ROXANNE

Real time network, text, and speaker analytics for combating organized crime

Ethical Approach to ROXANNE

Dr. Joshua Hughes (TRI)
Inputs from INTERPOL, CAPGEMINI, KEMEA, AIRBUS, all partners

This project has received funding from the European Union's Horizon 2020 Work Programme for research and innovation 2018-2020, under grant agreement n°833635

Why should we include ethical, legal, and privacy work in LEA projects?

- Investigative technologies invade privacy to expose hidden information about suspected criminals
- The impacts of AI tools can be harmful to people (e.g., automation bias, discriminatory effects, etc.)
- Impacts and harms need to be understood so they can be dealt with, mitigated, or avoided. Plus, opportunities can be seized.
- Contribute to understanding emerging issues that can affect technology use, especially in relation to regulation (e.g., AI Regulation).
- Contribution to making ROXANNE technologies the most ethical, privacy-aware, legally-compliant, and socially-acceptable versions they could be.

ROXANNE

High-level ethical approach to ROXANNE

Researching emerging ethics impacts and determining their applicability to ROXANNE

Highlighting ethical, legal, and societal impacts presented in ROXANNE, and sensitising the consortium to them

Assessing the extent and implications of impacts, and finding solutions

Incorporating Privacy-by-Design and Ethics-by-Design into technologies for law enforcement.

Offering design support to technical partners

ROXANNE

Research Ethics: TRI's TouchPoint Table™

- A key part of ethical AI development is ensuring that research is conducted ethically.
- Each task in the project was analysed for research ethics risks.
- Risks and mitigations were discussed with WP leaders, and agreed approaches were implemented.
- Any high-risk tasks, uses of human participants, or uses of sensitive personal data were monitored throughout the project.

Tasks	Task descriptions	Potential Ethical issues	Addressing these issues	Assessment of remaining risk Low/ Medium/ High
Tx.x	…	…	…	…

ROXANNE

Ethics Oversight structure

EC Ethics Reviewers and project reviewers

External Ethics Board

Partner data protection officers

Data Protection Working Group

Internal Ethics Board

USAAR ERB – reviewing uses of human participants

USAAR KEF – Security Review Commission

Data Management Plan – monitoring of ethical and lawful data use

WP3 – in-depth ethical, legal, and societal impact analysis and recommendations, Privacy and Ethics-by-Design

WP10 – EC imposed ethics requirements

ROXANNE

Ethical, Legal, Societal analysis: Impact Assessment

- Literature review to better understand issues relevant to ROXANNE technologies.
- Engaged in a comprehensive ethical, societal, fundamental rights, and applicable legislation impact assessment.
- Developed worst-case scenarios to expose additional issues for fundamental rights and societal values analyses.
- Developed checklist for compliance with applicable legislation.
- Shared briefing paper on societal values.
- Developed recommendations for making the ROXANNE project and technologies the most ethical, privacy-respecting, legally-compliant, and socially acceptable versions they could be.
- Shared analyses and worst-case scenarios with stakeholders and public for feedback. Results incorporated into later analysis.

ROXANNE

Ethics: High-level Requirements
1. Human agency, liberty, and dignity;
2..Technical robustness and safety;
3.Privacy and data governance;
4.Individual, societal and environmental well-being;
5. Transparency;
6. Diversity, non-discrimination and fairness;
7. Accountability

Phases
1. Requirement gathering;
2. Planning and designing;
3. Development;
4. Testing;
5. Evaluation;
6. Use.

Societal Values	
1. Citizen's Privacy;	6. Equality and Tolerance;
2. Trust and the perception of security;	7. Human Rights
3. Unintended consequences;	8. Respect for Human Life
4. Social Acceptability;	9. Rule of Law
5. Democracy and Solidarity;	

Key Fundamental Rights	
1. Human dignity;	7. Freedom of Expression and Information;
2. Right to the integrity of the person;	8. Freedom of assembly and association;
3. Prohibition of torture and inhuman or degrading treatment or punishment;	9. Non-discrimination +;
4. Right to liberty and security;	10. Right to an effective remedy and a fair trial;
5. Respect for private and family life;	11. Presumption of innocence and right of defence
6. Protection of personal data;	

Applicable Legislation
GDPR
Law Enforcement Directive
INTERPOL Rule of Processing of Data
CoE Convention 108+
Copyright Directive
Network and Information Security Directive

Selected Provisions
Lawfulness of Data Processing
Special Categories of Data
Data Processing Principles
Individual Rights
Accountability and Transparency
Data Security
Data Storage & Retention
Data Transfer

