# ROXANNE Ethical, Legal, and Societal analysis: Societal Values Aspects

ROXANNE

This project has received funding from the European Union's Horizon 2020 Programme under grant agreement n°833635. The document reflects only the authors' views and the European Commission is not responsible for any use that may be made of the information contained therein.

ROXANNE is a research project funded by the European Union[1] that intends to combine new speech, text, video, and network analysis technologies into a new platform that will assist law enforcement agencies to identify criminals in organised crime investigations. A key part of this project is ensuring that the activities within the project, and the project results comply with ethical, legal, and societal standards. This is achieved through taking Privacy and Ethics-by-Design approaches to the research activities in the project that are investigating tools and methods to be incorporated into the new platform. To be able to fully engage in these approaches, partners from the project's ethics and legal team have conducted in-depth analysis into the ethical, societal, fundamental human rights, and applicable legislation (including data protection and rules concerning INTERPOL) aspects of the project, and the proposed platform. So that these analyses can be validated, the project's ethics and legal team are sharing a series of briefing papers with important stakeholders to gather feedback. A link to a survey will be provided separately where you will be able to share comments if you wish.

# 1. BRIEFING PAPER

This document includes requirements coming from the societal values analysis, in order to display them with those from the ethical and legal analysis. They will be removed from the briefing paper before it is disseminated.

## Introduction

ROXANNE (Real time network, text, and speaker analytics for combating organized crime) is an EU funded project, aiming to enhance the identification of suspected criminals and their networks in investigations of organised crime and terrorism. It aims to do this by developing novel speech, text, and video analysis technologies to speed up the process of identification and fusing these outputs with network analysis in order to improve the visualisation of how criminal groups communicate. These will be brought together into the ROXANNE platform.

There are ethical and legal issues raised when producing surveillance technologies. The ROXANNE project will include principles of privacy-by-design and ethics-by-design. Key parts of these processes are considering the impacts this technology could have on a societal level. This briefing paper outlines values that are important in European societies. Societal values are 'principles or moral standards held by a person or social group' and are 'generally accepted or personally held judgement of what is valuable and important in life'.[2] They are used in this paper to display the potential impacts that the use of the ROXANNE platform could have on society, and how these effects can be mitigated. Also included are scenarios that highlight potential issues in future uses of the platform; we welcome comments and suggestion on these scenarios.

## Societal Values

### Citizens' privacy:

Privacy means 'the right [for people] to keep their personal life or personal information secret or known only to a small group of people'.[3] This value is closely associated with the value of individual freedom which is defined as 'the condition or right of being able or allowed to do, say, think, etc. whatever you want to, without being controlled or limited'.[4] This value is critical in the sense that citizens need to believe that no aspect of this project will hamper their rights. Privacy can be impacted when technologies are not used as intended; when their intended use impacts inappropriately upon privacy, or when they are inadequately secured allowing others to inappropriately exploit them.

A fundamental issue with platforms that process biometric data is data-subject privacy and trust in the platform.[5]  There is a possibility of the platform being used for unintended purposes which might differ from what was initially envisioned. This is called 'function creep',[6] an obvious example would be a platform intended for targeted surveillance being used for mass surveillance.[7] "Chilling effects" occur when such platforms might impede citizens' freedom to act due to the fears of misappropriation of biometric data or of a totalitarian future.[8] One of the possible repercussions of such a scenario could include attempts to control the public behaviour by leveraging the fear of being monitored.[9] This shows how intrusions into public privacy can lead to serious consequences.

As part of including privacy and ethical concerns into the design process of the ROXANNE platform, technical partners in the project should first ensure a sound legal basis for processing of personal data, with a clear link between the design of the platform and the necessity for processing such data. This should be based upon scientifically valid causal models (e.g, we have good scientific reasons to believe that processing a particular form of personal data will lead to useful and effective analytic tools).  From a security perspective, they should ensure that the data is completely secured from any unauthorised access by implementing efficient data protection measures. They should further incorporate data-security in the system architecture by design and by default.[10] This includes measures such as conducting data protection impact assessments, writing apt privacy policies in easy-to-understand language, providing data-subjects information on how their data is used and who they can contact about it, ensuring that personal data is not automatically made publicly available to others etc.[11]

Further, ensuring that data processing in the ROXANNE platform follows data protection legislation applicable to law enforcement activities[12] would be the key to preventing unauthorized data sharing. With respect to collection of data during operational use, the onus of using lawfully collected data would be on the end-user of the platform. The platform should support use-logging and access control to allow its use to be appropriately audited. However, the consortium must be very careful with respect to the organisations' who would be given access to this platform by conducting due diligence to make sure that the platform will not be abused by the end-users and would only be used for law enforcement. End-users will only be responsible law enforcement agencies (LEAs), mostly likely in Europe. All these measures should help mitigate the concerns related to citizens' privacy.

