D7.5 DATA VISUALIZATION V2, ROXANNE PLATFORM V2

Grant Agreement:	833635
Project Acronym:	ROXANNE
Project Title:	Real time network, text, and speaker analytics for combating organised crime
Call ID: **Call name:**	H2020-SU-SEC-2018-2019-2020, Technologies to enhance the fight against crime and terrorism
Revision:	V2.0
Date:	31 May 2023
Due date:	31 May 2023
Deliverable lead:	AEGIS
Work package:	WP7
Type of action:	RIA

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

Disclaimer

The information, documentation and figures available in this deliverable are written by the "ROXANNE - Real time network, text, and speaker analytics for combating organised crime" project's consortium under EC grant agreement 833635 and do not necessarily reflect the views of the European Commission.

The European Commission is not liable for any use that may be made of the information contained herein.

Copyright notice

Project co-funded by the European Commission within the H2020 Programme (2014-2020)		
Nature of deliverable:	R	
Dissemination Level		
PU	Public	☒
CO	Confidential, only for members of the consortium (including the Commission Services)	□
EU-RES	Classified Information: RESTREINT UE (Commission Decision 2015/444/EC)	□
* R: Document, report (excluding the periodic and final reports) DEM: Demonstrator, pilot, prototype, plan designs DEC: Websites, patents filing, press & media actions, videos, etc. OTHER: Software, technical diagram, etc.		

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

Revision history

Revision	Edition date	Author	Modified Sections / Pages	Comments
V0.1	07/10/2022	AEGIS	Entire Document	Table of Contents added
V0.2	21/10/2022	AEGIS	Section 2	Updates compared to D7.3
V0.3	04/11/2022	AEGIS	Section 3	Updates compared to D7.3
V0.4	18/11/2022	AEGIS	Section 5	Updates compared to D7.3
V0.5	25/11/2022	AEGIS	Section 5	Updates compared to D7.3
V0.6	09/12/2022	AEGIS, ADDITESS	Section 5	Updates compared to D7.3
V0.7	16/12/2022	AEGIS, IDIAP	Section 4	Contributions added
V0.8	22/12/2022	AEGIS	Entire Document	Contributions added
V0.9	27/12/2022	HENS	Section 5	Internal Review performed
V0.95	29/12/2022	AEGIS	Entire Document	Reviewer's Comments addressed; Further contributions added
V1.0	30/01/2023	IDIAP	Entire Document	Final Version
V2.0	26/05/2023	AEGIS	Section 1, Section 4	External Reviewers' comments addressed
V2.0	31/05/2023	IDIAP	Sections 1, 2, 3, 4	Revised version per-reviewed

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

Executive summary

This deliverable (as a natural follow up of D7.4) provides a detailed description of the ROXANNE technical platform (Version V2 as further developed in the second half of the project), also referred to as Autocrime, its features, the related system architecture, as well as the target user requirements that were elicited and whose priority was continuously updated based on the interactions with Law Enforcement Agents (LEAs) representatives. Furthermore, it provides the differences between the original platform architecture (V1) and the current lightweight approach and discusses the advantages and disadvantages of each option.

The primary purpose of the ROXANNE platform is to assist law enforcement practitioners during the early phases of an investigation when deciding about the next steps to be taken by the team is important. The ROXANNE platform offers some unique features compared to existing commercial tools, especially with regards to advanced audio processing for performing speaker clustering and identification. The ROXANNE platform supports users in various ways, namely:

- while processing and combining data of various modalities (e.g., audio recordings from wiretapped phone calls and video files from seized devices) for automating the most time-consuming tasks during evidence extraction. Furthermore, the ROXANNE platform offers the ability to manage cases (e.g., creation, comparison and administration) and configure what processing should take place are key for efficiency.
- by automatically producing information on the entities (e.g., persons engaged in a phone call and/or appearing on a video footage) and their importance, the topics that are discussed, the languages used, etc. The user is offered with a wide range of visualisation schemes and filters to choose from so that different inferences can be formulated and tested. Furthermore, the user can assess the platform outputs and, whenever needed, to provide expert knowledge back to the platform.
- in general, an ability to process large-scale data (i.e., large number of audio, video or text files) which can be processed in a batch mode, and/or supporting parallelization in case of availability of GPU processing on the back-end.
- by exporting key outputs in other commercial platforms that can be shared with other colleagues and generating reports.

Nevertheless, it is anticipated that the ROXANNE platform will complement (rather than replace) other LEA software systems and tools that are responsible for maintaining integrity of collected evidence, establishing chain of custody during the full life-cycle of these evidences, performing forensics analysis based on scientific standards, as well as documenting and reporting these findings to be used in court proceedings.

The ROXANNE platform consists of four (4) families of technologies that are closely integrated: a) Multilingual speech-related technologies that include Speaker Recognition and Automatic Speech Recognition, as well as Voice Activity Detection, Voiceprint extraction for speaker segmentation and identification, and Language identification (extended for accent identification), b) Multilingual text-processing technologies (namely Named-Entity Recognition, Topic Detection, Mention Network and Relation Extraction), c) Video-processing technologies (in particular face characterization and scene, object characterization) and d) Network analysis methods that refer to Community Detection, Social Influence Analysis, Link Prediction, Cross-Network Merging and Outlier Detection. The LEA practitioner uses the Graphical User Interface to configure the processing that should take place, to provide any intelligence already available, to explore the automated results and, if and when needed, to provide expert knowledge back to the platform by validating and updating the outputs. The back-end technologies are apriori developed as multilingual (i.e., generated from large multilingual pretrained models) and thus can operate on multiple languages (i.e., ASR supports 9 languages,

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

Language and accent identification supports even other languages than planned, NLU technologies are generally supporting 8 languages, etc.). For more details about multilingual support, please refer to D5.3 document.

By building upon the thorough analysis of current Law Enforcement Agents' (LEAs) practices, pain points and feedback received during the three field test demonstrations, this report documents a set of 37 (thirty-seven) functional and non-functional requirements grouped into 5 (five) broad categories, namely platform management, user management, case management, evidence processing and case analysis. Each of those user requirements is then translated into one, or more, system requirement that the ROXANNE platform should support. In total we identified 88 (eighty-eight) distinct system requirements that provide technical specifications for the ROXANNE platform.

Eventually, a set of features for the ROXANNE platform has been defined and implemented in order to meet the most important user (and system) requirements. These features include the following: Cases Dashboard, Social Network Analysis, Timeline Analysis, Statistical Analysis, Details, Synced filters, Custom Workspace, Cross-case analysis, Reporting, Platform management and User management. For each of those ROXANNE platform features, a short description is provided along with selected screenshots that demonstrate key functionalities offered.

V2.0 update

This is an updated version of D7.5 (version 2.0) reacting on requests from the review comments provided to the ROXANNE consortium after the final review meeting. Majority of technologies were updated, including multilingual speech recognition, language and accent recognition, topic detection, NLP related technologies, authorship attribution, video technologies, geo-location, etc. Please refer to the most updated results addressed in sections below (while more detailed and latest performance analysis is given in D5.3). This document (in sections 4.1.1 and 4.1.2) also explains the differences between the original (V1) platform architecture and the current V2 approach and if this affects the platform's capacity or any other feature. We also update details about minimum of data required (Section 4.4) to operate the platform.

Table of contents

Disclaimer 2
Copyright notice 2
Revision history 3
Executive summary 4
Table of contents 5
List of Acronyms 8
List of Tables 9
List of Figures 10
1. Introduction 12
1.1. Purpose of the document 14
1.2. Document structure 15

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

2. User requirements 16
2.1 Platform management 16
2.2 User management 17
2.3 Case management 19
2.4 Evidence processing 21
2.5 Case analysis 24
3. System requirements 33
3.1 Platform management 33
3.2 User management 34
3.3 Case management 35
3.4 Evidence processing 36
3.5 Case analysis 37
4. The ROXANNE technical Platform 43
4.1 ROXANNE platform architecture 43
4.1.1 The Initial ROXANNE platform (V1) architecture 43
4.1.2 The final ROXANNE (V2) platform architecture 45
4.2 Integrated technologies 54
4.2.1 Speech technologies 54
4.2.2 Text technologies 55
4.2.3 Network Analysis technologies 55
4.2.4 Video technologies 55
4.3 Evaluation of technologies 56
4.4 Overall workflow 56
4.5 Further updates 61
5. Platform features 62
5.1.1 Cases Dashboard 62
5.1.2 Uploading files to an existing case 64
5.1.3 Configuring an existing case 64
5.1.4 Social Network Analysis 66
5.1.5 Timeline Analysis 70
5.1.6 Statistical Analysis 71
5.1.7 Details 71
5.1.8 Synced filters 75
5.1.9 Custom Workspace 76
5.1.10 Data export 77
5.1.11 Geospatial analysis 79
5.1.12 Cross-case analysis 80
5.1.13 Reporting 81
5.1.14 Platform management 83
5.1.15 User management 84

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

6. Conclusions......86
References......88
Annex A - State of the art on visualisation and exploratory analysis tools......89
4.1 Cellebrite Pathfinder......89
4.2 i2 Analyst's Notebook......91
4.3 Forensic Toolkit......92
4.4 Magnet Axiom......93
4.5 Hansken......95
4.6 BriefCam......96
4.7 Palantir Gotham......97
4.8 Comparing the ROXANNE platform against other commercial tools......98
Annex B – Mapping of System Requirements to User Requirements......102

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

List of Acronyms

Table 1: List of Acronyms

Acronym	Description
AI	Artificial Intelligence
ASR	Automated Speech Recognition
CA	Case Analysis
CCTV	Closed Circuit TeleVision
CDR	Call Detail Records
CM	Case Management
CPU	Central Processor Unit
CSV	Comma Separated Values
EIA	Enterprise Insight Analysis
EP	Evidence Processing
FTK	Forensic Toolkit
GPU	Graphics Processing Unit
IT	Information Technology
GUI	Graphical User Interface
HTTP	Hypertext Transfer Protocol
JSON	JavaScript Object Notation
LEA	Law Enforcement Agent
NER	Named-Entity Recognition
OSINT	Open Source Intelligence
PM	Platform Management
ROXSD	ROXANNE Simulated Dataset
SR	System Requirement
UM	User Management
UR	User Requirement
VAD	Voice Activity Detection
WP7	Work Package 7
XML	Extensible Markup Language

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

List of Tables

Table 1: List of Acronyms 8
Table 2: user requirements related to platform management 17
Table 3: user requirements related to user management 18
Table 4: User requirements related to case management 21
Table 5: User requirements related to evidence processing 23
Table 6: User requirements related to case analysis 30
Table 7: Mapping of System Requirements to User Requirements and Status 102

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

List of Figures

Figure 1: *The intelligence cycle (adapted from [1])* 12
Figure 2: Popularity of features for analysis or search tool as expressed by respondents to online questionnaire on user requirements 20
Figure 3: Most useful type of visualizations as expressed by respondents to online questionnaire on user requirements 25
Figure 4: Usefulness of criminal social networks that were the main visualization chart that was demonstrated during the first field test 26
Figure 5: Most useful features for a search tool as expressed by respondents to online questionnaire on user requirements 27
Figure 6: Major concerns with regards to analysis or search tools as expressed by respondents to ROXANNE questionnaire on user requirements 30
Figure 7: A high-level overview of the ROXANNE platform 43
Figure 8: The ROXANNE platform architecture 47
Figure 9: A sample file containing information about the evidence files to be processed by the ROXANNE platform 48
Figure 10: Screenshot of a sample “link” object that is part of the result.json generated by the ROXANNE platform 50
Figure 11: Screenshot of a sample “node” object that is part of the result.json generated by the ROXANNE platform 51
Figure 12: Screenshot of a sample “outliers” object that is part of the result.json generated by the ROXANNE platform 51
Figure 13: Screenshot of a sample “scores” object that is part of the result.json generated by the ROXANNE platform 52
Figure 14: Screenshot of a sample “social influence scores” object that is part of the result.json generated by the ROXANNE platform 52
Figure 15: Screenshot of a sample “communities” object that is part of the result.json generated by the ROXANNE platform 53
Figure 16: Screenshot of a sample “link predictions” object that is part of the result.json generated by the ROXANNE platform 53
Figure 17: The overall workflow 57
Figure 18: Screenshot of Cases Dashboard widget for user analyst.a 62
Figure 19: Screenshot of Cases Dashboard widget for user analyst.a 62
Figure 20: A screenshot of the log-in page provided by the Keycloak instance of the ROXANNE platform 63
Figure 21: Creating a new case 63
Figure 22: Uploading files to a case 64
Figure 23: Configuring the speech technologies of ROXANNE platform 65
Figure 24: Configuring the ROXANNE platform’s technologies related to text-processing 66
Figure 25: Configuring where the outputs of the ROXANNE platform will be stored 66
Figure 26: An overview of the Phone-based Network 67
Figure 27: Screenshot of Social Network Analysis widget 68
Figure 28: Detail of speaker-based graph where “cs07F_T” is predicted to interact with nodes “10” and “12” 69
Figure 29: Adding user annotations on the speaker-based network 69
Figure 30: Screenshot of the Timeline Analysis widget (with annotations for better readability) 70

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

Figure 31: Screenshot of the Statistical Analysis widget 71
Figure 32: Screenshot of the Details widget 72
Figure 33: Screenshot of Merging (left) and Splitting (right) Entities 73
Figure 34: Assisting the user when assessing the similarity of the voiceprints for a selected pair of audio channels (rows) to the rest of the audio files (columns) 74
Figure 35: Screenshot of Synced Filters feature 75
Figure 36: Synced Filters feature 76
Figure 37: Loading a previously saved dashboard view 77
Figure 38: A screenshot of the confirmation message notifying the user that a snapshot of the ROXANNE outputs has been exported to a format that can be used with I2 Analyst Notebook. 77
Figure 39: Screenshot of the file exported using the I2 Analyst Notebook format 78
Figure 40: Mapping the data exported from the ROXANNE platform into the internal data schema of I2 Analyst Notebook 78
Figure 41: Visualising the ROXANNE outputs using the I2 Analyst Notebook 79
Figure 42: Searching for events near a specific location (left) and viewing the events on a map (right) 79
Figure 43: A screenshot of the social network generated when two cases are merged 80
Figure 44: A screenshot of the inputs expected by the user when merging two cases 81
Figure 45: The first page of a sample report generated by the ROXANNE platform 81
Figure 46: A sample report generated by the ROXANNE platform containing information about the speaker-based network 82
Figure 47: Candidate deployment models of the ROXANNE/Autocrime platform 83
Figure 48: A screenshot of the Autocrime desktop in case of Virtual Machine instance 83
Figure 49: A screenshot of exemplary console outputs when starting the ROXANNE platform's backend 84
Figure 50: A screenshot of the User Management page provided by the Keycloak instance of the ROXANNE platform 85
Figure 51: A screenshot of the User Management page provided by the Keycloak instance of the ROXANNE platform 87
Figure 52: Most popular investigation platforms used by respondents to ROXANNE questionnaire on user requirements 89
Figure 53: Screenshot of Cellebrite Pathfinder for enabling analytics engines that identify key entities and names in text or objects in images and videos 90
Figure 54: Screenshot of Cellebrite Pathfinder for manually (instead of automatically) identifying individuals 90
Figure 55: Screenshot of Cellebrite Pathfinder showing the set of filters that can be applied on processed data 91
Figure 56: Screenshot of i2 Analyst Notebook showing a social network graph 92
Figure 57: Screenshot of Forensic Toolkit workspace 93
Figure 58: Screenshot of Magnet Axiom showing a social network graph 94
Figure 59: Screenshot of Magnet Axiom showing a timeline graph 95
Figure 60: Screenshot of Hansken showing a case dashboard 96
Figure 61: Screenshot of Briefcam showing a case workspace 97
Figure 62: Screenshot of Palantir Gotham showing a social network graph 98

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

1. Introduction

Today LEAs face increasing challenges when they are on to suppress organized crime activities. In general, it is at the utmost priority for them to be able to investigate a large number of criminal cases at the same time without losing situational awareness.

Criminals often employ sophisticated technologies in planning and executing their illegal activities, as well as concealing the revenues stemming from them. Suppressing organized crime remains an important but challenging task for Law Enforcement Agencies (LEAs) for several reasons, such as:

- The growing variety of heterogeneous data that need to be collected, filtered, analysed and secured. Given the omnipresence of digital technologies in our lives, virtually any type of crime has a digital forensics component[1]. This data can be structured (e.g., call detail records) or unstructured (e.g., audio from lawfully intercepted conversations, video from Closed Circuit TeleVision (CCTV) systems, text communications in social media groups, etc.).
- The velocity of this data also tends to be high; large cases involve tens of subjects and each one of them may use several devices, some of them only once ("burner phones").
- The size of data that need to be processed grows further as some cases may last for several years resulting in thousands of recorded communications.
- The large number of cases that need to be investigated and the fact that for each case several interception warrants, usually involve several individuals, are issued.

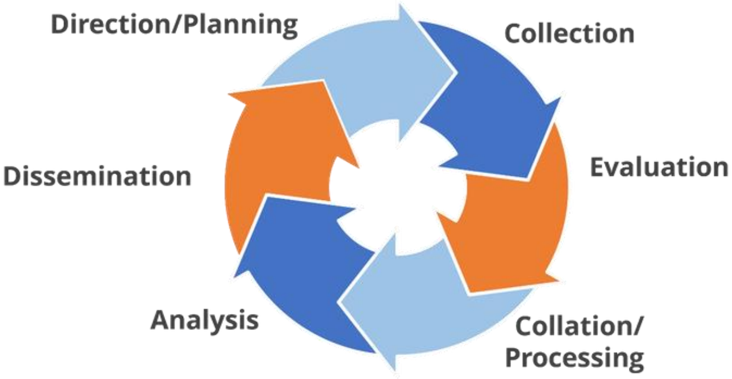


Figure 1: ***The intelligence cycle (adapted from [1])***

The lifecycle of criminal investigations can be modelled using the concept of the intelligence cycle of Figure 1, which includes the following phases:

- **Direction/Planning**: A new case is brought to law enforcement agency's attention and a set of questions are formulated. These questions express a requirement for an investigation focusing on a subject or a range of subjects of concern, a risk, or threat. Examples of questions are:
  - Who is central in this criminal organization?
  - Which communication links are most worth monitoring?

[1] Digital forensics is the scientific examination and analysis of data held on, or retrieved from, computer storage media in such a way that the information can be used as evidence in a court of law.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

  - How often are the interactions?
  - Who is the initiator of the interactions?
  - The removal or incapacitation of which individuals would sever a criminal group?
- **Collection**: After the set of questions are clearly defined, the law enforcement practitioners need to follow a data collection plan in which the specific data items needed, possible sources of information are identified and eventually the relevant data are obtained. There are several categories of data sources that are used by law enforcement agents, such as:
  - closed information sources (e.g., those collected for a specific purpose with limited access and availability to the general public such as a car license plate register),
  - classified information collected by specifically tasked covert means including use of human and technical resources (also known as Image intelligence and Signals Intelligence),
  - open information sources (i.e., those that are publicly available such as popular social networks), and
  - seized documents, biometrical data and digital media (e.g., smartphones and laptops).
- **Evaluation**: The validity and accuracy of the data obtained is assessed, which heavily depends on the reliability of the source. For example, open information sources need to be carefully examined as these can frequently be biased or inaccurate.
- **Collation/Processing**: The data that were found to be relevant and accurate is transferred into a digital storage system in a structured way (e.g., indexed and cross-referenced) that permits authenticated and timely access. Any further processing, such as use of translation software for dealing with data in foreign languages is also performed during this step, while maintaining chain of custody (i.e., a complete record of data lifecycle and each person who accessed it) for ensuring integrity.
- **Analysis**: In this phase the investigator examines the data by querying and filtering the digital storage system and performs temporal, subject, object and statistical analysis for identifying linkages between the data that convey useful information. By combining different pieces of information and information about the context the investigators obtain intelligence about trends and patterns.
- **Dissemination**: The intelligence developed during the analysis phase is disseminated to the rest investigative team using formal reports or presentations with supporting documentation. In this way, candidate answers to the questions posed are given, information gaps are highlighted and decisions on how to proceed are taken, e.g., new questions are posed, and responsibilities are assigned. This leads to a new intelligence cycle, where the scope may be broader or narrower, depending on the quality of intelligence already obtained.

