Public services, personal data and machine learning: prospects for infrastructures and ecosystems.

Proceedings Paper:

Keen, J orcid.org/0000-0003-2753-8276, Ruddle, R, Palczewski, J orcid.org/0000-0003-0235-8746 et al. (4 more authors) (2019) Public services, personal data and machine learning: prospects for infrastructures and ecosystems. In: Kaya, T, (ed.) 19th European Conference on Digital Government (ECDG 2019). European Conference on Digital Government, 24-25 Oct 2019, Nicosia, Cyprus. Academic Conferences Ltd . ISBN 9781510899018

Public Services, Personal Data and Machine Learning:
Prospects for Infrastructures and Ecosystems

Justin Keen[1], Roy Ruddle[2], Jan Palczewski[3], Georgios Aivaliotis[3], Muhammad Adnan[4], Anna Palczewska[4], Christopher Megone[5]

[1] Leeds Institute of Health Sciences
[2] School of Computing
[3] School of Mathematics
[4] Leeds Institute of Data Analytics
[5] School of Philosophy

University of Leeds, Leeds, England

j.keen@leeds.ac.uk
r.a.ruddle@leeds.ac.uk
j.palczewski@leeds.ac.uk
g.aivaliotis@leeds.ac.uk
m.adnan@leeds.ac.uk
a.palczewska@leeds.ac.uk
c.b.megone@leeds.ac.uk

Key words: machine learning, visualisation, governance, personal data, health care, social care

Abstract

There is a widespread belief that machine learning tools will improve decision-making in health and social care. Equally there are concerns that the new tools, used with large personal datasets, will jeopardise privacy and erode trust. We reflect on experiences gained in the course of the Quanticode research and development project in England. These suggest that the opportunities are real: it is possible to generate insights that are valued by health and social care planners. The concerns are also real, though, indicating that there is a need to address them, and to balance opportunities and risks. The terrain is also contested, with evidence of differences in values relating to the ownership of datasets in particular. We argue that developments in the governance of tools and datasets will be substantially shaped by the concerns and by debates over values.

Introduction

In the last few years a range of academics, policy makers and commercial interests have come to believe that a combination of machine learning tools and access to large personal datasets opens up new analytical opportunities. Some observers further argue that analyses will be produced at lower costs, and/or produce results with greater accuracy, than established methods (AHSN Network 2018, Future Advocacy 2019). Considerable sums of money are being invested, in many countries, to exploit the opportunities. The implied ideal is one where there are infrastructures that can hold a wide range of datasets, that can be linked together with one another, and on which new software tools can be trained, validated and used. The infrastructures would be supported by governance arrangements that facilitated analyses.

To set against these claims, a number of concerns have been raised. These include the possibility that government agencies and private firms will use the tools and datasets to understand and predict our behaviour, and hence invade our privacy and threaten our autonomy (O'Neil 2016, Ezrachi and Stucke 2016). In the health and social care domain there has been considerable debate – which predates the advent of new tool/dataset combinations – about privacy, confidentiality, and the implications for trust in institutions (O'Neill 2002, Kennedy 1994).

We report on work undertaken in the Quanticode machine learning and visualisation project, which has focussed on the possibility of developing what we have termed 'ethically aware tools'. Our approach in this paper is to combine our practical experiences with reflections on current academic thinking about the governance of machine learning and personal datasets (Dourish 2016, Dourish and Cruz 2018). We argue that developments in governance practices are likely to substantially shape developments in the health and social care domain. The next two sections discuss our experiences of acquiring datasets and of developing tools. We contrast our experiences with normative assumptions. Then, in the following sections, we reflect on the implications for the development of tools in the future, and more broadly on implications for information governance arrangements in health and social care.

Acquiring datasets

Quanticode is a research project which has developed machine learning and visualisation tools over the last 3 years. The tools are designed to enable end users to analyse quantitative and coded longitudinal data. Such data are common in many sectors. For example, health care data are classified using a hierarchy of hundreds of thousands of Read Codes (a thesaurus of clinical terms). Analysts need to provide business intelligence for decisions about the effective allocation of resources, and Quanticode has been working with partners, including managers and analysts in health and social care organisations, who have helped us to identify potential applications of tools and appropriate datasets. Methods for mining large health care, social care and other datasets have been developed, as have interfaces designed to allow managers and analysts to interrogate their own data in new ways, using new analytical tools. Key technical results are reported elsewhere (Adnan and Ruddle 2018, Adnan et al 2019, Palczewska et al 2017).

The proponents of machine learning implicitly assume that it is easy to agree on project objectives, to acquire datasets and develop tools. None of these assumptions held true for us. Acquiring a dataset involved negotiation across a patchwork of organisations including ethics committees, organisations holding datasets (which involved data sharing agreements and sometimes also data access applications) and accreditation bodies (eg for ISO27001 on data management), as well as our project partners. Negotiations proved to be time-consuming, and involve extensive form-filling and post-form checking and re-checking of details. Once we had worked our way through these processes some datasets arriving within weeks, but one took over a year from the time that (we understood that) release was approved.

