



ROXANNE Training Platform

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

Nikolaos Koutras ADDITESS Ltd

Campus Cyber, 29-11-2022



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

Summary

- **Web based Training (WbT)** in distance learning education/training is considered to be an innovative method of learning
- In the **security domain**, where the usage of innovative and state of the art tools becomes a necessity, the **continuous training** of end-users is a **challenge**.
- The ROXANNE project presents **a great potential** for WbT where the training of tools by the technology providers should take place **during and after** the implementation of the project
- Training is available to the end-users through the ROXANNE Training Platform (Based on Moodle)



Training Platform

<https://roxanne.kemea-research.gr/>

The screenshot displays the ROXANNE Training Platform dashboard. At the top left, there is a navigation menu with options: Dashboard, Site home, Calendar, Private files, My courses, and SLAGI. The main content area is divided into several sections:

- Recently accessed courses:** A row of six course cards with progress bars, all at 0% completion. The courses are: ROXANNE Forensics Visualisation Toolkit, ROXANNE Automatic speech recognition, ROXANNE Topic detection, ROXANNE Network analysis, ROXANNE Voice Biometry, and ROXANNE Named-Entity Recognition.
- Course overview:** A grid of course cards with progress bars. The visible courses are: ROXANNE Automatic speech recognition, ROXANNE Forensics Visualisation Toolkit, ROXANNE Named-Entity Recognition, ROXANNE Network analysis, ROXANNE Topic detection, and ROXANNE Voice Biometry. All progress bars are at 0%.

On the right side, there are several utility widgets:

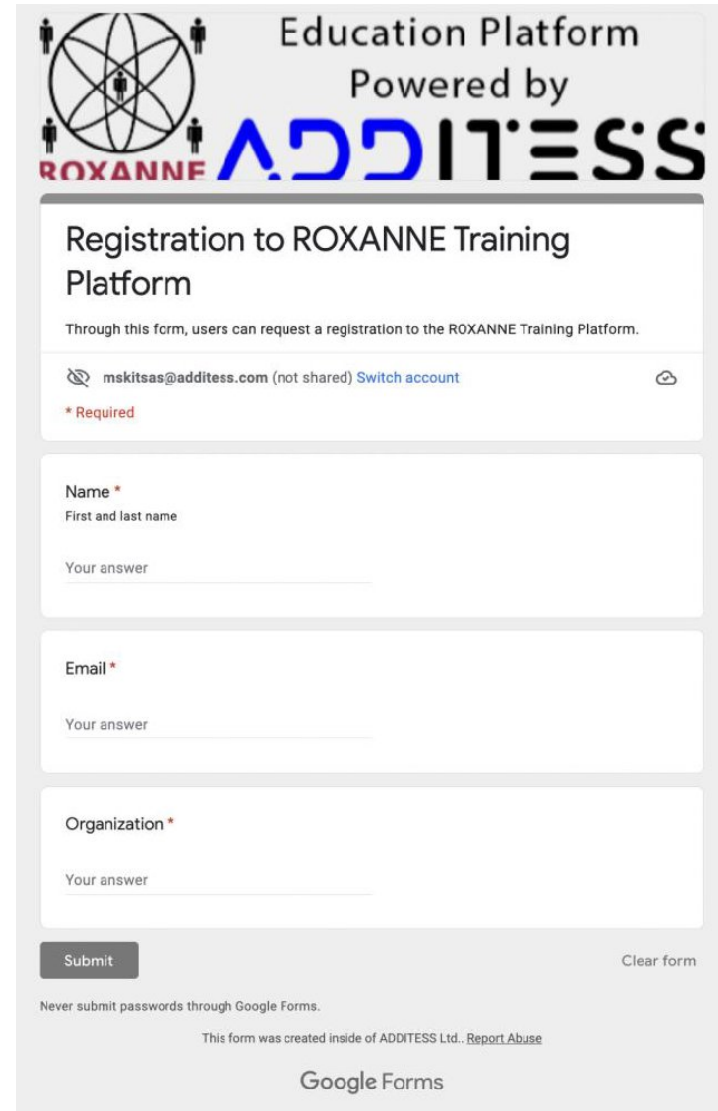
- Timeline:** Shows "No upcoming activities due".
- Private files:** Shows "No files available".
- Online users:** Shows "1 online user (last 5 minutes)" with a user icon for Michael Skitsas.
- Latest badges:** Shows "You have no badges to display".
- Calendar:** A calendar for September 2021.
- Upcoming events:** Shows "There are no upcoming events".