The ROXANNE technical platform (V2), also referred to as Autocrime, aims to support law enforcement practitioners in several phases of the investigation process outlined above. In particular, the ROXANNE platform supports with:

- **Collation/Processing** phase by processing and combining data of various modalities (e.g., audio recordings from wiretapped phone calls and video files from seized devices) for automating the most time-consuming tasks during evidence extraction. Furthermore, the ability to manage cases (e.g., creation, comparison and administration) and configure what processing should take place are key for efficiency.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

- **Analysis** phase by automatically producing information on the entities (e.g., persons engaged in a phone call and/or appearing on a video footage) and their importance, the topics that are discussed, the languages used, etc. The user is offered with a wide range of visualisation schemes and filters to choose from so that different inferences can be formulated and tested. Furthermore, the user can assess the platform outputs and, whenever needed, to provide expert knowledge back to the platform.
- **Dissemination** phase by exporting key outputs in other commercial platforms that can be shared with other colleagues and generating reports.

Note that the primary purpose of the ROXANNE platform is to assist Law enforcement practitioners during the investigation process, i.e., to speed up the investigation process by automatically analysing multi-modal evidence files, inferring relationships between suspect entities and events of high-priority, presenting these findings in an intuitive manner, enabling users to improve these findings based on their expertise and allowing the user to confirm/reject hypotheses. The ROXANNE platform is envisioned to be used in the early phases of an investigation when deciding about the next steps to be taken by the team is important. It is anticipated that the ROXANNE platform will complement (rather than replace) other LEA software systems and tools that are responsible for maintaining integrity of collected evidence, establishing chain of custody during the full life-cycle of these evidences (from collection till long-term archival), performing forensics analysis based on scientific standards, as well as documenting and reporting these findings to be used in court proceedings.

Having said that, we should note that the ROXANNE platform already features mechanisms that are important for ensuring case integrity (e.g., user authentication and role-based access control prior to interacting with a case), explain-ability of results (e.g. by giving hints to the user when the acoustic properties of two speakers' voices are similar, or not), smooth collaboration with colleagues (e.g., allowing the user to write a summary of findings that will be part of an informal report to be communicated to other practitioners). Furthermore, the ROXANNE platform is modular and can be extended to support additional user requirements that will emerge during the commercialisation phase, which could not be prioritised given the prototyping nature of a Research and Innovation Action project.

1.1. Purpose of the document

This deliverable responds to the requirements of Work Package 7 (WP7) that deals with integration and visualisation of results of the ROXANNE project. It provides a thorough description of the law enforcement officers' requirements related to the desired visual data exploration capabilities of the ROXANNE investigative platform, by:

- analysing the findings collected from the end-user requirements survey, and which were documented in [2] and
- analysing the feedback received during the three field test demonstrations of the ROXANNE platforms available at those times and the follow-up questionnaires that were shared with a closed set of users composed of LEA representatives from ROXANNE consortium and external stakeholder members (i.e., D8.4 [3], D8.9 [4]and D8.6 [5]).
- building upon and updating previous reports documenting the intermediate version of the platform, namely D7.3 [6] and D7.4 [7].

After eliciting the functional and non-functional user requirements, a set of system requirements are derived and mapped to a set of ROXANNE platform features. Furthermore, a detailed presentation of the ROXANNE platform architecture and a baseline workflow is provided.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

1.2. Document structure

This report has the following structure: Section 2 provides the set of user requirements, both functional and non-functional ones, that were extracted based on the views of ROXANNE LEAs and key external stakeholders, while Section 3 maps these user requirements into system requirements. Section 4 provides an overview of the ROXANNE platform architecture, a high-level overview of the underlying technologies related to speech processing, text processing, video analytics and network analysis. Consequently, Section 5 describes the main features of the ROXANNE platform, how these are expected to meet users' requirements and related screenshots for the supported features. Finally, Section 6 concludes this report with main findings and future steps.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

2. User requirements

The successful design of any information system strongly depends on correctly understanding and prioritising the user requirements. In this section we provide an update on the user requirements of the ROXANNE framework that were initially described in [6]. These requirements were elicited based on the responses of ROXANNE LEAs and key external stakeholders to an online survey as well as their feedback during and after the ROXANNE field tests. In particular [2] includes a detailed description and analysis of the current practices with regards to technologies used by LEAs and their pain points, while [3], [4] and [5] provide an overview and statistical analysis of the end-users' view on the evolutions of the ROXANNE platform.

User requirements can be categorised into functional and non-functional. The former ones are responsible for system behaviours or functions, while the latter refer to the properties that these functions should have, such as usability/user friendliness, performance or data security needs.

Each user requirement is identified using the following format: **<UR-FunctionalAreaX-ShortName>,** where:

- **UR** stands for User Requirement
- **FunctionalArea** refers to any of the following acronyms:
    - PM for Platform Management
    - UM for User Management
    - CM for Case Management
    - EP for Evidence Processing
    - CA for Case Analysis
- **X** is the incremental number of the functional area and
- **ShortName** is a distinctive name.

2.1 Platform management

UR-PM01-MultiOSplatform: As a law enforcement practitioner, I want to be able to run the ROXANNE platform on any computer, so that I can use it anywhere and anytime.

Users should be able to install the ROXANNE platform on any modern computer, regardless of its Operating System (e.g., Windows, Linux, MacOS). In fact, participants to the 2nd and 3rd field tests were using laptops with various Operating Systems and this is expected to be witnessed in LEA's premises.

Furthermore, while participants to the field tests were instructed to carry a personal laptop, the ROXANNE platform should be also deployable on any host type (e.g., server, desktop, laptop).

UR-PM02-EasyToMaintainPlatform: As a law enforcement practitioner, I want to be able to keep the ROXANNE platform up-to-date, so that I can use all the latest features.

Users should be able to automatically update the ROXANNE platform in order to make use of any features recently added, or stability improvements performed.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

Table 2 summarizes the user requirements related to platform management.

Table 2: user requirements related to platform management

Requirement ID	Short requirement description	Priority (final version)
UR-PM01-MultiOSplatform	As a law enforcement practitioner, I want to be able to run the ROXANNE platform on any computer, so that I can use it anywhere and anytime.	Must-have
UR-PM02-EasyToMaintainPlatform	As a law enforcement practitioner, I want to be able to keep the ROXANNE platform up-to-date, so that I can use all the latest features.	Must-have

2.2 User management

UR-UM01-UserAccountCreation: As a law enforcement manager[2], I want to be able to create a personal account for each law enforcement practitioner, so that the latter can use the ROXANNE platform.

ROXANNE handles sensitive data and thus only authorised law enforcement practitioners should be allowed to use it. Having individual user accounts following the organisation's security policy is important for any sensitive system at production for achieving:

- confidentiality - degree to which a product or system ensures that data are accessible only to those authorized to have access
- integrity - the degree to which a system, product or component prevents unauthorized access to, or modification of, computer programs or data
- non-repudiation - degree to which actions or events can be proven to have taken place so that the events or actions cannot be repudiated later
- accountability - the degree to which the actions of an entity can be traced uniquely to the entity and
- authenticity - the degree to which the identity of a subject or resource can be proved to be the one claimed.

UR-UM02-UserAccountView: As a law enforcement manager[3], I want to be able to view all existing accounts, so that I can check that the information of any individual practitioner is up-to-date.

Administrators should be able to check that users' personal information reflects their current status. While such a requirement is important in commercial systems, an experimental implementation should be available in the final ROXANNE version.

[2] In D7.3 this user requirement was associated to a law enforcement practitioner, but it was suggested that the administrator of the management team should be responsible for creating user accounts.

[3] In D7.3 this user requirement was associated to a law enforcement practitioner, but it was suggested that the administrator of the management team should be responsible for viewing details of user accounts for management purposes.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

UR-UM03-UserAccountEdit: As a law enforcement manager[4], I want to be able to edit an existing account, so that I can update its properties according to the current user's profile.

Users should be able to update their personal details, so that their colleagues can recognise them. Even though this requirement is usually met in commercial systems, we consider its priority to be moderate for the final version of the ROXANNE platform.

UR-UM04-UserAccountDelete: As a law enforcement manager[5], I want to be able to delete an existing practitioner's account, so that I can the latter can no longer use the ROXANNE platform.

Retired users should have no longer access to the final ROXANNE system for security purposes. Again, the priority of this requirement is not critical for the final version of ROXANNE platform.

UR-UM05-UserLogIn: As a law enforcement practitioner, I want to be able to log-in to the ROXANNE platform using my personal account, so that I can work on the cases assigned to me.

Even though the ROXANNE platform can be installed on a personal laptop, having a shared instance where several practitioners are using it means that users need to be able to log-in using their own account. Thus, this user requirement is of high importance.

The following table summarizes the user requirements related to user management.

Table 3: user requirements related to user management

Requirement ID	Short requirement description	Priority (final version)
UR-UM01-UserAccountCreation	As a law enforcement manager, I want to be able to create a personal account for each law enforcement practitioner, so that the latter can use the ROXANNE platform.	Must-have
UR-UM02-UserAccountView	As a law enforcement manager, I want to be able to view all existing accounts, so that I can check that the information of any individual practitioner is up-to-date.	Could-have
UR-UM03-UserAccountEdit	As a law enforcement manager, I want to be able to edit an existing account, so that I can update its properties according to the current user's profile.	Could-have

[4] In D7.3 this user requirement was associated to a law enforcement practitioner, but it was suggested that the administrator of the management team should be responsible for editing user accounts.

[5] In D7.3 this user requirement was associated to a law enforcement practitioner, but it was suggested that the administrator of the management team should be responsible for deleting user accounts.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

Requirement ID	Short requirement description	Priority (final version)
UR-UM04-UserAccountDelete	As a law enforcement practitioner, I want to be able to delete my personal account, so that I can no longer use the ROXANNE platform.	Could-have
UR-UM05-UserLogIn	As a law enforcement practitioner, I want to be able to log-in to the ROXANNE platform using my personal account, so that I can work on the cases assigned to me.	Must-have

2.3 Case management

UR-CM01-CaseCreation: As a law enforcement practitioner, I want to be able to create a new case, so that I can keep track of all relevant information in a single place.

Law enforcement agents are confronted with a growing volume and variety of digital evidence for each case they are involved in and at the same time they may be assigned to several ongoing investigations. Thus, it is important to be able to associate digital evidence to a single case/project when using the final ROXANNE platform.

UR-CM02-CaseEdit: As a law enforcement practitioner, I want to be able to edit information, about an existing case, so that I can keep its description up-to-date.

A criminal case can be broken down into four phases: creation, ongoing investigation, prosecution and closure. Law enforcement agents should be able to update the case status as well as its description (e.g., homicide). Nevertheless, this is not a critical requirement.

UR-CM03-CaseOverview: As a law enforcement practitioner, I want to be able to view a high-level description of an existing case, so that I can quickly pick it up from where I left off or associate it to other cases.

Law enforcement agents may be working on several open cases at the same time, or they may need to check an older case for similar patterns or suspect behaviour. Although, this is an important requirement for a mature investigative analysis platform, it is considered as optional for the ROXANNE platform.

UR-CM04-CaseRestrictedCollaboration: As a law enforcement practitioner, I want to be able to invite colleagues to explore an existing case in view-mode only, so that I can take advantage of synergies in experience and background.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

Law enforcement agents frequently share information and assertions with their colleagues for brainstorming, communicate their understanding for decision making or in case the latter have findings that are of interest. However, there are also times when security settings and processes suggest that other analysts and investigators cannot update information, or the latter cannot share the case with their colleagues. During the 1st field test one participant had explicitly mentioned that *"it would be appropriate to include an option in order to limit access to some cases"* [3]. Along the same lines, the chart below, which is taken from [1], suggests that the possibility of collaboration with others is a useful feature, but not the mandated by the majority of respondents.

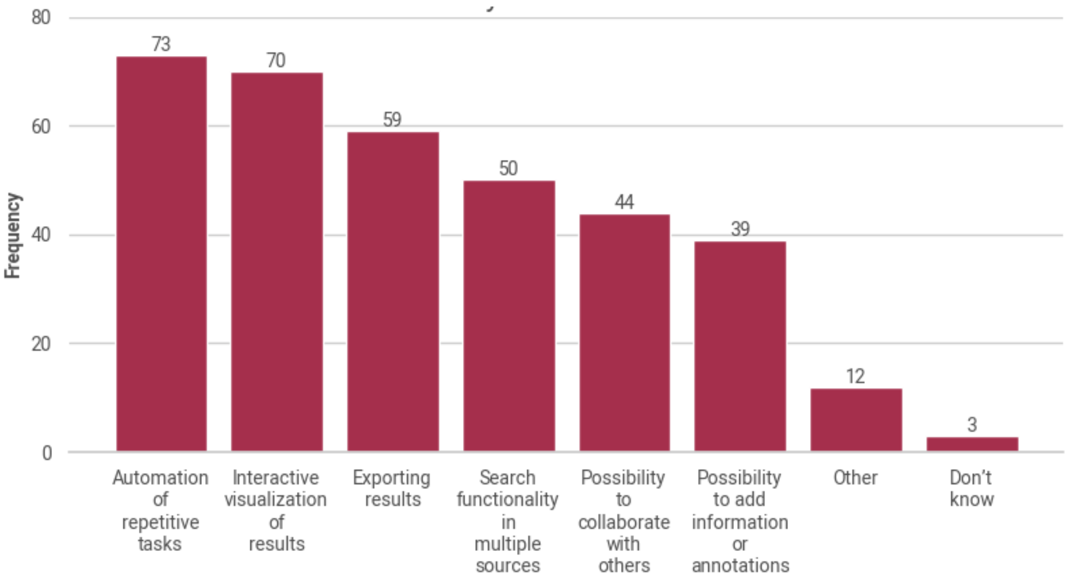


Figure 2: Popularity of features for analysis or search tool as expressed by respondents to online questionnaire on user requirements

UR-CM05-CaseParallelWorkspaces: As a law enforcement practitioner, I want to be able to invite colleagues to investigate an existing case in a separate workspace, so that we can work in parallel and at the same time avoid conflicts.

This requirement is similar to UR-CM04-CaseRestrictedCollaboration, but (selected) colleagues should be able to work on the same case, but on separate instances, for example, by testing their own hypotheses, adding notes, etc. As in the previous user requirement, it is deemed as useful for the final version of the ROXANNE platform.

UR-CM06-CaseIntegrity: As a law enforcement practitioner, I want to be able to keep the original data (raw files) and see a log of changes to a case so that I can track the history.

Persistently keeping the original input data in a structured way that permits rapid and accurate access (e.g., via indexes and cross-references) is considered important for integrity purposes. Furthermore, logging changes/refinements to the ROXANNE outputs is desirable for traceability purposes.

Nevertheless, the priority of this user requirement for the ROXANNE platform is reduced due to its experimental nature.

UR-CM07-CaseScenarios: As a law enforcement practitioner, I want to be able to create alternative scenarios of the same case, so that I can test alternative hypotheses.

The following table summarizes the user requirements related to case management.

Table 4: User requirements related to case management

Requirement ID	Short requirement description	Priority (final version)
UR-CM01-CaseCreation	As a law enforcement practitioner, I want to be able to create a new case, so that I can keep track of all relevant information in a single place.	Must-have
UR-CM02-CaseEdit	As a law enforcement practitioner, I want to be able to edit information, about an existing case, so that I can keep its description up-to-date.	Could-have
UR-CM03-CaseOverview	As a law enforcement practitioner, I want to be able to view a high-level description of an existing case, so that I can quickly pick it up from where I left off, or associate it to other cases.	Could - have
UR-CM04-CaseRestrictedCollaboration	As a law enforcement practitioner, I want to be able to invite colleagues to explore an existing case in view-mode only, so that I can take advantage of synergies in experience and background.	Should-have
UR-CM05-CaseParallelWorkspaces	As a law enforcement practitioner, I want to be able to invite colleagues to investigate an existing case in a separate workspace, so that we can work in parallel and at the same time avoid conflicts.	Should-have
UR-CM06-CaseIntegrity	As a law enforcement practitioner, I want to be able to keep the original data (raw files) and see a log of changes to a case so that I can track the history.	Should-have
UR-CM07-CaseScenarios	As a law enforcement practitioner, I want to be able to create alternative scenarios of the same case, so that I can test alternative hypotheses.	Must-have

2.4 Evidence processing

UR-EP01-EvidenceUpload: As a law enforcement practitioner, I want to be able to upload new files to one or more cases, so that key persons, their interactions, as well as entities (such as locations and

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

topics mentioned) are automatically recognized and I can spend more time on interpreting these results.

To build up a full picture of an investigation, the law enforcement agent should be able to upload input files that are available from a range of sources, such as: data retrieved from third-parties as part of ongoing targeting operations (such as call-detail records and wiretapped conversations from telecommunication service providers, CCTV footage from road operators, etc.), classified information collected by specifically tasked covert means (such as photographs of suspects), as well as open source intelligence (OSINT) that is publicly available (like discussion fora). Thus, some input files are following a specific format (e.g., call-detail records), while others are unstructured (e.g., text). According to the law enforcement representatives' feedback after the first field test, users should be able to upload multimodal input (images, video, audio and text). In particular, almost half of respondents to the online survey on LEAs' requirements reported that images are frequently fully used today, followed by text and video (around 40%) and then audio and geographic locations (around 33% of participants) [1]. Furthermore, users should be allowed to upload data that are already processed using other investigation platforms (such as those described in Section 4) or third-party software (e.g., for transcription of speech into text). Data should not be restricted to a single case. Automatically processing multimodal inputs are important for the final version of ROXANNE platform, given that automation of repetitive tasks was recognized by law enforcement agents as the most important feature to be delivered (see Figure 2 above).

UR-EP02-EvidenceAlert: As a law enforcement practitioner, I want to be able to get a warning before uploading a new set of raw files to an existing case, so that the integrity of that case is not jeopardized.

While some information may be available at the beginning of an investigation, new data may become available at later stages. Thus, it is important to be able to continuously upload files to existing cases. Seeing a warning when attempting to do so could avoid accidental data upload to the wrong case. Furthermore, during the first field test, one of the participants had suggested to see a warning every time mutual data occurs. Such warnings are acknowledged to be useful for mature investigation platforms but are considered as optional for the ROXANNE platform.

UR-EP03-MetadataUpload: As a law enforcement practitioner, I want to be able to upload additional metadata (especially call-detail records) to the ROXANNE platform, so that further sources of information are automatically added in the case work.

Users should be allowed to input metadata, especially data from trusted sources and the ones reflecting prior human knowledge. If a case includes metadata only (e.g., call detail records of some suspects) then one would perform an investigation based on the initial knowledge only. This requirement is deemed important.

UR-EP04-ConfigureWorkflow: As a law enforcement practitioner, I want to be able to configure the processing chain to be executed, so that I can choose the optional features to be enabled and in case of substitute components to choose the most appropriate to be used.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

Given the variety of input data available to law enforcement agents, a wide range of tools are needed for discovering new forensic artifacts or augmenting existing ones. Some of these tools may be optional and thus the user should be able to choose whether some of them will be excluded from the processing chain. For example, during the first field test it was mentioned that *"it would be beneficial to include an option to translate texts extracted from one language into another, including also in general more European languages"*, but translation may not always be required. In other cases, substitute tools may be present, and the user should be able to choose the one to be activated. Based on the responses of the participants to the field tests, this requirement is considered to be of high priority.

UR-EP05-MonitorWorkflow: As a law enforcement practitioner, I want to be able to monitor at real-time the status of the files under processing, so that I can plan my activities and be aware of any failures to complete a step.

Uploading or processing the evidence data is lengthy process, thus users should be able to see the status of the task. To this end, a law enforcement practitioner during the first field test mentioned that it would be useful to use an intuitive colour coding scheme, e.g., green colour indicates successful processing of a particular file, yellow suggests that hasn't started/finished yet, while red means that there were problems in processing that particular file. This requirement is of medium priority for the final version of the ROXANNE platform.

UR-EP06-ControlWorkflow: As a law enforcement practitioner, I want to be able to control the process on-the-fly, so that I can abort the processing if it takes too long, or wrong files/configuration have been used.

Given the abundance of input data to be analysed by law enforcement agents and the importance of gaining actionable insights as soon as possible, users should be able to abort the processing of data if they think that they realized that validating another inference will be beneficial. As in the previous case, this requirement is of medium priority for the ROXANNE platform.

UR-EP07-MonitorResources: As a law enforcement practitioner, I want to be able to see information on the utilisation level of hardware resources.

In order to be able to control the workflow a user should be capable of monitoring the hardware resources (e.g., server Central Processor Unit (CPU), Graphics Processing Unit (GPU), memory, etc.) used for that task. Given that all modern operating systems support this feature, it is useful only if processing takes place in servers. For this reason, this requirement is of low priority.

The following table summarizes the user requirements related to evidence processing.