Beyond these administrative costs, there were other issues which contributed to the significant time costs. Partners did not initially understand the governance of research projects, and in particular the need to draft data sharing agreements and other paperwork – such as ethics applications - carefully. Some versions of agreements referred only to staff – thus excluding students – and assumed that we would only need datasets for the duration of projects (and not for an extended period afterwards, to cover the acceptance of research articles).

To set against this there were also positive – if also time-consuming – aspects of these processes. We organised meetings with partners to determine project objectives (which went into more detail than had been possible in the research proposal). This included the research team writing the 'benefits of research' section, then asking partners to re-draft it in their own language, with us then checking the wording before submission. This helped to ensure that the potential benefits were both 'real', reflecting partners' expectations, and realistic, reflecting our knowledge of tools' likely capabilities.

Software tools

The possibilities and problems associated with machine learning and visualisation became apparent in practice. On the plus side it was possible to develop new tools, and to generate novel insights that were valued by health and social care partners (Ruddle and Hall 2019, Titarenko et al 2019). A key aim is automate the analysis of complex datasets. Model development, including the development of machine learning models, typically comprises four processes – prepare a dataset for analysis (including transformation and checking data quality), design the model, implement/train the model, refine and interpret the outputs (Sacha et al 2019). As noted above, the implicit assumption about machine learning is that tool development is relatively unproblematic. Our experiences offer mixed evidence for this assumption. On data preparation:

1. Novel visualization designs were instrumental in revealing previously undetected data quality issues across very high-dimensional data (116 variables) in multiple data extracts (Ruddle and Hall 2019);
2. A novel visualization tool allowed users to both detect and explain rarely occurring but important problems with missing values. It allowed the data provider to feed back to frontline staff, to improve future data quality;
3. Less positively, issues included 100,000 records that could not be used (there was no primary diagnosis), and un-linkable identifiers due to inconsistent encryption and widespread use of invalid characters (the lowest 2 out of 6 levels of detail in diagnostic codes could not be used);

4. Many hours of input from a domain expert were essential for reducing the complexity of one of the datasets so that modelling was feasible.

In model design and implementation the team:

1. Was able to identify factors that explained inconsistent recording of patients' diagnoses. Machine learning (random forests) was used to identify important features, and interactive visualization used to investigate those features to explain outputs;
2. Developed a framework for mining temporal patterns when data was known to have inconsistencies/errors in recording of time-stamps (Palczewska et al 2017);
3. Implemented an optimised algorithm for mining of frequent temporal patterns with errors in time-stamps and temporal constraints (Titarenko et al 2019).

At the same time, the characteristics of the datasets constrained our analyses. Some of these constraints were anticipated, including missingness, which has been a focus of our research. Others were more subtle. For example, in an analysis of likelihood of admission to a nursing home, we found that many people who were admitted were not judged to be at risk at the time of admission. We discovered that key information was not available, which concerned informal carers (typically spouses or children). Events in a carer's life, such as them breaking a leg and being unable to care, and that might therefore precipitate admission of the cared-for person, could simply not be predicted from the available data.

Implications: tools and datasets

The hopes for machine learning and visualisation are not – our experiences suggest – just hype. They should, though, be tempered in two ways. The first is that we used single datasets – datasets from single servers in host organisations, that did not require any linkage before or after release to us. We are aware that there is a general assumption that it will be possible and desirable to link hitherto separate health and social care datasets in future. To take the example just mentioned, if we were able to link datasets with information on both informal carers and cared-for people in a locality, we might be able to make more accurate and/or more timely inferences about them. The practical point here is that data linkage would be more burdensome, in terms of form-filling and negotiation – it would take yet more time. Indeed, the legal bases for linking datasets is not always clear, and desired linkage might not be possible, e.g. linkage of data on services to any data that might reveal the identities of staff providing services, as there is no legal basis for releasing National Health Service (NHS) or local authority staff data in England. Looking ahead, though, we also observe that our capacity to make inferences will only increase over time: we return to this point below.

Second, as Dourish (2016) observes, outputs are inscrutable. The nature of machine learning outputs meant that it was difficult to explain both the methods used and the confidence that we had in the results, even to technically knowledgeable partners. The problem would be multiplied for explaining to the people represented in the datasets that we used. This is an accountability problem in the straightforward sense of the term: it is very difficult to provide meaningful accounts.

Implications: governance

We were struck by the patchwork of regulatory organisations. Our interpretation of our experiences is that the organisations were managing their own risks, in significant part by requiring us to assume them. This might be due in part to the novel – and uncertain - nature of the analyses we were proposing, but we sense also to a broader, institutionalised aversion to risk. There have been a number of analyses of these problems, particularly by philosophers and socio-legal scholars, which have provided conceptual clarification of key terms such as consent, confidentiality and trust (Ezrachi and Stucke 2016, Mittelstadt and Floridi 2016, O'Neill 2002,). It appears that these have not, to date, penetrated very far into the world of practice.