## Trust and the perception of safety:

People have trust in others when they 'believe that someone is good and honest and will not harm you, or that something is safe and reliable'[13]. It is imperative for citizens to feel confident with respect to deployment of the ROXANNE platform, and that it will make them feel safer in their societies. This confidence will be closely correlated to the trust that citizens have in the organisations involved in using the ROXANNE platform.

Uses of the ROXANNE platform could misuse personal data in ways that abuse the trust of citizens. One possibility could be using the platform to run network analysis on individuals who are not directly associated with any known suspects. This would increase apprehensions of mass-surveillance and abuse of power by the state. Also, any bias in processing of data based solely on the difference in creed, colour, race or religion, which would be tantamount to discrimination by design, could increase distrust with respect to the platform. Algorithmic transparency is a crucial step to further the cause of garnering trust in the platform;[14] project partners should be able to explain how the platform works in order to give citizens an idea of how they can expect any data LEAs collect on them to be processed.

A lack of faith in LEAs deploying this platform, might also lead to suspicion over the intended or actual use of this platform. To mitigate these concerns, end-users (LEAs) should play an active role in trust-building actions regarding this platform.[15] In order to inform citizens and increase their trust in the platform, end-users should, as far as possible, provide information about how use of surveillance platforms is overseen and discuss with local populations; this should increase the security of society as a whole whilst reducing the scope for the abuse of power by end-users.[16] LEAs should also be open about data retention timespans (or criteria for determining whether to store personal data), and how data-subjects can be exercise their rights, along with training individuals to ensure ethical conduct while processing data. Moreover, every activity on the ROXANNE platform should be logged so that it can be audited. It should also be clear to citizens how end-users can be held accountable for cases of misusing surveillance platforms.

A recent whitepaper by the European Commission has suggested a comprehensive approach to build trust in AI systems through developing an 'ecosystem' of trust where all applicable laws are complied with and multiple entities have oversight of such systems.[17] End-users should consider implementing internal oversight measures to monitor deployment of such systems, but also external measures to evaluate processes for using the platform and its outputs, a key indicator for this having large scale pilots/trials.[18] Examples of such oversight include the West Midlands Police (UK) ethics board which has included lay members of the public and has been active in consideration of the development of predictive policing tools.[19]

## Unintended consequences of technological solutions:

Technologies are generally adopted for the benefits that they bring. However, there are often additional features of technologies that create consequences for their users and the public that are not beneficial. Understanding the implications of using a technology and the effects they can create is important for society as it helps all stakeholders get a better understanding of undesirable technical features of different platforms.There are also systems where benefits for end-users have negative (externalised) consequences for other groups, a problem that particularly affects marginalised and vulnerable populations, whose needs and circumstances are not taken into account in the design and deployment of a technology.[20] Externalised consequences can include impacts on other social values.

Biometrics based systems have some inherent limitations.[21] For instance, in voice/speech based biometric systems, a suspect's sample might actually sound different depending on person's health, time of the day and even depending on who the person is interacting with;[22] they could also be mimicked and fool a recognition algorithm.[23] Another example could be that of probabilistic outcomes, such as false-positives (highlighting an innocent citizen) or false-negatives (not recognizing a potential suspect),[24] which could cause issues for those individuals and the public. Further, if the end-users are not well trained, they might use the platform in a mistaken manner to get fallacious results.[25]

To tackle these issues in ROXANNE, technical partners should ensure that recognition technologies are accurate enough to identify targeted persons, but also have some variance to account for different circumstances that might affect the quality of data collected during operations. These technologies should also be thoroughly tested to ensure that the incidence of false-negatives and false-positives is not so great as to cause difficulties for impacted populations. As these issues can only be mitigated and not resolved, it is important that information about them is included in the training provision to be given to end-users so that they can understand the limitations of the platform and the implications of using it.

## Social acceptability:

Acceptability can be described as 'the quality of being satisfactory and able to be agreed to or approved of '.[26] Project partners are trying to develop the ROXANNE platform in a way that citizens trust it after appreciating the pros and cons associated with the platform. As public servants, LEAs need to use tools/technologies that are socially acceptable.

