ROXANNE

ROXANNE real time netwOrk, teXt and speaker ANalytics for combating orgaNized crimE

Technologies to enhance the fight against crime and terrorism

Project results

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 „Disclaimer: the document reflects only the author's views and the European Commission is not responsible for any use that may be made of the information contained therein."

What's inside

Letter from the Project Coordinator	2
About the project	3
Privacy and Ethics in ROXANNE	4
Law Enforcement Agencies and ROXANNE project	5
ROXANNE Field Tests	6–7
Autocrime platform	8–9
ROXSD: ROXANNE Simulated Dataset	10
Speech, text and video in ROXANNE	11–12
Network analysis in ROXANNE	13–14
ROXANNE Training Platform	15
Collaboration with sister projects	16
LEA Cluster	16

Letter from the Project Coordinator

Dear Readers,

The ROXANNE project has been envisaged with the objective to provide law enforcement agencies (LEAs) with a platform integrating advanced speech, text, video, and network analysis technologies that could be used lawfully to uncover criminal networks whilst minimising the investigation time and effort. The technical development was centred around an innovative platform with the intention not to replace humans but to automate time-consuming tasks, and to support LEA decision-making.

After 3 years of collaborative work, we can present the results. The main outcome is **the Autocrime platform** which, in comparison to other tools used by LEAs, stands out from the crowd as a unique solution to LEA problems and which will be made available to interested European LEAs free of charge. The second outcome that we are particularly proud to present is the **ROXSD** - a synthetic, but highly-realistic, dataset of communication data in a fictional organized crime network, that will be made available to other researchers in FCT. Please find all results of the ROXANNE project elaborated in the next pages. I hope you will find it interesting.

Yours faithfully,
Dr. Petr Motlicek
Project Coordinator

About the project

ROXANNE (Real time network, text, and speaker analytics for combating organized crime) is an EU-funded collaborative research and innovation project, aiming to unmask criminal networks and their members as well as to reveal the true identity of perpetrators by combining the capabilities of speech/language technologies, visual analytics and network analysis.

ROXANNE collaborated with Law Enforcement Agencies (LEAs), industry and researchers to develop new tools to speed up investigative processes and support LEA decision-making. The end-product is the Autocrime platform, which uses new tools to uncover and track organized criminal networks, underpinned by a strong legal and ethical framework.

The project consortium comprises 25 European organizations from 15 countries, including 11 LEAs from different countries.

Coordinator

Academia and Research Centers

End users

Industry

Privacy and Ethics in ROXANNE

As part of H2020, **the ROXANNE project has taken ethics, legal compliance and privacy into account at all stages of its work.** This ensures that the work of the consortium and the ROXANNE platform is **compliant with applicable ethical and legal standards to maximum possible extent.** As a result, ROXANNE has been developed to support LEAs to gather evidence using the project tools in a court of law. **We have adopted a privacy-by-design and ethics-by-design approach.** This means that the project itself, and the envisaged ROXANNE platform, were subject to legal and ethics frameworks, in particular those related to data protection. **To ensure a successful ethics-by-design approach in ROXANNE, the consortium has been advised by both an Internal and External Ethics Board.** Project partners have collaborated with ethical and technical experts **to conduct research on the platform to make it ethical, privacy-aware, and legally optimal.** Such an approach reduces many of the issues raised by LEA data analysis during the design state and alleviates their impact during its use. ROXANNE partners have also considered several possible scenarios regarding exploitation and have developed a **risk assessment to minimize risks of the ROXANNE tools falling into the wrong hands.**

Law Enforcement Agencies and ROXANNE project

INTERPOL: "The impact of new technologies on criminal activity can no longer go unnoticed. LEAs must remain at the forefront of innovation to stay ahead of criminals. To enhance the use of innovative tools such as ROXANNE by law enforcement, INTERPOL is participating in research projects into their development based on end-users needs and legal implications. To ensure an effective and lawful implementation of these new police technologies, they require an appropriate legal framework based on the principle of the rule of law and fundamental human rights. ROXANNE is an example that emphasizes strengthening security while minimizing impacts on fundamental rights, such as privacy."

