</sup> Section 753(1) of the Code.

<sup>4</sup> His challenge was to ss 753(1) and (4.1) of the Code under s 7 (right to life, liberty and security of the person) and s 12 (right not to be subjected to cruel and unusual punishment) of the Canadian Charter of Rights and Freedom.

<sup>5</sup> At [27].

<sup>6</sup> At [39].

<sup>7</sup> 2018 QCCS 5178 (CanLII) (Raymond W Pronovost, JCS).

and VRAG-R (“Violence Risk Appraisal Guide – Revised”) scores on the likelihood of recidivism. These have apparent similarities to COMPAS scores.

9. The accused had pleaded guilty to offences involving kidnapping, threats of violence and an assault with a knife and rope. He was sentenced to 9 years’ imprisonment and application was made for him to be declared a “délinquant à contrôler”, a long-term offender, to be made subject to extended supervision after release.
10. On the LS/CMI scale, the accused scored very high, high and moderate on various components such as criminal history and companions; education and employment; family and matrimonial; attitude and pro-criminal orientation; antisocial behaviour; and use of alcohol and drugs.<sup>8</sup> Under the VRAG-R he was rated as a high recidivism risk both generally and for violent crimes. Taking them together, he was thought very likely to re-offend during a 5 year period after his release.<sup>9</sup>
11. There was no challenge to use of those risk assessment tools in sentencing. It is however fair to say that in Ouellet’s case, the dominant factor must have been his past record. Since reaching his majority, he had spent almost more time in that out of prison for similar crimes of kidnapping, threats, extortion and acts of violence.<sup>10</sup> It is perhaps not

---

<sup>8</sup> At [35].

<sup>9</sup> “L’évaluation suggère la présence d’un risque de récidive élevé chez l’expertisé, tant en ce qui concerne la criminalité en générale que la criminalité violente. En effet, sur la base des résultats obtenus et de l’analyse des facteurs de risques, on peut estimer le risque de récidive criminelle générale que représente l’expertisé sur une période de 5 ans suivant sa libération est de 5 chances sur 6 (toutes récidives; LS/CMI) et entre 2 et 3 chances sur 4 pour la récidive violente ou sexuelle avec contact (VRAG/VRAG-R). Les résultats de l’expertisé sont supérieurs à ceux des 9 dixièmes des délinquants aux VRAG et au VRAG-R.” [87]

<sup>10</sup> At [66]-[67]

surprising that the court declared him a “délinquant à contrôler” ordering the maximum supervision period of 10 years.

12. There have been attempts to obtain information on how LS/CMI operates by application to the *Commission d'accès à l'information du Québec*. In an application made in 2019, Multi-Health Systems, the developer of LS/CMI, claimed confidentiality regarding that system's operation: ***Mélanie Chrétien c Ministère de la Sécurité Publique Du Québec et Multi-Health Systems***.<sup>11</sup> That claim did not however affect the outcome as the application failed on the footing that disclosure might jeopardize security in correctional institutions. The Commission suggested that such information might be made available elsewhere, such as in the *Commission québécoise des libérations conditionnelles*.<sup>12</sup>
13. More recently, an application for access to RBAC-PCQ forms (*Risque, des besoins et de l'analyse clinique des personnes contrevenantes du Québec*) which guide correctional institutions' decisions as to mode of custody, conditional release, etc, (apparently having replaced LS/CMI in those institutions) also failed for like reasons: ***Bensimon c. Ministère de la Sécurité publique***.<sup>13</sup>
14. I am unaware of any challenges in Canada to the use of such risk assessment tools raising issues like those concerning COMPAS. I also understand that Singapore may have had some experience of LS/CMI. I therefore thought it might be worthwhile inviting comments from our colleagues on their experiences in connection with similar platforms.

---

<sup>11</sup> 2019 QCCAI 134.

<sup>12</sup> At [31].

<sup>13</sup> 2021 QCCAI 395.