The willingness to accept key aspects of innovation among all stakeholders can be subdivided into two broad segments: (a) acceptance of the creation of the socio-economic conditions needed for implementation and (b) acceptance of all consequences of the innovation. The latter refers to the ways in which implementation will affect and change current practices in society.[27] Further, social acceptability is a result of citizens' attitude towards the overall proposition (use of the ROXANNE platform in this case). This attitude could be influenced by awareness about perceived risk/uncertainty, values or beliefs of the citizens, trust in the users and developers of the platform, participation in decision making process, potential benefits from the project etc.[28]

Literature on the technology industry suggests that citizens are overwhelmingly more likely to trust organisations with strong privacy policies, and those who are transparent about how they use data.[29] Assuming that people view public and private organisations in similar ways when it comes to trusting that they use data in compliance with ethical and legal standards, then this indicates that having LEAs be open about their data processing and a strong privacy policy should enhance citizen's trust of LEAs.

Indeed, citizens are unlikely to find biometrics based platform such as ROXANNE acceptable where: they fear it could be used for mass surveillance, or to encroach upon their privacy; when they do not trust the police;[30] or when they are uncomfortable with an organisation holding sensitive data about them.[31] Thus, providing citizens a complete picture of the platform, its policies and fairness of process becomes imperative. For citizens to be able to trust that LEAs use their data properly, LEAs need to be able to demonstrate that the use of ROXANNE would in no way affect the security or freedoms of innocent people. Steps toward this can include raising awareness about the accuracy and data security of the platform and taking citizens feedback into account wherever possible.

Citizens are usually only indirectly involved in the development of novel technologies. They shape the innovation process by voicing their opinions or by displaying actions that support or resist a technology, both after and before market introduction. However, the overall public acceptance for a such a technology can be gauged through opinion polls that represent aggregated attitude of citizens.[32] This feedback is key to making citizens part of the decision-making process and raising their confidence in this platform. This input will help guide the design, dissemination as well as exploitation of this platform as a whole, which in turn encourage greater social acceptability.

Further, by engaging in a continuous effort towards creating a platform which is built keeping in mind all other societal values, we can increase the probability of social acceptability for this project. This includes raising awareness about the platform as a whole. The consortium should inform the citizens about the extent of data security to make citizens feel safe about their data. The consortium should also highlight the extent to which this platform will help prevent crime while ensuring swift identification of suspected criminals. However, it is equally significant to inform citizens about the possibility of false-positives and false negatives and how the project is dealing with this; oversight mechanisms, and the process set forth to rectify errors in such a situation. The consortium should also spread awareness about legal measures that protect citizens from unjust effects of processing of their personal data.

## Democracy and solidarity

Democracy is a popular method of collective decision-making, particularly in political systems. Key to the implementation of democracy is that the people who participate in the decision-making are treated equally and have the necessary liberties to engage in it.[33] This is an important societal value as it allows people to group together in solidarity with others to pool their collective power for common causes (for example, political movements).

There is potential for the ROXANNE platform to affect democratic expression. For example, a person is less likely attend a political rally if they believe that they will be subject to surveillance by state agents and this will lead to unfavourable treatment by the state; this is an example of a 'chilling effect'.[34] ROXANNE poses a particular issue if its users identify people under surveillance and then use its network analysis capabilities to identify other people in the networks of political activists. This increases the likelihood of such chilling effects as people will be further disinclined to partake in particular activities so as not to implicate their friends and family. If this effect is realised, it is likely to lead to less political participation from the public and an acceptance of the status quo to protect their acquaintances, despite not being in favour of it.

These risks can be mitigated through preventing sales of the ROXANNE platform to authoritarian states, or actors who might engage in repression of persons who attend political events. Further, the implementation of decision-making processes that require ethical and legal compliance in order for the platform to function should prevent the platform being abused where these processes are followed. The legitimate use of ROXANNE is somewhat dependent on proper training of the end-users, and the incorporation of end-user training as part of the ROXANNE project should contribute to this.

## Equality and tolerance for other cultures

Equality is a societal value that holds all people to be equal whatever their differences. This is an important value as it enables all people to be treated fairly,[35] no matter what their status. This is a key principle of International and European political and legal systems.[36]
ROXANNE has the potential to affect people from different social groups in a disproportionate way. Bias in the outputs of a platform can be caused where: a data set used to train the models is biased toward or against a particular group; the dataset is not representative of the environment it will be used in, or the population it will be used with; where the system is not measuring representative data.[37]

For example, members of a group might be treated differently by a facial recognition algorithm due to the colour of their skin where the model has been trained on more pictures of people from one ethnic group than another.[38] This is an issue of particular relevance to policing. Where existing police data is biased and provides a skewed view of a particular group, then that affects how the outputs of data-analysis systems are assessed. If this influences future policing, it can lead to a compounding of bias.[39] However, it is complicated further by factors specific to criminality such as the greater prevalence of more crimes being committed by young men in comparison to other groups.[40] The impact of producing a system to specifically targeted these people is that biases are reproduced and such persons are at significant risk of being discriminated against.

