ROXANNE PROJECT ETHICS AND LEGAL TRAINING - FOR PUBLIC DISSEMINATION

Grant Agreement:	833635
Project Acronym:	ROXANNE
Project Title:	Real time network, text, and speaker analytics for combating organised crime
Call ID: **Call name:**	H2020-SU-SEC-2018-2019-2020, Technologies to enhance the fight against crime and terrorism
Revision:	V1.0
Date:	17th October 2022
Type of action:	RIA

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019-2022 – All rights reserved.

Disclaimer

The information, documentation and figures available in this deliverable are written by the "ROXANNE - " Real time network, text, and speaker analytics for combating organised crime" project's consortium under EC grant agreement 8833635 and do not necessarily reflect the views of the European Commission.

The European Commission is not liable for any use that may be made of the information contained herein.

Copyright notice

Executive summary

This document is an edited version of the ethical and legal training materials for the ROXANNE project that has been made available for public dissemination. It is intended to be provided to end-users of the ROXANNE tools, who are expected to be investigators working for law enforcement. Section 1 deals with artificial intelligence (AI) ethics issues as they related to technologies used by law enforcement agencies (LEAs). Section 2 deals with cross-European legal issues relevant to use of AI tools by LEAs. The intention of making this document available for public dissemination is, firstly, so that the LEAs can learn about issues relevant to their use of AI in investigations. Second, it is important that the public and other researchers can learn from our experience and understand issues relevant to technologies like those researched in ROXANNE. The document has been edited for public dissemination by removing references to specific ethical and legal questions for the ROXANNE technologies. As the project is still ongoing, some of these questions might be answered in the remaining months. A complete ethics and legal training document will be included in our final deliverables and will be made publicly available in full.

Table of contents

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

1. Ethical and Legal Training

Below is training material on ethical and legal issues applicable to the ROXANNE platform. The below training was mostly developed by Trilateral Research whose staff work across both ROXANNE and INSPECTr projects, and developed common parts of the below training across both projects. The training materials benefitted from inputs from ROXANNE partners, especially KEMEA.

1. Ethical issues in the use of AI technologies by LEAs

Policing is a high-risk domain. To make the wrong decision in an investigation means that perpetrators will be free, victims will be without justice, and the public will be at risk, for longer than they need to, or an innocent person could be accused of a crime, arrested, and subject to intrusive investigation and potential prosecution. AI technologies are powerful tools to help in investigations. But, despite the hype, they are not perfect, and so use of these powerful tools in an incorrect, reckless, or irresponsible way can create negative impacts for investigations and the people impacts by them.

So, it is important that LEA technologies are properly understood, it is clear how they should be used, and their impacts are appreciated. As police are subject to significant scrutiny, it is important that LEA officers can explain all of this to their superiors, to a court, and potentially to a victim or their family as well.

Using LEA technologies in compliance with ethical, societal, legal, and particularly human rights, standards, is important because this helps to build trust from the public. Where police enforce the law with the consent of the public, then it is important that LEAs act within what is societally acceptable. Otherwise, LEAs risk investigations being in violation of applicable standards. This also creates a risk that an investigation will collapse and offenders will go free.

We are trying to avoid these problems through responsible use of LEA technologies, and we can do this through understand the ethical and legal frameworks and how they apply to LEA technologies.

High-Level AI Ethics Requirements

SHERPA requirements and sub-requirements
1 Human agency, liberty and dignity: Positive liberty, negative liberty and human dignity
2 Technical robustness and safety: Including resilience to attack and security, fall back plan and general safety, accuracy, reliability and reproducibility
3 Privacy and data governance: Including respect for privacy, quality and integrity of data, access to data, data rights and ownership
4 Transparency: Including traceability, explainability and communication
5 Diversity, non-discrimination and fairness: Avoidance and reduction of bias, ensuring fairness and avoidance of discrimination, and inclusive stakeholder engagement

6 Individual, societal and environmental wellbeing:
Sustainable and environmentally friendly SIS, individual well-being, social relationships and social cohesion, and democracy and strong institutions

7 Accountability:
auditability, minimisation and reporting of negative impact, internal and external governance frameworks, redress, and human oversight

The seven high-level requirements for AI ethics come from the SHERPA project and the EU's High-Level Expert Group on Artificial Intelligence. They are used in the development phase to ensure ethical design of AI systems, but they are also useful in the use phase to ensure that the AI systems are used appropriately.

