Ethics

It is necessary to distinguish between two ethical concerns relating to this project. Firstly, the project must in itself be conducted in such a way that it lives up to generally accepted ethical standards for scientific research. Secondly, however, it is important to consider the potential consequences of successful implementation of a model of the type that the project seeks to construct. While the two are naturally intertwined, they are nevertheless distinct issues, and will consequently be analysed as such below.

# Research ethics

In addition to the general significance of sound research ethics, the subject matter of the present project makes it even more important to consider ethical questions. It would not be ethically defensible to work with statistics on stops and searches if it was possible to identify individuals from the information gathered. While it is common practice for states to keep criminal records, and for employers to require job applicants to submit their criminal record, it is generally accepted that the sharing of criminal records always happens with the consent of the relevant citizen. The project does not contain any information that could be used to directly identify the individual subjected to the stop and search.

Without containing information directly identifying the individual being searched, the data could, however, contain information that with some significant precision could be used to infer the identity of the individual. Upon thorough consideration, it has been concluded that this is not the case. The data provider, i.e. the UK police, has anonymised the location data before making it publicly available. According to the API documentation, this means that all coordinates given fulfil the following criteria:

1. Appears over the centre point of a street, above a public place such as a Park or Airport, or above a commercial premise like a Shopping Centre or Nightclub.

2. Has a catchment area which contains at least eight postal addresses or no postal addresses at all.[[1]](#footnote-1)

If no such anonymous coordinate is found within 20 kilometres of the actual site of the stop and search, no location data will be given. This is found to be sufficient that the identity of the suspect cannot with any reasonable accuracy be deduced from the location data.

Finally, it should be noted that the UK police itself has made a conscious effort to anonymise the published data and more generally assure that it lives up to the relevant privacy standards. The API documentation explains that both the Information Commissioner’s Office and Data Protection specialists at the Home Office were consulted heavily in preparation for the launch of the API.[[2]](#footnote-2) While this does not alleviate the present study of any ethical responsibility, it does add to the confidence with which it can be concluded that the project lives up to the highest ethical standards.

# The ethics of crime prediction software

This project can be challenged on ethical grounds not only for its methods but also for the result that it is trying to achieve. While it has been established above that the means employed are ethically sound, it remains to be seen whether the ends are defensible on ethical grounds. It could rightly be objected that the present project fails to produce a workable piece of crime prediction software. This is undisputedly true, but the project does contribute to a growing literature on the field and does thus help to further the stated ends in a more general sense. On the basis hereof, it is still necessary to consider the ethicality of this type of crime prediction software.

It should firstly be noted that this ethical discussion draws heavily on a number of well-established academic debates. Most prominently, the literature on racial profiling in applied ethics must necessarily appear in the present discussion. Furthermore, the discussion must draw on various legal debates.

The aim of this project is to develop machine learning based software that allows the police to efficiently employ *group-based patrol[[3]](#footnote-3)* (Machine Learning-based Group-based Patrol, MLGP)*.* This terminology follows the framework set out in Applbaum (2014)[[4]](#footnote-4). Importantly, however, the literature in applied ethics focuses on racial profiling as employed by human law enforcement officers. As such, an assessment of the ethicality of MLGP must answer two questions: (1) To what extent is racial profiling *simpliciter* ethically defensible, and (2) Are there any relevant features that makes MLGP ethically different from human racial profiling (HRP)?

1. <https://data.police.uk/about/#location-anonymisation> [↑](#footnote-ref-1)
2. <https://data.police.uk/about/#anonymisation> [↑](#footnote-ref-2)
3. One of three pure cases of racial profiling. The other two are “group-based enforcement” and “group-based identification”. [↑](#footnote-ref-3)
4. Arthur I. Applbaum (2014) “Bayesian Inference and Contractualist Justification on Interstate 95”. In Cohen and Wellman, eds. *Contemporary debates in applied ethics*, 2nd edn., 232-243 (Wiley-Blackwell). [↑](#footnote-ref-4)