The ROXANNE consortium should do all that it can to alleviate risks of this happening through evaluating all the data sets which it is using to train the platform on to ensure that they are not biased for or against different groups, and ensuring that the platform is measuring data that is representative. This should, therefore, reduce the risk of the ROXANNE platform having a discriminatory effect when it is used.

## Human rights

Human rights are legal rules that require states to respect and protect people. They also provide a framework where state actors can infringe upon rights in situations where this is necessary and proportionate. Privacy rights are well known, and these can be lawfully infringed upon in some situations such as where law enforcement needs to know private information in order to prevent or investigate serious crimes.[41]

There is always a risk with data-analysis technologies that they could be used in a way that is arbitrary, meaning that it is not necessary or proportionate to use in a specific situation. However, ROXANNE poses particular risks as it analyses not only at the individual who police are interested in but also at whom they communicate with; it could be arbitrary to include their associates in the surveillance activities.

In order to mitigate this risk, the ROXANNE platform should only be sold to law enforcement agencies in states with a good human rights record. The platform could be built to include decision-making process that require law enforcement officers to take a decision on whether to include or exclude the data of associates from a network analysis, the decision-making processes will incorporate the human rights legal framework in order to facilitate compliance.

## Respect for human life

Recognising that all people have an inherent dignity is a societal value that underpins human rights, equality, and fair treatment of others.[42] Where people ignore the dignity of others, this is a process of dehumanisation and people are treated as less than human, resulting in systematic atrocities at its worst extent.[43] Data processing about people can lead to a less dramatic form of dehumanisation where people are treated as mere data points, leading people to forget that the outputs of algorithmic systems have real consequences for other human lives.

With ROXANNE, this could be particularly problematic where, for example, the platform is used to analyse police surveillance data and operational decisions are made based on the outputs of the platform, rather than a police officer evaluating the person under investigation. Or an investigator trusts an algorithm, rather than making the decision themselves. For example, this could lead to a citizen being subject to further investigation and analysis of their sensitive data even though their actions are perfectly innocent, and this would have been understood had a human evaluated the original results in a meaningful way.

These risks can be mitigated in the ROXANNE platform through structuring the relationship between human and machine to avoid (or minimise) issues of blindly following machine outputs (automation bias), and to prioritise human decision-making. Technical partners should structure the human-machine relationship so that the benefits of machine analysis are used to complement human decision-making.

## The rule of law

Having all people and institutions subject to legal rules and legal frameworks is a key aspect of democratic systems as it prevents different parts of the system from gaining excessive power. This is a key societal value as it allows people to trust their institutions.[44]

ROXANNE does not necessarily pose a direct risk to this system or trust in institutions. But, human beings often give greater weight to the outputs of advanced technologies over themselves or other human beings.[45] In the context of a criminal trial, this could pose an issue whereby evidence from the surveillance data analysed by the ROXANNE platform is given greater weight than evidence from other surveillance technologies would normally be given. This could, potentially, mean that the results of data analysis are seen as more conclusive than they should be, and this could lead to misunderstandings in court. Potentially, this could affect how the guilt or innocence of a defendant is viewed in court.

A way to mitigate this would be for ethical/legal partners to disseminate information about this risk to highlight this issues so that it can be properly understand that whilst ROXANNE and similar technologies are advanced, this should not mean that evidence generated from them should be given significant weight in a criminal trial. Potential recipients could include groups representing judges and lawyers.

## Emerging themes

This paper discusses the issues that could be raised from the potential use of the ROXANNE platform in terms of societal values. Some values place importance on independent oversight of LEAs using the ROXANNE platform with accountability measures to increase compliance with applicable standards. These are important features about the organisations that will use ROXANNE. In terms of the platform itself, ensuring transparent processing and un-biased algorithms are important as this should result in fair treatment of citizens by the platform, and an ability for LEA officers to understand what is happening inside the platform to enable them to make fully informed decisions for their investigations. Another important theme is that LEAs should only use the ROXANNE platform in a way that is lawful and appropriate for the investigation at hand. These themes, amongst other issues, show that whilst violations of privacy are undesirable for society, they can be carried out in conformity with societal values where they are fair, lawful, and subject to accountability measures. In the specific case of using ROXANNE, ensuring that a human being is in control of deciding how to use machine outputs also seems to be a key requirement for compliance with societal values.

