# AI model Foresight sparks NHS data privacy debate amidst healthcare innovation



An emerging trend in healthcare is casting a long shadow over data privacy concerns as artificial intelligence (AI) increasingly shapes the field. The recent development of an AI model named Foresight, which has been trained on medical data from 57 million individuals within England’s National Health Service (NHS), epitomises both the potential benefits of AI in healthcare and the significant ethical dilemmas that accompany such initiatives. Researchers, including Chris Tomlinson from University College London's Institute of Health Informatics, aim to pioneer what they describe as the world's first "national-scale generative AI model of health data," leveraging an impressive array of approximately 10 billion health events recorded from late 2018 to the end of 2023.

Foresight, built upon Meta’s open-source Llama 2 model, incorporates a vast spectrum of datasets, from outpatient appointments to vaccination records. Its architects posit that the model could revolutionise treatment by enabling healthcare professionals to predict disease complications before they arise, thereby facilitating earlier interventions. Tomlinson underscored the transformative potential of Foresight, stating that “the real potential… is to predict disease complications before they happen, giving us a valuable window to intervene early,” promoting a shift towards preventative healthcare at scale.

However, employing such extensive datasets has raised considerable alarm among privacy advocates and researchers alike. While developers assert that the data used is “de-identified,” experts remind us that the intricacies inherent in large datasets heighten the risk of re-identification—a problem well-documented in the realm of data privacy. Luc Rocher, a researcher at the University of Oxford, articulated the challenges, suggesting that the very richness of the data makes it notoriously difficult to anonymise. He advocated for maintaining strict control over such models by NHS authorities to ensure patient safety.

Moreover, transparency has emerged as another critical point of contention. Caroline Green from the University of Oxford highlighted that using extensive datasets without adequate notification to individuals could erode public trust. “Even if it is being anonymised, it’s something that people feel very strongly about…because people usually want to keep control over their data and they want to know where it’s going,” she said. This sentiment is echoed in recent surveys, revealing that 56% of UK citizens harbour distrust regarding the NHS’s use of AI in patient data analysis, citing privacy and security as paramount concerns.

The matter is further complicated by existing legal frameworks. Current opt-out provisions applicable to NHS datasets do not extend to the data utilised by Foresight, as it is deemed de-identified and thus not classified as personal data under GDPR guidelines. However, guidance from the UK Information Commissioner’s Office (ICO) cautions against equating "de-identified" data with anonymity, flagging potential ambiguities in the law. The Foresight initiative currently operates under specific exceptions that arose during the COVID-19 pandemic, permitting a context where patient data is entangled with public health research.

Doubts surrounding the ethical governance of AI systems also return to the fore. Critics of the UK government's AI action plan, which proposes creating a National Data Library facilitating AI innovation, underscore potential misuse of anonymised health records. Concerns linger regarding the robustness of existing data protection laws and the implications these developments may hold for patient confidentiality and trust in the NHS. Sam Smith from medConfidential cautioned against the “covid-only AI” narrative, maintaining that data embedded in model training could compromise patient autonomy further.

Despite these concerns, the government has expressed ambition to embrace AI technology as a means to redefine healthcare delivery. Health Secretary Wes Streeting framed this as an effort to minimise unnecessary hospital visits and expedite diagnoses, while Science and Technology Secretary Peter Kyle hailed it as a critical step toward healthcare innovation. NHS transformation lead Vin Diwakar echoed sentiments on the potential of AI in predictive healthcare, stating, “AI has the potential to change how we prevent and treat diseases, especially when trained on large datasets.”

Looking ahead, the developers aim to expand Foresight's capabilities, potentially integrating even richer data sources, such as blood tests and clinical notes. With objectives that encompass identifying health disparities and supporting population-wide risk assessments, early indications suggest Foresight could predict over 700 health conditions with a lead time of up to five years. As this project unfolds, the journey navigates significant tension between leveraging technology for public good and ensuring the sanctity of patient rights.