At the bottom of the dashboard, a status bar indicates: "You are logged in as Michael Skitsas (Log out)".



Training Platform: How to register?

- Self-registration is NOT allowed
 - Request registration through:
 - ✓ Email to contact@additess.com
- Subject: [ROXANNE] Request for registration or
- ✓ Google Forms
- <https://forms.gle/ppZfFj72vkasUHZr7>




The screenshot shows a Google Form titled "Registration to ROXANNE Training Platform". At the top, it says "Education Platform Powered by ADDITESS" with the ROXANNE logo. The form includes a header with the title and a description: "Through this form, users can request a registration to the ROXANNE Training Platform." Below this is a user identification section showing "mskitsas@additess.com (not shared)" with a "Switch account" link and a red asterisk indicating a required field. The form has three main input fields: "Name*" (with a sub-label "First and last name"), "Email*", and "Organization*", each with a "Your answer" label and an input line. At the bottom, there are "Submit" and "Clear form" buttons. A footer note states "Never submit passwords through Google Forms." and "This form was created inside of ADDITESS Ltd.. Report Abuse". The "Google Forms" logo is at the very bottom.




Training Platform: How to access?


<https://roxanne.kemea-research.gr/>



Education Platform
Powered by



Remember username

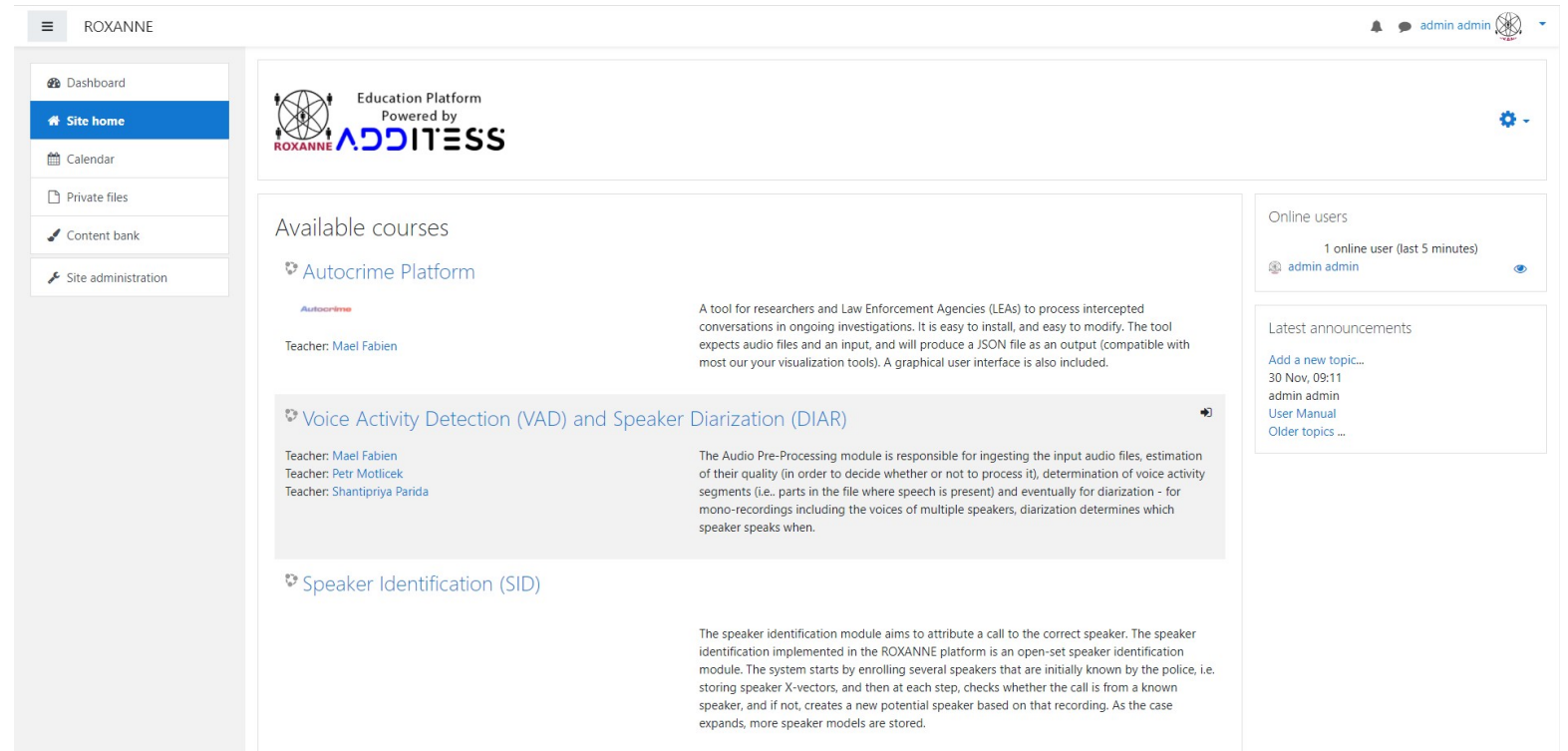
[Forgotten your username or password?](#)
Cookies must be enabled in your browser 