These societal requirements may appear to misalign with the project objective of increasing the speed at which organised crime investigations can take place: increasing the human role and oversight can slow down uses of automated systems. Yet, this need not be an issue for the use of ROXANNE as the overall speed of an investigation, even with the necessary human input, might well progress faster than the current tempo of investigations. Further, when the appropriate permissions and authorisations are in place, the investigation time will reduce. As such, human oversight both of the platform and the process of using the platform should not be sacrificed simply to increase the speed of investigations.

# 2. SCENARIOS

The below scenarios are entirely fictional; they present situations where tools like those in the ROXANNE platform could be used and things go wrong. The intention of doing this is to highlight issues that need to be raised and considered (with the focus here on societal issues). Once issues are highlighted, we can focus on developing solutions so that these incidents do not happen with ROXANNE. We would appreciate any feedback you wish to provide, especially if you could provide answers to the questions that are asked. You will be able to respond to these questions via a link to a survey that is available on the ROXANNE project website.

## Scenario 1 – suspected child abuse

An LEA (A) of a European nation receives intelligence from an LEA (B) in a neighbouring country regarding a possible perpetrator of child exploitation. The LEA (B) has used the ROXANNE platform to recognize voices in phone calls which match to those recorded in previous investigations from 2019. One of the calls is traced to Mark's house. His house falls under the jurisdiction of LEA (A). Mark lives with his wife, two children, and his father in a sophisticated area of the city. According to the tip, several infrequent telephone calls have been made to known child abusers from the house owned by Mark. These child abusers are associated with uploading homemade content to a dark web site.

| |
|---|
| Question: Would you (LEA A) require any information about the use of a data-analysis platform by LEA (B) upon receipt of intelligence ? Would you need to know specific results of the data analysis? Would you want to know which analysis platform was used? |
| Answer: |

LEA (A) officers request to place Mark under surveillance. After considering multiple documents, including the results of the ROXANNE platform, a judge gives the required permission for surveillance. This includes intercepting the voice calls from the landline phone in Mark's house, and footage from the CCTV cameras near Mark's house.

LEA (A) officers request to place Mark under surveillance. After considering multiple documents, including the results of the ROXANNE platform, a judge gives the required permission for surveillance. This includes intercepting the voice calls from the landline phone in Mark's house, and footage from the CCTV cameras near Mark's house.

| |
|---|
| Question: Is it likely that a judge would authorize surveillance in your country based primarily off the results of a data-analysis platform? Would other corroborating evidence be required? |
| Answer: |

After three days of surveillance, the speakers in a voice call from Mark's house are matched by the ROXANNE system. The caller speaks very little on the call, and the ROXANNE system suggest it is more likely than not that the caller is Mark. The LEA officers assume that the caller is Mark and the short voice samples are the reason that the match is not more definitive. The call recipient speaks a lot on the call and their voice is matched by the ROXANNE system to a known child abuser. The LEA (A) officers conclude from this that Mark is in direct contact with known child abusers.

On another call, Mark heard discussing a business trip to another city and the investigators are concerned that he might meet other child abusers. Officers begin to look into putting Mark under surveillance for the duration of his business trip.

| |
|---|
| Question:  Could you foresee a situation where LEA officers make decisions just based off the results of a data-analysis platform, rather than also using their intuition and experience? Would this concern you? |
| Answer: |

In the days before Mark's business trip new video content is uploaded to the dark web site used by the child abusers. The video metadata shows that the video was recorded one day earlier. The face of one of Mark's children is recognised in the video content by the ROXANNE platform that is comparing the video with CCTV images.

Owing to the child protection risks, LEA (A) officers raid Marks house. Mark, his wife and father are arrested, and his children are taken into temporary care by the authorities.

Question: It is likely that you would incorporate two streams of evidence from an investigation (e.g. video files gathered from CCTV and the dark web) for analysis? Or, would you only compare evidence with data in a verified database, for example?

Answer:

During questioning, it is shown that Mark and his wife were shopping all day when the abuse content was filmed. Mark's father, Simon, was staying in Mark's house and is shown to have a very similar voice to Mark. Upon further investigation, it is determined that Simon made the calls to child abusers from Mark's house and filmed the abusive content.

LEA officers take voice samples from Simon's police interviews and analyse them using the ROXANNE platform. Simon's voice matches with several samples from previous voice recordings associated with child abuse where the speaker was unknown.

Question: Would you need special permissions to process (biometric) data gathered in one case for another investigation? If so, what permissions would you require?