Hellenic Police: "The field tests prove that ROXANNE project can benefit LEAs by increasing accuracy and saving time in the investigations by providing the crucial information about the perpetrators' identity."

Ministry of Public Security - Israel National Police: "The ROXANNE vision is to provide a self-contained environment with AI investigative capabilities. It includes speech recognition, speaker recognition, criminal network analysis, named entity recognition from text, as well as face and scene recognition from video. All these capabilities are packed in a self-contained environment, easy to install, considering data sensitivity and supporting offline installation. The architecture is mostly open source, so internal adaptations can be easily handled and new trained models can be fetched and plugged into the system at any time."

Police of Czech Republic: "We see a lot of innovative ideas on this platform. We also recognize the improvements in speaker identification by adding text analysis technologies."

Romania's Ministry of Internal Affairs: "When joining the ROXANNE project, Romania's Ministry of Internal Affairs sought out new and innovative ways to fight organized crime.Perpetrators strive to find new ways to conduct their operations, they learn, and they adapt to the newest trends and opportunities. In order to be able to compete with them, we must do the same. That is where ROXANNE project comes in play, as it gathers academia, industry and law enforcement agencies under the same roof and allows them to share knowledge and new technologies. The ROXANNE project has capitalized a lot of aspects of AI technologies."

ROXANNE Field Tests

To ensure that the developed solutions are tested under realistic conditions and use cases, ROXANNE partners performed 3 Field Tests. During the events LEAs, as end users of the platform, could meet with ROXANNE technical partners, discuss and experiment with the platform functionalities during a hands-on session. After each Field Test, feedback from the participants was collected and all recommendations have been taken into account while developing the final solution.

During the 1st Field Test, the first version of ROXANNE platform was presented, accompanied by an online educational platform. The technologies included voice processing, text processing, and preliminary network analysis. The platform was supported by a forensic visualization toolkit. The participants reported that the platform could lead to creation of a new research system combining various modules. The platform could minimize time and efforts of LEAs allocated to investigative procedures that involved a large volume of data. The participants found the platform easy to understand, user-friendly, easy to navigate and intuitive even for people with a non-technical background. The main concerns were about the accuracy of speaker identification, speech analysis and language identification. Finally, for future updates, they recommended the capability to set keywords on topic detection by the users.

During the 2nd Field Test, multiple scenarios were presented. The technologies included were speaker identification, network mention, and co-reference network, named entity recognition (performed on manual and automatically obtained transcripts), improved automatic speech recognition, topic detection, link prediction, community detection and visualizations of the results. The participants stated that the functions given are widely applicable. According to them, the primary issue is that some functions are language dependent. Thus, some potential end users will not be able to use these features. They noted that they could incorporate ROXANNE's data into the investigation tools they already use. The main concerns in this field test were the visualization of GPS coordinates, if available, the separation of groups in network analysis and some add-ons to make the platform more interactive by providing quick and easy edits to the data.

During the 3rd Field Test, the platform's mature capabilities were presented alongside its performance and usefulness for solving criminal cases through streamlined data analysis. The technologies presented included speaker recognition and diarization, named entity recognition,

mention disambiguation and topic detection. In addition, video processing technologies, social influence analysis, outlier detection, community detection and link prediction were also presented. Moreover, police use case scenarios were demonstrated and a hands-on session was offered to the participants to experiment with the platform functionalities. The participants stated that the platform was fully understandable and stressed that they considered the functionalities presented being widely applicable and considered as innovative. At the same time, they indicated that more technologies would be of their interest if integrated in the platform and if it was available to process more languages. The participants also highlighted that ROXANNE platform is unique considering that it combines multiple technologies from multiple modalities and at the same time reduces time of the police investigations which would otherwise be done manually.