Human Agency, Liberty, and Dignity

It is important that end users engage with LEA technologies in a meaningful way so that they really understand what they are doing, and what the impacts are. It is also important to be appropriately sceptical of AI systems. They are not perfect, and their outputs need to be interrogated to ensure that they are robust. Consider if a defence lawyer could ask difficult questions about the way you use the tools.

Automation bias is where people trust algorithms, or computer systems, more than their own judgement and so just 'follow the algorithm'. This should not be done in investigations because algorithms cannot be used to understand the complexity of the world and so can only provide simplified outputs that need to be understood in context. Falling victim to automation bias can mean that people do not engage in difficult questions that they should do. This can lead to a loss, or atrophy, or moral decision-making ability, and so should be avoided. Dealing with difficult decisions is a key part of being a good investigator, and so this ability needs to be retained.

Following on from automation bias, where investigators see their work as purely technical, perhaps even seeing themselves merely as a technician, or machine operator, rather than a fully engaged investigator with insights about cases beyond those that come from the outputs of the AI system. It is important that investigators remain fully engaged in their work so that they do not lose their intuition. Intuition, or a 'police officer's nose' is a key skill for being a good investigator. An experience investigator with a feeling that someone isn't right, or needs further examination should not be prevented from pursuing further investigative leads simply because the results of an AI system suggest that the part of an investigation in question is completed.

Thinking now about the people whose data is being processed, in this case it is the victims, suspects, and witnesses relevant to the investigation. Where technology users analyse data about people frequently and for extended periods of time, there is a risk that they begin to think only of data, and see the people behind the data as mere data points. This is wrong to do as it means that the victims, suspects, and witnesses are not being treated as the people they are. This is a form of dehumanisation, and it should be avoided as crime is a human-centred issue and the people involved should be treated in this way.

As a consequence of all this, it is important that AI systems are not used to make decisions on behalf of investigators. Rather, investigators should take the outputs of AI system they use and incorporate them into their thinking about the investigation as a whole. The AI systems are one set of tools in the investigatory tool box and are not a replacement for the experience of the investigator themselves, nor the experts that can be called to assist them.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

Technical Robustness and Safety

Sometimes, we see people interacting with AI systems who expect something similar to the systems shown in a Hollywood movie. This can lead to an expectation that the AI system can perform a lot of tasks beyond its intended use, might have hidden knowledge from other data-sources that aren't visible to end-users, or that they might even be sentient and self-aware. The truth is that all AI systems are technologies that process input data and provide an output, and this activity can sometimes enable particular and narrowly defined tasks to be delegated to the system that would otherwise take a long time for people to do. For example, a named entity recognition algorithm might highlight names, locations, or other key information in text very quickly whereas a person would need to spend a lot of time reviewing documents. So, investigators should treat AI systems like any other investigative technology, and be aware of how it works so they can use it properly and report results to a court.

So that chain of custody is maintained, it is crucial that AI systems for LEAs are used in a safe and secure way. You should only use tools in an environment that has been approved by your organisation's information security team, and should be aware of the risk of harms that could come from malicious files in evidence, or harms that could arise from connecting technologies to outside systems.

It is also important to remember that these tools are new and, whilst they have been subjected to a wide range of testing, there might be situations where errors or bugs arise. This is especially the case where tools are exposed to the difficulties and complexities of real-world investigations. There could be situations that the tools are not able to deal with effectively. Investigators should be prepared to double-check results where they might look odd, or are unexpected.

Technical colleagues do their best to develop technologies that work well and do the job they are intended to be used for. However, sometimes we see end-users becoming frustrated with tools because they do not give concrete answers they expect. For example, an image analysis tool might give a confidence rating for a possible match that then needs to be assessed by the investigator, rather than a definitive answer. So, it is important that investigators pay attention to the other training modules to ensure that they fully understand what the technologies are intended to be used for, what they can do, what they cannot do, and what the limits of it are. If an investigator is unsure about this, they should revisit the relevant training module, or ask questions of technical partners.

