Ministry of Justice — Written evidence (NTL0053)

Background and context

1. The Ministry of Justice (MoJ) has a largely centralised analytical function that provides wide ranging analytical support to all of the Ministry's activities. The analytical function has a strong culture of governance and assurance that is structured around the principles found within the Treasury's AQuA book as the key guidance for providing quality analysis in government. This guidance not only addresses the technical aspects of developing analysis, but also focuses on the wider governance and processes needed to ensure that quality decisions are made with the right analysis. As such it provides the basis for much of our response.

Analytical governance and production of quality analysis

2. All analytical projects, including those involving algorithms and advanced data analytics, must follow the Treasury's [AQuA book](#) guidance. This provides proportionate assurance that analysis is robust, fit for purpose and appropriately deployed. Within the MoJ analytical community we have a specific AQA log which ensures analysis is assured by an appropriate senior analyst. Our data scientists develop projects with reference to our internal [coding standards](#) as well as cross-government standards such as the [Government Digital Service (GDS) data science cookie cutter](#), and most projects follow a [gitflow](#) process using GitHub – a workflow tool for digital projects – to support effective collaboration and ensure regular peer review of code development. Projects are generally undertaken on the [MoJ Analytical Platform](#), which provides a secure cloud environment for analysts to access and share data, and modern tools to analyse data and develop algorithms and advanced data analytics. Many of our projects are available on our public GitHub site, including our open source machine learning [data linking algorithm](#).

3. We have processes in place to govern the extraction and movement of data from operational systems into the analytical platform, which include: securing appropriate permissions from Information Asset Owners on data protection requirements; ensuring cyber security protections are in place for data movement and storage; and governing the use of that data, particularly for individual-level data access and algorithm development.

4. We currently reference the GDS [published guidance](#) on data ethics to support use of individual-level data to develop of algorithms. This helps us consider the fairness, accountability, sustainability and transparency of the approach. We recognise the need to continue to upskill developers and data users in considerations around ethics, as well as provide practical tools to guide them through the process, and are working to develop these tools.

Governance and oversight of operational usage

<u>Guidelines and approval processes for employing and deploying data analytics to meet business needs</u>

5. Our approach to identifying business needs and the appropriateness of analytical approaches is closely aligned to the AQuA book and the responsibilities of the Analytical and Business Senior Responsible Officers it outlines. There are not formal checkpoints in every project. Instead, we take a proportionate approach in each case – the case study below, on the use of risk predictors for sexual reoffending, outlines the structured approach we take for significant changes that will have a big impact for service users. The identification of business need and the appropriateness of the data, analysis and outputs to meet that need are specifically captured and signed off (at Deputy Director level) in the AQA log for each project.

<u>Guidelines and processes for the procurement of data analytics and algorithms</u>

6. MoJ routinely contracts with third parties such as consultants, contractors, SMEs (small and medium enterprises) as well as those from the third sector. Procurement specialists within the commercial team ensure that each procurement gives due consideration to the nature of the service being provided and in particular the collection, use, storage and retention of data, engaging with the relevant policy and operational teams to ensure that we are compliant with the existing General Data Protection Regulations (GDPR). Data protection is considered at the outset of any new programme of work and Data Protection Impact Assessments (DPIA) are undertaken for any new project involving the processing of personal data (particularly where new technologies are being used and may result in high risk processing) and these are overseen and approved by both the Data Privacy team and the Cyber Security team.

7. MoJ takes a 'Data Protection by Design' approach when developing new initiatives/new technologies which involve the processing of personal data, meaning that data protection considerations are made central to and fundamentally inform the design phase. For example, our Prison Video Call service makes use of AI in its built-in technology to detect an unauthorised person joining a video-call after it has commenced. We are currently engaging the market to identify innovative solutions to reoffending as part of our Prison Leaver Innovation Challenge – this may well require us to consider the appropriate use of AI and machine learning in some of the proposed solutions.

<u>Oversight of the way in which tools are deployed, including training, educational requirements, sign-off procedures and risk mitigation</u>

8. Ensuring that data and insight is properly used and understood is again a key element of the AQuA book and MoJ's internal approach to AQA, and is specifically captured and signed off in the AQA log for each project by the relevant analytical Deputy Director.

9. Where our analytical teams develop tools which are used to directly support decisions about individuals in our system, we strive to ensure that the tool is designed to be intuitive. Moreover, clear guidance is developed on when the data should and should not be used, and support (and sometimes training) is made available to staff using the tool. This guidance and related training are reviewed and revised based on experience, informed by a range of quality assurance and control activities to support continual improvement. A specific example is given, at the end of this evidence, in the context of the development of risk predictors.

Combating bias reflected in analytical tools

10. The MoJ Analytical team follows the guidance on data ethics including reference to the GDS published guidance. This helps us consider the fairness, accountability, sustainability and transparency of the algorithm. Through this process we assess algorithms built from record level data against protected characteristics to help uncover bias within the data and the algorithms. If biases are found we revisit the algorithm and data to understand why they exist, whether they are removable and, if not, whether the tool remains suitable for use in the intended setting. The case study below explores this question and how we review algorithms that have been deployed.