Autocrime platform

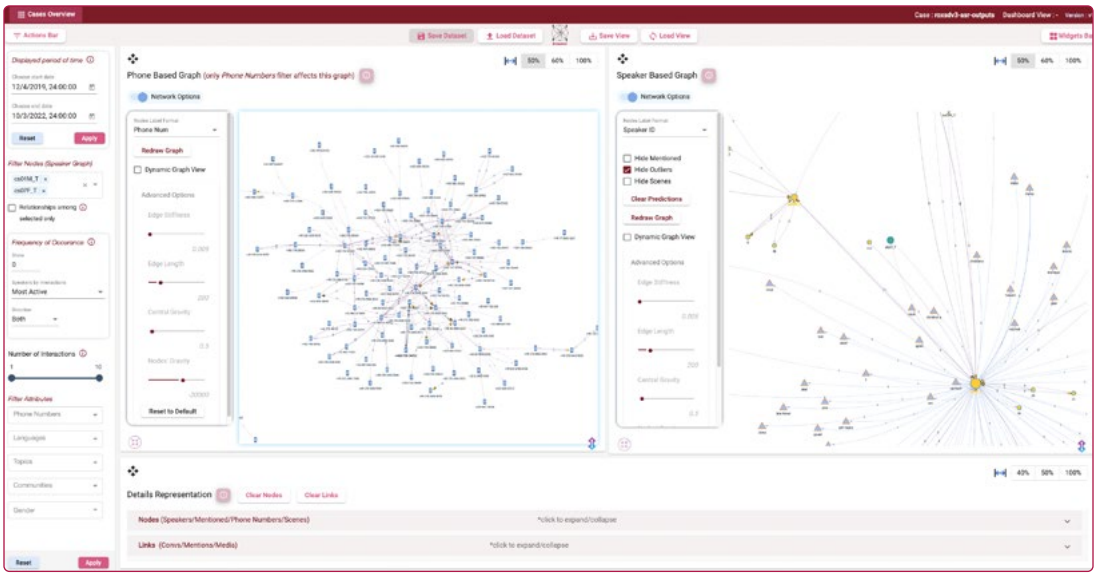


Figure 1 Screenshot of the Autocrime platform focusing on the comparison of the telephone-based network and the one based on speaker voiceprints

Autocrime is a platform that automates criminal investigations by ingesting multimodal data (e.g., wiretapped calls, metadata, seized videos) and providing insights such as:

- Which unique speakers participated in the calls despite usage of shared and/or burner phones?
- What is the topic of their conversation?
- Does any speaker appear in seized videos?
- How they are connected?
- And more information.

Here is a comparison showing why Autocrime is better than standard tools used by LEAs:

	Standard tools used by LEAs*	Autocrime platform
General features		
Ability to combine 2 (or more) cases in the same time	Usually yes	**Yes**
Incorporation of several different techniques into a single final product	Usually yes	**Yes**

Speech processing		
Voice Detection	Usually no	**Yes**
Speaker Clustering and Identification	Usually no	**Yes**
Gender Identification	Usually no	**Yes**
Language and Dialect identification	Usually no	**Yes**
Automatic Speech Recognition	Usually no	**Yes**
Keyword-Spotting	Usually no	**Yes**
Natural Language Processing (NLP)		
Named Entity Recognition	Usually yes	**Yes**
Topic Detection	Usually no	**Yes**
Identification of unknown 2nd party in calls	Usually no	**Yes**
Detection of Mentions of 3rd parties	Usually no	**Yes**
Exploratory analysis		
Timeline Analysis	Usually yes	**Yes**
Geospatial Analysis	Usually no	**Yes****
Details On Entities and Events	Usually yes	**Yes**
Advanced Filters	Usually yes	**Yes**
Visual analysis		
Facial Similarity Search	Usually no	**Yes****
Scene Similarity Search	Usually no	**Yes**
Network analysis		
Link Prediction	No	**Yes**
Social Influence Analysis	Usually no	**Yes**
Community Detection	Usually no	**Yes**
Outlier Detection	No	**Yes**
Cross Network Analysis	No	**Yes**