The debate surrounding Foresight is reflective of broader concerns within the tech landscape, where instances of data misuse have raised alarms. The use of NHS patient records by Google's DeepMind, for instance, was controversially classified as "legally inappropriate" due to insufficient consent, highlighting the paramount importance of safeguard mechanisms when handling sensitive health information. Such precedents serve as cautionary tales for anyone advocating for data-driven innovations.

As healthcare continues to advance towards a technology-driven future, it is imperative that ethical considerations rise alongside innovation. With evidence suggesting public enthusiasm for AI adoption exists—evident in a recent survey revealing 54% support for AI in patient care—there remains a gap in trust, particularly among older demographics. Ensuring adequate communication around data usage and bolstering privacy protections will be vital in nurturing public confidence as we tread this delicate path.

### Reference Map

1. Paragraphs 1-2, 4-5, 8-9
2. Paragraphs 3, 8-9
3. Paragraphs 3, 8
4. Paragraphs 4, 5
5. Paragraph 4
6. Paragraph 6
7. Paragraph 7

Source: [Noah Wire Services](https://www.noahwire.com)

## Bibliography

1. <https://winbuzzer.com/2025/05/12/nhs-health-data-used-for-ai-model-training-causing-privacy-concerns-xcxwbn/> - Please view link - unable to able to access data
2. <https://www.theguardian.com/society/2025/jan/13/ministers-mull-allowing-private-firms-to-make-profit-from-nhs-data-in-ai-push> - In January 2025, UK ministers considered allowing private companies to profit from NHS data to advance AI in healthcare. The plan aimed to create a National Data Library for startups and researchers to develop new treatments and diagnostic tools. While officials emphasized public sector control over data, concerns arose about privacy and ethical implications. Experts highlighted the risks of re-identification and the need for robust data protection measures to maintain public trust in the NHS's data-sharing practices.
3. <https://www.theguardian.com/politics/2025/jan/13/labour-ai-action-plan-nhs-patient-data-why-causing-concern> - The UK's Labour Party expressed concerns over the government's AI action plan, particularly regarding the use of NHS patient data. The plan proposed creating a National Data Library to facilitate AI innovation, including potential commercial use of anonymized health records. Critics worried about privacy risks, re-identification possibilities, and the adequacy of existing data protection laws. The debate underscored the need for transparent and ethical data governance to balance innovation with patient confidentiality.
4. <https://www.newscientist.com/article/2131256-google-deepmind-nhs-data-deal-was-legally-inappropriate/> - In 2017, Google's DeepMind accessed 1.6 million NHS patient records without proper consent to test a kidney disease monitoring app. The UK's National Data Guardian deemed the data-sharing agreement 'legally inappropriate,' as it violated data protection laws. This incident highlighted significant concerns about patient privacy and the necessity for explicit consent when sharing sensitive health information for research purposes.
5. <https://www.the-scientist.com/googles-deepmind-uks-nhs-criticized-for-sharing-data-31284> - Reports criticized Google's DeepMind and the UK's NHS for sharing patient data without proper consent. The Information Commissioner's Office found that the Royal Free NHS Trust breached the Data Protection Act by providing patient information to DeepMind. This raised significant privacy concerns and underscored the importance of adhering to data protection regulations when handling sensitive health data.
6. <https://archpublichealth.biomedcentral.com/articles/10.1186/s13690-024-01414-1> - A scoping review published in the Archives of Public Health examined perceived threats to patient rights and safety due to AI in healthcare. Findings indicated concerns about unpredictable errors, inadequate policies, and data security breaches. The study emphasized the need for robust data protection measures and ethical guidelines to ensure patient privacy and safety in AI applications.
7. <https://www.infosecurity-magazine.com/news/uk-nhs-ai-privacy/> - A 2023 survey revealed that 56% of UK citizens distrust the NHS's use of AI to analyze patient data, citing security and privacy concerns. Additionally, 25% of respondents opposed the NHS using AI for processing their health information. These findings highlight public apprehension regarding data privacy in AI applications within the healthcare sector.