11. In addition, we also follow all guidance for data privacy. DPIAs are completed and signed off when using data about individuals. The DPIA and associated processes mean that; our analysts only access data when there is a use case for the tool, only the data that is relevant for the project is accessible, and access is removed when no longer needed under strict controls.

Transparency processes: keeping track of data analytics used across the organisation

12. The Ministry of Justice carries out a Business Critical Model audit every year, with its remit covering analytical algorithms – only a small subset of these involved data and decisions about individuals. The review covers our five executive agencies (HMCTS, HMPPS, the Office of the Public Guardian, the Legal Aid Agency and the Criminal Injuries Compensation Authority), but not our other arm's length public bodies. The review requires all analysts to make a return on any model in use, in which analysts must detail the quality assurance process that has been carried out against best practice criteria derived from the AQuA book.

13. All data tools and algorithms built by suppliers / external bodies will be required to meet standards as part of the procurement phases (see procurement section). Any procurement of data analytics would be expected to meet the standards set out for internal tools and follow the AQuA Book principles. This would include being able to inspect the algorithm and the training, validation and test results.

14. The central MoJ analytical community is responsible for developing and improving MoJ AQA guidance and ensuring standards are defined and met. This includes responsibility for extending and expanding guidance and standards to apply to increasingly complex algorithms and advanced analytical tools, and providing advice to relevant governance forums to inform their deployment and continuing use in operations. We are currently working with the Alan Turing Institute to develop ethical frameworks for the application of AI to accurately reflect the justice context, and developing practical tools and processes to support and govern those deliberations.

15. More broadly, the use of AI and Machine Learning has been considered in a number of departmental groups. The exploration of opportunities to use machine learning techniques is considered from a wide range of perspectives, and ethical considerations are fundamental. Issues are generally considered in the context of a specific proposal by the business and analytical SRO; and we are considering how best to provide more structure and external influence.

16. Since autumn last year we have had in operation a (shadow) Senior Data Governance Panel comprised of civil servants, judiciary, and civil society to consider novel, sensitive or contentious court and tribunal data access issues. This panel derives from a recommendation of Dr Natalie Byrom's report on court data published in 2019. We expect the panel to have an ongoing role in advising on these matters, for which AI use is a natural consideration and in providing advice to ministers and relevant governance forums.

Case study: OASys Sexual reoffending Predictor

17. A recent example of procurement is the OASys Sexual reoffending Predictor (OSP), which was approved by HMPPS's Sexual Offending Management Board.

18. HMPPS's Sexual Offending Management Board considered the impacts on all areas of the organisation prior to approval, including information on the extent to which deploying the new instrument would improve HMPPS's ability to identify those at greater risk of sexual recidivism, alongside the impact on individual offenders' sentences (i.e. when it is necessary to replace a risk score from an obsolete tool with a score from the new tool, which may indicate higher or lower risk) and the impact on business processes (e.g. the need to update IT systems, and brief assessors and stakeholders such as the Parole Board).

19. Decisions that have a substantial impact on an individual offender are not made on the basis of an actuarial risk score alone. For example, probation practitioners and other assessors use their professional judgement to rate the

individual offender's Risk of Serious Harm by following a multi-step process to synthesise OSP scores and multiple other sources of information.

20. Quality Assurance is essential. Before approval to implement OSP was granted, the research study that demonstrated its validity as a risk predictor was peer reviewed by both internal experts and an external panel composed of international experts on sexual offending risk assessment.

21. Once the tool was approved for use, staff were provided with guidance materials, which explain both the step-by-step process of scoring the tools (e.g. guidance on scoring each question within OSP) and purpose (e.g. what OSP should be used for). Staff also have access to ongoing support. Guidance is improved and reissued in response to operational feedback and lessons learnt. The guidance includes explanations of the strengths and limitations of the tools, and how they should be used proportionately to support and inform (not automate) decision-making. An active programme of quality assurance and control is promoted, including briefings for practitioners to ensure they understand how to use the tool to inform decision-making. Guidance is reviewed and revised based on feedback and learning through implementation.

22. It is important to recognise that operational decisions are informed by analytical tools rather than being automatic consequences of tool outputs. The operational use and application of OSP, for example, is set out in an overarching Policy Framework which supports consistency of use.

23. Equality Analyses are produced as a mandatory part of the development of Policy Frameworks to support new algorithms. These will consider differences in the distributions of scores between certain groups of people, and whether any observed differences are consistent with equalities principles as well as the purpose of the algorithm.

24. For example, reoffending risk typically varies systematically by age and gender, after controlling for other relevant risk factors, and so these factors are usually included in the algorithms in order to enhance prediction and therefore contribute to the protection of the public. Other group characteristics are not included in the algorithm, but differences are sometimes observed. In the Equality Analysis for OSP, some differences in risk scores by religion were observed – these appeared to be due to the lower average age of those belonging to some faith groups.

25 October 2021