Table 5: User requirements related to evidence processing

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

Requirement ID	Short requirement description	Priority (final version)
UR-EP01-EvidenceUpload	As a law enforcement practitioner, I want to be able to upload new files to one or more cases, so that key persons, their interactions, as well as entities (such as locations and topics mentioned) are automatically recognized and I can spend more time on interpreting these results.	Must-have
UR-EP02-EvidenceAlert	As a law enforcement practitioner, I want to be able to get a warning before uploading a new set of raw files to an existing case, so that the integrity of that case is not jeopardized.	Could-have
UR-EP03-MetadataUpload	As a law enforcement practitioner, I want to be able to upload additional metadata (especially call-detail records) to the ROXANNE platform, so that further sources of information are automatically added in the case work.	Must-have
UR-EP04-ConfigureWorkflow	As a law enforcement practitioner, I want to be able to configure the processing chain to be executed, so that I can choose the optional features to be enabled and in case of substitute components to choose the most appropriate to be used.	Must-have
UR-EP05-MonitorWorkflow	As a law enforcement practitioner, I want to be able to monitor at real-time the status of the files under processing, so that I can plan my activities and be aware of any failures to complete a step.	Could-have
UR-EP06-ControlWorkflow	As a law enforcement practitioner, I want to be able to control the process on-the-fly, so that I can abort the processing if it takes too long, or wrong files/configuration have been used.	Could-have
UR-EP07-MonitorResources	As a law enforcement practitioner, I want to be able to see information on the utilisation level of hardware resources.	Could-have

2.5 Case analysis

UR-CA01-TemporalAnalysis: As a law enforcement practitioner, I want to be able to visualise the evolution of events in time, so that I can identify patterns or focus on certain key periods.

Events containing temporal data (i.e., timestamps) are considered important as these represent dates and times of phone call, meetings, interactions, transactions etc. between entities. Depending on the type of case, e.g., a homicide, the timing of the event can be the starting point of the investigation. In other cases, the law enforcement agent would need to "travel back in time" and identify groups of events that occurred close together in time. Establishing regular patterns of activity is important in

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

identifying suspicious behaviour. As evidenced from the chart below, timeline visualizations were ranked the third most useful visualization feature. Thus, we consider that this is an important feature that must be supported by the ROXANNE platform.

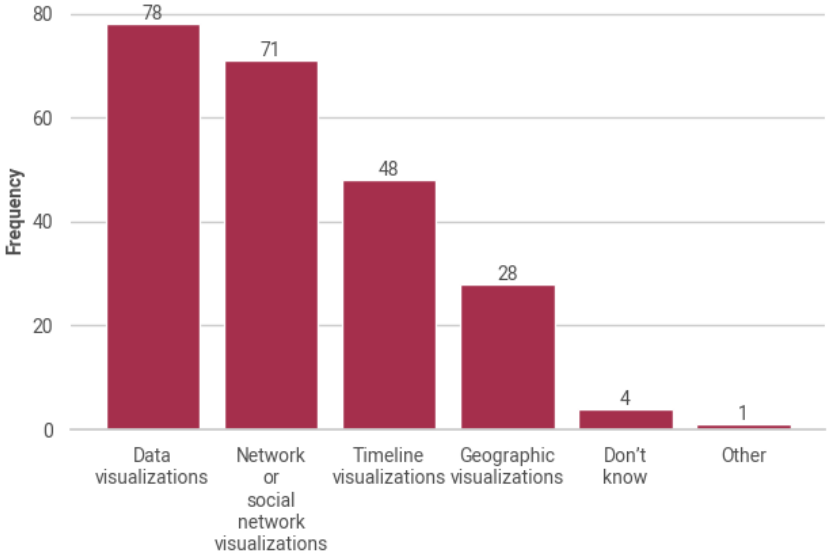


Figure 3: Most useful type of visualizations as expressed by respondents to online questionnaire on user requirements

UR-CA02-GeoSpatialAnalysis: As a law enforcement practitioner, I want to be able to visualise the location of events, so that I can identify patterns related to the spatial dimension or focus on certain geographical areas.

Like timing of events, the location where a person was photographed, a meeting/conversation took place can be important for the investigation process. Viewing the entities and events on a map helps law enforcement practitioners in correlating events that took place close to each other, movement patterns of suspects, etc. As shown in Figure 3 above, geographic visualizations are useful to some respondents and thus the priority of this user requirement is reduced compared to other visualization schemes. With regards to the final ROXANNE platform, geospatial analysis is of medium priority.

UR-CA03-EntityRelationshipsAnalysis: As a law enforcement practitioner, I want to be able to visualise the interactions of key entities, so that I can identify patterns related to the tactics, techniques and procedures of the suspects.

Investigators think in terms of entities (e.g., people), events and their relationships, thus viewing such associations can help them discover similarities and ultimately solve criminal network investigations. This can be done by seeking answers to the following indicative set of questions:

- Who is central in this case?
- Who forms a bridge or liaison between distinct organizations?

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

- Who is the initiator of the interactions?
- What communities are formed?
- Are there any links between distinct communities?

As noted by a participant to the first field test, *"structures typically evolve and new ones emerge during the investigation, requiring continuous re-structuring of entities and their associations".* Given that these associations are dynamic, understanding how these communities evolve over time is important for answering questions like:

- Have the links changed over time?
- Are the links changing in strength or centrality?

Criminal network visualizations were demonstrated during the first field test and found to be a very powerful scheme for most participating law enforcement agents (see Figure 4 below). Thus, this requirement is of high priority.

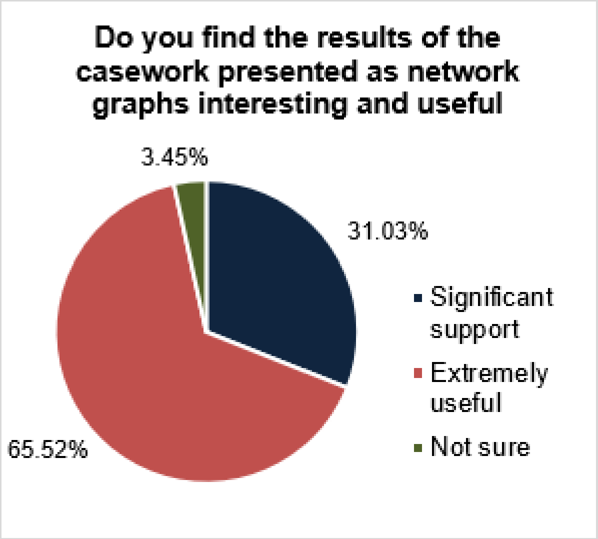


Figure 4: Usefulness of criminal social networks that were the main visualization chart that was demonstrated during the first field test

UR-CA04-AdvancedFiltering: As a law enforcement practitioner, I want to be able to filter the results using several properties at the same time, so that I can test a wide range of hypotheses.

Law enforcement agents participating in the field tests highlighted the importance of filters in data-reduction during visualization, as more hiding unnecessary entities or events can help during investigations. For example, a participant to the first field test recognised the need to "alter" the chart by hiding those entities that the investigator believed (or ROXANNE algorithms suggest) to be unrelated and/or irrelevant to the case. Furthermore, some law enforcement practitioners proposed that ROXANNE platform allows users to create their own custom filters. Such filters could be on entity properties, event properties or even input data used. Filters should be applied on all properties of entities and when multiple filters are defined only those entities that pass all active entity filters shall appear. Similarly, when multiple filters are activated on events, only elements that meet all constraints shall be visible. When both entity and event filters are activated, results that do not meet any of the constraints should be hidden.

Advanced filters were also considered important by participants of the following field tests also. For example, a practitioner involved in the 3rd field test explicitly requested the ability to hide from the graphs any files that were used for enrolling users (as these do not involve a second party).

Advanced filters were recognized as important by the majority of the respondents to the online user requirements survey and thus needs to be supported in the final ROXANNE platform release. Furthermore, the police representatives during the 3rd field-test stated that the implemented filters are necessary for the investigation [5].

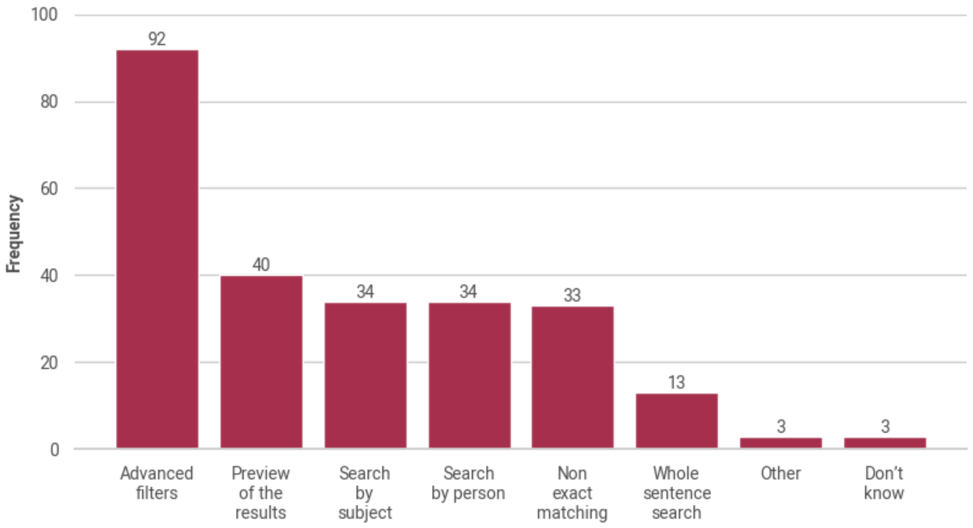


Figure 5: Most useful features for a search tool as expressed by respondents to online questionnaire on user requirements

UR-CA05-StatisticalAnalysis: As a law enforcement practitioner, I want to be able to visualise statistics about events, so that I can identify patterns based on data attributes

Visualising statistics about entities and events using pie charts, bar charts, histograms etc is important for analysts in identifying:

- Patterns related to a particular case.
- Emerging issues in an area and by acting promptly to prevent a situation from deteriorating or developing.
- Longer-term/strategic issues, as well as the scope of criminality.

According to Figure 3 above, the "data visualisation" requirement was the most popular visualisation feature and thus is highly prioritised.

UR-CA06-MultidimensionalAnalysis: As a law enforcement practitioner, I want to be able to see several diagrams at the same time, so that I can combine different types of data referring to the same case and along several dimensions.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

The basic problem for intelligence analysts is putting information together in an organized way so the difficult task of extracting meaning from the assembled information is made easier. For example, during the first field test it was mentioned that criminal networks are dynamic and some of the methods used for visualisation and analysis of the networks can be improved by combining the temporal dimension. This approach could be extended to take into consideration the geospatial information (being able, for example, to pinpoint the exact location of a suspect phone call), as well as the statistical view. Such interactive visualisation of results was found to be acknowledged as an important feature by some 70 participants (ranked second) to the user requirements survey of ROXANNE (see Figure 3 above). Thus, this requirement is interpreted as an important one and needs to be included in the final release of the ROXANNE platform.

UR-CA07-UserAnnotation: As a law enforcement practitioner, I want to be able to add annotations, so that I can keep notes of important details and include information derived from external sources.

Law enforcement officers need to be able add information or annotations based on their prior knowledge about a case, or when drawing interim hypothesis(es). This ability has been identified as useful by almost 33% of the participants to the user requirements survey (39 responses out of 121 in total as can be seen in Figure 3 above).

UR-CA08-OutputRevision: As a law enforcement practitioner, I want to be able to revise any analysis outputs I find erroneous, so that ROXANNE platform improves over time based on my expertise.

Law enforcement practitioners are aware that no technology is error-free and thus they expect that revisions are allowed by users that are authorised to do so. For example, during the first field test it was mentioned that *"it would be great if the entity representation could be changed as to allow a better distinction between the entities involved"*. These revisions should be communicated to the respective ROXANNE backend module(s), if AI (Artificial Intelligence) technology is used so that respective models are retrained. This requirement is of high priority and needs to be supported by the ROXANNE platform.

UR-CA09-CustomRepository: As a law enforcement practitioner, I want to be able to add proprietary keywords and any complementary information/metadata, so that I can improve the accuracy of transcription, topic detection, etc. with my expertise on the matter.

This requirement is similar to the previous one, but backend modules are directly (rather than implicitly) instructed so that accuracy metrics are improved. For example, during the first field test it was suggested that "topics detection" would be significantly improved if LEAs could supply their own topics and link them with their specific keywords allowing ROXANNE system to detect and categorize them. This is required as every police force has different needs and cases and have to take into consideration the different languages spoken, jargons or nicknames frequently used.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

UR-CA10-OutputCharacterisation: As a law enforcement practitioner, I want to be able to judge the quality of predicted properties, so that I can assess the accuracy of ROXANNE outputs.

Given that technologies are not error-free, law enforcement officers should be able to judge the quality of ROXANNE outputs, especially the ones that employ AI techniques (e.g., for topic extraction). This requirement is widely recognized as an important one and has been brought up during the first field test (a participant had asked *"Why we see what we see?"*). Thus, it should be supported by the ROXANNE platform.

UR-CA11-OutputComparison: As a law enforcement practitioner, I want to be able to compare two different charts of the same case, so that I can easily validate my hypotheses and identify any patterns.

Law enforcement practitioners would appreciate the ability to compare different graphs of the same case. This requirement is considered as optional.

UR-CA12-OutputMerging: As a law enforcement practitioner, I want to be able to combine insights from previous cases, so that I can associate other cases with the main one.

Law enforcement officers expect that the ROXANNE platform will have the ability to retain what it learns. For example, the following question was raised during the first field test: *"Is the information analysed in case X used in the analysis of the case Y, or in case Y the platform starts from the beginning?"* Thus, this requirement is considered to be important.

UR-CA13-AutomatedSuggestions: As a law enforcement practitioner, I want to be able to get system suggestions, so that I can uncover any hidden relationships or patterns.

Law enforcement practitioners could make use of system suggestions regarding, for example, the importance of each individual identified, or any interactions and relations among individuals in a network that are hidden or not observed in the existing dataset. ROXANNE will offer several ways a) to quantify the influence that individuals have over the others within a social network, b) to measure the probability that two nodes are connected, c) to identify cohesive subgroups of individuals/entities who interact more often among them than with other individuals in the network. For example, a participant to the 3rd field test, recommended that suggestions should be presented, even if the odds are low, as it would be better to check a relationship that is probably not relevant than to have missing information. Nevertheless, this requirement is considered to be optional given that incorrect ROXANNE results was the biggest concern among the participants to the user requirements survey (see Figure 6, below).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

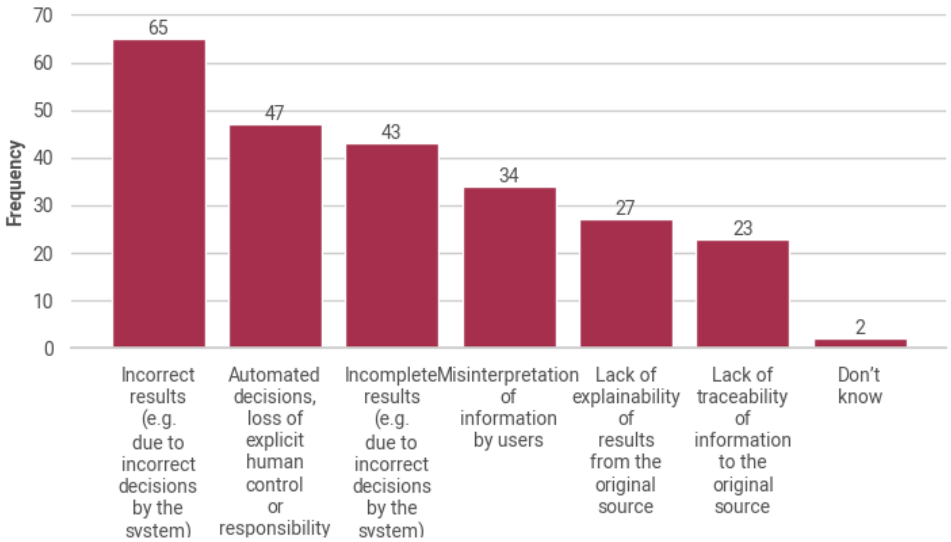


Figure 6: Major concerns with regards to analysis or search tools as expressed by respondents to ROXANNE questionnaire on user requirements

UR-CA14-ExportedOutputs: As a law enforcement practitioner, I want to be able to export results of the analysis to other formats, so that I can import these to other investigative tools.

Participants to the first field test indicated that it would be of their benefit if the interface could link with existing platforms within an environment that they as an LEA have control over, or even if the ROXANNE solution could work as a plugin, rather than a platform on its own. This requirement was found to be an important one for any future commercialisation activities of the ROXANNE platform.

UR-CA15-Reporting: As a law enforcement practitioner, I want to be able to generate a report of the analysis, so that I can use it offline.

As soon as the law enforcement officer is confident enough that the insights obtained during the analysis phase are actionable, they typically compile reports that include graphs, tables or charts with statistical information on current and potential criminal activity. These reports are useful for communicating the findings to other colleagues so that next steps are decided. Supporting this user requirement is important.

The following table summarizes the user requirements related to case analysis.

Table 6: User requirements related to case analysis

Requirement ID	Short requirement description	Priority (final version)
UR-CA01-TemporalAnalysis	As a law enforcement practitioner, I want to be able to visualise the evolution of events in time, so that I can identify patterns or focus on certain key periods.	Must-have

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

Requirement ID	Short requirement description	Priority (final version)
UR-CA02-GeoSpatialAnalysis	As a law enforcement practitioner, I want to be able to visualise the location of events, so that I can identify patterns related to the spatial dimension or focus on certain geographical areas.	Should-have
UR-CA03-EntityRelationshipsAnalysis	As a law enforcement practitioner, I want to be able to visualise the interactions of key entities, so that I can identify patterns related to the tactics, techniques and procedures of the suspects.	Must-have
UR-CA04-AdvancedFiltering	As a law enforcement practitioner, I want to be able to filter the results using several properties at the same time, so that I can test a wide range of hypotheses.	Must-have
UR-CA05-StatisticalAnalysis	As a law enforcement practitioner, I want to be able to visualise statistics about events, so that I can identify patterns based on data attributes	Must-have
UR-CA06-MultidimensionalAnalysis	As a law enforcement practitioner, I want to be able to see several diagrams at the same time, so that I can combine different types of data referring to the same case and along several dimensions.	Must-have
UR-CA07-UserAnnotation	As a law enforcement practitioner, I want to be able to add annotations, so that I can keep notes of important details and include information derived from external sources.	Should-have
UR-CA08-OutputRevision	As a law enforcement practitioner, I want to be able to revise any analysis outputs I find erroneous, so that ROXANNE platform improves over time based on my expertise.	Must-have
UR-CA09-CustomRepository	As a law enforcement practitioner, I want to be able to add proprietary keywords and any complementary information/metadata, so that I can improve the accuracy of transcription, topic detection, etc. with my expertise on the matter.	Must-have
UR-CA10-OutputCharacterisation	As a law enforcement practitioner, I want to be able to judge the quality of predicted properties, so that I can assess the accuracy of ROXANNE outputs.	Should-have
UR-CA11-OutputComparison	As a law enforcement practitioner, I want to be able to compare two different charts of the same case, so that I can easily validate my hypotheses and identify any patterns.	Could-have
UR-CA12-OutputMerging	As a law enforcement practitioner, I want to be able to combine insights from previous cases, so that I can associate other cases with the main one.	Must-have
UR-CA13-AutomatedSuggestions	As a law enforcement practitioner, I want to be able to get system suggestions, so that I can uncover any hidden relationships or patterns.	Could-have

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

Requirement ID	Short requirement description	Priority (final version)
UR-CA14-ExportedOutputs	As a law enforcement practitioner, I want to be able to export results of the analysis to other formats, so that I can import these to other investigative tools.	Must-have
UR-CA15-Reporting	As a law enforcement practitioner, I want to be able to generate a report of the analysis, so that I can use it offline.	Must-have

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

3. System requirements

In this section, we present the transformation of the user expectations and requirements to technical specifications for the ROXANNE platform, i.e., a detailed description of how the functional and non-functional requirements listed in Section 2 should be implemented. These system requirements span the following categories:

- Platform management (analysed in Section 3.1)
- user management (elaborated in Section 3.2)
- case management (analysed in Section 3.3)
- evidence processing (elaborated in Section 3.4), and
- case analysis (described in Section 3.5).