The domain is also contested. We came across genuine differences of opinion, particularly about consent, privacy and confidentiality. If, for example, you believe that you have property rights over a personal dataset, you may not believe that you need to obtain anyone's consent to analyse it. This belief appears to have driven high profile health care dataset releases in England (Powles and Hodson 2017, Health and Social Care Information Centre 2014). If you don't assert any rights, on the other hand, you may feel that you need to go out of your way to ensure that stakeholders are aware of, and support, your proposed analyses.

Discussion

Enthusiasts for machine learning envisage infrastructures that hold a wide range of datasets, and support linking and analysing those datasets, making ever-more sophisticated inferences from them. The infrastructures will be supported by governance arrangements – platform ecosystems – that facilitate analyses. We are currently a long way from this state of affairs, at least in health and social care. It has been possible to develop tools that are valued by partners, but our experiences have also highlighted institutional barriers that have constrained our work.

The current arrangements appear, to us, to have two important weaknesses. The first concerns accountability. Current arrangements are essentially bureaucratic, with organisations negotiating with one another, and little with wider publics. If trust in these organisations to manage personal data is at risk, as it seems to be in England, then future governance arrangements will need to be more outward-looking. They might need to involve active engagement with a wide range of stakeholders – including the people who are represented in datasets - and the development of effective ways of explaining methods, findings and their interpretation. Arrangements of this kind have been mooted, but there is no immediate prospect of them being implemented across England (O'Neill 2002, Ezrachi and Stucke 2016). One implication of these changes would be increased time costs associated with the various negotiations. While the new tools and datasets might lead to new insights, therefore, it is not obvious that they will do so at lower cost than more classical statistical methods.

The second weakness arises from an increased capacity to make inferences. The challenge will be to regulate that capacity, so as to facilitate some inferences and block others. We are not aware of practical solutions to this governance problem. We speculate that the problem, together with embedded governance arrangements – with many risk-averse actors - and the contested nature of the terrain, will limit the rate of progress towards the enthusiasts' ideal.

References

Adnan M, Nguyen P, Ruddle R, Turkay C. (2019) Visual analytics of event data using multiple mining methods. *Proceedings of the EuroVis Workshop on Visual Analytics* (EuroVA) 2019.

Adnan M, Ruddle R. (2018) A set-based visual analytics approach to analyze retail data. *Proceedings of the EuroVis Workshop on Visual Analytics* (EuroVA) 2018.

AHSN Network (2018) Accelerating artificial intelligence in health and care: results from a state of the nation survey. (https://bit.ly/30K8AHM - accessed 6th June 2019)

Dourish P. (2016) Algorithms and their others: algorithmic culture in context. *Big Data and Society* July-December 1-11.

Dourish P, Cruz E. (2018) Datafication and data fiction: narrating data and narrating with data. Big Data and Society July-December 1-10.

Ezrachi A, Stucke M. (2016) Virtual competition: the promise and perils of the algorithm-driven economy. Cambridge MA: Harvard University Press.

Future Advocacy (2018) Ethical, social and political challenges of artificial intelligence in health. London: Wellcome.

Kennedy I. (1994) Between ourselves. *Br J Med Ethics* 20:69-70.

Mittelstadt B, Floridi L. (2016) The ethics of big data: current and foreseeable issues in biomedical contexts. *Sci Eng Ethics* 22:303-41.

O'Neil C. (2016) Weapons of math destruction: how big data increases inequality and threatens democracy. London: Penguin.

O'Neill O. (2002) Autonomy and trust in bioethics. Cambridge: Cambridge University Press.

Palczewska, A., Palczewski J., Aivaliotis G., Kowalik L. (2017) RobustSPAM for inference from noisy longitudinal data and preservation of privacy. *Proceedings of 6th IEEE International Conference on Machine Learning and Applications* (ICMLA) 2017.

Palczewska, A., Palczewski J., Marchese Robinson, R., Neagu, D. (2014) Interpreting random forest classification models using a feature contribution method, *Integration of Reusable Systems, Advances in Intelligent Systems and Computing* 263, 193-218.

Powles J, Hodson H. (2017) Google DeepMind and healthcare in an age of algorthms. *Health and Technology* 7:351-67.

Ruddle R; Hall M, (2019)*Using Miniature Visualizations of Descriptive Statistics to Investigate the Quality of Electronic Health Records.* Proceedings of the 12th International Joint Conference on Biomedical Engineering Systems and Technologies - HEALTHINF 5:230-238.

Sacha D, Kraus M, Keim D, & Chen M. (2019) VIS4ML: An Ontology for Visual Analytics Assisted Machine Learning. *IEEE Transactions on Visualization and Computer Graphics* 25(1): 385-395.

Health and Social Care Information Centre (2014) *Data release review* (chair: Sir Nick Partridge). Leeds: HSCIC.

Titarenko S, Titarenko V, Aivaliotis G, Palczewski J. (2019) Fast implementation of pattern mining algorithms with time stamp uncertainties and temporal constraints. *Journal of Big Data* 20196:37.