

ROXANNE ROXANNE ETHICAL, LEGAL, AND SOCIETAL ANALYSIS: ETHICAL 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 Union¹ 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.

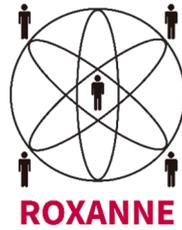
1. INTRODUCTION

The ROXANNE project is committed to respecting and fulfilling ethical principles. To achieve this, ROXANNE partners with ethical expertise assessed the activities of the project and proposed uses of the technologies to be developed against emerging ethical principles for the development of technologies such as artificial intelligence (AI).

Ethics principles

The ethics principles used in the analysis come from the EU-funded SHERPA project, which adopted the seven principles devised by the EU's High-Level Expert Group on AI and adapted them to take into account 24 other similar ethics codes for smart/intelligent technologies.²

Each of the seven SHERPA ethical requirements is subdivided into subsidiary or associated principles. The following table lists the seven ethical requirements as well as the subsidiary principles associated with each of the main guidelines. These ethical requirements come from SHERPA deliverable D3.2 (Guidelines for the development and use of SIS).³



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 smart information systems, 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



Ethical analysis and governance

ROXANNE partners with ethical expertise conducted an initial analysis of project tasks for ethical risks; this involved use of a 'touchpoint table', which is a tool developed by Trilateral Research for the assessment of ethical, privacy, and societal risks.⁴ These risks were discussed with other partners leading work in the project to determine if their concerns and mitigation strategies were realistic. This work was then used as a starting point for discussion of ethical issues between the ethics and legal specialists in the project, and to begin a deep analysis of these issues in order to develop benchmark requirements as part of an impact assessment.

The project is currently at its mid-point and as the project progresses and the ROXANNE tools become better defined, the ethical analysis will become more granular, both in terms of how project tasks progress, and what the proposed technologies will look like toward the end of the project. However, some initial considerations about potential use of the ROXANNE tools are incorporated into this document.

Detailed analysis of the ROXANNE project was grouped into phases, in line with the guidelines from the SHERPA⁵ project. The phases considered are:

Phase 1: Requirement Gathering

Phase 2: Planning and Designing

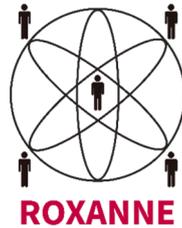
Phase 3: Development

Phase 4: Testing

Phase 5: Evaluation

Phase 6: Use

This briefing paper now gives a summary of some of the ethical analysis conducted of the ROXANNE project across each of the 6 phases of the project using each of the SHERPA ethical requirements.



2. PHASE 1: REQUIREMENT GATHERING

A major task of the ROXANNE project is to gather end-user requirements (i.e. technical, operational, legal, and training) from the end-user community in order to integrate them into the system design and development of the ROXANNE platform. This will enable the ROXANNE consortium to develop a solution tailored to the experiences and needs of law enforcement. To comply with the diversity, non-discrimination and fairness requirements, it is key to target and collect input from a diverse and representative pool of end-users. To this end, in addition to the 10 partner law enforcement agencies (LEAs) in the project, the end-user requirements survey was circulated to the Stakeholder Board and INTERPOL's global network of law enforcement contacts in 194 member countries. This enabled a diverse range of people from different professions, countries, and background to participate.

Owing to the use of human participants in this research activity, issues of human agency, liberty, and dignity are raised by how those who are providing the requirements are treated; these mostly overlap with principles of research ethics, particularly those related to informed consent. Participants' agency to choose whether to participate was respected, their positive liberty to make choices about their participation was enabled, and their negative liberty to be free to leave their participation at any time was respected by giving respondents full information about their participation and allowing them to make a free choice about whether they wish to participate. These requirements were met as partners provided detailed information sheets with the survey that explained what was being asked of respondents, and, importantly, that they were in control of their participation at all times and could skip questions, or only partially complete the survey if they wished. Further, partners' answers were only evaluated where they consented. The information sheets also provided details on how personal data would be processed and so also contributed providing transparency.

