Using Artificial Intelligence (AI) in Criminal Justice
Mike Nellis

On April 18th (2024) I attended a four-hour online Westminster Legal Policy Forum seminar on "Next Steps for AI technology in the Criminal Justice System" in England and Wales. It was an excellent overview of current official thinking on artificial intelligence (AI) and criminal justice, and confirmed that major digital innovation is being anticipated in this sector (and in public services more generally), about which many people are underinformed. While international scientific debate on AI is 70 years old, operational use of it in governance (military applications apart) only began in earnest, worldwide, in 2018, and was then accelerated in 2022 by the surprise advent of "large language models", aka "generative AI", epitomised by ChatGPT and its imitators.

The seminar flagged up the "AI-powered changes" (to use the jargon!) that are coming to criminal justice, especially the police – the leader in the field - the judiciary and the legal profession. The 55 participants were largely senior people from a range of state agencies and professional bodies tuning in for "training" on developments which will imminently affect them all in some degree. Influential speakers were drawn from the Centre for Justice Innovation, the Police Foundation, the National Audit Office, the Criminal Bar Association (and its international equivalent), techUK (a trade organisation), the enchantingly(?) named Alchemy Machines (a tech start-up) - all of whom described their work in progress on AI - as well as the human rights defenders, Justice and Big Brother Watch, who urged caution.

My heart tended to be with the latter organisations. My own thoughts on infusing AI into criminal justice (and public service generally) tend to be more critical than the mostly mainstream views expressed here. I don't think AI can be plausibly be regarded as a "neutral tool" which possesses easily identified "benefits" and "risks" that can with due care be mitigated and minimised. Yet this is the near universal, official Western view of what AI is. It is not neutral because the governments and corporations which sponsor AI have forceful agendas of their own about what our collective western futures should look like, and limitless resources for seducing consumers into thinking that this represents the common good. But that's another story: here I will provide a condensed but honest summary of what was said in the seminar.

Although the term "AI" was used as a kind of shorthand throughout the seminar the point was regularly made that there are several variants of it, and multiple potential uses in government and criminal justice, some arguably more contentious than others. Sometimes "advanced digitalisation" may more aptly describe the developments addressed here. But the more extensive datasets become and the greater the computing power used to process them the more "digitalisation" segues into territory that can legitimately be called 'AI', because it replicates human cognitive processes - but at incomparably vast scales and speeds. "Machine learning" perhaps better names software which becomes progressively more autonomous at "pattern recognition" across multiple datasets, which is essential for identifying risks, making predictions and then offering recommendations in specific contexts. While the seminar mostly addressed this "traditional" AI, new "generative AI" is posing its own challenges, and the Criminal Bar Association referenced developing responses to AI-generated crime, notably deepfake alibis and nudifcation apps.

The Centre for Justice Innovation reported on a survey of attitudes towards using AI in criminal justice settings, among both relevant professionals, the general public and crime victims, informed by the perspectives it developed in its own (very impressive) *JustTechnology* report in 2018. All professionals are understandably curious about AI's likely implications for them, but

there are as yet relatively few developments on the ground – which is not to say there are none. Some police forces have experimented for a while with enhanced behavioural risk assessments (which shade into "predictive policing"), while their contentious use of facial recognition cameras is becoming more widespread, all of which require vast, speedily searchable data bases.

Questions about making AI trustworthy, transparent and accountable pervaded the seminar, alluded to by most speakers. The "ethics of AI" are being extensively debated across the world. There is widespread international agreement that informal guidelines and probably formal regulation are needed in all settings where it is used – although there is a British twist on this emanating from Rishi Sunak's "AI Safety" Summit in 2023 (see below). AI-informed systems need to be effective, reliable and fit for purpose if they are not to do harm, and win public trust. AI systems also need to be transparent in the sense that the algorithmic and statistical techniques used for making recommendations (eg on risk, or for sentencing) must not be inaccessible, especially when retrospective questioning is required. Not all are transparent: some commercial providers demand proprietal confidentiality. Above all, the agencies using AI systems and/or the companies selling them must be accountable if things go wrong - "the computer made me do it" must never be an operative's excuse.

The National Audit Office (NAO) - which assesses "value for money" in government projects - has recently published a lucid and insightful report on existing, piloted and planned uses of AI in central government (but not local government or police forces). It sees huge potential benefits in efficiency and productivity from AI - mostly cost-savings and "doing more for less" in public administration and service delivery. It encourages both the improvement of existing services and crucially, the development of entirely new services - eg new, *automated* ways of delivering once-personal services, via chatbots. To this end it identifies strategic and operational obstacles to the expansion and intensification of AI use across government, namely lingering investment in outdated "legacy systems"; the absence (as yet) of a coherent AI plan; anxiety in government of where accountability should lie; and continuing ambivalence about which standards, guidance and ethics to apply.

Although technically a spending watchdog, the NAO was foursquare with the bold ambitions of the junior minister and senior civil servant from the Ministry of Justice who spoke at the seminar. Both drew on the AI Safety Summit in 2023 in wanting "guardrails" rather than "tramlines" for AI's development, ethical guidelines which do not stifle the technological innovation which will make Britain a "world-leader" in this field. The minister believed that merely using AI to improve "archaic practices" would sell it short and wanted something new: he was up for challenging risk aversion in criminal justice and proud that LawTech UK had been deliberately set up to stimulate innovation at arm's length from government itself. Watch this space …….

Justice reiterated long-established human rights warnings about continuing risks of discrimination in AI systems because human prejudices can still be designed-in to, and reproduced by, the databases on which they draw. Big Brother Watch averred that, with digital technology, "just because you can, you don't have to" – not, I think, a widely held view in government. Bug Brother Watch further insisted on attending to what preserving "the human" might now mean in an AI-infused world. **Don't just watch this space, Friends – work in it**.