Each system requirement is identified using the following format: **<SR-FunctionalAreaX-ShortName>,** where:

- **SR** stands for System Requirement
- **FunctionalArea** refers to any of the following acronyms:
  - PM for Platform Management
  - UM for User Management
  - CM for Case Management
  - EP for Evidence Processing
  - CA for Case Analysis
- **X** is the incremental number of the functional area and
- **ShortName** is a distinctive name.

3.1 Platform management

SR-PM01-NativeInstallation: The ROXANNE platform should be deployable by executing some scripts. This non-functional requirement is related to UR-PM01-MultiOSplatform.

SR-PM02-VirtualMachine: The ROXANNE platform should be deployable by opening an image using third-party Virtual Machine software. This non-functional requirement is an alternative to SR-PM01-NativeInstallation above and thus is related to UR-PM01-MultiOSplatform.

SR-PM03-UpdatedPlatform: The ROXANNE platform, upon user confirmation, should be updated to the latest version, including the latest documentation. This non-functional requirement is related to UR-PM02- EasyToMaintainPlatform.

SR-PM04-UpdatedDatasetsUsedByPlatform: The ROXANNE platform and upon user confirmation should download newly available datasets from the default data repository. This non-functional requirement is related to UR-PM02- EasyToMaintainPlatform.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

3.2 User management

SR-UM01-NewUserAccountRequest: The ROXANNE Graphical User Interface (GUI) shall allow an eligible user to create an account for new users, by providing the appropriate administrative details. This non-functional requirement is related to UR-UM01-UserAccountCreation user requirement.

SR-UM02-NewUserAccountRegistration: The ROXANNE platform will create a new user and store the relevant details. This non-functional requirement is related to UR-UM01-UserAccountCreation user requirement.

SR-UM03-UserAuthenticationRequest: The ROXANNE GUI shall allow a registered user to log-in to the system by providing the appropriate security details. This non-functional requirement is related to UR-UM05-UserLogIn.

SR-UM04-UserAuthenticationResponse: The ROXANNE platform will authenticate the user. This non-functional requirement is related to UR-UM05-UserLogIn.

SR-UM05-UserDetailsRequest: The ROXANNE platform will supply the details of a particular user. This non-functional requirement is related to UR-UM02-UserAccountView.

SR-UM06-UserDetailsView: The ROXANNE GUI shall allow a registered user to view details about her account. This non-functional requirement is related to UR-UM02-UserAccountView.

SR-UM07-UserDetailsUpdateRequest: The ROXANNE GUI shall allow a registered user to update her account details. This non-functional requirement is related to UR-UM03-UserAccountEdit.

SR-UM08-UserDetailsUpdateConfirmation: The ROXANNE platform will update the user's details. This non-functional requirement is related to UR-UM03-UserAccountEdit.

SR-UM09-UserDeleteRequest: The ROXANNE GUI shall allow a registered user to delete her account. This non-functional requirement is related to UR-UM04-UserAccountDelete.

SR-UM10-UserDeleteConfirmation: The ROXANNE platform will delete an existing user. This non-functional requirement is related to UR-UM04-UserAccountDelete.

3.3 Case management

SR-CM01-NewCaseCreationRequest: The ROXANNE GUI shall allow a registered user to create a new case by supplying relevant details, such as title, type of investigation, description and access control information, while other properties (e.g., case identifier, creator, date created) are automatically populated. This functional requirement is a response to UR-CM01-CaseCreation.

SR-CM02-NewCaseCreationConfirmation: The ROXANNE backend will create a new case and store the relevant details. This functional requirement is a response to UR-CM01-CaseCreation.

SR-CM03-CaseDescriptionUpdateRequest: The ROXANNE GUI shall allow a registered user to update the description of an existing case, e.g., its status, only if she is authorised to do so. This functional requirement is a response to UR-CM02-CaseEdit.

SR-CM04-CaseDescriptionUpdateConfirmation: The ROXANNE backend will store the updated case details. This functional requirement is related to UR-CM02-CaseEdit.

SR-CM05-UpdateActionLogging: The ROXANNE backend will log the update action. This functional requirement is a response to UR-CM02-CaseEdit.

SR-CM06-CaseDescriptionUpdateRejection: The ROXANNE GUI shall deter an authorised user from updating non-editable description fields of an existing case, e.g., its creation date. This functional requirement is a response to UR-CM02-CaseEdit.

SR-CM07-FailedUpdateActionLogging: The ROXANNE backend will log the failed update attempt. This functional requirement is related to UR-CM02-CaseEdit.

SR-CM08-UserCasesRequest: The ROXANNE GUI shall request a list of all existing cases a registered user has access to. This functional requirement is a response to UR-CM03-CaseOverview.

SR-CM09-AuthorisedCasesRetrieval: The ROXANNE backend will list the cases she is authorised to view. This functional requirement is related to UR-CM03-CaseOverview.

SR-CM10-AuthorisedCasesListing: The ROXANNE GUI shall present all existing cases she has access to, along with a high-level description of each one. This functional requirement is a response to UR-CM03-CaseOverview.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

SR-CM11-LimitedAccessGrantingRequest: The ROXANNE GUI shall enable an authorised user to grant "view-only" access to other registered users. This functional requirement is a response to UR-CM04-CaseRestrictedCollaboration.

SR-CM12-LimitedAccessGrantingConfirmation: The ROXANNE backend will update the user privileges, so that that s/he can formulate inferences using the ROXANNE GUI, but not revise any evidence or adjust any ROXANNE output. This functional requirement is related to UR-CM04-CaseRestrictedCollaboration.

SR-CM13-FullAccessGrantingRequest: The ROXANNE GUI shall enable an authorised user to grant full access to other registered users for an existing case. This functional requirement is a response to UR-CM05-CaseParallelWorkspaces.

SR-CM14-FullAccessGrantingConfirmation: The ROXANNE backend shall create a new, personalised instance of an existing case for each user who has full access on that case. This functional requirement is a response to UR-CM05-CaseParallelWorkspaces.

SR-CM15-LoggedEventsRequest: The ROXANNE GUI shall allow an authorised user to request a log of changes related to the case under investigation. This functional requirement is a response to UR-CM06-CaseIntegrity.

SR-CM16-LoggedEventsResponse: The ROXANNE backend will compile the log. This functional requirement is a response to UR-CM06-CaseIntegrity.

SR-CM17-LoggedEventsListing: The ROXANNE GUI shall allow an authorised user to see a log of changes on a certain case. This functional requirement is related to UR-CM06-CaseIntegrity.

SR-CM18-AlternativeScenarios: The ROXANNE GUI shall allow the user to work on various scenarios at the same time. This functional requirement is a response to UR-CM07-CaseScenarios.

3.4 Evidence processing

SR-EP01-AudioFileUploading: The ROXANNE GUI shall allow an authorised user to upload a new audio file or batch of new audio files. This functional requirement is related to UR-EP01-EvidenceUpload.

SR-EP02-VideoFileUploading: The ROXANNE GUI shall allow an authorised user to upload a new video file or batch of new video files. This functional requirement is related to UR-EP01-EvidenceUpload.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

SR-EP03-TextFileUploading: The ROXANNE GUI shall allow an authorised user to upload a new text file or batch of new text files. This functional requirement is related to UR-EP01-EvidenceUpload.

SR-EP04-EvidenceFilePersistentStorage: The ROXANNE backend will persistently store the new file or batch of files. This functional requirement is related to UR-EP01-EvidenceUpload.

SR-EP05-IncrementalEvidenceWarning: The ROXANNE GUI shall show a warning if change to an existing case is attempted. This functional requirement is related to UR-EP02-EvidenceAlert.

SR-EP06-MetaDataFileUploading: The ROXANNE GUI shall allow an authorised user to upload a file containing additional information about the files to be processed, e.g., phone numbers present in Call Detail Records (CDR). This functional requirement is related to UR-EP03-MetadataUpload.

SR-EP07-EvidenceWorkflowConfiguration: The ROXANNE GUI shall allow an authorised user to configure the processing that will take place. This functional requirement is related to UR-EP04-ConfigureWorkflow.

SR-EP08-RealTimeEvidenceProcessingStatusUpdate: The ROXANNE GUI shall show information on the activities that are completed, pending, ongoing or failed. This functional requirement is related to UR-EP05-MonitorWorkflow.

SR-EP09-RealTimeEvidenceProcessingControl: The ROXANNE GUI shall allow an authorised user to stop the processing. This functional requirement is related to UR-EP06-ControlWorkflow.

SR-EP10-EvidenceProcessingControlWarning: The ROXANNE GUI shall show a warning that process will be interrupted. This functional requirement is related to UR-EP06-ControlWorkflow.

SR-EP11-RealTimeInfrastructureStatusUpdate: The ROXANNE GUI shall show real-time information on the utilisation level of server resources. This functional requirement is related to UR-EP07-MonitorResources.

3.5 Case analysis

SR-CA01-ProcessedEventsRequest: The ROXANNE GUI shall request the events related to the case under investigation from ROXANNE backend. This functional requirement is related to UR-CA01-TemporalAnalysis, UR-CA02-GeoSpatialAnalysis, UR-CA03-EntityRelationshipsAnalysis and UR-CA06-MultidimensionalAnalysis.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

SR-CA02-ProcessedEventsResponse: The ROXANNE backend will serve the processed results based on user preferences, including any quality scores such as likelihood ratios, to ROXANNE GUI in JSON format. This functional requirement is related to UR-CA01-TemporalAnalysis, UR-CA02-GeoSpatialAnalysis, UR-CA03-EntityRelationshipsAnalysis and UR-CA06-MultidimensionalAnalysis.

SR-CA03-PresentEventsEvolution: The ROXANNE GUI shall be able to present the evolution of events based on timestamp information. This functional requirement is related to UR-CA01-TemporalAnalysis.

SR-CA04-ShowEventOverview: The ROXANNE GUI shall be able to present high-level information about individual events on mouse hover. This functional requirement is related to UR-CA01-TemporalAnalysis, UR-CA02-GeoSpatialAnalysis, UR-CA03-EntityRelationshipsAnalysis and UR-CA06-MultidimensionalAnalysis.

SR-CA05-ShowEventDetails: The ROXANNE GUI shall be able to present detailed information about selected events. This functional requirement is related to UR-CA01-TemporalAnalysis, UR-CA02-GeoSpatialAnalysis, UR-CA03-EntityRelationshipsAnalysis and UR-CA06-MultidimensionalAnalysis.

SR-CA06-PresentEventsLocation: The ROXANNE GUI shall be able to show the location where different events took place, based on geographic information. This functional requirement is related to UR-CA02-GeoSpatialAnalysis.

SR-CA07-FinetuneAlgorithmParameters: The ROXANNE GUI shall allow the user to adjust the level of sensitivity for clustering biometrical data (e.g., voiceprints) and to select network analysis algorithm and to configure its parameters. This functional requirement is related to UR-CA03-EntityRelationshipsAnalysis.

SR-CA08-PresentSocialNetwork: The ROXANNE GUI shall be able to present a social network[6] graph containing persons and other entities (as nodes), as well as their interactions (as directed edges). This functional requirement is related to UR-CA03-EntityRelationshipsAnalysis.

SR-CA09-ShowCommunitiesOfEntities: The ROXANNE GUI shall be able to group entities that appear to be part of the same community. This functional requirement is related to UR-CA03-EntityRelationshipsAnalysis.

SR-CA10-AdjustSocialNetworkAppearance: The ROXANNE GUI shall be able to adjust the appearance of the social network graph (e.g., reposition nodes). This functional requirement is related to UR-CA03-EntityRelationshipsAnalysis.

[6] Even though some of those individuals may be suspects of criminal activities, innocent entities can also appear. For this reason, in this report, we use the term Social Network graphs or Social Network Analysis, rather than Criminal Social Network, as a way to understand the interactions and relationships of individuals that are included or referenced in the files to be processed.

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

SR-CA11-ChooseSocialNetworkZoomLevel: The ROXANNE GUI shall be able to focus on certain parts of the social network graph by zooming-in/out. This functional requirement is related to UR-CA03-EntityRelationshipsAnalysis.

SR-CA12-SocialNetworkScrolling: The ROXANNE GUI shall be able to focus on certain parts of the social network graph by moving in all four directions. This functional requirement is related to UR-CA03-EntityRelationshipsAnalysis.

SR-CA13-ShowEntityOverview: The ROXANNE GUI shall be able to present high-level information about individual entities on mouse hover. This functional requirement is related to UR-CA01-TemporalAnalysis, UR-CA02-GeoSpatialAnalysis, UR-CA03-EntityRelationshipsAnalysis and UR-CA06-MultidimensionalAnalysis.

SR-CA14-ShowEntityDetails: The ROXANNE GUI shall be able to present detailed information about selected entities. This functional requirement is related to UR-CA01-TemporalAnalysis, UR-CA02-GeoSpatialAnalysis, UR-CA03-EntityRelationshipsAnalysis and UR-CA06-MultidimensionalAnalysis.

SR-CA15-TranscribedAudioPlayback: The ROXANNE GUI shall allow users listen to audio files and view automatically generated transcription. This functional requirement is related to UR-CA01-TemporalAnalysis, UR-CA02-GeoSpatialAnalysis, UR-CA03-EntityRelationshipsAnalysis and UR-CA06-MultidimensionalAnalysis.

SR-CA16-VideoPlayback: The ROXANNE GUI shall allow users to see a video footage where key entities appear. This functional requirement is related to UR-CA01-TemporalAnalysis, UR-CA02-GeoSpatialAnalysis, UR-CA03-EntityRelationshipsAnalysis and UR-CA06-MultidimensionalAnalysis.

SR-CA17-FilterEntitiesByProperty: The ROXANNE GUI shall support filtering on entities and their properties. This functional requirement is related to UR-CA04-AdvancedFiltering.

SR-CA18-FilterEventsByTiming: The ROXANNE GUI shall support filtering on the timing of events. This functional requirement is related to UR-CA04-AdvancedFiltering.

SR-CA19-FilterEventsByLocation: The ROXANNE GUI shall support filtering on the location of events. This functional requirement is related to UR-CA04-AdvancedFiltering.

SR-CA20-FreeTextSearch: The ROXANNE GUI shall support free-text search on processed text, e.g., phone call transcriptions. This functional requirement is related to UR-CA04-AdvancedFiltering.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

SR-CA21-ActivateSyncedFilters: The ROXANNE GUI shall be able to adjust the content of visible charts based on all activated filters. This functional requirement is related to UR-CA04-AdvancedFiltering.

SR-CA22-DeactivateFilters: The ROXANNE GUI shall be able to reset all activated filters. This functional requirement is related to UR-CA04-AdvancedFiltering.

SR-CA23-ShowStatisticalResults: The ROXANNE GUI shall present statistical charts (e.g., pie charts, bar charts, etc.) for different data attributes (such as language, gender, age etc.). This functional requirement is related to UR-CA05-StatisticalAnalysis.

SR-CA24-ChooseVisibleCharts: The ROXANNE GUI shall allow charts to be deactivated for saving screen estate or improving responsiveness. This functional requirement is related to UR-CA06-MultidimensionalAnalysis.

SR-CA25-ResizeVisibleCharts: The ROXANNE GUI shall allow charts to be resized according to user preferences/needs and technical constraints. This functional requirement is related to UR-CA06-MultidimensionalAnalysis.

SR-CA26-RelocateVisibleCharts: The ROXANNE GUI shall allow charts to be relocated for customising screen real estate. This functional requirement is related to UR-CA06-MultidimensionalAnalysis.

SR-CA27-SaveDashboardConfiguration: The ROXANNE GUI shall be able to save the configuration of dashboard. This functional requirement is related to UR-CA06-MultidimensionalAnalysis.

SR-CA28-LoadDashboardConfiguration: The ROXANNE GUI shall be able to load a previously saved dashboard configuration. This functional requirement is related to UR-CA06-MultidimensionalAnalysis.

SR-CA29-AnnotateOnSocialNetworks: The ROXANNE GUI shall allow annotations to be saved on the social network graph. This functional requirement is related to UR-CA07-UserAnnotation and UR-CA03-EntityRelationshipsAnalysis.

SR-CA30-MergeEntities: The ROXANNE GUI shall allow two or more entities (e.g., speaker A and speaker B), who are believed to be the same entity A, to be merged into one, unless the user is not authorised to do so. Entity B will be deleted and any events that were initially associated to B will be linked to entity A. This functional requirement is related to UR-CA08-OutputRevision and UR-CA03-EntityRelationshipsAnalysis.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

SR-CA31-SplitEntities: The ROXANNE GUI shall allow an event (e.g., audio channel), which is believed to be associated with another entity (e.g., B instead of A), to be attributed to entity B, unless the user is not authorised to do so. If entity B does not already exist, a new one will be created with user-supplied properties (e.g., name, age, etc.). This functional requirement is related to UR-CA08-OutputRevision and UR-CA03-EntityRelationshipsAnalysis.

SR-CA32-AddEntities: The ROXANNE GUI shall allow a new entity to be manually added. This functional requirement is related to UR-CA08-OutputRevision and UR-CA03-EntityRelationshipsAnalysis.

SR-CA33-AddEvents: The ROXANNE GUI shall allow a new event to be manually added. This functional requirement is related to UR-CA08-OutputRevision and UR-CA03-EntityRelationshipsAnalysis.

SR-CA34-UpdateEntityProperties: The ROXANNE GUI shall allow entity properties to be updated by authorised users. This functional requirement is related to UR-CA08-OutputRevision.

SR-CA35-EntityPropertiesUpdateRejection: The ROXANNE GUI shall deter an authorised user from updating the properties of an entity. This functional requirement is related to UR-CA08-OutputRevision.

SR-CA36-UpdateEventProperties: The ROXANNE GUI shall allow event properties to be updated by authorised users. This functional requirement is related to UR-CA08-OutputRevision.

SR-CA37-EnventPropertiesUpdateRejection: The ROXANNE GUI shall deter an authorised user from updating the properties of an event. This functional requirement is related to UR-CA08-OutputRevision.

SR-CA38-AddContextSpecificKeywords: The ROXANNE GUI shall allow context-specific keywords to be added or tagged. Furthermore, for each keyword a couple of sentences along with pronunciation information (like in vocabularies). This functional requirement is related to UR-CA09-CustomRepository.

SR-CA39-SupportQualityAssessment: The ROXANNE GUI shall document confidence levels on predicted properties of entities and/or events, such as language, age, etc. This functional requirement is related to UR-CA10-OutputCharacterisation.

SR-CA40-CompareDifferentInferences: The ROXANNE GUI shall allow the user to make a side-by-side comparison of charts that are produced from the same case but using two different sets of filters. This functional requirement is related to UR-CA11-OutputComparison.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

SR-CA41-AssociateDifferentCases: The ROXANNE GUI shall allow the user to make a side-by-side comparison of charts that are produced from different cases. This functional requirement is related to UR-CA12-OutputMerging.

SR-CA42-HighlightImportantEntities: The ROXANNE GUI shall be able to indicate entities that appear to be highly active. This functional requirement is related to UR-CA13-AutomatedSuggestions and UR-CA03-EntityRelationshipsAnalysis.

SR-CA43-ShowPredictedRelationships: The ROXANNE GUI shall indicate missing links between entities. This functional requirement is related to UR-CA13-AutomatedSuggestions and UR-CA03-EntityRelationshipsAnalysis.

SR-CA44-ExportResultsUsingPopularFormats: The ROXANNE GUI shall be able to export analysis results to popular formats that can be imported to other investigative tools. This functional requirement is related to UR-CA14-ExportedOutputs.

SR-CA45-GenerateCaseAnalysisReport: The ROXANNE GUI shall be able to export analysis results as a report. This functional requirement is related to UR-CA15-Reporting.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

4. The ROXANNE technical Platform

This section describes the overall ROXANNE platform, which aims at enhancing criminal network analysis capabilities of Law Enforcement practitioners. The ROXANNE platform is a suite of multi-modal technologies for extracting evidence and actionable intelligence and an effective forensics visualisation scheme providing an environment that streamlines the presentation, discovery and elaboration of these insights.

4.1 ROXANNE platform architecture

Figure 7 presents the high-level overview of the ROXANNE platform. The ROXANNE platform consists of four (4) families of technologies that are closely integrated: a) Speech-related technologies, b) Text-processing technologies, c) Video-processing technologies and d) Network analysis methods. The LEA practitioner uses the Graphical User Interface to configure the processing that should take place, to provide any intelligence already available, to explore the automated results and, if and when needed, to provide expert knowledge back to the platform by validating and updating the outputs.

