Serious Fraud Office — Written evidence (NTL0034)

1. Do you know of technologies being used in the application of the law? Where? By whom? For what purpose?

1.1. The use of machine learning and natural language processing in 'e-discovery' (the earliest stage of reviewing documents in a case) is of critical importance for organisations' ability to digest and understand the large volumes of materials that are commonplace in many cases. The Serious Fraud Office (SFO) does not yet use machine learning or natural language processing in its cases (except on an ad hoc basis) but intends invest in such technologies in the next financial year.

1.2. The SFO has used machine learning to identify which materials in three cases are subject to legal professional privilege (LPP). A programme was fed several thousand documents, each marked as either 'LPP = yes' (i.e. 'this document is subject to legal professional privilege) or 'LPP = no' to train the programme to recognise LPP. In the Rolls Royce Deferred Prosecution Agreement, the SFO benefitted hugely from the technology, saving up to two years and significant costs. In all cases, it is necessary for a human to check the programme's output, but this is significantly less time-consuming than had the technology not been used.

1.3. The use of machine learning and natural language processing in e-discovery appears to be common in private sector law firms. As set out in paragraph 42(e) of the Airbus Deferred Prosecution Agreement Statement of Facts, Airbus's lawyers used such technologies to identify previously unknown wrongdoing, which was subsequently referred to the SFO.

1.4. The SFO also uses technology in a wide variety of circumstances, but the technology does not involve a high degree of artificial intelligence. For example, staff at the SFO have used Visual Basic for Applications (VBA) to allow the automated creation of case chronologies and disclosure schedules for unused materials in Microsoft Excel. In these circumstances, as above, a human is required to check the product. As a second example, investigators use machine translation to translate large volumes of documents at an early stage of an investigation. While machine translated documents cannot be used in court, this aids investigators in understanding whether a document is broadly relevant, which saves time for investigators, as well as reducing costs spent on professional translators.

2. What should new technologies used for the application of the law aim to achieve? In what instances is it acceptable for them to be used? Do these technologies work for their intended purposes, and are these purposes sufficiently understood?

2.1. Technology should be used to replace tasks in which it is more effective than a human, but should not ultimately make decisions. The role of technology is to make the delivery of justice more efficient, benefitting victims and freeing up more resource for a larger number of cases to be taken on by both public sector agencies and private sector firms.

2.2. It is critical that technologies are applied by those who understand them. Organisations wishing to make use of new technologies must hire staff with the required skillsets, or train existing staff. (However, it must be noted that people with relevant skillsets can command extremely high salaries in the private sector, which will limit the public sector's ability to hire the required number of staff, as the application of new technologies becomes more commonplace).

2.3. Courts will be responsible for ensuring that technology has been applied appropriately, and it is necessary that both prosecution and defence counsel can explain its application in an effective way.

3. **Do new technologies used in the application of the law produce reliable outputs, and consistently so? How far do those who interact with these technologies (such as police officers, members of the judiciary, lawyers, and members of the public) understand how they work and how they should be used?**

3.1. Machine learning and natural language processing programmes can only be successful if they are 'trained' and tested effectively. See the example in paragraph 1.2. In this case, the programme was trained extensively and the outputs tested on a number of occasions. It is worth noting, however, that humans are fallible. Our current (human-based) systems are not 100% accurate—especially in SFO cases, where a high number of people are working on a data set with millions of items—and we believe that greater application of technology will reduce the margin of error

3.2. As set out above, it is critical that these technologies are used by those who understand them, and can explain how and why that technology was applied in a particular case. However, as set out in our proposed third principle (in response to question 10), this approach should be proportionate and should not hamper the application of technology. As suppliers will wish to protect their intellectual property, there are limits to (a) what those who apply a programme will know about the underlying technology and (b) how much can be understood about the underlying technology, given how complex these technologies are to develop. Instead of explaining the underlying technology, rather it is critical that those who apply the technology understand (and can explain) the impact that they will have by applying the technology and the reasons for taking, or not taking, various steps in applying that technology.

3.3. It is critical that the judiciary understands the potential applications of various technologies in a case, so that agencies or firms that make use of new technologies can be used to their full potential, and not hampered in their range of applications.

4. **How do technologies impact upon the rule of law and trust in the rule of law and its application? Your answer could refer, for example, to issues of equality. How could any negative impacts be mitigated?**

4.1. The interests of justice in criminal proceedings will, of course, require checks and balances to ensure that the defendant receives a fair trial. Technology, when used effectively, can improve the rule of law, by making its application more efficient, and speeding up criminal justice outcomes. The

interests of victims and wider society will also be better protected by ensuring that issues with disclosure can be identified—through the effective use of technology—and resolved more quickly without the collapse of cases.

4.2. It is critical that artificial intelligence is not an ultimate decision-maker in inappropriate circumstances. Maintaining the role of human decision-makers in the application of new technologies can allow for the identification and mitigation of any negative impacts.

4.3. It will also be critical to ensure the narrative around the application of any new technology in the public sector is truthful, in order to maintain trust in those institutions.

5. **With regards to the use of these technologies, what costs could arise? Do the benefits outweigh these costs? Are safeguards needed to ensure that technologies cannot be used to serve purposes incompatible with a democratic society?**

5.1. New technologies will bring with them high (monetary) costs to users. In many circumstances, the financial saving brought on by the efficiencies created by that technology will outweigh the cost.