Answer:

## Scenario 2 – suspected drug dealing

Frank is a member of an ethnic minority and lives in a community that has recorded a high crime-rate for a long time. He is seen interacting with known leaders of criminal organisations who are under video surveillance by officers investigating gang violence. Surveillance images are analysed using the ROXANNE platform which suggests a high-probability that Frank is actually William, the former leader of a drug gang who left the area several years ago. LEA officers who remember William think that Frank looks similar to, but not exactly like, their memories of William. They put the difference down to the years that have passed and trust the algorithm.

| |
|---|
| Question: How should the ROXANNE platform present the results of components that can recognise an individual? Display the most probable match? List the 10 most probable matches? List all those with a probability match above a certain percentage? Something else? |
| **Answer:** |

| |
|---|
| Question: Should LEA officers be allowed to 'trust the algorithm'? Should algorithmic solutions only be used to inform an LEA officer's judgement? Should investigators corroborate data-analysis results they want to use? |
| **Answer:** |

The determination that William has returned to the area is included in intelligence reports to a new regional anti-drug squad who are investigating a large and well-organised drug gang. Owing to William being observed interacting with criminal leaders, and William's extensive criminal record, investigators show the information they have to a judge who is also convinced that Frank is William and obtain a warrant to place William under surveillance by monitoring his phone calls, text messages, and emails.

Question: How should information about the results of recognition technologies be reported within and by LEAs? Reporting who was recognised? Reporting the probability of recognition? Something else?

Answer:

Question: If multiple people are recognised with a high probability, should all these possible recognitions be included in reports?

Answer:

Officers record several phone calls where William is heard telling the leaders of drug gangs that they should 'work for him'. Investigators use the ROXANNE platform to visualise the connections between people whose communications are monitored; this shows William as a key node in a network with known criminals. William's emails also reveal that he manages a community organisation campaigning for better political representation of ethnic minorities. Owing to the strength of communications with many criminals, investigators theorise that the community organisation could be a front for hiding a criminal network run by William. They decide to investigate the community organisation further.

Question: How should the context of data analysis be conveyed? Should suspects, known criminals, and innocent people all be highlighted in some way?

Answer:

In their expanded investigation, LEA officers use the ROXANNE platform to analyse the seemingly innocent communications William has with his staff at the community organisation. The text analysis part of the platform outputs that staff members regularly use slang terms for drugs typical of criminal organisations, and the voice recognition part of the platform recognises several staff members of staff who are from ethnic minorities as having criminal records in an LEA database.

Question: If data from innocent persons is captured by LEA surveillance, how should these people's privacy be protected during data-analysis? What safeguards should be implemented?

Answer:

Question: Should data analysis systems have access to historical LEA databases even if those databases contain data generated by discriminatory policing practices from the past? What safeguards should be implemented?

Answer:

From all of these data, investigators conclude that William is overseeing a major drug dealing operation with several local gangs working for him. LEA officers decide to raid the community organisation for evidence of drug dealing. They find no evidence, but determine that Frank is not William and was in contact with criminals in order to try and convince them to leave their criminal activities and 'work for him' at the community project. They also discover that the prevalence of criminal records and use of slang typical of criminal organisations is due to the community organisation hiring ex-prisoners as an example of rehabilitation.

Question: Is it likely that arrests could be made based purely on the results of a data-analysis platform? Would corroborating evidence be required?

Answer:

Owing to the sensitive nature of the investigation, LEA officers are unable to explain their actions in detail. This results in a loss of trust between the community and LEAs. It also deters people from engaging in legitimate political activism as some locals feel the community organisation was targeted for its political activities. Owing to the complexity of the algorithms used, LEA officers are also unable to explain why the platform made the determinations that it did.

| |
|---|
| Question: Should LEAs be open with the public about what surveillance tools they are using? How open should they be? How should they explain surveillance and data-analysis tools to the public? |
| Answer: |

| |
|---|
| Question: If possible, would you like to know why data analysis platforms produce the results that they do? How much detail would be beneficial? |
| Answer: |

# REFERENCE

[1] This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 833635

[2] See ‘Value’ 6d, Oxford English Dictionary, OUP, UK, 3rd edn. 2011.