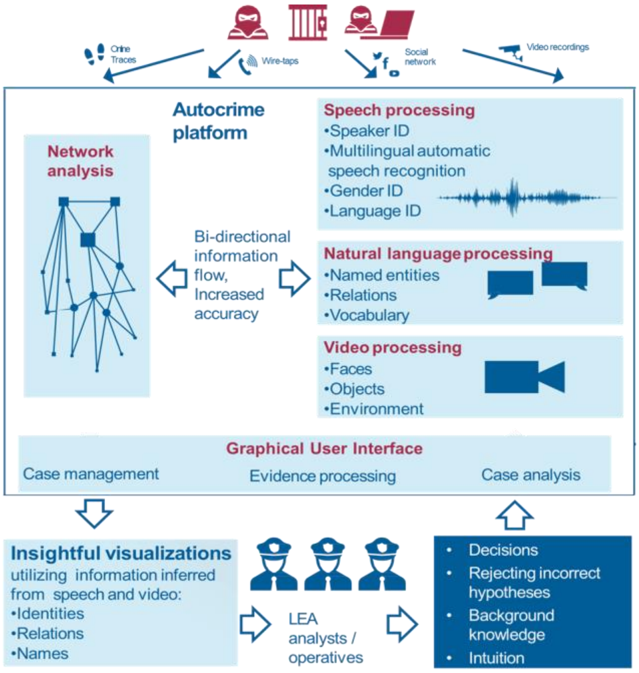


Figure 7: A high-level overview of the ROXANNE platform

4.1.1 The Initial ROXANNE platform (V1) architecture

The initial architecture had adopted two recent architectural paradigms; namely Microservices, and Event-Driven Architecture (EDA).

The former suggests that complex applications are broken into several independent and loosely coupled services, in the sense that each one: 1) is specialized in a single task and thus can be assigned

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

to smaller teams, 2) is autonomous and self-sufficient as it has its own codebase, state (if needed) and runs in its own container that improves fault tolerance and resilience, 3) can communicate with each other using APIs either synchronously or asynchronously and 4) is scalable as each individual services can be scaled horizontally and automatically based on its specific needs.

The latter paradigm focuses on handling and processing events to enable decoupled and asynchronous communication between components. These events can range from user actions to system notifications and external stimuli. Instead of direct interactions or method calls, components publish events to event brokers. Other components that act as consumers, can subscribe to these events, and perform any required actions asynchronously. By maintaining a log of events that have occurred, the system state can be reconstructed. This asynchronous nature together with persistent logging, allows autonomous operation while enabling scalability and fault tolerance.

As evident from above, these two paradigms can be combined. For example, Event brokers acting as a central communication subsystem for event collection and asynchronous distribution can run as a Microservice for automated scalability.

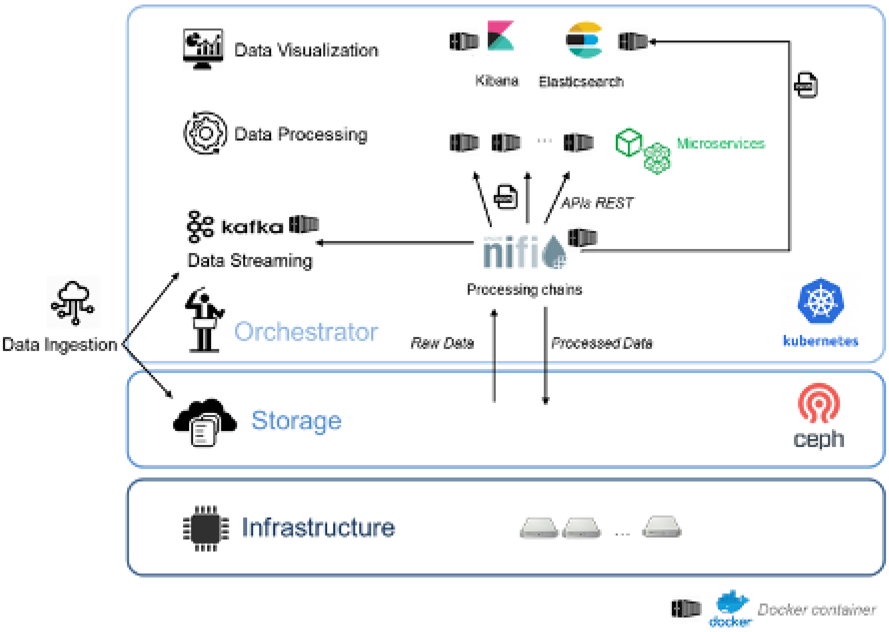


Figure 8: The initial architecture of the ROXANNE platform

The initial, cloud native ROXANNE platform used the following technologies (more details can be found in D7.1 and D7.2):

- Docker for providing a way to run individual ROXANNE technologies in a securely isolated manner, by packaging all required dependencies and libraries in a container following the microservices paradigm.
- Kubernetes for managing the lifecycles of docker containers, especially in large, dynamic environments.
- Apache Kafka serves as the messaging backbone where individual ROXANNE technologies can publish their outputs and/or retrieve results from other technologies.
- Nifi is used for orchestrating the data processing by listening to Kafka input topics, invoking the appropriate microservices to process / enrich data and finally by pushing processed data in Elasticsearch indices, which is periodically retrieved for visualization purposes.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

- Ceph for providing persistent data storage to the stateful technologies (e.g., those that need to maintain system state regardless of the scaling up/down of the rest technologies) such as Elasticsearch,Apache NiFi and Apache Kafka.
- Elasticsearch as a real-time indexing, search and analytics engine, enabling efficient searching and analysis of the data ingested through NiFi and Kafka by visualization frameworks (such as Kibana).

While combining these two approaches offers numerous advantages, it can also introduce delays in software development process and maintenance. For example, the ROXANNE platform was originally conceived as **a bi-directional system** between multimodal technologies (such as speaker identification, automatic speech recognition, entity recognition and resolution, and face/place/background identification) and criminal network analysis algorithms. This was for the consortium considered as the most innovative idea to be researched, developed and eventually integrated. Given the Research and Innovation nature of the ROXANNE project, significant effort has been placed in conceptualising how different technologies can interoperate, building prototypes and integrating these so that each proof-of-concept is validated. As expected, by introducing a change in one service, changes in several other services due to cross-service dependencies were frequently triggered. At the same time, technologies were co-developed by several technical partners (e.g., see Table 1 of D5.3), which added further complexities to the development and maintenance process.

Following the microservices paradigm in the first half of the project, when cross-service dependencies were many and were keep evolving, resulted in increased coordination and development effort to keep these services independent and autonomous. When an issue or bug was identified, tracing the flow of data and events across multiple services increased the time and effort required to diagnose and resolve issues, leading to maintenance delays. Furthermore, as ROXANNE evolved over time, event schemas had to be updated. Adding new fields, modifying existing fields, or deprecating old fields in event payloads while supporting backward compatibility (whenever needed) as a direct consequence of following the Event-Driven Architecture, created further challenges.

4.1.2 The final ROXANNE (V2) platform architecture

In order to streamline the continuous research, development and integration process and simplify the continuous testing by end-users, it was decided to rather follow a traditional layered architecture that could be later ported to the two modern paradigms mentioned above in order to be able to scale up/down horizontally (i.e., to increase/decrease the number of instances per individual component according to demand), as well as to scale-up/down vertically to easily support batch processing of input data, or parallelization of the input data (in case a cluster computing HW is availale). This would allow to process a larger amount of data that are typically exist in large cases.

According to the layered architecture paradigm, an application is organised into distinct layers, each responsible for specific functionality. Typically, these layers include presentation, business logic, and data storage layers. Figure 9 presents the ROXANNE platform architecture that includes the frontend for interacting with the user, the backend that generates the outputs, as well as the filesystem that stores all required inputs, outputs and configurations.

The Graphical User Interface of ROXANNE is a web application in JavaScript that is offered as a set of docker images and builds upon a wide set of state-of-the-art technologies, namely:

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

- the Angular[7] framework for developing interactive applications (either single-page or large ones) in a maintainable manner, e.g., by including custom implementations of visualisation libraries such as graph-based visualizations, time-dependent data, etc.
- the NodeJS[8] open-source server environment that performs business logic activities for producing dynamic web page content and interfacing with the local filesystem.
- the nginx[9] web server that is deployed for serving static or dynamic web pages, as well as a proxy for all requests.
- the Keycloak[10] Identity and Access Management solution for defining and enforcing privileged access to the ROXANNE platform and relevant case data.

The ROXANNE backend is composed of a set of technologies for a) Speech processing, b) Text-processing, c) Network analysis and d) Video processing. Furthermore, it includes a number of python scripts for:

i. integrating the core ROXANNE technologies,
ii. exporting the outputs of ROXANNE platform to the format used by external commercial tools (i.e., I2 Analyst Notebook),
iii. exposing the backend capabilities as a set of Hypertext Transfer Protocol (HTTP) methods that can be consumed by any frontend application. These methods provide the following capabilities:
    a. running a particular case on the ROXANNE platform based on the user's configuration related to the input files to be processed, the technologies to be enabled and their parameters, as well as the scenario where the results will be stored. Special care is taken if a user reruns a particular case so that any changes are taken into consideration.
    b. running the link prediction feature, which provides the three most probable predicted links for a target node.
    c. clearing link prediction results, which removes all predicted links from the case output.
    d. merging the outputs of two cases that the user believes these share some common entities (suspects or victims). The user supplies pairs of nodes that are common across the two cases and the Autocrime platform will suggest additional pairs of common nodes based on the properties of the two networks. The merged cases will be saved as a new scenario (of the working case) and thus the original cases will not be affected.

[7] https://angular.io/

[8] https://nodejs.org/en/

[9] https://www.nginx.com/

[10] https://www.keycloak.org/

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

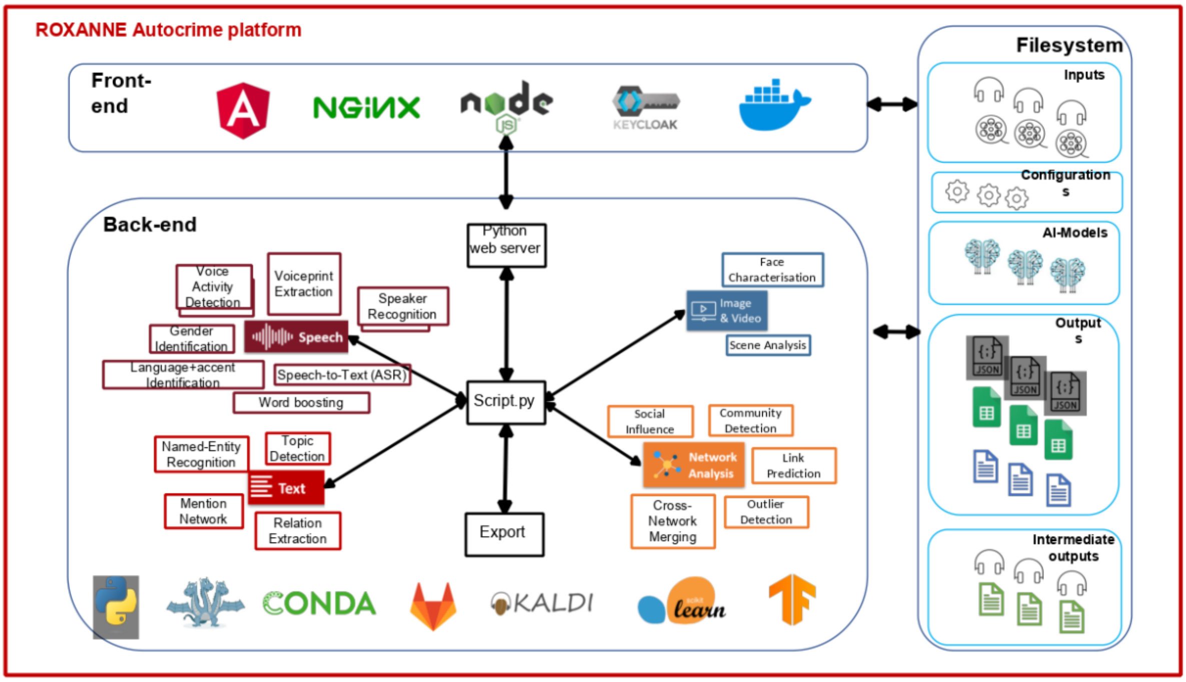


Figure 9: The final ROXANNE (V2) platform architecture

The backend also makes use of state-of-the-art technologies for

- supporting several configurations dynamically using the Hydra[11] framework
- supporting several virtual environments, each one with its own packages and their dependencies, using Conda[12] system
- managing the software development and integration process using Gitlab[13] with Gitlab's CI/CD
- AI-models using scikit-learn and TensorFlow[14]
- Speech recognition using Kaldi[15] and Pytorch
- Sphinx to generate documentation

Both the frontend and backend interact with the local filesystem for storing and retrieving files to be processed, configurations, AI-models, intermediate outputs as well as the main results of the ROXANNE platform. There are several folders; the "core" ones, which contain the main algorithms, are "network", "speech" and "nlp". Each of these folders contains the individual ROXANNE technologies. For example, the "speech" folder includes the following subfolders:

- "gender" for gender identification
- "asr" for automatic speech recognition
- "lid" for language identification

[11] https://hydra.cc/

[12] https://docs.conda.io/en/latest/

[13] https://about.gitlab.com/

[14] https://www.tensorflow.org/

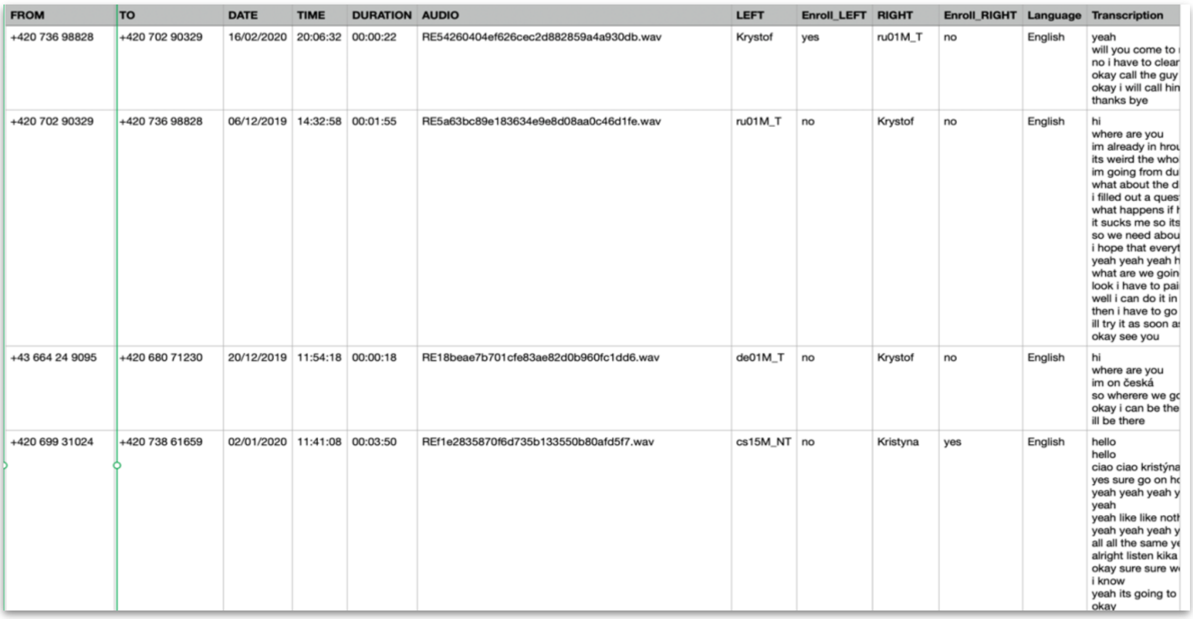
[15] https://kaldi-asr.org/

- vad_diar for voice activity detection and diarisation
- sid for speaker identification

Furthermore, there is a common data folder in which users can upload their data via the Graphical User Interface. Input data supported are:

- calls: audio in all formats are supported, but are downsampled to 8 kHz. In addition, files that are not in WAV format (e.g. mp3, amr, etc.) are converted to single-channel files.
- video and images
- and a Comma Separated Values (CSV) file describing the evidence files along with any additional information from intelligence sources.

Figure 10 presents a screenshot of a sample CSV file that includes key information about the evidence files.



FROM	TO	DATE	TIME	DURATION	AUDIO	LEFT	Enroll_LEFT	RIGHT	Enroll_RIGHT	Language	Transcription
+420 736 98828	+420 702 90329	16/02/2020	20:06:32	00:00:22	RE54260404ef626cec2d882859a4a930db.wav	Krystof	yes	ru01M_T	no	English	yeah will you come to no i have to clear okay call the guy okay i will call hin thanks bye
+420 702 90329	+420 736 98828	06/12/2019	14:32:58	00:01:55	RE5a63bc89e183634e9e8d08aa0c46d1fe.wav	ru01M_T	no	Krystof	no	English	hi where are you im already in hrou its weird the who im going from du what about the d i filled out a ques what happens if h it sucks me so its so we need abou i hope that everyt yeah yeah yeah h what are we goin look i have to pai well i can do it in then i have to go ill try it as soon a okay see you
+43 664 24 9095	+420 680 71230	20/12/2019	11:54:18	00:00:18	RE18beae7b701cfe83ae82d0b960fc1dd6.wav	de01M_T	no	Krystof	no	English	hi where are you im on česká so wherere we go okay i can be the ill be there
+420 699 31024	+420 738 61659	02/01/2020	11:41:08	00:03:50	REf1e2835870f6d735b133550b80afd5f7.wav	cs15M_NT	no	Kristyna	yes	English	hello hello ciao ciao kristýna yes sure go on ho yeah yeah yeah y yeah yeah like like noth yeah yeah yeah y all all the same ye alright listen kika okay sure sure w i know yeah its going to okay

Figure 10: A sample file containing information about the evidence files to be processed by the ROXANNE platform

The headers that appear in Figure 10 are explained in the list below:

- "FROM": the telephone number of the caller
- "TO": the telephone number of the receiver
- "DATE": date of the recording in YYYY-MM-DD hh:mm:ss format
- "TIME": the time of the recording in hh:mm:ss format
- "DURATION": length of the recording in seconds
- "AUDIO": name of the audio file. e.g. recording1.wav
- "LEFT": speaker identity of the left channel in the audio
- "Enroll_LEFT": yes/no depending on whether the speaker on the left channel in the audio is already known and will be provided by the user (see “LEFT” header)
- "RIGHT": speaker identity of the right channel in the audio

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

- "Enroll_RIGHT": yes/no depending on whether the speaker on the right channel in the audio is already known and will be provided by the user (see "RIGHT" header)
- "Language": language of the recording. e.g., English, Czech
- "Transcription": transcripts of the recording

Additional headers may be present, such as the following geolocation information:

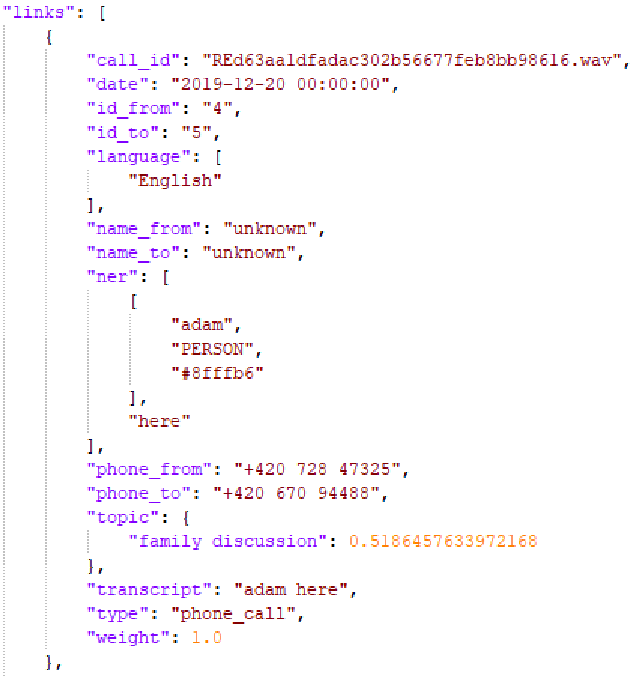
- "CALLER_START_BTS_LAT", that includes latitude information of the base station that serviced the caller
- "CALLER_START_BTS_LON", that includes longitude information of the base station that serviced the caller
- "RECEIVER_START_BTS_LAT", that includes latitude information of the base station that serviced the receiving party
- "RECEIVER_START_BTS_LON", that includes longitude information of the base station that serviced the receiver

The outputs of the ROXANNE platform are stored in the "output" folder[16] on a file called "result.json". The JSON file has seven top-level keys:

1. links: a list of objects with the following keys (see Figure 11 for an example)
    - "call_id", a unique identifier of the evidence file (e.g., call)
    - "date", providing a timestamp of the file (e.g., when a call took place)
    - "id_from", describing the node id who placed a call
    - "id_to" , describing the node id who received a call
    - "language": the language used in the call
    - "observed": a Boolean value describing whether the link is actual (true) or predicted (false)
    - "ner": a list of words detected by NER
    - topic: the most probable topic automatically detected (if topic detection module was run).
    - "transcript": either the original transcript copied from the CSV file, or the transcript generated by ASR (the latter overrides the former)
    - "type": providing a description of the link type (e.g., "phone_call", "image" for appearing on a video/image file, or "mention" for referring to a third party.
    - "weight": a floating-point value representing the number of occurrences of a link (typically 1 for phone calls, but can be any positive value in case of mentions)

[16] That can be found under ~/autocrime/data/{case_name}

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

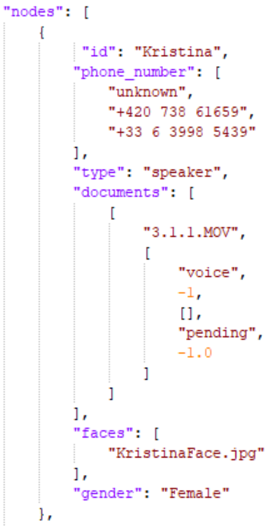


```
"links": [
    {
        "call_id": "REd63aa1dfadac302b56677feb8bb98616.wav",
        "date": "2019-12-20 00:00:00",
        "id_from": "4",
        "id_to": "5",
        "language": [
            "English"
        ],
        "name_from": "unknown",
        "name_to": "unknown",
        "ner": [
            [
                "adam",
                "PERSON",
                "#8fffb6"
            ],
            "here"
        ],
        "phone_from": "+420 728 47325",
        "phone_to": "+420 670 94488",
        "topic": {
            "family discussion": 0.5186457633972168
        },
        "transcript": "adam here",
        "type": "phone_call",
        "weight": 1.0
    },
```

Figure 11: Screenshot of a sample “link” object that is part of the result.json generated by the ROXANNE platform

2. nodes: a list of objects with the following keys (see Figure 12 for an example)
    - “id”: a unique identifier of the node that can be the enrolled name (if the person is known) or a number that is automatically assigned by the ROXANNE platform
    - “phone_number”: a list of phone numbers that a speaker was found to be using (empty if node refers to a scene, object or mentioned entity)
    - “type”: providing information about the type of the node, being "speaker" (for persons whose voice was identified), "mention" (for entities whose name was mentioned) or “media” (in case of a scene or object)
    - “documents” a list of image/video files that were associated with this file (e.g., due to voice heard)
    - “faces”, which points to an image file that was enrolled by the user and contains a known face of the node
    - “gender”: that can be "male", "female" or “unknown” (e.g., if node refers to a scene, object or mentioned entity)

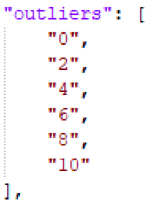
This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.



```
"nodes": [
    {
        "id": "Kristina",
        "phone_number": [
            "unknown",
            "+420 738 61659",
            "+33 6 3998 5439"
        ],
        "type": "speaker",
        "documents": [
            [
                "3.1.1.MOV",
                [
                    "voice",
                    -1,
                    [],
                    "pending",
                    -1.0
                ]
            ]
        ],
        "faces": [
            "KristinaFace.jpg"
        ],
        "gender": "Female"
    },
```

Figure 12: Screenshot of a sample “node” object that is part of the result.json generated by the ROXANNE platform

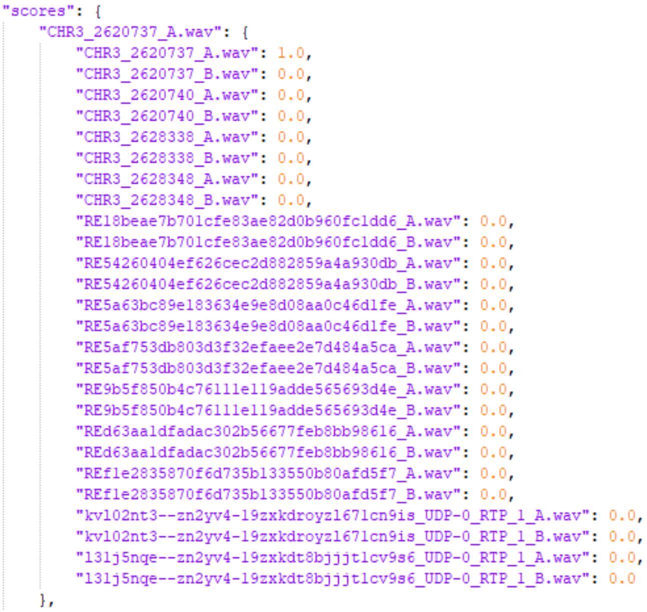
3. outliers: a list of node id’s that were found to be an outlier (see Figure 13Figure 12 for an example)



```
"outliers": [
    "0",
    "2",
    "4",
    "6",
    "8",
    "10"
],
```

Figure 13: Screenshot of a sample “outliers” object that is part of the result.json generated by the ROXANNE platform

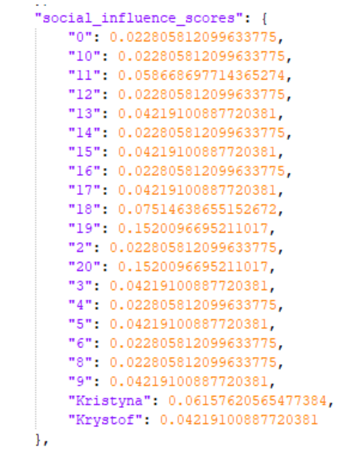
4. scores: a list of pairwise similarities of audio channels (see Figure 14Figure 12 for an example)



```
"scores": {
    "CHR3_2620737_A.wav": {
        "CHR3_2620737_A.wav": 1.0,
        "CHR3_2620737_B.wav": 0.0,
        "CHR3_2620740_A.wav": 0.0,
        "CHR3_2620740_B.wav": 0.0,
        "CHR3_2628338_A.wav": 0.0,
        "CHR3_2628338_B.wav": 0.0,
        "CHR3_2628348_A.wav": 0.0,
        "CHR3_2628348_B.wav": 0.0,
        "RE18beae7b701cfe83ae82d0b960fc1dd6_A.wav": 0.0,
        "RE18beae7b701cfe83ae82d0b960fc1dd6_B.wav": 0.0,
        "RE54260404ef626cec2d882859a4a930db_A.wav": 0.0,
        "RE54260404ef626cec2d882859a4a930db_B.wav": 0.0,
        "RE5a63bc89e183634e9e8d08aa0c46d1fe_A.wav": 0.0,
        "RE5a63bc89e183634e9e8d08aa0c46d1fe_B.wav": 0.0,
        "RE5af753db803d3f32efaee2e7d484a5ca_A.wav": 0.0,
        "RE5af753db803d3f32efaee2e7d484a5ca_B.wav": 0.0,
        "RE9b5f850b4c76111e119adde565693d4e_A.wav": 0.0,
        "RE9b5f850b4c76111e119adde565693d4e_B.wav": 0.0,
        "REd63aa1dfadac302b56677feb8bb98616_A.wav": 0.0,
        "REd63aa1dfadac302b56677feb8bb98616_B.wav": 0.0,
        "REf1e2835870f6d735b133550b80afd5f7_A.wav": 0.0,
        "REf1e2835870f6d735b133550b80afd5f7_B.wav": 0.0,
        "kvl02nt3--zn2yv4-19zxkdroyzl671cn9is_UDP-0_RTP_1_A.wav": 0.0,
        "kvl02nt3--zn2yv4-19zxkdroyzl671cn9is_UDP-0_RTP_1_B.wav": 0.0,
        "l31j5nqe--zn2yv4-19zxkdt8bjjjtlcv9s6_UDP-0_RTP_1_A.wav": 0.0,
        "l31j5nqe--zn2yv4-19zxkdt8bjjjtlcv9s6_UDP-0_RTP_1_B.wav": 0.0
    },
```

Figure 14: Screenshot of a sample "scores" object that is part of the result.json generated by the ROXANNE platform

5. Social Influence scores: a list of influence score of each person identified (see Figure 15Figure 12 for a sample output)

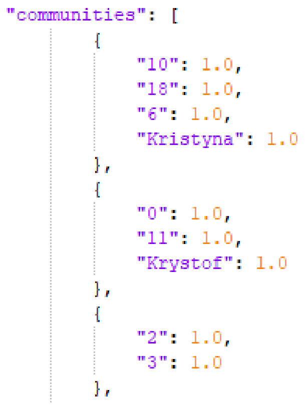


```
..
"social_influence_scores": {
    "0": 0.022805812099633775,
    "10": 0.022805812099633775,
    "11": 0.058668697714365274,
    "12": 0.022805812099633775,
    "13": 0.04219100887720381,
    "14": 0.022805812099633775,
    "15": 0.04219100887720381,
    "16": 0.022805812099633775,
    "17": 0.04219100887720381,
    "18": 0.07514638655152672,
    "19": 0.1520096695211017,
    "2": 0.022805812099633775,
    "20": 0.1520096695211017,
    "3": 0.04219100887720381,
    "4": 0.022805812099633775,
    "5": 0.04219100887720381,
    "6": 0.022805812099633775,
    "8": 0.022805812099633775,
    "9": 0.04219100887720381,
    "Kristyna": 0.06157620565477384,
    "Krystof": 0.04219100887720381
},
```

Figure 15: Screenshot of a sample "social influence scores" object that is part of the result.json generated by the ROXANNE platform

6. Communities: a list of communities identified and their members (see Figure 16 for an example)

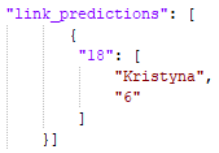
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.



```
"communities": [
        {
            "10": 1.0,
            "18": 1.0,
            "6": 1.0,
            "Kristyna": 1.0
        },
        {
            "0": 1.0,
            "11": 1.0,
            "Krystof": 1.0
        },
        {
            "2": 1.0,
            "3": 1.0
        },
```

Figure 16: Screenshot of a sample "communities" object that is part of the result.json generated by the ROXANNE platform

7. Predicted Links: a list of candidate links that were predicted by the ROXANNE platform (see Figure 17 for an example)



```
"link_predictions": [
        {
         "18": [
             "Kristyna",
             "6"
         ]
       }]
```

Figure 17: Screenshot of a sample "link predictions" object that is part of the result.json generated by the ROXANNE platform

Note that the ROXANNE platform can be ported to a cloud-native solution as underlying technologies are modern ones and well-maintained by the relevant communities. With regards to the ROXANNE backend technologies, a python 3.7 base image could be used as the starting point of each container, in which their individual dependencies (e.g., python libraries, frameworks, and other external components) that are already documented will be installed. Then the python code of each technology will be copied into the Docker container, followed by setting up the environment variables, entry points and exposed ports. When it comes to the ROXANNE frontend, it is already dockerised and the latest version includes the following images:

- nginx[17] that can be deployed using the following command: docker pull nginx
- fvt that can be deployed using the following command: docker pull andreasalexop/roxanne_new:fvt2022v17
- Frontend middleware that can be deployed using the following command: docker pull andreasalexop/roxanne_new:serv2022v17
- Keycloak that can be deployed using the following command: docker pull andreasalexop/roxanne_new:keycloak

Finally, the remaining middleware technologies that were present in the initial architecture (especially kafka, nifi, kubernetes) will be employed, as well as ceph for making sure that raw data (e.g., evidence files, case results, etc) are available across all technologies' instances.

In the following subsections, more details about each of the four core backend technologies are provided.

[17] https://hub.docker.com/_/nginx

4.2 Integrated technologies

Sections below give a brief description of integrated technologies in the ROXANNE final (V2) platform. For more details about performance of each technology, please refer to D5.3.

4.2.1 Speech technologies

The following speech-related technologies are available in the current version of the ROXANNE platform:

- Voice Activity Detection (VAD) that distinguishes speech from non-speech in an audio signal. The output of a VAD module can be used to decide which parts of a recording should be processed by other technologies, such as speech recognition or topic detection. In the ROXANNE platform, several modules for VAD are available:
  - Energy-based VAD that assumes that regions of the signal with high energy are speech and regions with low energy are non-speech (and thus is not robust against noise).
  - Neural network VAD that builds upon a neural network trained on audio data, which was carefully labelled for speech and non-speech segments by human operators.
  - Multilingual-based VAD that leverages a multi-lingual ASR system to detect voice activity being robust to noise and language variabilities.
- Voiceprint extraction that uses signal processing and machine learning techniques for converting an audio signal to a fixed size vector (i.e., a collection of numbers), which is usually referred to as a *voiceprint* or *embedding*.
- Speaker Recognition, which refers to the process where a computer algorithm attributes a voice segment to a particular speaker based on the acoustic properties of the audio signal. Speaker recognition is a key task of the ROXANNE platform as criminal investigations usually involve recordings of several unknown parties. Speaker recognition can range from "speaker verification" (i.e., decide whether a test utterance is spoken by the specified enrolled/registered speaker or not), to "closed-set speaker identification" (i.e., similar to speaker verification but the test utterance should be attributed to any of a set of enrolled speakers), to "open-set speaker identification" (i.e., similar to "closed-set speaker identification" but a test utterance may be associated to a non-enrolled speaker), to "speaker clustering" (i.e., the system should infer the set of unique speakers and attribute them to the calls based on a set of unlabelled recordings). While performing speaker recognition, Speaker Diarisation is often involved (e.g., in mono recordings) to segment the recording into regions such that only one person speaks in a region and cluster the regions according to speaker identity.
- Gender identification that involves the classification of the audio signal to a gender based on the speech patterns.
- Language identification, which refers to automatic detection of a target language given a test segment, (further extended by accent identification as well to support language variations).
- Automatic Speech Recognition (ASR) that uses Artificial Intelligence techniques to transform human speech into readable text. The ASR module available supports English and German languages, while other models for several other languages, including Greek, Hebrew, Spanish, Arabic, Lithuanian. Czech and Dutch was added. This is made possible by the fact that our acoustic models are per-language fine-tuned from the XLSR-53 model which was self-trained with 53 languages. Furthermore, boosting of specific (i.e., highly informative) words is supported, which aims to significantly improve recognition of custom words (or word sequences) that the law enforcement practitioners frequently encounter. The words to be boosted are supplied by the user in the configuration file.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

4.2.2 Text technologies

The ROXANNE (V2) platform includes the following text-related technologies:

- Named-Entity Recognition (NER) module, which automatically extracts useful information – the named entities in particular – from text documents (e.g., speech transcripts) and thus assist law enforcement practitioners to quickly focus on key entities found, such as names, locations, dates, etc. without having to read vast amounts of textual data.
- Topic Detection that automatically infers the topic of a document (e.g., a transcription of a wiretapped phone call) so that users can identify it is about a particular theme, such as drug trafficking, Work, Family discussion, Meeting, etc.
- Mention Network automatically, which resolves the person names mentioned in the phone call and disambiguates these mentions into "Party" (i.e., caller or receiver) or "Third Party" (e.g., friends of both parties).
- Relation Extraction that extracts semantic relationships between a pair of entities, e.g., Alice in Paris.

All these technologies are developed as multilingual, supporting all 8 major languages (Lithuanian is excluded here, and topic detection was for the moment developed only for 5 languages, as we were not able to identify 3 other languages for fine-training – which are planned to be updated during the exploitation phase of the project).

4.2.3 Network Analysis technologies

Network analysis is a core feature of the ROXANNE platform that allows to generate and analyse a network of caller-callees, where nodes represent persons (e.g., speakers) or locations found on video/images and edges indicate a relationship (e.g., calls between two speakers). The following network analysis techniques are currently supported by the ROXANNE platform:

- Community Detection for identifying cohesive groups that may not be apparent to a human investigator (e.g., using the Modularity Maximization algorithm).
- Social Influence Analysis for ranking entities/persons based on their "importance", which is computed by taking into account the network structure (e.g., the centrality of a node and the centrality of its neighbours using the Pagerank algorithm).
- Link Prediction in order to identify any links for particular persons that although not evident in the files processed, these could be inferred based on the mathematical properties of the network (e.g., using the Jaccard Coefficient algorithm).
- Cross-Network Merging for suggesting nodes that appear to be common across different cases due to the similarity of their surrounding network.
- Outlier Detection for determining inactive entities, which could be either removed (non-target persons) or be analysed further (highly ranked criminals that rarely communicate with the rest members).

4.2.4 Video technologies

The ROXANNE platform, also referred to as Autocrime, includes two image/video analysis technologies that aims at enriching the insights gained from the previous backend technologies i.e.,

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

to add further edges and nodes on the social network to support investigations in an ethical way. These technologies are the following:

- Face characterization that combines existing and open-source face detection and embedding extraction model with new face clustering and face cluster summarization algorithms for summarising each input video in a set of representative "signatures". Eventually, these identifiers are used for performing face similarity evaluation between a limited set of manually enrolled face pictures and all ingested images and videos. By limiting the usage of facial technology only on key persons of interest (e.g., victims and suspects), mass surveillance is avoided while investigator needs are supported in a proportionate manner.
- Scene or object characterization for identifying, clustering and summarizing locations and objects found in the processed video/image files. As in the case of faces described above, some target locations and objects that are manually enrolled/associated with a particular victim or suspect, will be searched in all other ingested images or videos to try to find this location/objects in other documents. Eventually, new findings will be included in the network graph as new edges that connect a location/object with other persons found on the processed files even if no prior knowledge allows to associate this location to any person.

4.3 Evaluation of technologies

To ensure reproducibility of the results presented by each module provider, the backend contains scripts used to obtain the evaluation results on the ROXANNE Simulated Dataset (ROXSD). The relevant documentation is provided in the README file of the backend. This was the minimum requirement posed for all module providers — to provide meaningful evaluation of their modules, and relevant description in the README file.

Scripts are also provided in order to reproduce the case demonstrated during the Hands-on of Field Test-3. The following actions are made possible:

1. A script is provided to download and extract ROXSD dataset from the ROXANNE collaborative workspace — this script also downloads relevant models (ASR acoustic models, decoder, pre-computed outputs, etc.).
2. A script is provided to generate the CSV file to run the case — it modifies the original scenario used to collect ROXSD by making sure that (a) no telephone number is assigned to more than one speaker, (b) no speaker has more than two telephone numbers.

4.4 Overall workflow

Figure 18 presents a typical workflow that is supported by the ROXANNE platform.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

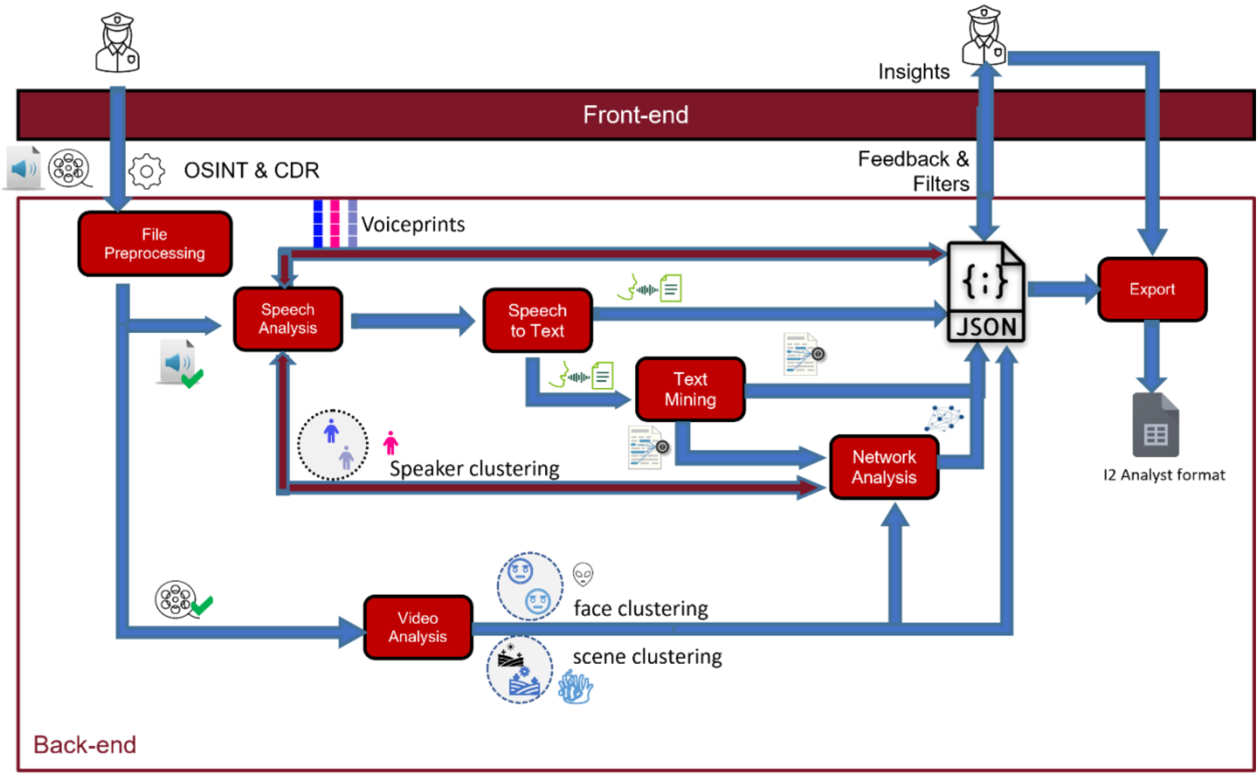


Figure 18: The overall workflow

The workflow is the following

1) The user creates a new case using the Graphical User Interface (frontend).
2) Then the user uploads the data to be processed using the Graphical User Interface and configures which technologies should be activated, supplies parameters and chooses where the outputs should be saved. The data includes the audio and video/images, as well as any metadata that may be available from Call Detail Records (CDR), intelligence available. Nevertheless, we assume that video files from seized phones will be uploaded later in the investigation process. The following table summarises the inputs that are expected by the ROXANNE platform and characterises their importance.