With regard to technical robustness and safety, partners used reliable tools for data collection, including the EUSurvey platform, only processed data in accordance with the GDPR, and implemented appropriate technical and organisational measures for data protection. To ensure accuracy, reliability, and precision, each part of the surveys (end-user, end-user training, and legal requirements) were drafted and reviewed by different partners with relevant expertise, and tested in a pilot-study with internal LEAs. Further, the surveys were developed in English and then translated into Arabic, Spanish, and French to ensure greater accuracy from people who might not speak English as a first language.

In terms of access to survey responses, these were pseudonymised by removing names and personally identifying information before being provided to partners to analyse. Having informed participants about this process and only collecting their data where they consent, their right to privacy has been respected. Further, participants were given control over their data through being able to limit how their responses are used on their consent form, and request that their data is removed from the research data; thus contributing to compliance with the standards of research ethics,⁶ and the use of consent as the legal basis for processing.⁷



3. PHASE 2: PLANNING AND DESIGNING

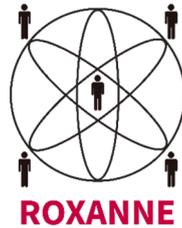
In this phase, ethical risks are raised in relation to partners being treated with respect and dignity when planning and designing the ROXANNE platform. This includes complying with the European Code of Conduct on Research Integrity, as required by the ROXANNE Grant Agreement. Partners have been acting in compliance with this, and respecting the dignity of colleagues during the project.

To ensure technical robustness and safety, the work of partners in the planning and designing phase is safe, secure, accurate, reliable, and precise, the platforms used for consortium work need to meet these requirements. Partners use secure collaborative platforms for joint work and secure systems for individual work.

The planning and designing phase is one of the most crucial phases with respect to transparency as an ethical factor. All major decisions related to project work are discussed openly, and with oversight boards where necessary. In terms of ensuring algorithmic transparency of the technologies, the platform will be developed in such a way that it can be understood firstly so that the scientific results can be properly assessed, and so that LEAs will be able to comprehend how the tools work and how its results were generated. In terms of being transparent with the public, the project will be posting public deliverables and information about its work on the project website. Further, the project will provide opportunities for feedback from the public on its work.

The design of the ROXANNE system architecture represents a key opportunity for fulfilling diversity, non-discrimination and fairness requirements. Partner shave, and will continue to gathering feedback from a wide range of stakeholders across the life of the project. so that the project results cater for different needs and uses.

With respect to environmental wellbeing, having partners avoid excessive travel for meetings, and design the ROXANNE tools so that they are energy efficient would seem to fulfil this requirement. A reasonable amount of meetings and travel were planned, although these have been reduced during the Coronavirus pandemic-induced lockdowns.



4. PHASE 3: DEVELOPMENT

During development of the ROXANNE platform, respect for human agency, liberty, and dignity would seem to be most relevant to the persons whose data is being processed in order to develop the platform. This relates both to initial data collection by the project, and re-using datasets from previous research activities. Compliance with these requirements involves treating research participants according to research ethics standards.⁸ The use of human participants was reviewed and approved by the ethics review board at one of the ROXANNE university partners, thus demonstrating compliance with research ethics standards.

With regard to re-using datasets, ethical compliance involves respecting the persons whose data is contained in datasets to be reused and only using datasets that were created using proper safeguards. Partners are currently only using datasets that were developed by researchers at universities or research institutes where it is most likely that data was collected under the oversight of an ethics committee.

Another aspect of re-using datasets is only using them in the ways that data-subjects would expect. As the datasets being re-used come from research organisations and the ROXANNE partners are using these datasets for research, this seems to meet the requirement. Additionally, there will be no direct effects for data-subjects; their data will be used for developing algorithms and no decisions will be made about them using their data.

The development phase is clearly the prime opportunity to build the ROXANNE platform so that it is the best that it can be. Thus, choices about data quality are very relevant to avoid high-levels of false positives/negatives, for example, as this could result in people being wrongly highlighted or missed by the ROXANNE platform. ROXANNE partners consider data accuracy in all datasets to minimise the risks of this happening.

In terms of transparency in this phase, dissemination of results and progress made by the project are key opportunities to provide information to the public. Algorithmic transparency is also key here as it facilitates partners being able to demonstrate how the tools work to any reviewers. Following the project, this will also facilitate potential LEA users to be able to demonstrate how they use the platform in court, for example. It would also enable citizens to challenge uses of the platform where they feel aggrieved by the effects of the ROXANNE tools being used. Therefore, partners will be working on ensuring that the ROXANNE tools, and platform, are transparent in such a way as to facilitate these actions.