*We compared general tools used by different LEAs, including i2 Analyst's Notebook, i2 iBase, Tovek, Hansken, SIVE, BATVOX, ACU - EXPERT LAB, ADOBE AUDITION, Watson NLU, Google Cloud Natural, VoiceGain, Speechmatics, Newton Technologies

**In the process of integration

Speech, text and video in ROXANNE

Speaker recognition

The speaker recognition scenario in criminal investigations differs from most other speaker recognition scenarios in two significant ways. First, the expected scenario in criminal investigations is more general than standard verification or identification scenarios. In the ROXANNE platform, the user can provide a set of recorded phone calls and optionally provide the speaker identity for some speakers in the calls (enrolled speakers). The system then groups the voices in all calls so that voices belonging to the same speaker are in the same group. Note that this means that speakers that have not been enrolled can also be detected. Second, additional information can be used to help the speaker recognition system, for example the telephone numbers.

Multilingual Automatic Speech Recognition (ASR)

The speech recognition technology we have used is based on Time Delay Neural Network Factorization[1] model which is reasonably small compared to the end-to-end models and is based on Kaldi ASR toolkit[2]. We also experimented with boosting the language and grammar models using a lattice-based approach, which is a representation of the alternative word-sequences that are „sufficiently likely" for a particular utterance as suggested in previous research on semi-supervised learning[3]. The lattice generation algorithm ensures that the word-sequence is present in the lattice and has the best possible alignment and weight. Also, each word-sequence must be present in the lattice only once. It is achieved by first creating a state-level lattice, appropriately pruning it and then determinizing it using an algorithm that picks the most likely alternative path. For lattices written to disk, each word-sequence has only one path through the lattice. This word sequence becomes the prediction for given input.

Natural Language Processing (NLP)

During criminal investigations, a large volume of textual data is processed to track and analyze criminal activities. Inspecting documents line by line is time-consuming

[1] Khonglah, B., Madikeri, S., Dey, S., Bourlard, H., Motlicek, P., & Billa, J. (2020, May). Incremental semi-supervised learning for multi-genre speech recognition. In ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 7419-7423). IEEE.

[2] https://github.com/kaldi-asr/kaldi

[3] Juan Zuluaga-Gomez, Iuliia Nigmatulina, Amrutha Prasad, Petr Motlícek, Karel Veselý, Martin Kocour, Igor Szöke: Contextual Semi-Supervised Learning: An Approach to Leverage Air-Surveillance and Untranscribed ATC Data in ASR Systems. Interspeech 2021: 3296-3300

and error-prone. We developed different NLP technologies in ROXANNE which assist the user to process text efficiently. One of the main components we provide is a Named Entity Recognition (NER) module. It detects important "entities" appearing in the text (e.g., person names, locations and times) automatically. We note that our modules are specifically designed to work well on daily conversations, unlike standard NER modules, which expect formal text as input. We also make the models work robustly with the (potentially noisy) input from ASR systems. Another module we provide is called Mention Network (MN), which extends the NER module specifically for person names detected in the conversation. In a phone call consisting of two parties (the caller and the receiver), MN distinguishes whether a mentioned name refers to the caller/receiver or a third party. This is helpful for building networks that account for contacts and acquaintances in a criminal setting.

Geospatial analysis

Call records of a cell phone are a solid source of information about its user's whereabouts and, in many cases, helps identifying its user. Call records typically do not contain the geolocation of the device itself but the identity of the serving cell tower, which indicates that the device is in a particular service area. Although they cannot be used to pinpoint a device's location, they can be obtained for any device, even for a period in the past and even without access to the device. This makes them suitable for identification when a suspect carries a „private" phone and there is a second device whose use is disputed. Given a set of cell phones subject to investigation, we measure co-location for any two phones. Co-location can be quantified in the form of a likelihood ratio, which can be forensically interpreted and combined with other information such as speech, if available.