5.2. Safeguards will be critical to ensure the proportionate and lawful application of technology. Effective safeguards will also improve the public's trust in the application of technology, as it can demonstrate that any technology applied is within what is proportionate and lawful. In particular—considering the SFO's specific interest in e-discovery—safeguards should consider the role of humans as decision-makers; sample testing of initial results; updating the programme with results from the sampling then subsequently re-testing; the use of multi-disciplinary teams with the right skillsets; recording processes and protocol; and early communication with the other side to the litigation to agree protocol. These safeguards have emerged in civil case law in the UK and overseas.

5.3. We note three key examples of the need for safeguards:

5.3.1. In August 2020, the Court of Appeal ruled that the use of an automatic facial recognition tool by South Wales Police was proportionate but ultimately unlawful. The Court determined that there was a lack of guidance on its application and an incomplete data protection impact assessment, and concluded that "as AFR [automatic facial recognition] is a novel and controversial technology, all police forces that intend to use it in the future would wish to satisfy themselves that everything reasonable which could have been done had been done in order to make sure that the software used does not have a racial or gender bias".[1]

5.3.2. The Home Office announced in August 2020 that it would no longer use an algorithm which separated visa applicants in to three channels, ahead of a Judicial Review from the Joint Council for the Welfare of

[1] Judgement here: https://www.judiciary.uk/wp-content/uploads/2020/08/R-Bridges-v-CC-South-Wales-ors-Judgment.pdf

Immigrants, which said that the algorithm prioritised white applicants from rich countries.

5.3.3. In October 2018, it was reported that Amazon would cease to use artificial intelligence to review job applications, after it became clear that the AI was often rated CVs from men over those from women. It is critical that artificial intelligence is not taught in a way that enforces bias—it was reported that Amazon's software was 'trained' using CVs from the previous ten years, of which many successful ones, especially in the early days, belonged to men.

6. **What mechanisms should be introduced to monitor the deployment of new technologies? How can their performance be evaluated prior to deployment and while in use? Who should be accountable for the use of new technologies, and what accountability arrangements should be in place? What governance and oversight mechanisms should be in place?**

6.1. No comment.

7. **How far does the existing legal framework around new technologies used in the application of the law support their ethical and effective use, now and in the future? What (if any) new legislation is required? How appropriate are current legal frameworks?**

7.1. The SFO has not considered this in detail, however it will be necessary for the legislation and/or guidance to be updated. For example, the Attorney General's Guidelines on Disclosure[2] have been recently updated (in 2020) to allow for the application of technology in disclosure. It may be necessary to further update these to accommodate for more advanced technology. However, it may be that these guidelines require updates as the application of technology increases, and there is greater need for clarity on what technologies may or may not be used. The ability to apply new technologies in relation to disclosure is critical for the SFO.

7.2. The existing data protection framework does allow for new technology to be deployed for law enforcement purposes, when implemented with safeguards in place to mitigate privacy risks. The Information Commissioner's Office has issued a toolkit designed to support law enforcement agencies with data analytics[3] and have issued guidance on the use of artificial technology in a GDPR context.[4]

---

[2] "Where there is a large volume of material, it is perfectly proper for the investigator and/or disclosure officer to search by sample, key words, or other appropriate search tools or analytical techniques to locate relevant passages, phrases and identifiers…Technology that takes the form of search tools which use unambiguous calculations to perform problem-solving operations, such as algorithms or predictive coding, are an acceptable method of examining and reviewing material for disclosure purposes"
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/946082/Attorney_General_s_Guidelines_2020_FINAL_Effective_31Dec2020.pdf

[3] https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2020/12/ico-launches-tool-to-help-police-forces-using-data-analytics/

[4] https://ico.org.uk/for-organisations/guide-to-data-protection/key-data-protection-themes/guidance-on-ai-and-data-protection/

7.3. Delays to updating legislation and/or guidance may hamper the application of new technology.

8. **How can transparency be ensured when it comes to the use of these technologies, including regarding how they are purchased, how their results are interpreted, and in what ways they are used?**

8.1. There needs to be a balance between transparency in how technologies are purchased, to protect the commercial interests of the purchaser and supplier, while also protecting the rights of those are subject to the application of the technology from inappropriate activity.

8.2. The application of advanced technologies should be explained in court, including how their results were interpreted. In the same way that an expert witness may appear in court to explain how, for example, the angle that a bullet hits a wall means the gun was fired from a certain location, it will be necessary to explain how technologies were applied in a way that led to the specified outcome. This must take in to account the need to protect the supplier's intellectual property. This, however, must be proportionate, and it may not be necessary to explain the application of simple or common technologies.

9. **Are there relevant examples of good practices and lessons learnt from other fields or jurisdictions which should be considered?**

9.1. There is civil case law approving the use of predictive coding to identify relevant documents; the case law highlights the benefits of doing so. These cases contain guidance on how the technology should be used, and refer to case law in other jurisdictions (e.g. US, Ireland, Australia) where the use of such technology has also been approved. This case law also provides guidance on safeguards, as referred to in our response to question 5.

10. **This Committee aims to establish some guiding principles for the use of technologies in the application of the law. What principles would you recommend?**

10.1. The SFO proposes the following guiding principles:

10.1.1. Technology should be embraced, and used in a way that improves the delivery of effective criminal justice outcomes;
10.1.2. Technology should not be the ultimate decision-maker in circumstances where a wrong decision could result in an unjust criminal justice outcome;
10.1.3. The application of advanced technologies—including how their results are interpreted—must be explained in court, to ensure an appropriate level of transparency. This approach must, however, remain proportionate and more work will be required to define how any criteria will be applied. This principle should not be used to hamper the application of technology.
10.1.4. There must be sufficient safeguards in place.

10 September 2021