

NHS Blueprint for Testing Bias in AI Models



AI BIAS

The UK's National Health Service (NHS) is trialling a programme designed to identify and eradicate algorithmic biases in systems used to administer healthcare. The NHS Artificial Intelligence (AI) Ethics Lab commissioned the Ada Lovelace Institute to develop an Algorithmic Impact Assessment (AIA) tool. This tool is intended to provide an algorithmic accountability mechanism to support researchers in assessing potential risks and biases of AI systems and algorithms before they can access NHS data.

While AI can support health and care workers in delivering better care for people, it could also exacerbate existing health inequalities if algorithmic biases remain unaddressed. This work complements ongoing work from the NHS AI Lab ethics team on ensuring datasets for training and testing AI systems are diverse and inclusive. As a condition of accessing NHS data, this forces developers to explore and address the legal, social, and ethical implications of their proposed AI systems. However, this is not intended as a complete solution and should be complemented by algorithmic accountability initiatives, like audits or transparency registers. Furthermore, AIA should not replace existing regulatory frameworks but complement those already used in the UK.

The NHS press release was accompanied by an Ada Lovelace research that details step-by-step processes for using AIAs in the real world. One example of AIA provided in the report is that of Canada, which was created to manage public sector AI delivery and procurement standards.

Source: [UK Department of Health and Social Care](#), [Ada Lovelace Institute](#)

Published on : Mon, 14 Mar 2022