Is this your first time here?

For full access to this site, you first need to create an account.

List of courses

- ROXANNE User Manuals
- Autocrime Platform
- Voice Activity Detection (VAD) and Speaker Diarization (DIAR)
- Speaker Identification (SID)
- Voiceprint Extraction
- Gender Identification
- Automatic Speech Recognition
- Topic Detection
- Named Entity Recognition
- Mention Network
- Network Analysis
- Forensics Visualisation Toolkit



The screenshot displays the ROXANNE Education Platform interface. On the left is a navigation sidebar with options: Dashboard, Site home (selected), Calendar, Private files, Content bank, and Site administration. The main content area is titled 'Available courses' and features three course cards:

- Autocrime Platform**: A tool for researchers and Law Enforcement Agencies (LEAs) to process intercepted conversations in ongoing investigations. It is easy to install, and easy to modify. The tool expects audio files and an input, and will produce a JSON file as an output (compatible with most of our visualization tools). A graphical user interface is also included. Teacher: Mael Fabien.
- Voice Activity Detection (VAD) and Speaker Diarization (DIAR)**: The Audio Pre-Processing module is responsible for ingesting the input audio files, estimation of their quality (in order to decide whether or not to process it), determination of voice activity segments (i.e., parts in the file where speech is present) and eventually for diarization - for mono-recordings including the voices of multiple speakers, diarization determines which speaker speaks when. Teachers: Mael Fabien, Petr Motlicek, Shantipriya Parida.
- Speaker Identification (SID)**: The speaker identification module aims to attribute a call to the correct speaker. The speaker identification implemented in the ROXANNE platform is an open-set speaker identification module. The system starts by enrolling several speakers that are initially known by the police, i.e. storing speaker X-vectors, and then at each step, checks whether the call is from a known speaker, and if not, creates a new potential speaker based on that recording. As the case expands, more speaker models are stored.

On the right side of the interface, there are two widgets: 'Online users' showing 1 online user (last 5 minutes) and 'Latest announcements' with a link to 'Add a new topic...' and a recent announcement from 30 Nov, 09:11.

Course Management

Course categories

Create new category

ROXANNE 👁 ⚙ 12🎓

Sorting

Selected categories ▾

Sort by Category name ascending ▾

Sort by Course full name ascending ▾

Sort

Move selected categories to

Choose... ▾ Move

ROXANNE

Create new course

Sort courses ▾
Per page: 20 ▾

<input type="checkbox"/>	Autocrime Platform	⚙ 📄 🗑 👁 ⬇
<input type="checkbox"/>	Voice Activity Detection (VAD) and Speaker Diarization (DIAR)	⚙ 📄 🗑 👁 ⬆ ⬇
<input type="checkbox"/>	Speaker Identification (SID)	⚙ 📄 🗑 👁 ⬆ ⬇
<input type="checkbox"/>	Voiceprint Extraction	⚙ 📄 🗑 👁 ⬆ ⬇
<input type="checkbox"/>	Gender Identification	⚙ 📄 🗑 👁 ⬆ ⬇
<input type="checkbox"/>	Automatic Speech Recognition	⚙ 📄 🗑 👁 ⬆ ⬇
<input type="checkbox"/>	Topic Detection	⚙ 📄 🗑 👁 ⬆ ⬇
<input type="checkbox"/>	Named Entity Recognition	⚙ 📄 🗑 👁 ⬆ ⬇
<input type="checkbox"/>	Content Network	⚙ 📄 🗑 👁 ⬆ ⬇
<input type="checkbox"/>	Network Analysis	⚙ 📄 🗑 👁 ⬆ ⬇
<input type="checkbox"/>	Forensics Visualisation Toolkit	⚙ 📄 🗑 👁 ⬆ ⬇
<input type="checkbox"/>	Facial Biometry	⚙ 📄 🗑 👁 ⬆ ⬇

Showing all 12 courses

courses to...