With regard to non-discrimination, a fundamental concern in data-driven analytical tools such as the ROXANNE platform is the potential reliance on biased datasets to build, improve and/or test the technologies under development as this would result in a skewed product. Whatever the



motivations of end-users, use of biased data can create biased tools which have biased effects during use.⁹ Therefore, partners will be implementing measures to review the composition of datasets, and how this could create effects in the technologies, and mitigate harmful consequences.

With regard to accountability, technical partners' decisions on how to develop the technology, such as data formats or processing methods, are recorded and this gives an opportunity for them to be audited and, if necessary, challenged. Most documents are made public by default in order to provide public accountability; however, some are confidential to the consortium due to concerns of commercial sensitivities or platform security, for example. Still, non-public decisions are subject to accountability within the project through the various boards and compliance mechanisms.

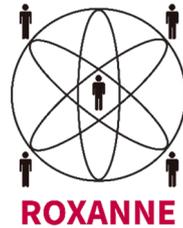
Further, the development phase is a key point at which accountability measures for future use can be embedded into the system. Partners are working on how measures such as access controls, authorization and authentication, logging, encryption, and more could be implemented in a final platform. The partners intend to develop an architecture that contributes to, and facilitates, accountability for users.

5. PHASE 4: TESTING

In the testing phase, issues of human agency, liberty, and dignity relate to the use of human participants who engage in testing of the platform. As outlined above, the use of human participants complies with principles of research ethics and has been approved by an ethics committee at a university partner from the consortium.

It is unlikely that any platform can function perfectly, we have already noted the risks of false negatives and false positives, and in, conjunction with human dignity, there would seem to be an ethical obligation to communicate to users in a meaningful manner the level of accuracy achievable in the platforms outputs, across different contexts of use. Project partners should devote some effort in trials to understand levels of accuracy and how this could affect use of the platform.

Partners hope to be able to test the ROXANNE tools on data from real closed cases in order that the testing situation can be as realistic as possible. Partners will not seek to use data from real ongoing or cold cases, due to the clear privacy risks and potential for incidental findings.¹⁰ Even so, data from real closed cases does pose privacy risks. Partners should, therefore: (1) justify why synthetic data is inadequate and (2) justify why data needs to come from a real case; (3) ensure necessary approvals are collected; (4) ensure the data was lawfully collected; (5) assess the privacy implications, of using the data and consider gaining consent; (6) ensure any infringements on privacy are benign; (7) ensure diversity of and in the data to ensure adequate testing; (8) keep real case data with LEAs; (9) log all persons who access and use LEA data, and what the purposes are.



No ROXANNE LEA partners have determined that they would like to process personal data under the data protection regime for law enforcement agencies; all personal data is expected to be processed under the GDPR. This is the most ethical approach as the ROXANNE project is a research activity and should not be confused with a law enforcement activity. Any consideration of processing personal data under the Law Enforcement Directive¹¹ would need a very strong justification and potentially approval from a national data protection authority before it could go ahead.

In terms of transparency, the testing phase is similar to the development phase in primarily requiring active dissemination of technical progress made in this project. In particular, the project should inform the public about how the platform is tested and what the efficacy of such a platform is in terms of the project objectives. Further, algorithmic transparency is also key in terms of understanding how the platform works during testing. It is important to enable persons testing the platform to see how the platform works and make suggestions about how it can be improved, it is also key for technical partners so that they can understand how to make improvements to the platform.

With respect to diversity, test datasets should contain varied data so that the way in which the ROXANNE tools analyse different data can be understood. In terms of the parties who are testing the platform, it is important that partners involve people from a wide range of backgrounds to collect different views and perspectives as this can lead to improvements in the tools that might not have been noticed by people with similar backgrounds.