Some LEA technologies have very obvious applications. For example, a text analysis algorithm would not provide good results if image data were inputted into it. However, other applications might be very similar. For example, two algorithms might analyse a piece of text, and one might highlight important information in the text (names entity recognition), and another might highlight key information relevant to the purpose of the text (sentiment analysis). So that results are accurate, reliable, and reproducible, it is important the end-users are clear on which tools should be used for which tasks, and do not confuse them.

Forecasting algorithms work by extrapolating from recorded data about the past and applying them to current situations. This means that they are not intended to work on the individual level. You should not expect to use a forecasting algorithm to predict the actions of an individual. Because the algorithms work on trends, it is also not possible for them to deal with unexpected events that might alter behaviour suddenly (e.g., a security incident that means people cannot access their work or homes in their usual way due to an apparent danger, or afterwards people might gather to memorialise victims). Predictive algorithms should not be used to deal with policing of unique events, as the algorithms are unlikely to be able to deal with the individualised nature of the event.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

Privacy and Data Governance

Privacy is a fundamental right, so it is important that it is protected. As an end-user, information is collected about investigators using tools via access controls and logging. So privacy concerns apply to both the people being investigated and the people using the technologies. Society recognises that infringing on privacy is an acceptable thing for LEAs to do during a legitimate investigation, as secrets about crime need to be uncovered so that they can be prevented and punished. However, investigations that go beyond the criminal acts of a suspect, and into their private life, or the private lives of others involved in the case would not seem to be appropriate and so would generally seem illegitimate. LEA technologies should not be used to pry into private information that is not relevant to an ongoing investigation.

To maintain the autonomy and dignity of persons whose private lives are examined during an investigation, as much as possible, investigations should focus on whether the acts of the suspect under investigation are unlawful. The investigation should not focus on why they are as a person. For example, an investigation might use location data to determine that a suspect was present at the site and time of a violent assault during the evening. But, the same data file might also demonstrate that the suspect also visited their workplace, their home, a political club, their lover, or anywhere else they might have been. Such data should not be probed further unless there is a pressing investigative reason to do so. Such information is likely to be immaterial to an investigation, and so infringing on that person's privacy is likely to be illegitimate. As is explained in the next section, applicable legal frameworks on appropriate limits to investigations should be followed at all times.

Transparency

Transparency is a key issue with LEA technologies. As all parts of the criminal justice system can be examined in court, it is important that the tools used can be explained to the court to demonstrate that any investigation has been conducted fairly and lawfully; for example, forensic experts might be called upon to explain their techniques and demonstrate that they meet high scientific standards and this can also often happen with LEA technologies. So, it is important that investigators using the technologies can properly understand what the tool does (i.e., the purpose of the technology), how it works (i.e., how the tool processes data), and how it produces an output (i.e., what processes are used to analyse the input data and transform it into a result). Being able to understand and communicate how a technology works is referred to as explainability.

So that end-users properly understand the tools, they should thoroughly engage with the training programme to ensure they understand these issues. It is also important to recognise that all uses of tools should be tracked and logged to present chain of custody. This means that how the tools are used by end-users can be presented in reports, or explained before a court if needed. Being able to demonstrate how a tool was used is referred to as traceability.

As mentioned previously, being able to communication about how technologies work when they have been used in an investigation is a key part of explainability. In the case of LEA technologies, being able to communicate how a tool works to a court and the defence is of paramount importance. You might also need to explain the tool to senior investigators, and, in some situations, victims or their families as well.

Diversity, Non-Discrimination, and Fairness

Unfortunately, no dataset is a perfect representation of the population from which the relevant sample is drawn. Sometimes this is due to simple inaccuracies in data collection, and sometimes this is due to the impact of discriminatory policies that have caused unequal treatment of people in the past and this has

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635.

been recorded and reflected in historical data. The datasets used in the INSPECTr project have been reviewed so that they are as representative as possible. The INSPECTr partners have done further work to minimise the impact of biases on how the tools work and on the results they produce. However, some biases could remain and some of these biases might not be recognised until they are used with datasets where particular demographics expose such a bias. Therefore, end-users should be conscious that biases can arise in the use of tools, especially those trained using machine learning techniques.