Choose... ▾ Move

Autocrime Platform

View Edit Enrolled users Delete Hide Backup Restore



Full name	Autocrime Platform
Short name	Autocrime Platform
ID number	
Category	ROXANNE
Groupings	0
Groups	0
Role assignments	Teacher: 1 Student: 1
Enrolment methods	Self enrolment (Student)
Format	Topics format
Sections	General Structure Installation Run the ROXSD case Running your own case
Modules used	Forum Page URL



Course Enrollment

- Self Enrollment is enabled

Enrolment options

 Autocrime Platform 

Autocrime

Teacher: Mael Fabien


A tool for researchers and Law Enforcement Agencies (LEAs) to process intercepted conversations in ongoing investigations. It is easy to install, and easy to modify. The tool expects audio files and an input, and will produce a JSON file as an output (compatible with most our your visualization tools). A graphical user interface is also included.

▼ Self enrolment (Student)


No enrolment key required.


[Enrol me](#)


-  Autocrime Platform  
-  Voice Activity Detection (VAD) and Speaker Diarization (DIAR)  
-  Speaker Identification (SID)  
-  Voiceprint Extraction  
-  Gender Identification  
-  Automatic Speech Recognition  
-  Topic Detection  

Autocrime Platform 

[Dashboard](#) / [My courses](#) / [Autocrime Platform](#)

You are enrolled in the course. 

Your progress 

 Announcements

Course Structure

Add an activity or resource

Search

All Activities Resources

Assignment	Book	Chat	Choice	Database	External tool
Feedback	File	Folder	Forum	Glossary	H5P
IMS content package	Label	Lesson	Page	Quiz	SCORM package
Survey	URL	Wiki	Workshop		

Automatic Speech Recognition

Dashboard / My courses / ASR

Update 10.09.2021: A new topic "The ASR Component in the Autocrime Platform", is now online.

ASR: Setting, Benefits and Limitations

- ROXANNE Blog Post on ASR

The ASR Component in the ROXANNE Platform

- Introduction
- Installation Manual
- User Manual

The ASR Component in the Autocrime Platform

- Introduction
- Installation Manual
- User Manual

The Language Model Toolkit

- Language Model Toolkit - YouTube Video

Autocrime Platform

Dashboard / My courses / Autocrime Platform

Announcements

Structure

- Structure

Installation

- Hardware Specifications
- Install and downloads
- Video Tutorial: Installation Guide

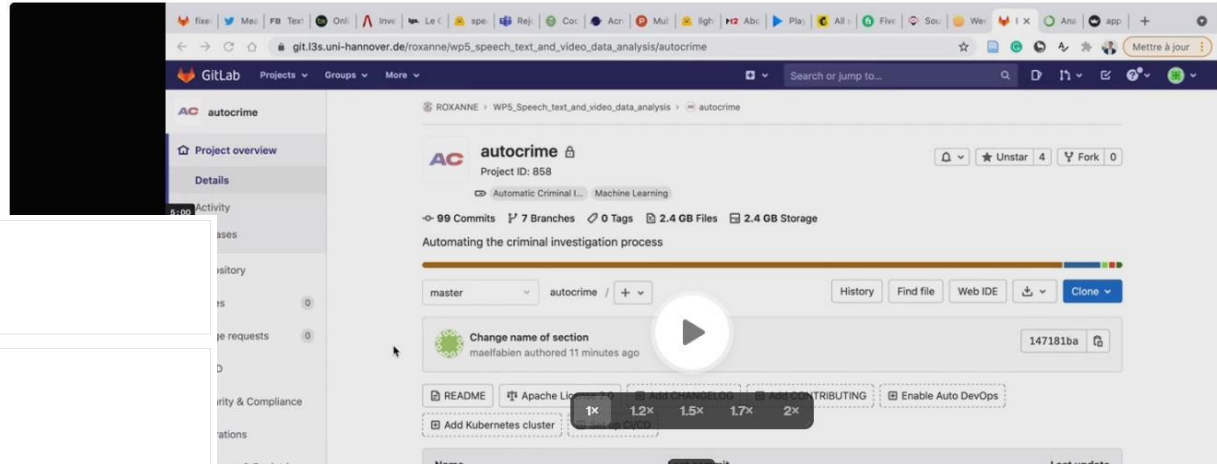
Run the ROXSD case

- Input Data
- Output Data
- Choosing Technologies to Run
- Run and Visualize the Results
- Video Tutorial: Getting Started
- Video Tutorial: The platform