Individual wellbeing is clearly relevant during testing as this is the point where participants will partake in using the platform and providing their views on it in interviews/workshops. This relates to research ethics, as discussed above. However, other issues can arise with artificial intelligence technologies where there users begin to see human traits in them and anthropomorphise them. For example, DeepMind's AlphaZero system that plays the board game Go has been described with human qualities of 'insight' and having a 'breed of intellect'.¹² This system is of course, simply a machine. Anthropomorphising machines can create an emotional connection between the user and machine which is illusory and therefore it would be unethical for the ROXANNE partners to allow this to happen. The ROXANNE partners do not, and should not, refer to the platform in human terms; this should also be made clear in the training provision.

With regard to accountability, the discussions and decisions about how the project tools will be tested have been, and should continue to be deliberated among the consortium and where possible with the involvement of a wider stakeholder group (i.e. Stakeholder Board, Ethics Boards). This should include collective validation of decisions following assessment of associated risk and benefits. Furthermore, the corresponding deliverables will document the arguments and reasoning behind the decisions made by the consortium.

6. PHASE 5: EVALUATION

The evaluation phase will be focussed on acquiring and understanding feedback from field-test participants. Human agency and liberty is enabled where participants/partners are able to contribute to a feedback process. Dignity is respected where all responses to the feedback process are treated fairly and equally.

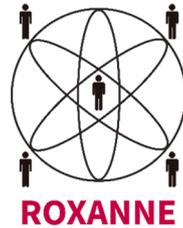
A fundamental issue in the development of computer models is that they will include assumptions made about the data and how the model should use these data. This will affect the accuracy, reliability, and precisions of the models. Whatever the choices and assumptions made when creating the models, these will affect the outputs of the algorithms.¹³ To some degree, this is inevitable; people will always make different assessments. Yet, there is a particular risk in relation to technical partners creating models for use by LEAs as technical partners might bring inaccurate assumptions to their work.

Following the requirements gathered in Phase 1 will go some way to dealing with this issue. However, as the needs of the models will change as the project progresses, additional measures should be taken. Ideally, models would be reviewed by external researchers.¹⁴ However, this would not be possible in the context of ROXANNE due to the potential for commercial exploitation in the future, and the risks of criminal organisations viewing any publicly available code. As such, having technical partners to review the work of others and discuss the needs of various models within the consortium would be beneficial

A key privacy issue in the evaluation phase is that persons who test the ROXANNE platform will be interviewed. As mentioned above, where interviews follow research ethics principles, it is likely that ethics risks would be avoided.¹⁵ To ensure data quality and data integrity, partners should incorporate best practices when formulating interview questions and methods so that the data collected is relevant, accurate, complete, and reliable. Additionally, participants should have control over their data and it should only be used where, and for as long as, the participants consent.

In terms of transparency in this phase, the consortium should disseminate information about the platform and project progress, including any setbacks or issues with the platform or project. This will allow the scientific and research community to better understand the results of the project, and learn from any challenges that have been overcome.

At this stage, the technical partners should carefully evaluate the extent of the algorithmic transparency and judge whether it is sufficient. The need is to ensure that the functioning of algorithms, at least at a basic level, is clear to the LEAs. Further, it will be difficult for partners to evaluate the system if they do not adequately understand how it works. This is particularly relevant to the ROXANNE project, where different technological modules are being put together in order to build an integrated platform. Consequently, the technical partners should build the platform in such



a way as to enable comprehension of how the data-processing modules and operations work both individually, and in combination, in order that they can adequately evaluate the platform. Based on the analysis of field-test results and the evaluation of continuous testing results, the consortium should be in a position to enhance the ROXANNE platform's operation and present end-users with a robust and reliable product that responds to their needs. To this end, partners must treat equally, impartially and openly all the results and feedback received from partners and external stakeholders.

With regard to accountability, the consortium is collectively responsible to build a technically robust, legally compliant and efficient ROXANNE platform with acceptable errors rates. Designated WP leaders are in charge of leading efforts to this end, with the support of other partners, and should work toward minimised levels of false positive/negative results, human errors, algorithm bias and malicious interference with the results. Evaluating these results should be properly documented, and shared with stakeholders to validate the partner's interpretation and the sound functioning of the platform.

7. PHASE 6: USE

As the ROXANNE platform is still in development, it is not possible to give a detailed assessment of how the platform will be used and the implications for its use on individuals and society. However, we can highlight some issues that might affect the use of the proposed platform. These will be further refined in future work.