Video analysis

Autocrime's image and video ingestion chain extracts for each image and video a set of signatures that encode the visual aspect of the observed scenes and detected faces. These signatures are used to match new ingested documents with enrolled faces or scenes of manually associated suspects or victims by investigators. In this way, additional edges can be discovered between voice cluster nodes initially built from telephone calls only such as two voices heard in a same video, or same video observing a face (resp. scene) of an enrolled person (resp. scene) and containing the voice of another speaker. The additional edges suggested by this image/video ingestion chain significantly accelerate the identification of information of potential interest on documents (especially videos), which would require hours of tedious assessment without this capability.

Network analysis in ROXANNE

Network analysis assesses the relations among nodes (e.g., people, companies, countries, places). In law enforcement activities, network analysis is traditionally used to analyze the communications among suspects based on, e.g., wiretapped telephone calls or to map other known attributes of the suspects (e.g., financial, family ties, affiliation to the same gang). On the ROXANNE platform, the user can exploit several network analysis tools to gain a better insight into an investigation, including:

Social influence analysis

Inquiry on the most influential individuals in a criminal group has been one of the leading drivers of network analysis applications for organized crime. In this line, social influence analysis aims to quantify the influence of individuals on other individuals within a social network. On the ROXANNE platform, therefore, we focus on the overall influence of each individual within the network. In doing so, we assign each individual a relative importance score that measures the individual's influence compared to the other individuals.

Community detection

Very often, individuals within a network tend to form communities. In most cases, however, this community structure is hidden as communities are usually not well defined, and individuals do not publicly reveal their ties to groups (i.e., membership), making identification particularly difficult. To overcome this hurdle, we need to find unknown hidden structures and identify cohesive subgroups of individuals that interact more frequently with each other than with other individuals in the network using clustering methods. Consequently, the ROXANNE platform deploys several well-established community detection methods to identify these groups so investigators can focus their attention on the interactions within and between the identified communities, thus providing added value to LEAs.

Link prediction

In reality, individuals interact and communicate through many channels that are not always observable. This is due to the disclosure of hidden interactions in the past, the

availability of new data sources, or because individuals form new relationships with new individuals. In this task, we aim to uncover these missing or hidden interactions in the network and further predict the most likely ones. In the ROXANNE platform, this may assist investigators by providing clues about potential connections that have not (yet) been observed in the investigation and thus prompt specific actions (e.g., increased surveillance on two individuals due to their high likelihood of relatedness).

Outlier detection

An outlier is defined as an observation that deviates considerably from other observations in a way that raises suspicion. Because of their proven usefulness, outlier detection methods have been widely applied for fraud detection and crime investigation. In the ROXANNE context, outliers relate to individuals exhibiting abnormal attributes or interactions with other individuals, as well as peculiar behaviors that certain subjects do not often engage in. We aim to identify such individuals by providing a range of anomalies for individuals in the network and highlighting the ones with the highest scores (i.e., node-level outlier detection). Based on this, a criminal event detection and prevention alert system could be produced to reduce the manual effort of investigators by shrinking the size of the network to those more interesting nodes and prioritizing the most suspicious cases using topology-based metrics.

Cross-network entity matching

Cross-network entity matching finds nodes identified as the same entity in different graphs. These graphs can be constructed from various data sources by correlating different data sets and their corresponding networks. In the ROXANNE platform, this is done using information based on the description of the network entities that deviate from the observed and recorded properties (i.e., location, name, time) across the different data sets or on the structure of the graphs. LEAs often collect different modalities of data, each of which contains (partially) useful information for criminal investigations. In addition, criminals may use different communication channels to organize their activities (e.g., social media), as well as the dark web to transfer money. Therefore, isolated network-based investigations may not have a complete picture of the criminal network, so different modalities need to be merged to obtain relatively helpful information for further investigations. Node matching, or user identity linkage, can be beneficial for this end.