Biases might affect data, and the people who are represented in the data, in a variety of ways. Biases and discrimination might arise based on race, sex, gender, sexuality, social class, age, different abilities, relationship status, pregnancy/maternity status, religion or beliefs, or could arise in other ways. Sometimes, these biases are compounded where multiple protected characteristics are present (intersectionality).

Where biases in tool outputs are detected, then these should be taken into account by investigators when assessing all the information available to them about how investigations should progress. Conclusions about the results, and any investigation, should not be drawn directly from biased results. In some cases, biased results can still be useful but they must only be used in a way that mitigates the effects of those biases. Not doing so would result in an investigation that could not only have discriminatory effects, but also take the wrong investigative leads. This is not appropriate for LEAs to do, and so the potential for bias and discrimination as a result of tool outputs should be considered very carefully during investigations.

Individual, Societal, and Environmental Wellbeing

Fundamentally, law enforcement is supposed to be about ensuring that rules decided as a result of societal debate are followed for the betterment of all. Therefore, LEAs should make sure that they stay not only within the law, but where law is unclear or missing that they act in ways that society would accept. New technologies might not be subject to the appropriate level of regulation yet, and so it is important that they are used in ways that do not contravene the expectations of the public.

So that the LEA use of technology can be tested against societal expectations, the public should be made aware of the technologies being used as much as is practicable taking into account the sensitive nature of criminal investigations. This would allow the society that is being policed to determine what is acceptable to them, and if they wish to be policed in such a way.

Where forecasting algorithms are used to highlight likely criminal occurrences in future, this is based on crime trends. Where such a technology is used properly, it can allow police to target their resources at locations expected to exhibit the most criminality. However, information about those trends is not neutral. They can reflect the policy choices of the police forces that collect the data, and this can result in entrenching policy choices in a negative feedback loop. For example, if an area with a high population of people from ethnic minorities has had many historical crimes recorded due to over-policing as a result of racialised policing policy, then a forecasting algorithm will suggest that large amounts of crime happen in this area and one interpretation of this could be that police should increase their presence there, thus leading to more over-policing of an ethnic minority. Rather than preventing or punishing crime on a large scale, this can simply result in racial harassment. If over-policing worked, then the crime trends would point toward less crime taking place. However, more petty crime is often recorded, leading to a feedback loop that entrenches posting significant amounts of law enforcement resources at minor criminality and can cause increased tensions with the local community. As such, where forecasting algorithms are used wrongly, they can have a negative impact at a community level. End-users should be acutely aware of this if they choose to employ such technologies, and be critical of results that are outputted when they are developing their policing strategies using the results of such algorithms.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

Further, AI systems can process a lot of data and use a lot of energy to do so. Energy should not be wasted when such technologies are used. You should only use the tools you need for an investigation, and only use them where they are needed.

Accountability

Logging of how AI tools are used is important for accountability purposes so that LEA uses of tools can be checked by courts and professional standards units. These are important mechanisms to provide oversight so that the public can maintain trust in the manner that LEA technologies are used in investigations.

Whilst there are many oversight mechanisms for LEA officers to ensure lawful and ethical behaviour in line with policing ethics, the emergence of AI systems in policing creates a potential gap where standards designed to govern human behaviour are not specifically applicable to governing the use of machines. So there could be challenges about how AI systems should be used in line with existing standards, or whether they need to be adapted. Research projects that contain ethics work streams do go some way to dealing with ethical issues that could arise during LEA operation, but cannot anticipate all situations that could arise. Where possible, LEAs should consult AI ethics experts to ensure that they are using AI systems in an appropriate way. Ideally, this would involve setting up an expert advisory board with ethical expertise. Alternatively, consulting experts on AI ethics would also be beneficial.

It is important to also remember that LEAs are accountable to the public as well. An LEA that has lost the trust of the public is unlikely to be able to adequately enforce legal standards. Investigators should therefore ensure that they use tools in a way that would be expected of those holding public office.

AI Regulation

The current draft of the AI Regulation means that many AI systems designed for LEA use will be 'high-risk'. Some technologies, such as live facial recognition are banned, but can be used by law enforcement in some situations. This means that the use of such technologies are especially risky, and should not be used unless they are definitely needed. Just because an exemption exists does not mean that it should be relied upon readily or easily. Indeed, any use of an exemption where the technology in question poses such risks should be strongly justified.