In discussing human agency, liberty, and dignity in the context of criminal network analysis technologies being used, there are two key areas for discussion. The first is LEA officer themselves: how does the use of advanced technology affect them and their role? The second area is the people who are subject to (or potentially subject to) analysis via these means, how does it impact on their ability to live their lives and act freely?

A major issue in relation to human agency and advanced technologies is automation bias. This is the process whereby human beings trust the outputs of machines more than themselves and so follow what the machine suggests, even when it goes against their own knowledge.¹⁶ A classic example is where drivers follow the instructions of their satellite navigation system into a dangerous situation.¹⁷ Automation bias causes significant issues for human agency as the person making decisions is not acting with true agency, but is essentially acting as the agent of the machine. The implication of this is that, if affected by automation bias, the users of ROXANNE would be removing their own critical reflection about the outputs of the machine and the machine would, functionally, be acting as an autonomous agent itself. This is particularly concerning in the use of ROXANNE where, for example, an individual could be communicating with a criminal for innocent reasons but be included in the network analysis and potentially be included an intrusive investigation without an LEA officer taking a meaningful decision about whether they should be included.



The Law Enforcement Directive places a general prohibition on the use of automated individual decision-making.¹⁸ However, Member State law can provide an exemption to this as long as there are appropriate safeguards and, at the very least, the possibility for human intervention.¹⁹ Further, the use of automated decision-making on special category data is prohibited.²⁰ Still, the use of data-analysis technologies can still lead to automation bias even when they are not used for decision-making. For example, an LEA officer who might blindly follow an assistance tool as if it were making decisions rather than providing assistance; consider that a network analysis tool might potentially highlight persons of interest as an assistance tool, but, if this advice is blindly followed due to automation bias, then the tool is functionally being used for decision-making.

For the human being using the machine, this can result in them being 'alienated' from their work.²¹ For LEA officers, this could lead to de-skilling, possibly resulting in: a subsequent loss of their 'intuition' about what is suspicious in criminal investigations; an atrophy of moral skills in deciding what the right course of action is in relation to law enforcement operations;²² officers becoming less able to determine when a technological system has made a mistake, and assess the outputs of a system.

For the people whose data is being analysed by an end-user suffering the effects of automation bias, this can result in them being treated as a mere data-point rather than a person. This dehumanisation would be an affront to dignity in the Kantian sense where treating individuals as mere objects violates their dignity.²³

One solution to dealing with issues of automation bias, the alienation of end-users, and dehumanisation of data-subjects is to keep end-users at the very centre of critical decision-making and making them aware of issues that could be generated by their data-analysis and its effects. To ensure human-centric operation of ROXANNE, an electronic decision-making component is currently in development. This asks end-users a series of questions when beginning a new case with the ROXANNE platform, when uploading new data, and when analysis has been completed.

This is structured so that LEA officers can see a series of questions, and the rationale behind them, and should give meaningful answers to justify the decisions they make. This facilitates deeper engagement with the issues by officers and a better understanding of the implications of their work using tools such as those incorporated into ROXANNE. The need for human beings to engage directly with decision-making should also be made clear to end-users during training.

Another potential impact of ROXANNE on the individuals who will be subject to criminal network analysis is the potential for its use to interfere with negative liberty (meaning the freedom to do things without interference), in particular political freedoms. Surveillance technologies that already exist are known to create 'chilling effects',²⁴ meaning that people who know or suspect they will be surveilled act in more restrained ways to what they would otherwise do if they were not subject to surveillance.

A potential effect of the ROXANNE platform is that the impact of criminal network analysis does



not just affect the individuals whose data is analysed, but also that of the people whom they communicate with. Thus, where people perceive their friends and family to be at risk of surveillance, or data-analysis by LEAs using ROXANNE, due to their actions, they would likely be less inclined to participate in activities that could raise the interest of LEAs. This is especially concerning in countries that suppress political opposition and the deterrent effect interferes with their perfectly innocent political activities and freedom of expression. To deal with this issue, the ROXANNE exploitation plan will detail that ROXANNE results should not be exploited to countries that are non-democratic, currently experiencing conflict, or have a poor track record of human rights compliance.