Inc. BizHR and comparative approaches

ROXANNE

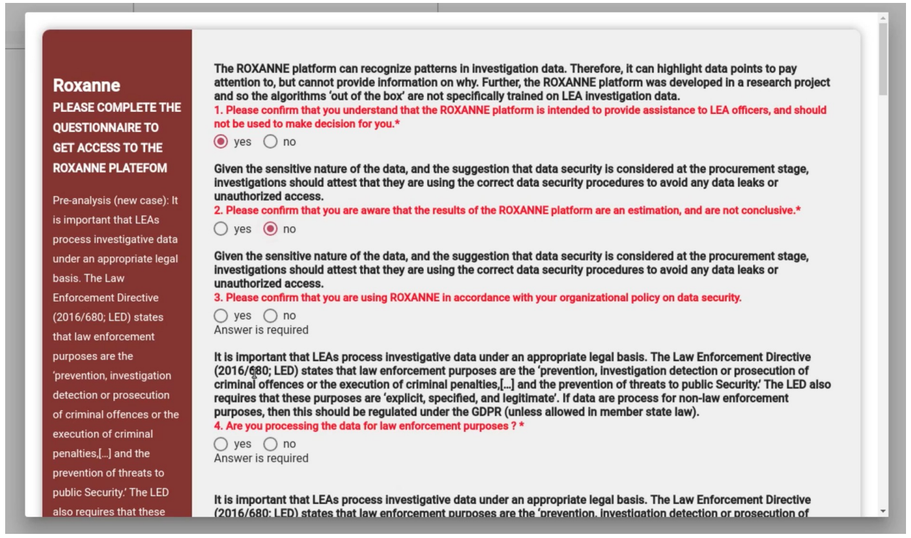
Ethics analysis: Applying recommendations

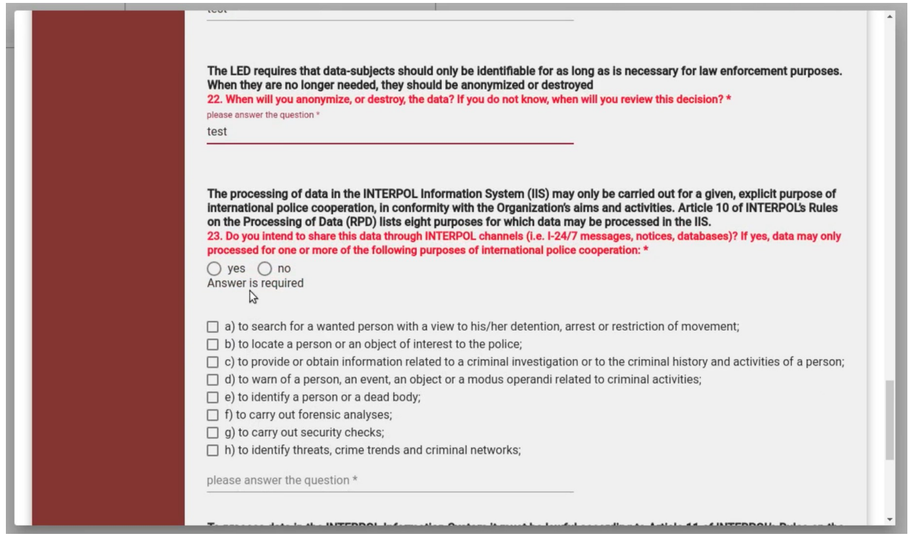
- 180+ ethics recommendations for across project, and beyond. Combined into 59 more easily implementable recommendations.
- Prioritised using a MoSCoW methodology
- Separated into specific requirements for direct implementation, and general requirements for implementation as needed

ELS issues	Implementation status
Data Protection and Privacy	17/17 recommendations completed
Transparency	6/7 recommendations completed – 1 ‘should have’ remaining
Responsible Research	6/7 recommendations completed – 1 ‘could have’ remaining
Avoid biases	7/7 recommendations completed
Environmental Concerns	2/2 recommendations completed
Technical Concerns	8/8 recommendations completed
Accountability mechanisms	2/2 recommendations completed
Training	4/4 recommendations completed
Exploitation	4/4 recommendations completed
Platform use	11 recommendations for use of Autocrime after project

Ethics tools: Decision-support tool

- Series of questionnaires for LEA senior officers, and investigators to assess ethical, legal, and societal impacts of their uses of ROXANNE-like technologies. Rationale for each question provided so that the reasoning is understandable.
- Enables LEA officers to meet most of the post-project requirements
- Separate questionnaires for:
  - Procurement of tools like ROXANNE
  - Beginning a new case
  - Including new data in an investigation
  - Assessing the results of analysis





ROXANNE

Applied legal research: Examining use of INTERPOL infrastructure

- Analysis of the potential to use INTERPOL's global communications infrastructure and data storage mechanisms for rapid sharing of electronic evidence.

