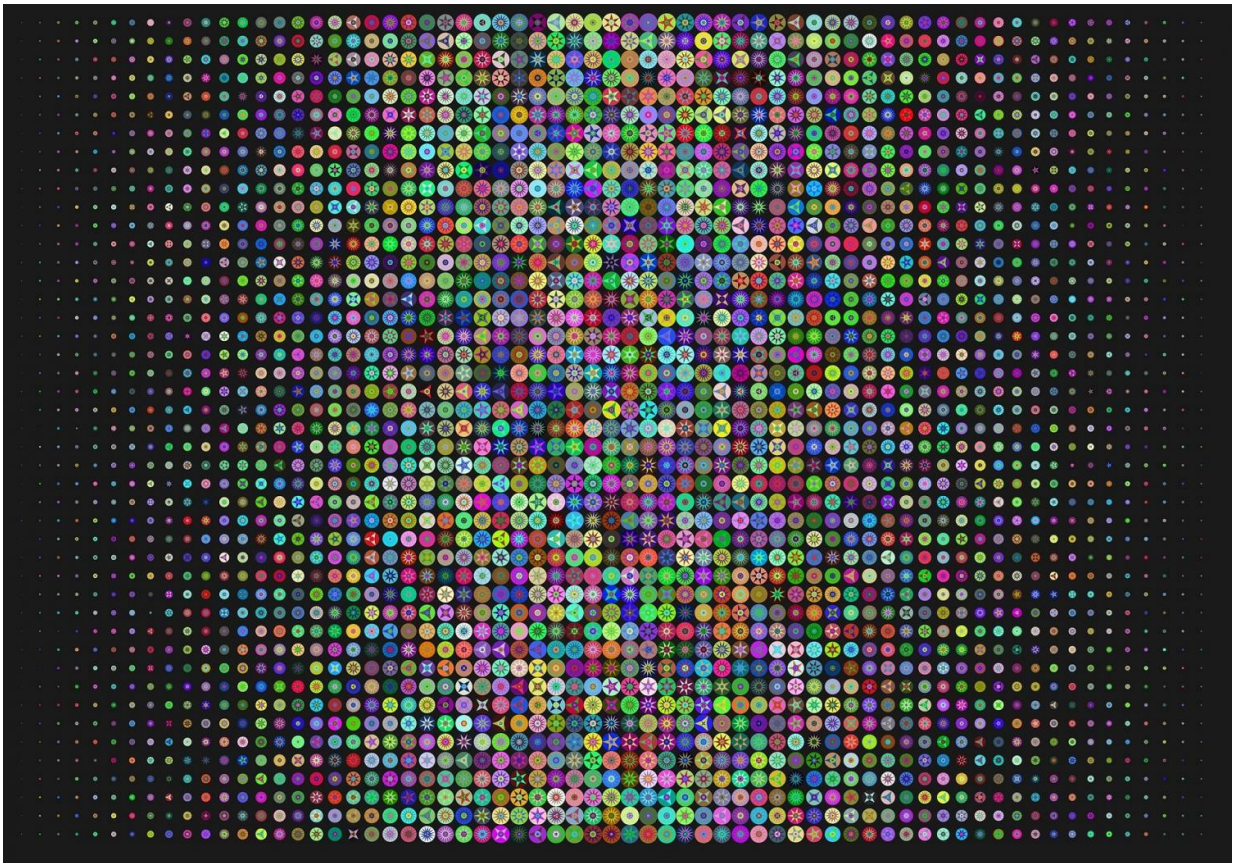


AI sentencing tools need to be closely scrutinised, says new study

June 9 2020



Credit: CC0 Public Domain

In a paper published by the *Behavioral Sciences & Law* journal, experts from the University of Surrey take a critical look at the growing use of algorithmic risk assessment tools, which act as a form of expert

scientific evidence in a growing number of criminal cases.

The review argues that because of several issues, such as the biases of the developers and weak statistical evidence of the AI's predictive performance, judges should act as gatekeepers and closely scrutinise whether such tools should be used at all.

The paper outlines three steps that judges should consider:

- Fitness, this is to consider whether using the AI [tool](#) is relevant to the case at all
- Accuracy, this is to understand whether the tool can truly distinguish between reoffenders and non-reoffenders
- Reliability, this would require the judges to scrutinise the trustworthiness of tool's outcomes in practice. This stage would not be required if the [judge](#) found the AI lacking in the one of the first two steps.

Dr. Melissa Hamilton, author of the paper and Reader in Law and Criminal Justice at the University of Surrey's School of Law, said: "These emerging AI tools have the potential to offer benefits to judges in sentencing, but close attention needs to be paid to whether they are trustworthy. If used carelessly these tools will do a disservice to the defendants on the receiving end."

More information: Melissa Hamilton, Judicial gatekeeping on scientific validity with risk assessment tools, *Behavioral Sciences & the Law* (2020). [DOI: 10.1002/bsl.2456](https://doi.org/10.1002/bsl.2456)

Provided by University of Surrey

Citation: AI sentencing tools need to be closely scrutinised, says new study (2020, June 9)
retrieved 5 May 2024 from <https://phys.org/news/2020-06-ai-sentencing-tools-scrutinised.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.