ROXSD: ROXANNE Simulated Dataset

Developing a cutting-edge solution to help LEAs perform efficient investigations through sophisticated use of biometric technologies within the boundaries of a privacy-first legal framework was a significant challenge for the evaluation of the ROXANNE solution, as the real-world data could not be utilized due to privacy, GDPR and other ethical constraints. As a consequence, the consortium decided to prepare its own research dataset: **ROXANNE Simulated Dataset (hereinafter, ROXSD).**

ROXSD is a multimodal and multilingual dataset of communication in organized crime, based on a fictional but realistic story that takes into account the constraints and challenges of a real investigation. It consists of around **18.5 hours** of intercepted telephone conversations in **13 different languages, 1.5 hours video recordings** and over **400 messages in textual format. 104 speakers and 42 authors** have contributed to the dataset with valid consent for the use of their biometric data. Ground-truth information, such as anonymized telephone and IMEI numbers, speaker labels, age and gender, transcription of the calls, date and time of the calls/messages, are also provided together with the raw data. **Thanks to ROXSD, LEAs as well as interested researchers are now able to test the performance of their algorithms, systems and solutions using realistic organized crime communication data without having to deal with privacy and ethics issues.**

ROXANNE Training Platform

An online educational platform has been developed for the purpose of the training in the use of the ROXANNE platform and its features. The implementation of the ROXANNE Training Platform is based on the Moodle Framework[1], which is widely used by several Universities and other educational organizations worldwide. Training is available to the end-users through the ROXANNE Training Platform at https://roxanne.kemea-research.gr/.

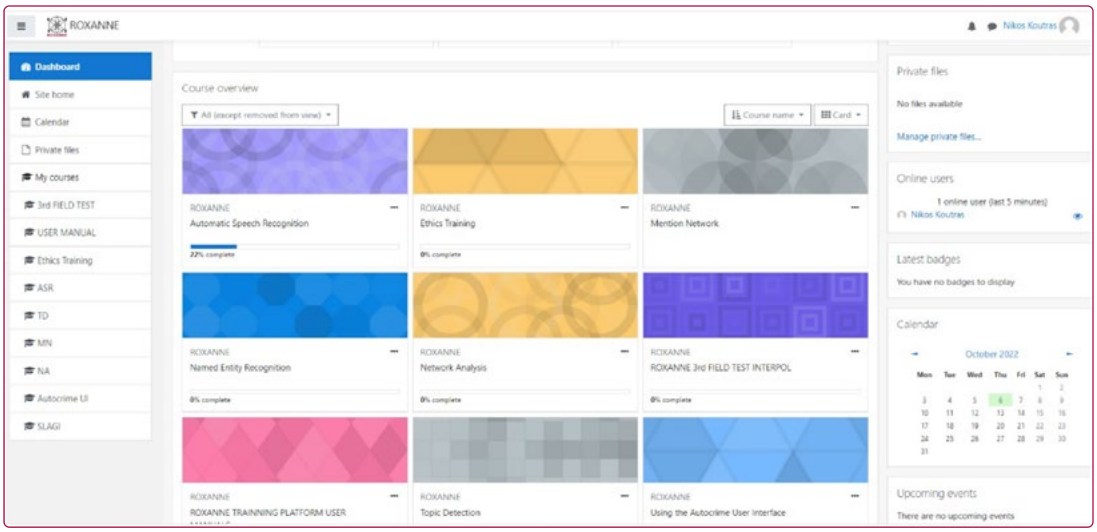


Figure 2 Screenshot from ROXANNE Training Platform

How to access the training platform?

Access to the platform is **allowed only to registered users**. Self registration is NOT allowed. To register, a request must be submitted by sending an email to contact@additess.com, with the Subject: [ROXANNE] Request for registration at the Training Platform.
Alternatively, registration can be requested through the Google Form available here: https://forms.gle/ppZfFj72vkasUHZr7.

Scan QR code to visit Training Platform website

Request for registration at the Training Platform

[1] https://moodle.org/

Collaboration with sister projects

Partners working on legal and ethical issues in ROXANNE have developed close links with colleagues working on similar issues in the LOCARD and FORMOBILE projects. Together they have discussed and collaborated on common issues. They are also in the process of finalizing a joint article on handling of eEvidence in compliance with data protection law and privacy rights.