Whilst the exact requirements that will arise from the AI Regulation are unclear until a finalised Regulation is agreed, it will likely mean that high-risk AI systems will need to undergo conformity assessments and will be subject to further scrutiny when they are in use by LEAs. Much of the work that Trilateral has done in the INSPECTr project as part of an Ethics and Privacy-by-Design approach has already contributed to dealing with the risks and generating information that could be included in the expected documentation.

Summary

To summarise this ethics training, there are clear ethical issues related to the use of AI tools in investigations by LEAs. These relate to a wide range of issues, from the agency of the tool user, to the impact on citizens being investigated.

Some harm to citizens could be inevitable. For example, you will need to invade the privacy of an offender in a lawful investigation to prove a case against them. However, invasions of privacy should not be more than is required for the investigation. So you should take care to protect the privacy of the people you are investigating as much as possible. Where other harms could be created for people you are investigating, you should also try to minimise harms as much as you reasonably can whilst still carrying out an effective investigation.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635.

We have discussed how following the algorithm is not appropriate for LEA officer during investigations, so you should make sure to critically engage with any AI system or tools you use.

- You should think about your role in using each tool: are you using it as an investigative aid, or are you being guided by the tool outputs?
- You should think about what each output from a tool means and whether those outputs need to be checked manually: Are you confident about the output, and is it what you expected? Which outputs will you need to check to maintain a robust investigation?
- You should think about how the results from each tool could impact on investigations: What opportunities or risks could the outputs create for your investigation? How do the outputs change your investigative decisions, and how will they affect people in your investigation?

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

2. Legal training

In this section of the Ethical and Legal training, we will explore the legal framework applicable to use of AI tools by LEAs in Europe and look at some of the key provisions for use of such tools.

Legal framework

This is the general legal framework that will regulate the use of AI systems in LEA investigations.

- European Convention on Human Rights

First, the European Convention on Human Rights from the Council of Europe provides the minimum standards about how public authorities, in this case LEAs, should treat people they interact with. Most of the rights present in the Convention are mirrored in the EU's Charter on Fundamental Rights that applies to EU countries applying Eu legislation.

- Convention 108+

Second is the modernised Convention 108 (known at Convention 108+) from the Council of Europe that provides an overarching approach to data protection and information privacy.

- Law Enforcement Directive

This is the EU's Law Enforcement Directive. This provides a common legal approach to the processing of personal data by law enforcement across the EU. This is sister legislation to the GDPR, which regulates the processing of personal data in general situations. AI systems that process data about people are processing the personal data of those people, so it is imperative that specific data protection law is respected when using AI tools for law enforcement.

- National law

Finally, national law regarding criminal procedures and data protection also applies to the use of AI systems so that their use is regulated by the standards of the population and society in which they are to be used.

Law Enforcement Directive

Personal data is any information related to an identifiable natural person. This is a very wide definition, and includes data where the data-subject (i.e., the person whose data the information relates to) is indirectly identifiable, i.e., where it has been pseudonymised. Claiming that a person is not identifiable is a really difficult threshold to reach in legal terms, and you should not assume data is anonymous unless you have discussed this with a data protection officer or legal adviser.

Data Protection legislation has several principles that underpin all the other rules that are applicable to the processing of personal data.

These are paraphrased from Article 4 of the Law Enforcement Directive and are the principles applicable to the processing of personal data for law enforcement purposes:

Personal data processed for LEA purposes, shall be:

(a) processed lawfully and fairly;

(b) collected for specified, explicit and legitimate purposes …;

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

(c) adequate, relevant and not excessive in relation to the purposes …;

(d) accurate and … kept up to date…;

(e) kept in a form which permits identification of data subjects for no longer than is necessary …;

(f) processed in a manner that ensures appropriate security of the personal data, … using appropriate technical or organisational measures.

Human Rights

All human rights are important to consider when interacting with the public. However, the use of AI systems by LEAs touch on rights to privacy, a fair trial, and the prohibition on discrimination most clearly. So, we will focus on these in this training module.