Running your own case

- Running your own case

Course Material



Autocrime Platform

[Dashboard](#) / [My courses](#) / [Autocrime Platform](#) / [Installation](#) / [Hardware Specifications](#)

Hardware Specifications

The platform currently supports:

- Linux Ubuntu
- macOS (BigSure with M1 will not work)

Windows integration is a work in progress. However, running the platform in a virtual machine has already been tested. The following requirements should be respected:

- 16Gb of RAM
- 50Gb of disk space

Last modified: Thursday, 7 October 2021, 11:44 AM

[Structure](#)

Jump to...

Autocrime Platform

[Dashboard](#) / [My courses](#) / [Autocrime Platform](#) / [Run the ROXSD case](#) / [Run and Visualize the Results](#)

Run and Visualize the Results

You now have two ways to run the algorithm and visualize the results:

1. Process a **batch** of data (fully automatic), which will simply generate the JSON output. In the `config.yaml`, change `method` in `run` to "batch".
2. Process the data **file by file** via a user interface provided with an embedded network visualization tool, that lets you correct errors of the models. In the `config.yaml`, change `method` in `run` to "file".

This is what it looks like:

A screenshot of the Autocrime web interface. On the left, there is a sidebar with a 'Select your page' dropdown menu set to 'Ongoing case', a 'Display CSV file' checkbox, and a 'Process conversation' section with 'From file:' and 'To file:' dropdown menus. The main area displays a network graph with nodes and edges. Nodes are labeled with mentions like '#Hu', '#Michael', '#Christ', '#Kuba', '#Gordon', '#Rom?', '#Paul', '#Ron', '#Ronald', '#Adam', '#Mesi', and '#Ramsay'. The graph shows connections between these nodes, with some nodes being larger and more prominent.



Course Material

Autocrime Platform

Dashboard / My courses / Autocrime Platform / Installation / Hardware Specifications

Hardware Specifications

The platform currently supports:

- Linux Ubuntu
- macOS (BigSure with M1 will not work)

Windows integration is a work in progress. However, running the platform in a virtual machine has already been tested. The following requirements should be respected:

- 16Gb of RAM
- 50Gb of disk space

Last modified: Thursday, 7 October 2021, 11:44 AM

The screenshot shows the GitLab interface for the 'autocrime' project. The project overview includes details like 'Project ID: 858', 'Automatic Criminal L...', and 'Machine Learning'. It shows 99 commits, 7 branches, 0 tags, 2.4 GB files, and 2.4 GB storage. A recent commit 'Change name of section' is highlighted. Below, there are options to add a Kubernetes cluster or enable Auto DevOps. A table lists repository contents:

Name	Last update
data	V1.0.0 - Major update of the UI 4 hours ago
network	Correcting numpy version, modifying UI, S... 4 days ago
nlp	V1.0.0 - Major update of the UI 4 hours ago
pictures	Change images 2 hours ago
speech	V1.0.0 - Major update of the UI 4 hours ago
temp	Temp 2 months ago

Autocrime Platform

Resources / Autocrime Platform / Run the ROXSD case / Run and Visualize the Results

Run and Visualize the Results

You now have two ways to run the algorithm and visualize the results:

1. Process a **batch** of data (fully automatic), which will simply generate the JSON output. In the `config.yaml`, change `method` in `run` to `'batch'`.
2. Process the **data file by file** via a user interface provided with an embedded network visualization tool, that lets you correct errors of the models. In the `config.yaml`, change `method` in `run` to `'file'`.

This is what it looks like:

The screenshot shows the Autocrime web interface running on localhost:8502. On the left, there are controls for 'Ongoing case' (a dropdown menu) and 'Display CSV file' (a checkbox). Below that, 'Process conversation' is shown with 'From file:' set to '1' and 'To file:' empty. On the right, a network graph visualizes mentions. Nodes represent mentions like '#Adam', '#Paul', '#Ronald', '#Ramsay', '#Ramon', '#Kuba', '#Christ', '#Gordon', '#Mesli', and '#Adam'. Edges connect these nodes, showing relationships between different mentions.

