

CLIVER SHERLOCK



The DeepMind debacle demands dialogue on data

Mishandling of patient information shows how governments and companies must become more worthy of trust, says **Hetan Shah**.

Without public approval, advances in how we use data will stall. That is why a regulator's ruling against the operator of three London hospitals is about more than mishandling records from 1.6 million patients. It is a missed opportunity to have a conversation with the public about appropriate uses for their data.

This month, the UK Information Commissioner's Office declared that the hospital operator had broken civil law when it gave health data to Google's London-based subsidiary DeepMind (whose artificial-intelligence technology has defeated leading human players of the board game Go). DeepMind was to develop an app to check test results for signs of acute kidney injuries. But the arrangement failed to consider how patients expect their data to be used, and by whom. (There was no ruling against DeepMind).

This episode is disheartening for groups that promote the power of data for the public interest, such as the Royal Statistical Society, which I lead. I hope for a world where data is at the heart of understanding and decision-making. To achieve this we need better public dialogue.

The United Kingdom has not shied away from innovation. It was the first country to explicitly allow 'three-parent baby' technology, which enables women with mitochondrial disease to have healthy children by inserting a nucleus from one of their eggs into a denucleated egg from a healthy donor. The process was approved after years of debate, open calls for evidence and fact-finding. The Royal Free London hospital group and DeepMind were much less open, although DeepMind said in a statement that its focus on its work, rather than accountability, had been wrong and it had improved transparency.

Society already suffers from a 'data trust deficit'. A 2014 poll commissioned by the Royal Statistical Society found that 32% of respondents had low levels of trust in Internet companies in general but 54% had low trust in them to use personal data appropriately. It found similar patterns across nearly all institutions.

What can be done to address this deficit? Beyond meeting legal standards, all relevant institutions must take care to show themselves trustworthy in the eyes of the public. The lapses of the Royal Free hospitals and DeepMind provide, by omission, valuable lessons.

The first is to be open about what data are transferred. The extent of data transfer between the Royal Free and DeepMind came to light through investigative journalism. In my opinion, had the project proceeded under open contracting, it would have been subject to public scrutiny, and to questions about whether a company owned by Google — often accused of data monopoly — was best suited to create a relatively simple app.

The second lesson is that data transfer should be proportionate to the task. Information-sharing agreements should specify clear limits. It is unclear why an app for kidney injury requires the identifiable records of

every patient seen by three hospitals over a five-year period.

Finally, governance mechanisms must be strengthened. It is shocking to me that the Royal Free did not assess the privacy impact of its actions before handing over access to records. DeepMind does deserve credit for (belatedly) setting up an independent review panel for health-care projects, especially because the panel has a designated budget and has not required members to sign non-disclosure agreements. (The two groups also agreed a new contract late last year, after criticism.)

More is needed. The Information Commissioner asked the Royal Free to improve its processes but did not fine it or require it to rescind data. This rap on the knuckles is unlikely to deter future, potentially worse, misuses of data. People are aware of the potential for over-reach,

from the US government's demands for state voter records to the Chinese government's alleged plans to create a 'social credit' system that would monitor private behaviour.

Innovations such as artificial intelligence, machine learning and the Internet of Things offer great opportunities, but will falter without a public consensus around the role of data. To develop this, all data collectors and crunchers must be open and transparent. Consider how public confidence in genetic modification was lost in Europe, and how that has set back progress.

Public dialogue can build trust through collaborative efforts. A 14-member Citizen's Reference Panel on health technologies was convened in Ontario, Canada in 2009. The Engage2020 programme incorporates societal

input in the Horizon2020 stream of European Union science funding.

Last month, the Nuffield Foundation in London — which has established an independent council on bioethics issues such as mitochondrial donation — announced plans to establish an independent convention on data ethics. The Foundation will work with the British Academy, the Royal Society, the Royal Statistical Society and the Alan Turing Institute for data science to bring society, policymakers, industry and researchers together for urgent conversations about the role of data in our society.

There are many questions to address. How can one understand and respect standards of privacy and consent when devices continuously communicate data? How can decisions made by algorithms be held to account for bias? What ethics codes should 'data scientists' work to?

Done differently, the data sharing between the Royal Free and DeepMind could have sparked such discussions. We need to begin this conversation soon, or the data trust deficit will grow and the opportunities for data to improve society will diminish. ■

Hetan Shah is executive director of the Royal Statistical Society in London and is on Twitter @HetanShah.
e-mail: h.shah@rss.org.uk

INNOVATIONS
OFFER GREAT
OPPORTUNITIES
BUT WILL FALTER
WITHOUT
PUBLIC
CONSENSUS.