Issues of robustness and safety are significant in terms of use, as it is the crucial moments where effects might be created for the public if ethical risks manifest. In terms of safety and security, there are risks that an insecure platform could be attacked by criminals; if they were to gain access to the platform this might not only disrupt investigations, but also ongoing criminal trials and, potentially, previously secured convictions through evidence tampering. As such, it is imperative that LEAs only use the ROXANNE platform on secure systems.

However, other aspects related to the accuracy, reliability, and precision of the system, and how this is perceived, can generate significant ethical risks during the use of the platform. A significant issue with the use of algorithmic processes to assess real life activities is the fact that it can only comprehend quantitative methods. As Malik writes, we may be able to engage in many mathematical analyses of different situations or individuals, but these types of analysis do not lend themselves to understanding what they mean.²⁵ The ROXANNE platform might highlight a particular activity as unusual but cannot explain why. There is, therefore, a risk that an end-user could misunderstand the outputs of the platform and view an innocent activity as worthy of further investigation, thereby unnecessarily intruding on an innocent individual. Therefore, the ROXANNE technical partners need to work on the explainability of the tools and their results.

Linked to this is a need for end-users to understand the context in which data is gathered and what it means.²⁶ ROXANNE partners should be how their tools and how this could affect the processing of investigative data by end-users. For example, considering how people come to be included in, or excluded from, police datasets, and what social dynamics might affect this. In technical terms, this is also an issue that affects the generalisability of any analytic model. Project partners should explain the appropriate contexts for using the ROXANNE tools to potential customers, and end-users should consider how the data that was used to build the ROXANNE tools could affect data that is likely to be gathered in their circumstances.

In terms of use, privacy of data-subjects is clearly a major potential issue. End-users should ensure that any data they process using ROXANNE tools was gathered and processed lawfully, on secure systems, and by trained operators. It is important to note also that data quality can affect results of data analysis; for example, data from a low-resolution surveillance camera might be of limited utility if the ROXANNE tools do not work well on low-quality data. This would clearly raise an issue in terms of accuracy, completeness, and reliability of data. Thus, technical partners should determine an



appropriate minimum standard for which data is acceptable to be used with the ROXANNE platform. Further, LEA officers should be cognisant of the potential effects that poor-quality data could have for the results when they review them.

Logging the use of ROXANNE provides clear links with accountability, as it identifies individuals who can be held responsible for the actions that take place using ROXANNE. Further, this can also add to a culture of respecting privacy in organisations; where individuals know they must explain their actions regarding personal data, this would, presumably, cause people to consider whether they need to engage in the activity in question. Logging the actions of end-users who used the platform is incorporated into the electronic decision-making mechanism that is outlined above.

Transparency is a particular issue in today's technology-filled world. Hence, it is important to focus on people's knowledge of and familiarity with uses of their data, and in, the case of ROXANNE, how their data could be analysed. Whilst many of the tools and practices used by LEAs are lawful, there are risks of misuse which can result in distrust between the public and law enforcement.²⁷ Hence, partners and end-users should disseminate information about how the platform works, how it could be used, and what data could be processed about people. This should ensure that citizens are informed about how their data could be used if it were to be included in an ongoing investigation. However, it is important to note that such dissemination might need to keep specific details confidential in order to avoid providing criminals information that might allow them to 'game' the technology.

Algorithmic transparency is important in the use of data-analysing platforms, especially in law enforcement. As LEA activities are inextricably linked to the criminal justice system, LEA use of these technologies is subject to court scrutiny. If the ROXANNE platform is a 'black box', and its functioning is unknowable, this can pose a serious risk to due process, and accountability. Indeed, some authors suggest that 'black box' algorithms should be prohibited in the criminal justice process.²⁸ As mentioned above, it is important that partners create tools that can be understood not only by the scientific community and end-users, but also members of the criminal justice system and, potentially, lay people participating in juris at criminal trials.

With regard to potential issues of bias, the use of biased training data can create algorithms that process data in an unequal way, potentially leading to discriminatory effects in the real world.²⁹ As mentioned above, ROXANNE partners should work toward minimising risks of bias in the ROXANNE tools. However, it is also clear that end-users are products of their societies, and biases in those societies can, therefore, affect the work of LEAs.³⁰ Babuta and Oswald note that specific guidelines for operations using data-driven technologies should be provided to investigators, and these should complement existing professional practice and approaches, to help reduce biases that could be brought in by end-users. The electronic decision-making mechanism mentioned above asks questions of end-users about potential data and analysis tools that could, potentially, affect the results of the ROXANNE tools and their actions in an effort to do this.