- Conclusion that INTEPROL channels would good avenues for sharing ROXANNE outputs for cross-border cases.

INTERPOL Tools and Services
I-24/7 secure communications network
Notices
Databases

Use-Case
Child exploitation ICSE database

Analysis of Applicable Legal Requirements
Data Protection principles
Data management
Oversight mechanisms

ROXANNE

Global legal survey

Legal survey of INTERPOL National Central Bureaus. Responses from Europe, Asia and South Pacific, and Americas

Key results:

23/39 respondents countries do not have specific legislation for advanced data analysis tools for biometric evidence.

49% of respondents do not have any known process for LEAs to have ethical oversight of data analysis technologies.

14/39 respondents do not have ‘good practice’ standards for using biometric data

There is a wide variety of people needed to give authorisations for biometric data analysis across different countries (ranging from senior investigators, to police commissioners, to data protection officers)

22/39 respondent countries require biometric data analysis to be presented in court as expert evidence

Web scraping

Web scraping very difficult for research projects from a data protection perspective – few data–subjects would expect their data to be scraped and analysed.

Transparency requirements would require researchers to notify website users about scraping activities.

Where websites are of interest to LEAs, fulfilling transparency requirements could alert suspected criminals.

Partners opted to develop ROXHOOD dataset: fictitious, but realistic, dataset of messages sent via a web forum.

Facial Similarity Searching

Developing facial analytics very controversial issue: people cannot easily control their exposure to these technologies without drawing unwanted attention. But, clear utility for LEAs in investigations.

In ROXANNE, facial analytics are and ‘add-on’ to object and location analytics in video technologies.

Co-design between ethics and technical partners. Many options discussed for privacy protection, but some privacy-preserving limitations can remove any usefulness of the technologies.

Chosen privacy/ethics protections:

- All uses are logged for accountability purposes.
- Number of potential queries set to a low-number by default, but could be increased with (logged) authorisation from senior officer.
- Only pre-tagged persons of interest can be searched for.

Analytics tools posing a risk of bias are not applied by default, but available as optional filters.

ROXANNE

Bias, Transparency, and Robustness

Biases are present where technologies evaluate individuals from some groups in a different way to those from other groups. Where unmitigated, this can lead to discriminatory impacts.

Transparency is important so that end-users can adequately understand the tools they are using, what they can do, what they are intended to do, and what they cannot do.

Robustness is key in LEA technologies, as a tool that is not optimised for accuracy, precision, and reliability is not a tool that should be used in high-risk situations like LEA investigations.

Some issues to be expected with research results, rather than production-grade technologies.

Bias

Tool	Data evaluated for bias?	Tool evaluated for bias?	Steps taken to avoid bias?	Training input on end-users dealing with remaining biases?
Speech technologies	Yes, technical biases expected	Yes, but limited test data available	Additional data to be included in future.	Yes, to explain potential issues
Text technologies	Yes, technical biases expected. Limited impact due to sentence structures.	Yes, initial validations completed	Entities can be boosted, gender-queer pronouns included.	Yes, to explain potential issues
Geo-location tool	N/A, no machine learning	Yes, non expected	N/A	Yes, to avoid misinterpretation
Network analysis tools	N/A, tools do not analyse personnal characteristics	N/A	N/A	N/A
Video processing tool	Yes, diverse data used in model	To be done, test data selected.	Yes, biased algorithms not applied by default	Yes, end-users can request additional results
Visualisation tool	N/A	N/A	N/A	N/A
Case Management System	N/A	N/A	N/A	N/A

Transparency

Tool	Logging?	Understandable to end-users?	How is conspicious activity dealt with?	Adequate training manual?
Speech technologies	Yes	Yes, to investigators/experts depending on specific tool	By end-user	Some manuals to be updated.
Text technologies	Mostly yes	Yes, to investigators	By end-user	Some manuals to be updated.
Geo-location tool	Yes	Yes, to experts	By end-user	Yes
Network analysis tools	Yes, on platform level	Yes, to investigators	By end-user	Yes
Video processing tool	Yes	Yes, to investigators	By end-user	Manual to be updated.
Visualisation tool	Yes	Yes, to experts	N/A	Yes
Case Management System	N/A	Yes, to experts	N/A	Yes