Table: Minimum required inputs and optional inputs to ROXANNE platform

Inputs to ROXANNE	**Short description of inputs**	**Mandatory / Optional**
Audio files	Audio files referring to wire-tapped phone calls, covert recordings, etc. to be processed as candidate evidence.	**Mandatory** Audio files are considered the starting point of an investigation and are considered minimum requirement for analysing a case.
Images / Photos	Still images referring to photos including persons, objects, scenes (or combinations) that are collected from Open-Source Intelligence (such as online	**Optional** Images/Photos can be useful in identifying additional relationships amongst speakers, i.e., to complement the findings obtained from speech analysis. For example, if two persons appear on the same photo then it is highly likely that are known to each other.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

	social networks), undercover agents, seized phones, etc. as candidate evidence	Similarly, if a person appears on a photo extracted from a seized phone, then it is very likely to be acquainted. Nevertheless, images are not mandatory for a case to be analysed in the first place.
Videos	Videos that include persons, objects, scenes, etc. that are collected from various sources as candidate evidence.	**Optional** Similar to images/photos, videos can be useful for identifying further relationships amongst speakers. In contrast to Images/Photos, a video may include one or more voice channels that can be also processed by speech (and text) technologies. Furthermore, one video may include multiple viewpoints of the same face(s) and multiple scenes. Nevertheless, videos are not mandatory for a case to be analysed in the first place.
Metadata / Intelligence	Additional data, information and knowledge from various sources (including telecommunication operators) about evidence audio files, images/photos and videos.	**Optional** This metadata can be useful for providing information on the characteristics of a person (e.g., by associating an audio file with a name the user can enrol a voiceprint to a speaker, by associating a facial image to a person the user can enrol a face to a suspect), phone numbers participating on a phone call, the timing of an event (a phone call placed, a photo taken, etc.), the location (e.g., of a base station serving a phone number), as well as transcription of the audio part obtained manually or from third-party software. While this metadata is not prerequisite for processing a case comprised only of audio files (the only required inputs to ROXANNE platform), it is needed if images and videos are also uploaded. The reason for this is that in the case of images and video files, the user needs to supply the enrolment information, while in the case of audio file the ROXANNE platform associates these with virtual phone numbers. In order to confirm this, the user can check that a network of speakers automatically identified by the ROXANNE platform has been created (see Figure 19a) even though no information on the phone numbers was used, resulting in a phone-based network that looks like a dandelion flower (see Figure 19b).
Case configuration	Case name and configuration of the technologies to be used, as well as their parameters.	**Mandatory** Each case needs to be unique, thus the user needs to provide a unique name upon case creation. While case configuration may involve many aspects of a case, supplying a name is the only required information. The rest configuration details are optional. In particular, specifying and configuring the technologies to be used is not mandatory as the ROXANNE platform comes out-of-the-box with a default configuration that can be used.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

		A special type of configuration includes any user-supplied information that can improve the accuracy of pre-trained Artificial Intelligence / Machine Learning models, such as words to be boosted during the Speech-to-text process or topics that are of interest. This information is needed for Topic Detection to run and ASR Boosting to work well in any context, nevertheless these features are optional and not need to be always provided.
User permissions per case	Information about registered users and their access rights per case.	**Mandatory** Given the sensitive nature of the ROXANNE platform, users need to be authenticated and authorised. Furthermore, each user can be granted permissions to add evidence files and/or to configure a case and/or to perform an analysis on a case-by-case basis. This means that the Identity and Access Management platform (i.e., Keycloak) needs to be properly configured by the administrator of the ROXANNE platform.

3) As soon as the case processing starts, the files are preprocessed, e.g., for checking the number of channels present on the audio files and, in case of stereo files, splitting these into two mono channels and storing these intermediate results on a dedicated folder (i.e., calls_mono)[18]. Furthermore, the audio channels on the video files, if any, are extracted and stored locally.
4) All supported audio files, including these extracted from video files, are then processed by the selected speech technologies. Note that Speaker Diarisation is chosen automatically when using mono files as input data. Voiceprints are persistently stored on the filesystem and are subsequently processed by speaker recognition (e.g., for performing speaker clustering).
5) If Speech-to-Text engine is enabled, then a transcription will be generated for each audio file and stored on the filesystem.
6) The transcription will be consequently used by Text-processing technologies (if enabled), for identifying names, locations, dates etc. by Named Entity Recognition, detecting topic of discussion, characterising the names found as third parties or participants and extracting relations about these names.
7) If the user decided to activate video analysis pipeline, then videos and images uploaded are processed and embeddings are created for faces, scenes and objects. Furthermore, if any of the face clustering and scene clustering modules are activated, then similar faces, similar scenes and similar objects are clustered.
8) All previous results will be used by the (activated) Network Analysis modules for generating a network and identifying most influential and/or outlier nodes, revealing communities of nodes, predicting links between nodes, as well as identifying common nodes on two different cases. Network Analysis can also provide back to Speech Analysis module hints on how to perform speaker verification, e.g., based on shared phones used by several people (see the purple bidirectional arrow).

[18] The ROXANNE platform currently supports either stereo audio files (with only one speaker in one channel) or mono channel audio files with only two speakers in the recording.

9) The outputs of the activated backend technologies are collated by the backend python script in a single JavaScript Object Notation (JSON) file that follows the ROXANNE platform data scheme and which is being consumed by the ROXANNE platform Graphical User Interface (GUI).
10) The LEA practitioner explores Autocrime results via the Graphical User Interface by applying filters, saving views for future reference. Furthermore, the user can provide expert knowledge back to Autocrime (e.g., merge two persons that are the same according to his intelligence sources). Special care was taken so that these adjustments are not ignored, but considered by the Speech Analysis component when the user decides to re-run the case (e.g., due to additional evidence files made available).
11) Furthermore, the user can also export the results of the case to a file whose format is supported by other popular commercial tools (i.e., the I2 Analyst Notebook).

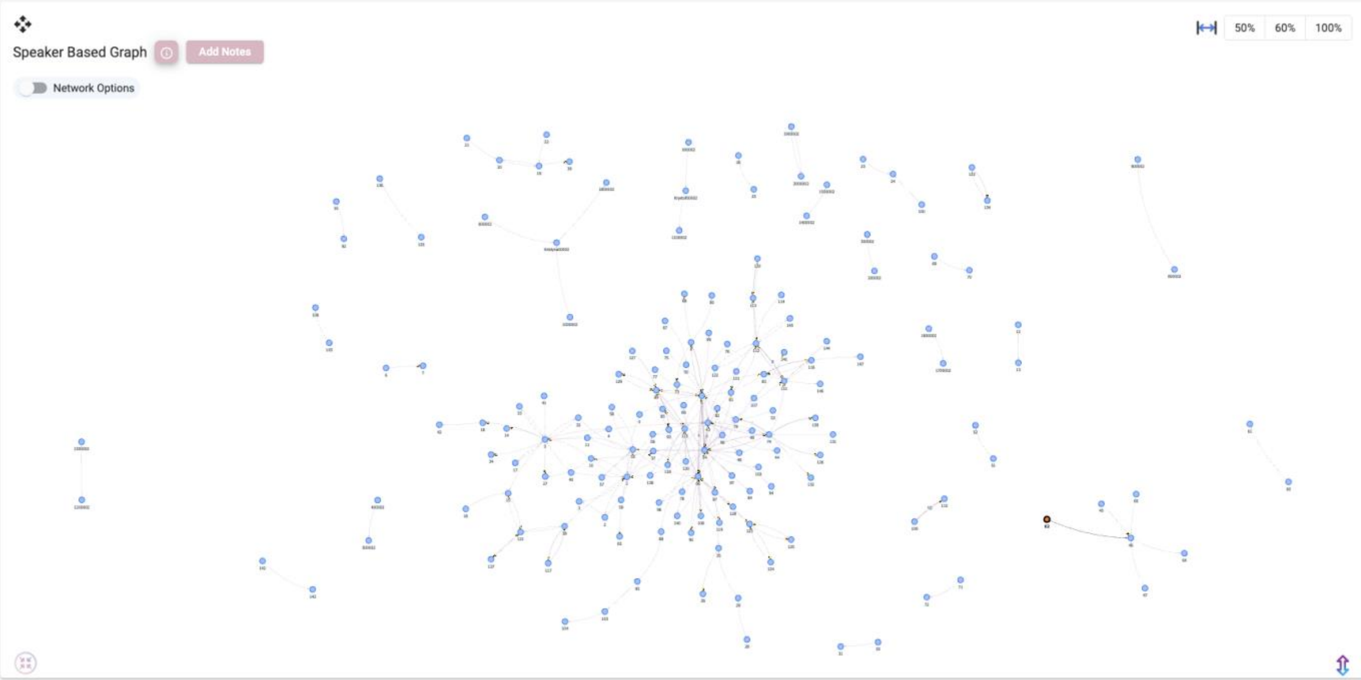


Figure 19a: The speaker-based graph produced by the ROXANNE platform when processing 481 audio files with no metadata about phone numbers (e.g., Call Detail Records)

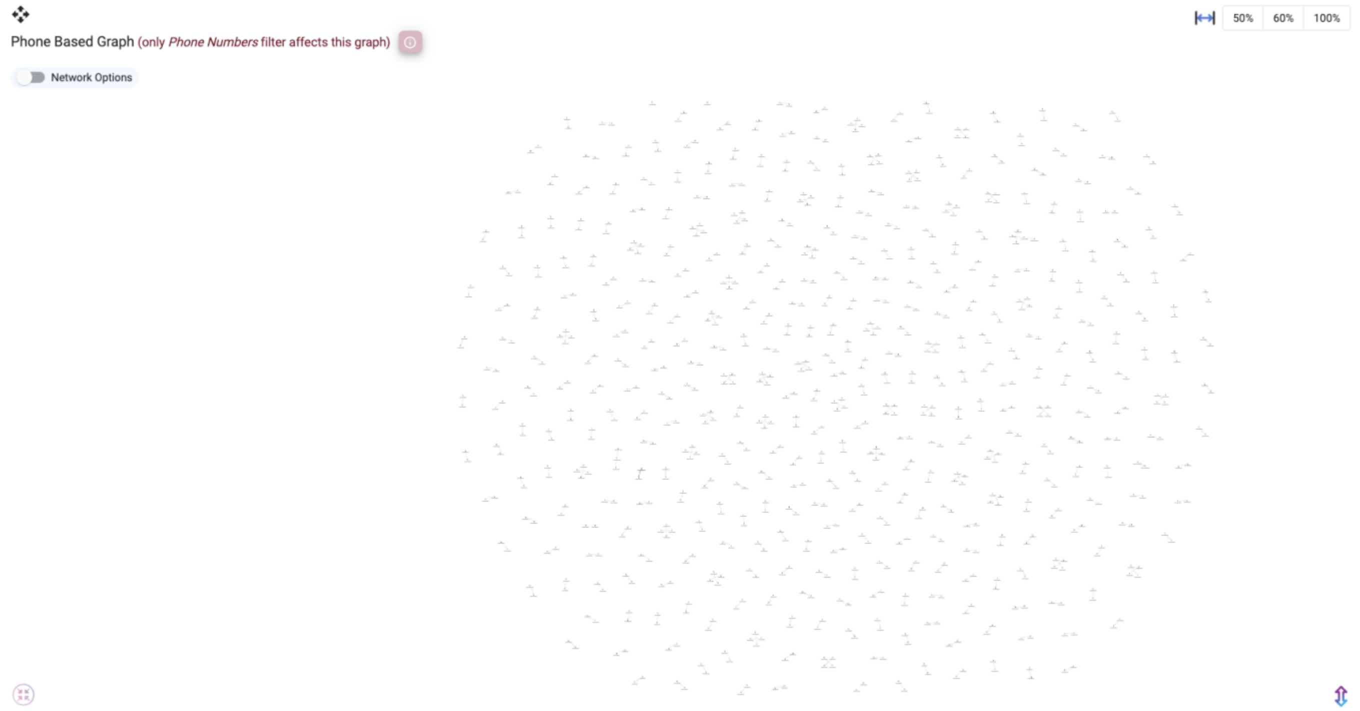

Phone Based Graph (only *Phone Numbers* filter affects this graph)
Network Options
50%
60%
100%

Figure 19b: The phone-based graph obtained with the ROXANNE platform when processing 481 audio files with no metadata about phone numbers (e.g., Call Detail Records)

4.5 Further updates

We are working on two further improvements to the backend (which are due to further support exploitation results of the project after its end) with the following updates in the pipeline:

1. Topic detection model (to support larger number of languages) to better respond to needs of the LEA partners.
2. Run-time performance optimization with caching: this new feature of the backend will help avoid running the technologies on data that was already processed. This is useful in the typical case where the user adds more data to the case over a time period (as opposed to adding all data at once, which will typically not be available on day one).

5. Platform features

This section will associate system requirements to features of the ROXANNE platform, also referred to as Autocrime, related to case analysis and data visualisation, as well as case management, workflow configuration and user management.

5.1.1 Cases Dashboard

The Cases Dashboard shows all available cases that a particular user can work on, with the case name and the number of files processed. Furthermore, the user has the option to create new cases. Figure 19 and Figure 20 present the cases dashboard for two different users, e.g., analyst.a and analyst.b, where several cases appear. Note that these users have different rights, in particular:

- analyst.a has full management rights (create, read, update) on the case "FinalConf", the ability only to add files to the case "video", as well as read-only rights to the case "roxhood"
- analyst.b has read-only access to the "FinalConf" case and only the ability to add files to the case "video"[19].

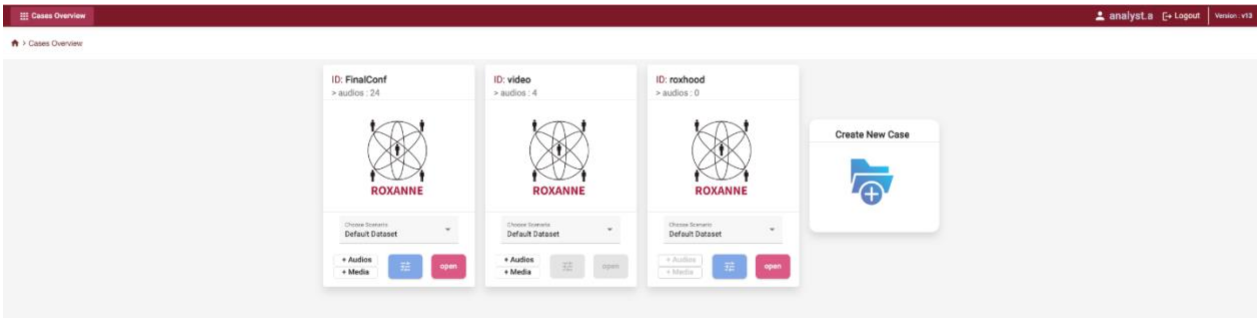


Figure 19: Screenshot of Cases Dashboard widget for user analyst.a

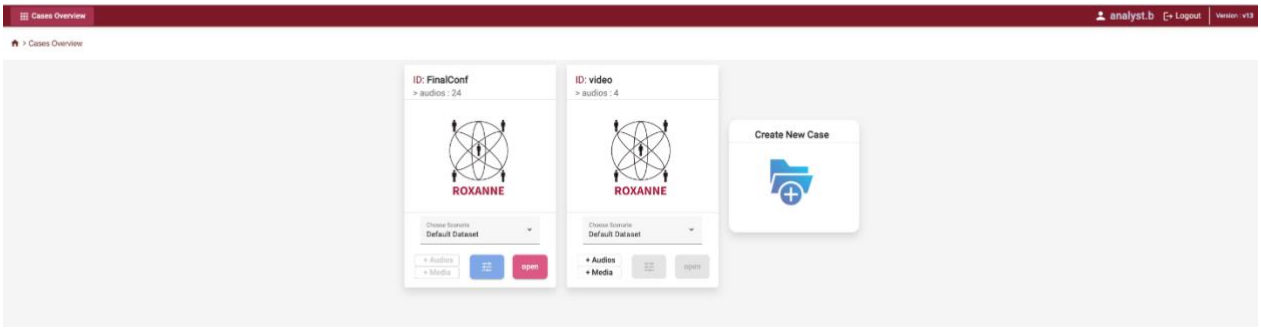


Figure 20: Screenshot of Cases Dashboard widget for user analyst.a

These features respond to the following system requirements:

SR-CM01-NewCaseCreationRequest (for triggering an authentication request to the User Management service that is offered by a Keycloak instance as expected based on SR-UM03-

[19] Analyst.b has no rights to the case "roxhood" and thus is not visible

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

UserAuthenticationRequest and SR-UM04-UserAuthenticationResponse system requirements; see Figure 21),

- SR-CM02-NewCaseCreationConfirmation (for authenticating user and authorising the request that, if successful, will generate an access token),
- SR-CM08-UserCasesRequest (for issuing an authentication request for listing all eligible cases),
- SR-CM10-AuthorisedCasesListing (for authenticating user and authorising the request by checking the users' privileges on the available cases)

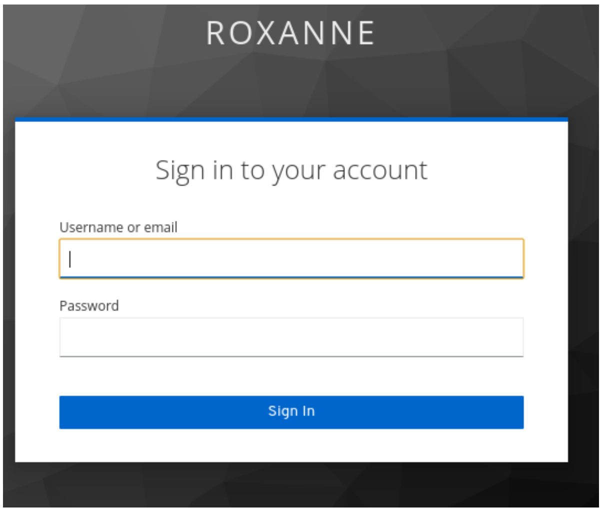


Figure 21: A screenshot of the log-in page provided by the Keycloak instance of the ROXANNE platform

By clicking on the respective button, the user initiates the new case creation process where s/he is asked to supply a case name (see Figure 22). When a case is created, a file is automatically created that contains information about the files to be processed, such as sequence order, filename, etc[20]. The user can also upload a metadata file that has been processed using external software (see SR-EP06-MetaDataFileUploading).

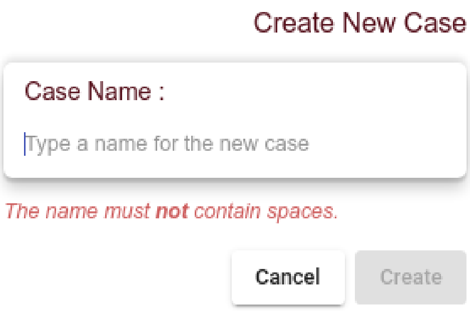

Create New Case
Case Name :
Type a name for the new case
The name must **not** contain spaces.
Cancel
Create

Figure 22: Creating a new case

[20] The file is called data.csv and is located on the folder: ~/autocrime/{caseName}/CSV, where {caseName} is a placeholder for the case name.