[3] See ‘Privacy’ B2, Cambridge Advanced Learner's Dictionary & Thesaurus, CUP, UK. Available at: https://dictionary.cambridge.org/dictionary/english/privacy

[4] See ‘Freedom’ B2, Cambridge Advanced Learner's Dictionary & Thesaurus, CUP, UK. Available at: https://dictionary.cambridge.org/dictionary/english/freedom

[5] Ashbourn, Julian, “Background paper for the Institute of Prospective Technological Studies”, European Commission DG Joint Research Centre, 2005. Available at: http://www.statewatch.org/news/2005/apr/jrc-biometrics-julian-ashbourn.pdf

[6] Dekkers, Dick, “Privacy or security? - 'Function Creep' kills your privacy”, Digidentity, 2016. Available at: https://www.digidentity.eu/en/article/Function-creep-kills-your-privacy/

[7] Pastukhov, Oleksandr and Els Kindt, “Voice Recognition: Risks to Our Privacy”, Forbes, 2016. Available at: https://www.forbes.com/sites/realspin/2016/10/06/voice-recognition-every-single-day-every-word-you-say/#2621e79e786d

[8] Zuboff, Shoshana, "The Surveillance Threat Is Not What Orwell Imagined", Time.com, 2019. Available at: https://time.com/5602363/george-orwell-1984-anniversary-surveillance-capitalism

[9] Ievdokymova, Iryna, "Surveillance and profiling: what's next?", Leidenlawblog, 2013. Available at: https://leidenlawblog.nl/articles/surveillance-and-profiling-whats-next

[10] Art.25, GDPR.

[11] Irwin, Luke, "What is data protection by design and default?", Itgovernance, 2019. Available at: https://www.itgovernance.co.uk/blog/what-is-data-protection-by-design-and-default

[12] Art.1(1), LED.

[13] See ‘Trust’ B1, Cambridge Advanced Learner's Dictionary & Thesaurus, CUP, UK. Available at: https://dictionary.cambridge.org/dictionary/english/trust

[14] Epic, “Algorithmic Transparency: End Secret Profiling”, Epic. Available at: https://epic.org/algorithmic-transparency/

[15] Swaminathan, Aravind and Antony P. Kim, "Biometrics: A Fingerprint for Privacy Compliance, Part I", Orrick, 2016. Available at: https://blogs.orrick.com/trustanchor/2016/03/04/biometrics-a-fingerprint-for-privacy-compliance-part-i/

[16] European Parliament, “A governance framework for algorithmic accountability and transparency”, EU, 2019. Available at: https://www.europarl.europa.eu/RegData/etudes/STUD/2019/624262/EPRS_STU(2019)624262_EN.pdf

[17] European Commission, “Whitepaper on Artificial Intelligence - A European approach to excellence and trust”, EU, 2020, pp.9-10. Available at: https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf

[18] Institute for Prospective Technological Studies, “Biometrics at the Frontiers: Assessing the Impact on Society”, European Commission DG Joint Research Centre, 2005. Available at: http://www.statewatch.org/news/2005/mar/Report-IPTS-Biometrics-for-LIBE.pdf

[19] Heubl, Ben, “West Midlands Police strive to get offender prediction system ready for implementation, E&T, the IET, September 24, 2019. https://eandt.theiet.org/content/articles/2019/09/ai-offender-prediction-system-at-west-midlands-police-examined/

[20] Noble, Safiya, Algorithims of Oppression, NYU Press, New York, 2018; Eubanks, Virginia, Automating Inequality, St Martins Press, New York, 2018, and Benjamin, Ruha, Race Against Technology, Polity Press, 2019

[21] Prabhakar, Salil, Sharath Pankanti, and Anil K. Jain, “Biometric recognition: security and privacy concerns”, IEEE Security & Privacy, Vol.1, No.2, March-April 2003, 33-42.

[22] Scheips, Derek, "Voice Recognition – Benefits And Challenges Of This Biometric Application For Access Control", Securityinformed. Available at: https://www.securityinformed.com/insights/co-3108-ga.4100.html ; Ahaskar, Abhijit, "Voice biometrics are cleverer now, but still need more work", Livemint, 2020. Available at: https://www.livemint.com/technology/tech-news/voice-biometrics-are-cleverer-now-but-still-need-more-work-11581011267941.html

[23] Panjwani, Saurabh and Achintya Prakash, “Crowdsourcing Attacks on Biometric Systems”, USENIX, 2014. Available at: https://www.usenix.org/system/files/conference/soups2014/soups14-paper-panjwani.pdf

[24] Penny, Wayne, "Biometrics: A Double Edged Sword - Security and Privacy", SANS Institute, 2020. Available at: https://www.sans.org/reading-room/whitepapers/authentication/biometrics-double-edged-sword-security-privacy-137