In terms of individual and societal wellbeing, the potential for ROXANNE to not only analyse lots of data about specific persons but also their connections to other people could provide significant insight into their lives and relationships. For Foucault, this can go beyond simply knowing what illegal activities a criminal might engage in, but to reveal insights as to who they are as a person.³¹ As such, the combination of analysis tools creates a potential risk that there could be a greater invasion of privacy that that associated with the technologies being used individually.

Whether this can be ethically justified depends upon a proportionality assessment. Therefore, end-users should consider whether their proposed methods of data analysis are necessary and whether they could complete their tasks using less tools. The ROXANNE partners can facilitate this by building the ROXANNE platform in such a way that end-users can use tools individually rather than needing to process data through all components in a processing chain.

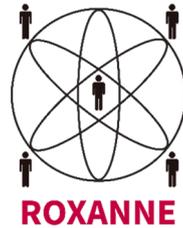
With respect to environmental wellbeing, the ROXANNE platform could use a lot of energy. ROXANNE partners should dedicate some time to considering if the platform could be made more energy efficient, and, ideally, if energy re-use might be possible (for example, water-cooled systems can repurpose heated water for other uses.³²)

Any use of the operational ROXANNE platform by end-users holds them accountable for the tool's use in accordance with applicable national legislation and/or organisational code of ethics. The technology integrated oversight and access control mechanisms should help ensure compliance and deter potential abuses by authorised system users, which can be detected through log verification (i.e. purpose of search, user details,). This logging system allows management to monitor the platform and ensure that it is used in a compliant manner. Additionally, training users on good practices prior to their first interaction with the ROXANNE technology is recommend for a sound understanding of the platform's functioning, accurate interpretation of results, and to remind users of associated ethical considerations.



REFERENCE

- 1 This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635
- 2 Philip Brey, Björn Lundgren, Kevin Macnish, and Mark Ryan, 'D3.2 Guidelines for the development and use of SIS', SHERPA project, 2019, p.10 (hereafter: 'SHERPA Guidelines'). Available at: https://dmu.figshare.com/articles/D3_2_Guidelines_for_the_development_and_the_use_of_SIS/11316833
- 3 SHERPA Guidelines, p.1
- 4 For more information, see David Wright, David Barnard-Wills, Leanne Cochrane, Krzysztof Garstka, Agata Gurzawska, Joshua Hughes, Richa Kumar, Anaïs Ressaygues, and Zuzanna Warso, "The TRI touchpoint table: A new methodology for ethical, privacy and societal impact assessment of projects" (unpublished)
- 5 SHERPA Guidelines, p.21. In the SHERPA guidelines, these phases are used to generally describe an 'Agile' methodology, as oppose to a CRISP-DM product development methodology that is the main focus of the guidelines; ethics partners determined that, as they are quite general, these phases would be more applicable to the ROXANNE project.
- 6 European Commission, European Textbook on Research Ethics, Directorate-General for Research Science, Economy and Society, Brussels, 2010, 35-47.
- 7 Art.6(1)(a), GDPR.
- 8 See, for example, European Commission, European Textbook on Research Ethics, Directorate-General for Research Science, Economy and Society, Brussels, 2010.
- 9 See, for example, Benjamin, Ruha, Race After Technology, Polity Press, Cambridge 2019, p.165; Valentine, Sarah, "Impoverished Algorithms: Misguided Governments, Flawed Technologies, and Social Control" Fordham Urban Law Journal, Vol.46, No.2, pp.364-427, pp.370-37; D'Ignazio, Catherine, and Lauren F. Klein, Data Feminism, MIT Press, Cambridge, Massachusetts, 2020, p.123.
- 10 Closed cases should be distinguished from cold cases. A closed case is where the investigation has finished with the legal process having been concluded; a cold case is an ongoing case where investigative leads have dried up. The ROXANNE partners will not seek to access data from real cases, whether they are ongoing or cold.