Robustness

Tool	Tool optimised for use-case?	False Negative and False Positive risks assessed?	Minimum data quality?	Interim human oversight needed?
Speech technologies	Yes	Generally low risk, FP favoured. Some outputs to be calibrated.	Yes	Not needed
Text technologies	Yes, and can be improved in future	No risks for transcript outputs. Low risk for analysis outputs	No, though noisy data = worse results	Yes, for assessing noisy data, optimise follow-on tasks, and approve connections in mention network.
Geo-location tool	Generalised, but can be calibrated to the use-case	No risk, outputs calibrated probability	Yes	Not needed
Network analysis tools	Yes, algorithms can be tuned for use-cases	Low risk	Yes, larger networks = greater accuracy	Not needed
Video processing tool	Yes, algorithms can be tuned for use-cases	Low risk, analysis focusses on clearest entities	No, as low-resolution entities could still be interesting	Not needed
Visualisation tool	N/A	N/A	N/A	N/A
Case Management System	Yes	N/A	Yes, data needs to follow pre-defined fields	Not needed

Analysis of anonymisation of LEA data

Examined novel data minimisation techniques used by LEA partner to determine if they met the GDPR standard of anonymisation. Full anonymisation required to allow data sharing.

GDPR standard is a very high-bar to reach (‘reasonably likely’ test applicable any data controller using all objective means)

Anonymisation techniques analysed in context so as to give more useful conclusions considering state-of-the-art

Anonymisation techniques determined to be successful, and so resulting data could be shared with partners.

Applied the ‘Anonymisation Decision-Making Framework’ to deal with resulting privacy risks.

- Some ethical perspectives argue that there is no such thing as anonymous data. So, we still needed to ensure ethical treatment of data.
- Even with anonymous data at the legal threshold, privacy risks are never zero

Knowledge gained has been used in standardisation activities (ISO/IEC FDIS 27559)

Ethics Training

- Training provision so that end-users can be aware of ethics issues relevant to the use of the Autocrime platform

Ethics:
High-level ethics requirements
AI Regulation

Legal:
Legal framework
Law Enforcement Directive
Human rights & considerations for when they might be impacted
Court-ready evidence
Good practices
Governance of AI tools in investigations

Specific considerations:
Web scraper
Speech analysis
Text analysis
Video analysis
Network analysis
Considerations about the integration of technologies

‘Know You Customer’ Exploitation Risk Assessment

- The ROXANNE technologies pose risks if they fall into the wrong hands.
- Analysed risks of bad actors acquiring and using ROXANNE technologies in terms of threats, vulnerabilities, and consequences.
- Draws from ‘know your customer’ approaches from risk-based approaches to providing financial services.
- Has questions aimed at the potential end-user, assesses legal and ethical risks, and has the technology provider reflect on how they (and their colleagues) feel about the end-user acquiring the technologies in question.
- Will be available to exploitation partners.

Risk Assessment Question Areas	
Organisation	Information security and data protection
Location	Misuse and mass surveillance
Use of technologies	External monitoring [optional]
Regulation of technologies	Political oversight and outside influence
Onward provision of technologies	Bringing risks home

ROXANNE

Example risk table

Organisation		
Questions to end-users: 1. What type of organisation do you represent? LEA, private security, university, research organisation, etc.? a. Please give a brief summary of your organisation, its history and what it does. Please mention if your organisation is one arm of a larger corporate group, or similar.		
Questions to the technology provider: Has this customer recently changed names or only recently been constituted as a company? If the organisation has done unethical things in their past, have they done enough to distance, or redeem, themselves from this? Is this the type of organisation that we want to be involved with?		
Consideration	**Details**	**Score**
Risk/failure mode		N/A
Potential risk impact and effects		N/A
Severity of risk		x/10
Likelihood of risk		x/10
Safeguarding measures		x/10
Overall risk		x/30
Actions/recommendation		

ROXANNE

Policy recommendations

LEAs intending to use AI tools for operations should have an ethics board, or ethics adviser(s).

Where AI tools are made available to LEA officer for investigative use, training should include considerations of ethical issues.

INTERPOL channels can be used for transmission of investigative data, and outputs of tools like those in ROXANNE.

The EC ethics check process should be more open and cooperative.

The GDPR should be the default data protection regime for research activities, rather than the Law Enforcement Directive.

Thank you!

Any questions?