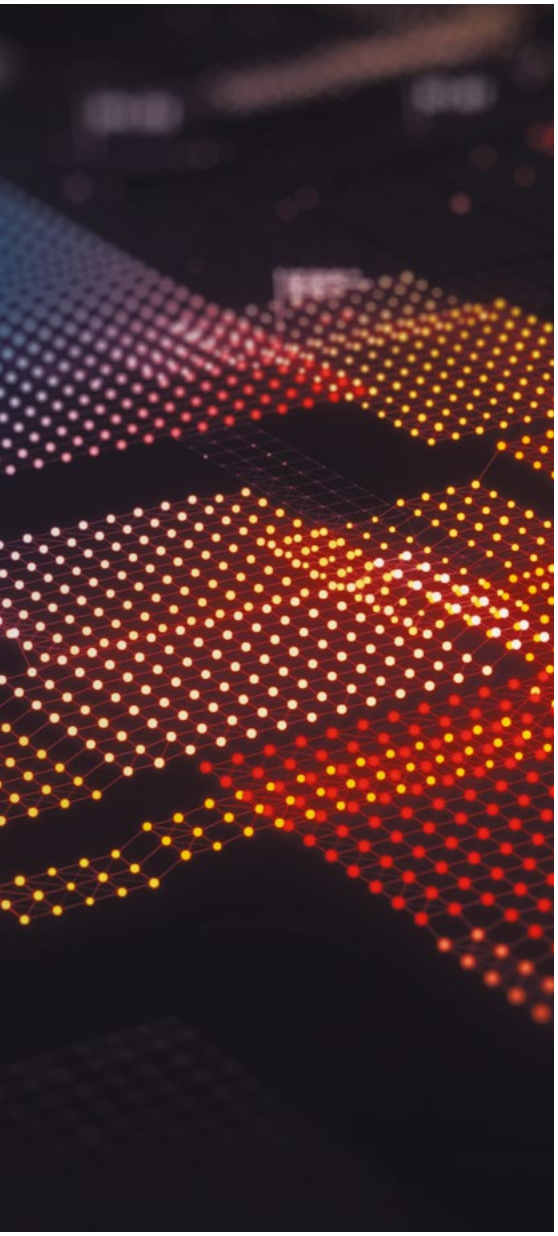
LOCARD FORMOBILE · FROM MOBILE PHONES TO COURT ·

LEA Cluster

Recognizing that our projects have common stakeholders and similar objectives in supporting law enforcement against organized crime and terrorism, we have joined a cluster with the following objectives:

1. To share knowledge to support law enforcement against money laundering, cybercrime, organized crime and terrorism, for example, through webinars prepared by the partners in the cluster.
2. To leverage our dissemination activities by mentioning the projects in the cluster on our websites, inviting articles from the cluster projects in our newsletter.
3. To ensure the coherence and complementarity of our recommendations to the EC, LEAs and other stakeholders, as far as possible.
4. To explore a degree of interoperability or compatibility between our technical platforms, modules and/or services.
5. To explore synergies, research opportunities and possible joint exploitation activities.

LEA Cluster consists of: CC-DRIVER, COPKIT, CYBERCRIME EXIT, CYBERSPACE, CYCLOPES, DARLENE, FREETOOL, HEROES, INSPECTr, LAW-GAME, LOCARD, NOTIONES, PREVISION, PROTAX, RAYUELA, STARLIGHT, Tools4LEAs, TRACE and ROXANNE.



Contact point

Petr Motlicek
Project Coordinator

petr.motlicek@idiap.ch

Idiap Research Institute Centre du Parc
Rue Marconi 19, CH-1920 Martigny, Switzerland

+41 27 721 77 49
www.idiap.ch

ROXANNE

Our website
ROXANNE-euproject.org

Follow us on Twitter
@ROXANNNE_Project

Check out our LinkedIn profile
ROXANNEProject

Our email
info@roxanne-euproject.org