11 European Parliament and Council, Directive (EU) 2016/680 of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, OJ L 119/89, Vol.59, 4 May 2016 (Law Enforcement Directive, hereafter: LED)

12 Strogatz, Steven, "One Giant Step for a Chess-Playing Machine", The New York Times, 26 Dec 2018. Available at: <https://www.nytimes.com/2018/12/26/science/chess-artificial-intelligence.html>

13 Silberzahn, R., "Many Analysts, One Data Set: Making Transparent How Variations in Analytic Choices Affect Results" *Advances in Methods and Practices in Psychological Science*, Vol.1, No.3, 2018 (hereafter: Silberzhan, 2018), pp.337-356.

14 Silberzahn, 2018, pp.337-356, 353-354.

15 See, for example, reference to interview standards in European Commission, *Ethics in Social Science and Humanities, Horizon 2020 Guidance*, October 2018. Available at: https://ec.europa.eu/research/participants/data/ref/h2020/other/hi/h2020_ethics-soc-science-humanities_en.pdf

16 See Skitka, Linda J., Kathleen L. Mosier and Mark Burdick, "Does Automation Bias Decision-Making?" *International Journal of Human-Computer Studies*, Vol.51, 1999, p.991.

17 See, for example, Milner, Greg, "Death by GPS" *Ars Technica*, 2016. Available at: <https://web.archive.org/web/20190602041744/https://arstechnica.com/cars/2016/05/death-by-gps/>

18 Art.11 (1), European Parliament and Council, Directive (EU) 2016/680 of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, OJ L 119/89, Vol.59, 4 May 2016 (Law Enforcement Directive, hereafter: LED).

19 Art.11 (1), LED.

20 Art.11 (2), LED.

21 Automation meaning the use of automatic process to replace human cognition, in contrast to mechanisation replacing physical labour. Pyke, Magnus, *Automation: Its purpose and future*, Scientific Book Club, London, 1946, p.38.

22 Brownsword, Roger, "In the year 2061: from law to technological management" *Law, Innovation and Technology*, Vol.7, No.1, 2015, pp.1-51, 35.



- 23 Kant, Immanuel (trans. H.J. Patton), *The Moral Law*, Hutchinson's University Library, London, 1947, pp.96-97.
- 24 Solove, 2006, 477-560, p.487.
- 25 Malik, Momin M., "A Hierarchy of Limitations in Machine Learning" arXiv.org, 29 February 2020 (hereafter: Malik, 2020). Available at: <https://arxiv.org/abs/2002.05193>, pp.6-8.
- 26 Borgman, C., *Big Data, Little Data, No Data: Scholarship in the Networked World*, Cambridge MA., MIT Press, 2015.
- 27 Douglas Heaven, Will, "Predictive policing algorithms are racist. They need to be dismantled", *Technology Review*, 17 July 2020. Available at:
- 28 Selbst, Andrew, and Solon Barocas (eds.), *AI Now 2017 Report*, AI Now Institute, 2017. Available at: https://ainowinstitute.org/AI_Now_2017_Report.pdf
- 29 See, for example, Benjamin, Ruha, *Race After Technology*, Polity Press, Cambridge 2019, p.165; Valentine, Sarah, "Impoverished Algorithms: Misguided Governments, Flawed Technologies, and Social Control" *Fordham Urban Law Journal*, Vol.46,, No.2, pp.364-427, pp.370-371
- 30 Williams, Patrick, and Eric Kind, *Data-driven policing: The Hardwiring of Discriminatory Policing Practices Across Europe*, European Network Against Racism, 2019. Available at: <https://www.enar-eu.org/IMG/pdf/data-driven-profiling-web-final.pdf>; Chowdhury, Areeq, *Unmasking Facial Recognition*, WebRoots Democracy, 2020. Available at: <https://webrootsdemocracy.files.wordpress.com/2020/08/unmasking-facial-recognition-webroots-democracy.pdf>, pp.8-9.
- 31 Foucault, 1995, p.208; Stoddart, Eric, "A Surveillance of Care: Evaluating Surveillance Ethically" in Kirstie Ball, Kevin D. Haggarty, and David Lyon (eds.) *Routledge Handbook of Surveillance Studies*, Routledge, London, 2012, pp.369-376, 372.
- 32 Jones, Nicola, "How to stop data centres from gobbling up the world's electricity" *Nature*, 2018. Available at: <https://www.nature.com/articles/d41586-018-06610-y>