Right to Privacy

This is the text of the Right to Privacy:

> *'1. Everyone has the right to respect for his private and family life, his home and his correspondence.*
>
> *2. There shall be no interference by a public authority with the exercise of this right except such as is in accordance with the law and is necessary in a democratic society in the interests of national security, public safety or the economic wellbeing of the country, for the prevention of disorder or crime, for the protection of health or morals, or for the protection of the rights and freedoms of others.'*

Note that the European Union Charter on Fundamental Rights mirrors this right, but refers to 'communications' rather than 'correspondence'.

As you can see, infringements on privacy for the prevention of disorder or crime are an exemption to the general rule. This should only be done in certain circumstances, as is explained on the next slide.

Infringing on the right to privacy

Where an investigation needs to uncover private details, this should only be done where it is necessary to achieve a legitimate aim (i.e., for the effective investigation of a crime), proportionate to that aim (i.e., is focussed on the criminal acts and not on uncover all aspects of a person's life), and in accordance with national law (i.e., the privacy infringement is also allowed for by the legal standards in your country).

The right to privacy, like all human rights, applies to everyone. So, when an investigator is weighing up whether and how to cover private details, they should consider infringements on the suspect, victim, witnesses, bystanders, and others in the same way.

When investigators engage in privacy infringements in legitimate investigations they should ensure not to examine information beyond the purposes of an investigation. This would violate the data minimisation principle we discussed earlier, it would also likely be illegitimate. When using AI tools, there are high risks of invading a person's privacy unnecessarily and illegitimately because they can uncover or infer previously hidden connections, patterns of information, and private details. AI tools can allow this to happen quickly and with ease, as the thought processes about whether the extent of privacy invasions are appropriate that might have been included when investigators examined information 'by hand' are removed. Therefore, it is important to consider privacy risks in advance of using such tools.

Right to a Fair Trial

This is paraphrased from the right to a fair trial:

> *'1. … everyone is entitled to a fair and public hearing within a reasonable time by an independent and impartial tribunal established by law. Judgment shall be pronounced publicly ...*
>
> *2. Everyone charged with a criminal offence shall be presumed innocent until proved guilty according to law…*
>
> *3. Everyone charged with a criminal offence has the following minimum rights:*
>
> *(a) to be informed … of the <u>nature and cause of the accusation against him</u>;*
>
> *(b) to have <u>adequate time and facilities for the preparation of his defence</u>;*
>
> *(c) to <u>defend himself</u> …*
>
> *(d) to <u>examine or have examined witnesses against him</u> …*
>
> *(e) to have the free assistance of an interpreter …'*

Key parts of fair trial rights are being able to know and understand the nature of the crimes the defendant is supposed to have committed. It could be necessary to explain how an AI system works so that the defendant can properly understand how and why there are being accused of a crime.

Further, it is important that the defendant can access the data and potentially the tools used for analysis, so that they can prepare an adequate defence, and properly defend themselves.

Finally, a defendant could ask investigators or AI experts to provide expert testimony about the AI systems used, whether they were used correctly, and what the inputs, processes, and outputs mean.

Right to a Fair Trial aspects

The right to a fair trial is one of the few absolute rights, as there is not legitimate reason to have an unfair trail. There are some circumstances where an open trial could infringe the rights of others and so reporting around a trial might be temporarily restricted, but that should not affect the use of AI systems in investigations.

As discussed on the previous slide, investigators might need to be able to present their results, and reason for using a tool to the court to ensure that their actions stand up to proper legal scrutiny.

Further, despite the large amounts of data that could be generated in modern-day investigations, and the complexity of the tools used to analyse that data, the legal requirements to disclose investigative files to the defence is still present. This might involve assisting defence counsel with their examination of relevant data, and also providing them with versions of the tools that were used in the investigation, and documentation so they can use the tools, so that they can determine whether the use of AI systems in the investigation is an appropriate area for their defence case (a key legal case here is Rook v Germany (App No: 1586/15) 2019).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

Prohibition on Discrimination

This is the text of the prohibition on discrimination:

> *'The enjoyment of the rights and freedoms set forth in this Convention shall be secured without discrimination on any ground such as sex, race, colour, language, religion, political or other opinion, national or social origin, association with a national minority, property, birth or other status.'*

Note that Article 1, Protocol 12 extends this to prohibiting discrimination under national law, or as a result of the acts of a public authority.