The screenshot shows the 'Gender and Age Identification' section of the documentation. It explains that Gender Identification (GI) technology can automatically distinguish between male and female speakers in the examined segment. It also mentions an alternative speaker identification technology (AGE) that estimates the speaker's age. Configuration parameters for voice-print extraction and age estimation are listed:

```

Voice-print extraction
output setting:
  -o, --out-file file           output file
  -D, --out-dir dir           output directory
  -e, --out-ext ext [wp]      extension of voice-print files

dissimilarity options:
  -total-speakers num         total number of speakers
  -max-speakers num [4]      maximal number of speakers
  -max-avg-dist num [0.75]   maximal average distance between speakers

calibnet options:
  --calibnet-total-chunks num [4] total chunks in calibnet voiceprint

Age estimation
output:
  -o, --out-file file           output more file
  -f, --out-ext columns        enable column output format, columns to print are
                                specified by string of the characters below, e.g. aic
  a                             age
  l                             speech length
  r                             record length
  c                             channel number
  -suppress-too-short          suppress the 'too short' output


Gender detection
  
```



Progress Tracking

Course overview


▼ All (except removed from view) ▼



ROXANNE
Autocrime Platform

45% complete

...



ROXANNE
Automatic Speech Recog

0% complete

Your progress ?

Announcements

Structure

- Structure

Installation

- Hardware Specifications
- Install and downloads
- Video Tutorial: Installation Guide

Run the ROXSD case

- Input Data
- Output Data
- Choosing Technologies to Run
- Run and Visualize the Results
- Video Tutorial: Getting Started
- Video Tutorial: The platform

User Profile

ROXANNE

- Dashboard
- Site home
- Calendar
- Private files
- My courses
- USER MANUAL
- Ethics Training
- ASR
- TD
- NA
- Autocrime UI
- SLAGI

 Nikos Koutras [Message](#)

[Dashboard](#) / [Profile](#)

User details [Edit profile](#)

Email address
nkoutras@additess.com

Country
Cyprus

City/town
Nicosia

Privacy and policies
[Policies and agreements](#)

Course details

Course profiles

- [ROXANNE USER MANUAL](#)
- [Ethics Training](#)
- [Automatic Speech Recognition](#)
- [Topic Detection](#)
- [Named Entity Recognition](#)
- [Network Analysis](#)
- [Using the Autocrime User Interface](#)
- [Voice Biometry](#)

Miscellaneous

- [Blog entries](#)
- [Forum posts](#)
- [Forum discussions](#)
- [Learning plans](#)

Reports

- [Browser sessions](#)
- [Grades overview](#)





Login activity

First access to site
Wednesday, 15 June 2022, 11:27 AM (86 days 3 hours)

Last access to site
Friday, 9 September 2022, 3:07 PM (12 secs)

Online users

2 online users (last 5 minutes)

-  [Nikos Koutras](#) 
-  [admin admin](#) 

You are logged in as [Nikos Koutras](#) ([Log out](#))
[Home](#)

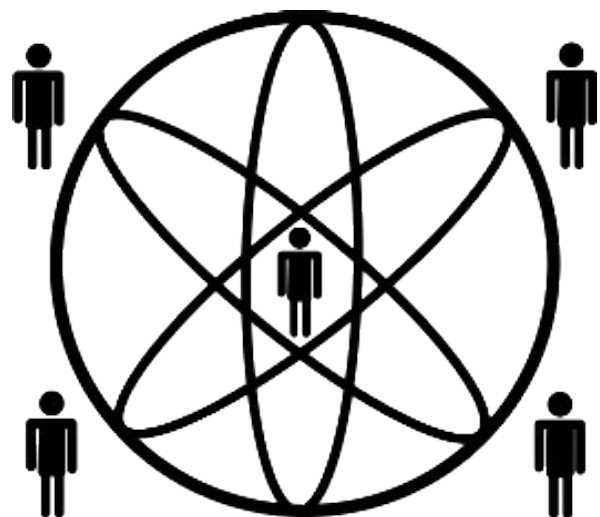
Conclusion

- The training material is documented in order to familiarized the non-experts to the usage of high-tech components.
- Future updates on training content will allow interactive training sessions and workshops
- The trainees will be able to “asynchronously” attend online courses related with the tools of ROXANNE project.
- Announcements, calendar and more tools are available...
- Blogs and direct messages will establish a communication channel between members of a course

Conclusion

- Live demo

<https://roxanne.kemea-research.gr/course/index.php>



ROXANNE

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

ROXANNE Training Platform

Thank You

Nikolaos Koutras ADDITESS Ltd

Campus Cyber, 29-11-2022



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