25 Zadelhoff, Marc van, "The Biggest Cybersecurity Threats Are Inside Your Company", Harvard Business Review, 2016. Available at: https://hbr.org/2016/09/the-biggest-cybersecurity-threats-are-inside-your-company ; SSE, "KnowBe4 Benchmarking Report: Untrained Users Pose The Greatest Risk To Your Organization", SSE. Available at: https://www.sseinc.com/cyber-security/knowbe4-benchmarking-report-untrained-users-pose-the-greatest-risk-to-your-organization/

26 See 'Acceptability', Cambridge Advanced Learner's Dictionary & Thesaurus, CUP, UK.
Available at: https://dictionary.cambridge.org/dictionary/english/acceptability

27 Wolsink, Maarten, "The research agenda on social acceptance of distributed generation in smart grids: Renewable as common pool resources", Elsevier, Vol.16, Issue 1, January 2012, 822-835.

28 Government of Quebec, "Social Acceptability", Quebec.ca, 2019. Available at: https://www.quebec.ca/en/government/policies-orientations/social-acceptability/

29 Fraser, Adam, "Is an ethical approach to customer data privacy your trust differentiator?", EY, 2020. Available at: https://assets.ey.com/content/dam/ey-sites/ey-com/en_au/topics/data-privacy/ey-is-an-ethical-approach-to-customer-data-privacy-your-trust-differentiator.pdf

30 Goldsmith, Andrew, "Police reform and the problem of trust", Sage Publications, London, 2005. Available at: http://www.slcdocs.com/ODHR/Website/Right%20to%20Safety/Literature/PoliceReformAndTheProblemOfTrust.pdf

31 Ada Lovelace Institute, "Beyond face value: public attitudes to facial recognition technology", Ada Lovelace Institute, 2019. Available at: https://www.adalovelaceinstitute.org/wp-content/uploads/2019/09/Public-attitudes-to-facial-recognition-technology_v.FINAL_.pdf

32 Rijnsoever, Frank J. van, Allard van Mossel and Kevin P.F. Broecks "Public acceptance of energy technologies: The effects of labeling, time, and heterogeneity in a discrete choice experiment", Elsevier, Vol.45, May 2015, 817-829.

33 Christiano, Tom, "Democracy", The Stanford Encyclopaedia of Philosophy, 2006. Available at: https://plato.stanford.edu/entries/democracy/

34 Solove, 2006, 477-560, 487.

35 See, for example, Rawls, John, Justice as Fairness: A Restatement, Belknap Press, United States, 2001.

36 See, for example, International Convention on the Elimination of All Forms of Racial Discrimination (adopted 7 March 1966, entered into force 4 January 1969) 660 UNTS 1; Directive 2006/54/EC of the European Parliament and of the Council of 5 July 2006 on the implementation of the principle of equal opportunities and

equal treatment of men and women in matters of employment and occupation (recast) OJ L 204, 26.7.2006.

[37] Woodie, Alex, Three Ways Biased Data Can Ruin Your ML Models, datanami, 2018. Available at: https://www.datanami.com/2018/07/18/three-ways-biased-data-can-ruin-your-ml-models/

[38] See, for example, EU Agency for Fundamental Rights, #BigData: Discrimination in data-supported decision making, FRA, 2018.

[39] Babuta, Alexander and Marion Oswald, Data Analytics and Algorithmic Bias in Policing, RUSI, 2019, pp.11-12. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/831750/RUSI_Report_-_Algorithms_and_Bias_in_Policing.pdf

[40] See, for example, Schwartz, Jennifer, et al. "Trends in the Gender Gap in Violence: Reevaluating NCVS and Other Evidence" Criminology, Vol.47, No.2, May 2009, pp.401-425; Devon, James, "Age and Crime" The Police Journal, Vol.65, No.3, July 1992, pp,268-273.

[41] See, for example, Article 8, European Convention for the Protection of Human Rights and Fundamental Freedoms (adopted 4 November 1950, entered into force 3 September 1953) 213 UNTS 221.

[42] See, for example, Preamble, Universal Declaration of Human Rights (3rd session, 10 December 1948) UN Doc. A/RES/217(III).

[43] Smith, David L, "Dehumanization, Essentialism, and Moral Psychology", Philosophy Compass, Vol.9, Issue 11, 2014, 814-824.

[44] Postema, Gary J, "Trust, Distrust, and the Rule of Law", in Paul B. Miller and Matthew Harding (eds.), Fiduciaries and Trust: Ethics, Politics, Economics and Law, CUP, Cambridge, Forthcoming. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3394978

[45] See, for example, Skitka, L.J., Kathleen L. Mosier and Mark Burdick, "Does Automation Bias Decision-Making?" International Journal of Human-Computer Studies, Vol.51, 1999, 991.