It is closely linked with the belief that all people are born free and equal with the same rights and dignity. What the prohibition means is that you where you are facilitating or fulfilling someone's rights, you should not treat people differently when they are in a similar situation due to one of the protected characteristics, without good justification. Protocol 12 to the Convention extends this prohibition from prohibiting discrimination when providing for the enjoyment of Convention rights, and extends it to activities under national law and the acts of public authorities.

Acts of Discrimination

You should not treat people differently if they are in similar situations as your investigation will engage rights to privacy at a minimum. So, you should make sure that where you make decisions based on INSPECTr outputs, especially if there is a risk of bias, that you would make the same choice if it were a person with different characteristics. For example, if a risk assessment algorithm is trained on historical police data then it is likely to include biased data that results from discriminatory policing policies from the past. Today, this could have the effect that persons from the ethnicity that were previously targeted are likely to have a higher risk assigned to them than they would if they were from another ethnicity. This might result in launching an unnecessary investigation into such persons, which would be an unjustifiable violation of their rights and a clear case of discrimination.

As mentioned, different treatment can be allowed where there is an objective and reasonable justification. Any differences in treatment must relate to pursuing a legitimate aim and be reasonably proportionate to that aim. An example could be requiring someone being interviewed by police who has a poor command of the language where they are being investigated to speak via an interpreter to ensure that they can communicate, and be communicated to, clearly.

Court-Ready Evidence

It is important that rules of procedure are followed, as you do not want the evidence you have worked hard to gather to be thrown out of court merely because you did not follow the correct procedures.

So that your actions can be followed by the defence and the court, it is crucial that you maintain contemporaneous notes about how you handle and process evidence data so that your actions can be replicated if needed.

To that end, it is essential to a successful prosecution that all data and devices that are analysed were lawfully seized. Anything that was unlawfully gathered, and evidence resulting from that line of enquiry is likely to be prevented from being used in court by the judge.

To maintain the integrity of the seized data or devices, it is imperative that your analysis work takes place on copies of the data, or images of the device, so that the original seized evidence is not harmed by your investigative activities. This also allows the defence to examine the evidence on the same basis.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

Good Practices

So that evidence is analysed by investigators with the appropriate training, you should only use tools where you have received the appropriate training and have the correct permissions to use them.

Linked with the discussion on human rights and the applicable legal framework, you should ensure that your use of a tool is proportionate and justified. I.e., you should not use an especially intrusive tool or investigatory technique where a less intrusive approach would achieve the same goal.

Further, you should choose the best tool for your investigative leads. This means that you should not use all tools just in case something important is discovered. Infringements on the rights of suspects, victims, and witnesses should only take place where they can be justified, and 'fishing expeditions' are not justifiable.

Chain of custody has been mentioned several times during this training, and it is important that this is considered during your use of the tool. You should not use tools in a reckless way that could endanger the integrity or viability of the evidence.

So that your use of tools follows national and organisational requirements, and potentially any relevant requirements applicable to the particular techniques used, you should always follow the standards that are applicable to the tools you are using.

Governance of AI investigation tools

As a responsible end-user, it is important that you maintain full control over the decision you make when analysing the result of AI tools. It is not appropriate for you to delegate decisions to the AI system, as it is not equipped to do this and, in any case, cannot deal with the complexities of the investigation, especially qualitative data.

So, it is crucial that you verify any results that generate investigative leads. How they should be verified will depend on the tool in question, but looking at the original data and seeing if the algorithmic results are logical is likely to be a good step.

Linked with the following of standards that we discussed on the previous slides, you should ensure that you use your tool in compliance with the security measures that are required by your organisation, and the legal framework that applies to you. A key part of demonstrating compliance is to have the uses of your tools logged so that they can be examined and prove that you used the tools responsibly.

Summary

To summarise this section, there is a clear legal framework that applies to the use of AI tools by investigators. LEA will have internal policies that clarify legal rules in the context of each LEAs work. However, it is important to be aware of the legal framework so that you know where the rules come from and what is why they are in place. Key legal rules to consider during investigations are the right to privacy, right to a fair trail, and the prohibition on discrimination.

In terms of actually using AI tools, Investigators should check all investigative leads generated by tools so that the investigative choices made can be justified to a court, and potentially to a victim as well.

So that evidence analysed by the AI tools can be presented in court, investigators should make sure that rules of procedure are properly followed.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.