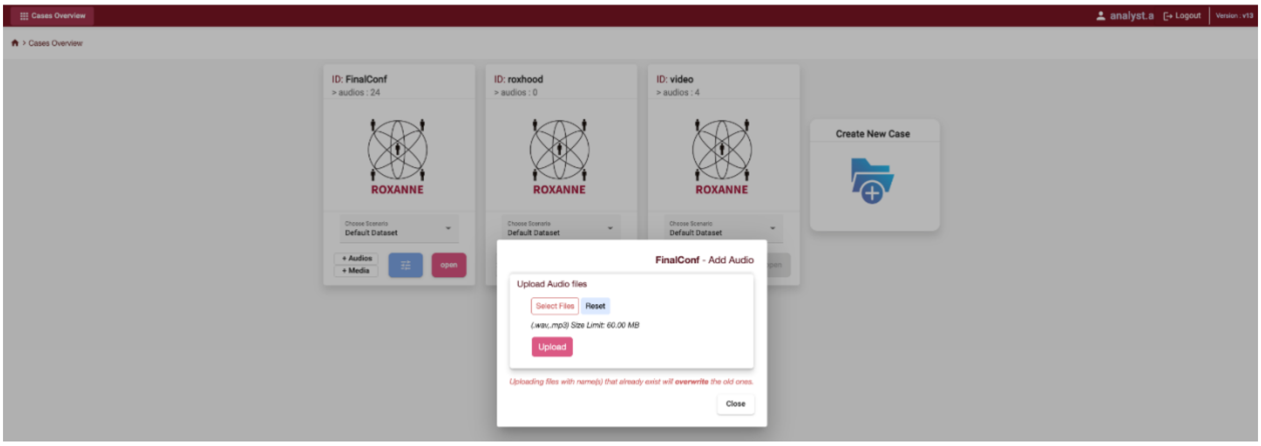
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

For the time being, existing cases cannot be edited or deleted. Thus, SR-CM03-CaseDescriptionUpdateRequest and SR-CM06-CaseDescriptionUpdateRejection are not currently supported.

5.1.2 Uploading files to an existing case

A user who has the right to add files can upload new evidence files to a case. By clicking on the buttons "+Audio" or "+Media", the user can choose the audio files (e.g., wav and mp3 files) and media files (i.e., images and videos supporting jpeg, png, mp4 and mov extensions) respectively to be uploaded to the platform. In particular, when files (that should be already available on a mounted hard drive) are uploaded, these are copied to the appropriate folder in the ROXANNE platform that is determined by the name of the case and their type. For the time being, the same filename is being used, but in future releases these filenames could also include information about the case and/or a timestamp so that the same evidence file can be part of several cases. Note that the user can upload files in batches, with total size limit up to 60 MB. These user actions are related to the following system requirements:

- SR-EP01-AudioFileUploading
- SR-EP02-VideoFileUploading
- SR-EP04-EvidenceFilePersistentStorage
- SR-EP07-EvidenceWorkflowConfiguration



ID: FinalConf
> audios : 24
Choose Scenario: Default Dataset

ID: roxhood
> audios : 0
Choose Scenario: Default Dataset

ID: video
> audios : 4
Choose Scenario: Default Dataset

FinalConf - Add Audio

Upload Audio files

(.wav,.mp3) Size Limit: 60.00 MB

Uploading files with name(s) that already exist will **overwrite** the old ones.

Figure 23: Uploading files to a case

Note that SR-EP03-TextFileUploading and SR-EP05-IncrementalEvidenceWarning are currently not supported.

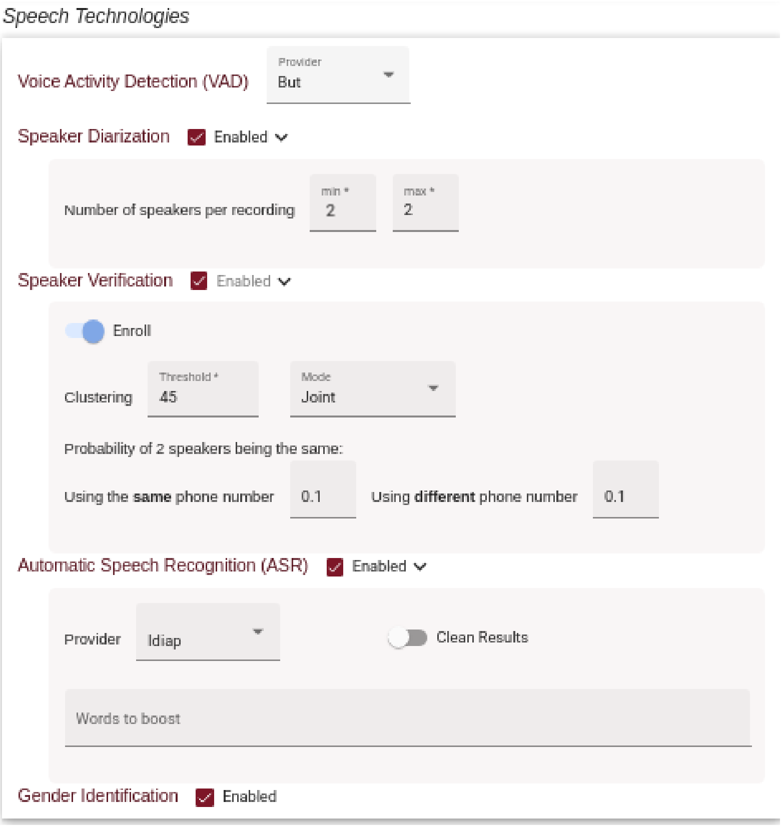
5.1.3 Configuring an existing case

After a set of files has been uploaded to the Autocrime, the user can click on the blue configuration button in order to choose a) the files to be processed, b) the Autocrime technologies to be activated along with any parameter values that can affect the sensitivity of the results obtained, and c) whether

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

the outputs will be stored on the default case or an alternative scenario. These features are related to the following system requirements:

- SR-CA07-FinetuneAlgorithmParameters. For example, in Figure 24 the user can finetune parameters of speech technologies (e.g., set a similarity threshold for clustering voiceprints, or instruct how likely it is that two voices which are heard in different calls made using a single telephone number belong to the same person two voices which are heard in different calls made using a single telephone number belong to the same person).
- SR-CA38-AddContextSpecificKeywords. For example, as shown in Figure 24, the user can specify custom words that should be recognised by the ROXANNE platform, even if these words (i.e., slang words) do not usually appear in its vocabulary. Furthermore, the user can provide a list of topics that should be taken into account when identifying the most probable discussion topic (see Figure 25).


Speech Technologies

Voice Activity Detection (VAD) Provider: But

Speaker Diarization ☑ Enabled

Number of speakers per recording min * 2 max * 2

Speaker Verification ☑ Enabled

● Enroll

Clustering Threshold * 45 Mode Joint

Probability of 2 speakers being the same:

Using the **same** phone number 0.1 Using **different** phone number 0.1

Automatic Speech Recognition (ASR) ☑ Enabled

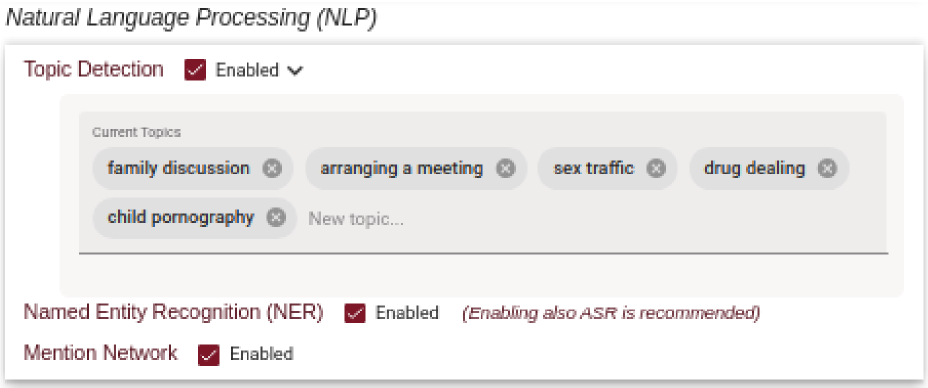
Provider Idiap ○ Clean Results

Words to boost

Gender Identification ☑ Enabled

Figure 24: Configuring the speech technologies of ROXANNE platform

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.



Natural Language Processing (NLP)

Topic Detection ☑ Enabled

Current Topics

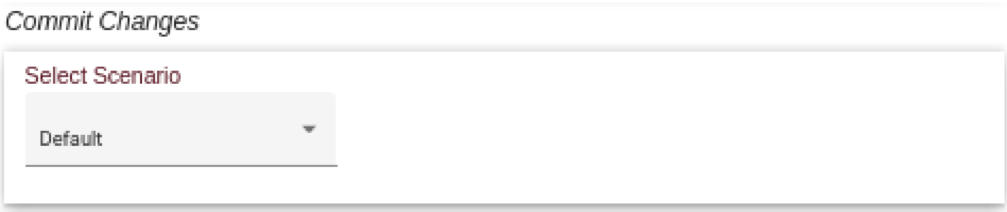
family discussion ⊗ arranging a meeting ⊗ sex traffic ⊗ drug dealing ⊗ child pornography ⊗ New topic...

Named Entity Recognition (NER) ☑ Enabled (Enabling also ASR is recommended)

Mention Network ☑ Enabled

Figure 25: Configuring the ROXANNE platform's technologies related to text-processing

Finally, the user can save the configuration and start the processing. During all phases of case processing (e.g., initiation, on-going, completion), the user can see notifications of success or warnings that an error occurred. This is associated to the SR-EP08-RealTimeEvidenceProcessingStatusUpdate system requirement. As soon as the processing is successfully completed, the outcomes of the ROXANNE platform will be saved on the local directory. Previously generated outcomes will be replaced unless the user had instructed the ROXANNE platform to save new results as a separate scenario (see Figure 26 implementing SR-CM18-AlternativeScenarios system requirement).



Commit Changes

Select Scenario

Default

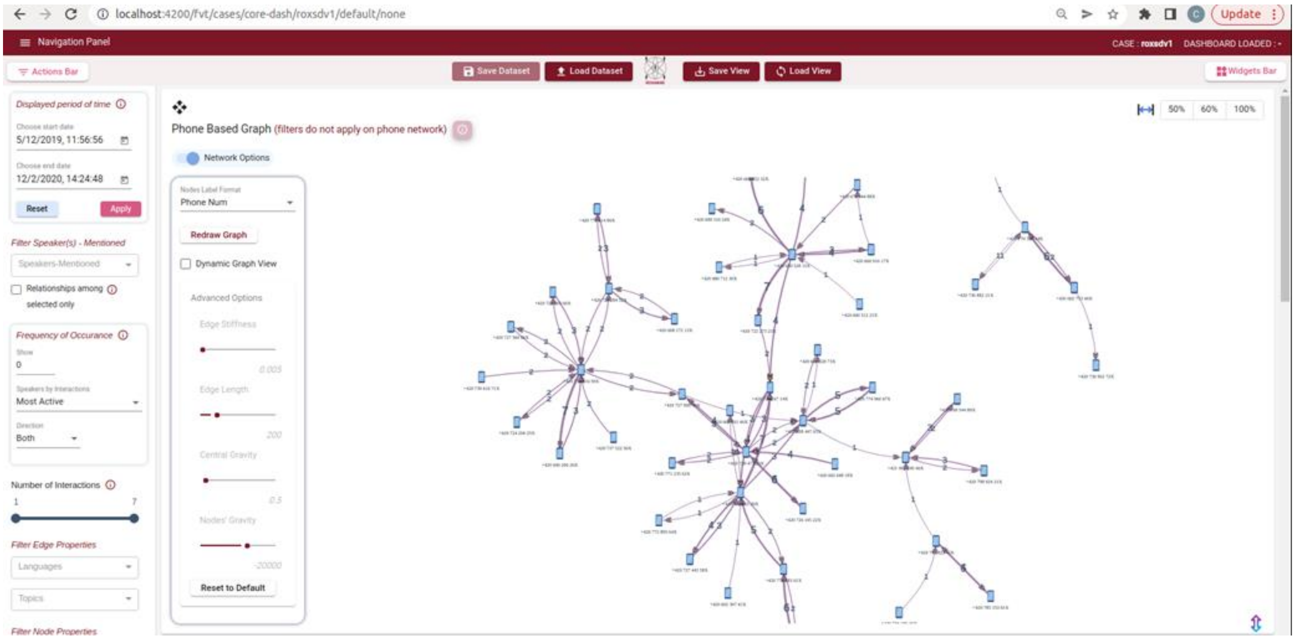
Figure 26: Configuring where the outputs of the ROXANNE platform will be stored

5.1.4 Social Network Analysis

The Social Network Analysis widget provides an efficient way to reveal groups and patterns, identify groups of entities that are interlinked, uncover paths that connect entities that appeared disconnected, etc.

The ROXANNE platform supports two types of networks:

1. A phone-based network that is based on Call-Detail Records (see Figure 27) and presents a network containing a set of phone numbers corresponding to the wiretapped phones (already known to the investigator e.g., via intelligence) and those were calls originated/received to/from. These numbers appear as nodes and their interactions (e.g., calls) appear as edges.



localhost:4200/fvt/cases/core-dash/roxsdv1/default/none

CASE : roxsdv1 DASHBOARD LOADED : -

Displayed period of time

Choose start date: 5/12/2019, 11:56:56

Choose end date: 12/2/2020, 14:24:48

Filter Speaker(s) - Mentioned

☐ Relationships among selected only

Frequency of Occurance

Show: 0

Speakers by Interactions: Most Active

Direction: Both

Number of Interactions: 1 – 7

Phone Based Graph (filters do not apply on phone network)

Nodes Label Format: Phone Num

☐ Dynamic Graph View

Advanced Options

Edge Stiffness: 0.005

Edge Length: 200

Central Gravity: 0.5

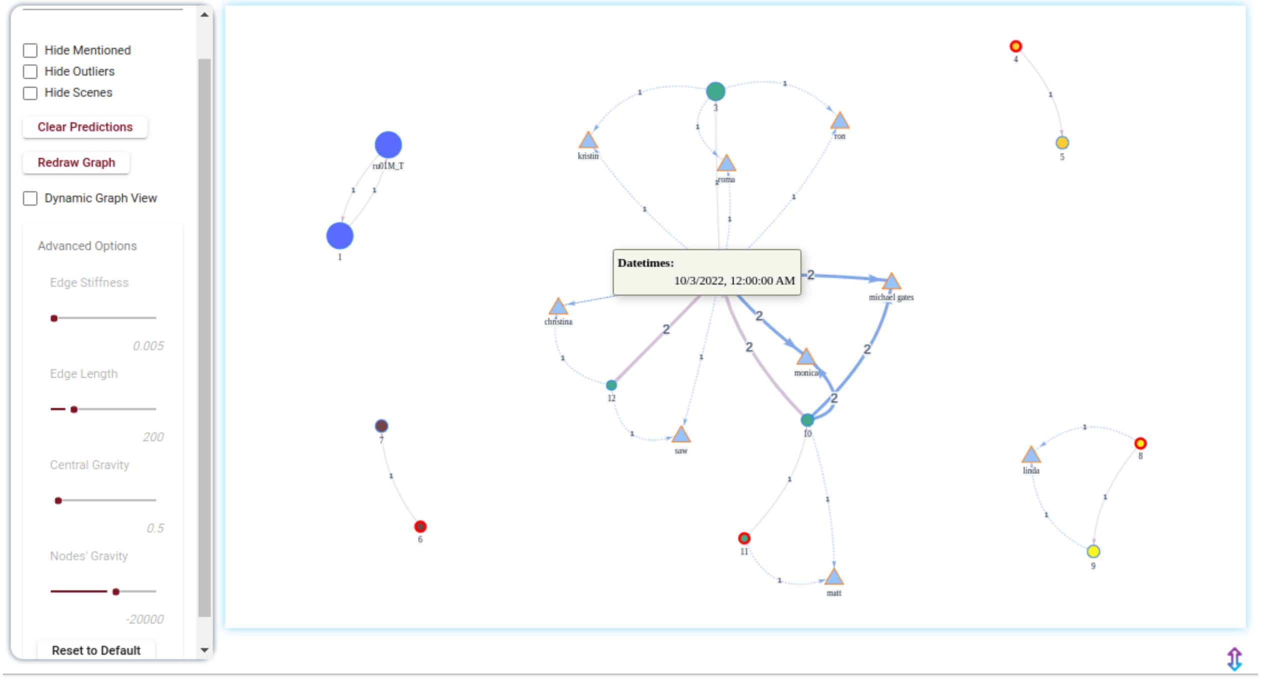
Nodes' Gravity: -20000

Figure 27: An overview of the Phone-based Network

2. A speaker-based network that is based on ROXANNE platform outputs (see Figure 28) and which contains
    a. the following three types of nodes:
        i. Circular nodes referring to individuals who were identified on raw audio files and may be associated to a particular face (if specified by the user on the metadata file)
        ii. Triangular nodes that refer to names mentioned in the conversations, and which were automatically identified
        iii. Media nodes (media icons or scene miniature) to group images or videos coming from a given source or to represent a location of interest
    b. directional edges between:
        i. a pair of persons that refer to phone calls and which are visualised with purple solid lines,
        ii. a person and name that was mentioned in some phone calls[21], or
        iii. any media or speaker node associated to a same image or video thanks to image or video modalities[22].

[21] Shown with blue solid lines.

[22] Rendered with red dashed lines.



Edge Stiffness 0.005
Edge Length 200
Central Gravity 0.5
Nodes' Gravity -20000

Datetimes:
10/3/2022, 12:00:00 AM

Figure 28: Screenshot of Social Network Analysis widget

In order to present the social network to the user, the following system requirements are realised:

- SR-CA01-ProcessedEventsRequest, so that all processed events related to the case in question are retrieved from ROXANNE backend.
- SR-CA08-PresentSocialNetwork, so that each individual that was identified (on raw audio files) appear as a single node and the events s/he initiated (i.e., phone calls placed) appear as (aggregated) directed edges.
- SR-CA04-ShowEventOverview, so that the number of calls initiated by an entity appears as edge label.
- SR-CA10-AdjustSocialNetworkAppearance, so that the user can adjust the social network appearance (e.g., manually reposition a node so that it is closer to another entity, change node labels, or adjust the edge length by using the "Dynamic Graph View" advanced options).
- SR-CA11-ChooseSocialNetworkZoomLevel, so that the user can zoom in/out on certain parts of the graph
- SR-CA12-SocialNetworkScrolling, so that the user can focus on certain parts of the graph.
- SR-CA13-ShowEntityOverview, so that high-level information about individual entities appears on mouse hover.
- SR-CA09-ShowCommunitiesOfEntities, so the community/communities[23] that each individual was automatically inferred to be part of, is rendered using a distinct colour. In this example, 2 communities were identified.
- SR-CA42-HighlightImportantEntities, so that popular entities appear with a node of bigger diameter.

[23] Communities are cohesive groups of individuals whose intra-group interaction is denser and more frequent than their interaction with the rest of the network.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

Furthermore, Figure 29 presents how the ROXANNE platform realises system requirement SR-CA43-ShowPredictedRelationships, which shall indicate to users any missing links that involve a particular entity, and which are highly likely to exist due to the structure of the topology. These predicted links and are shown with yellow colour.

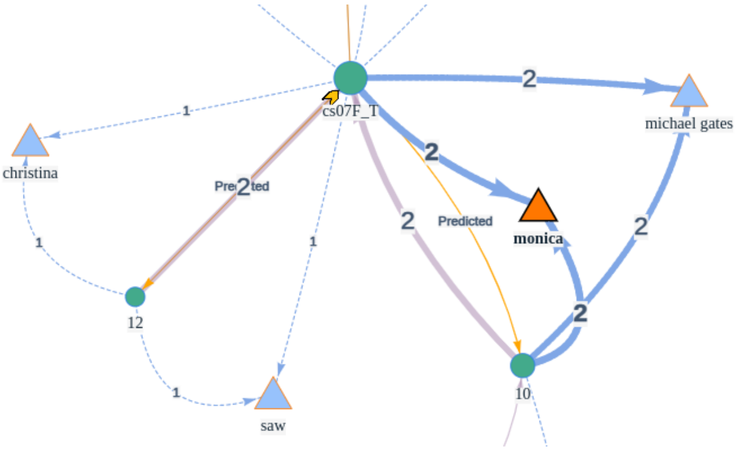
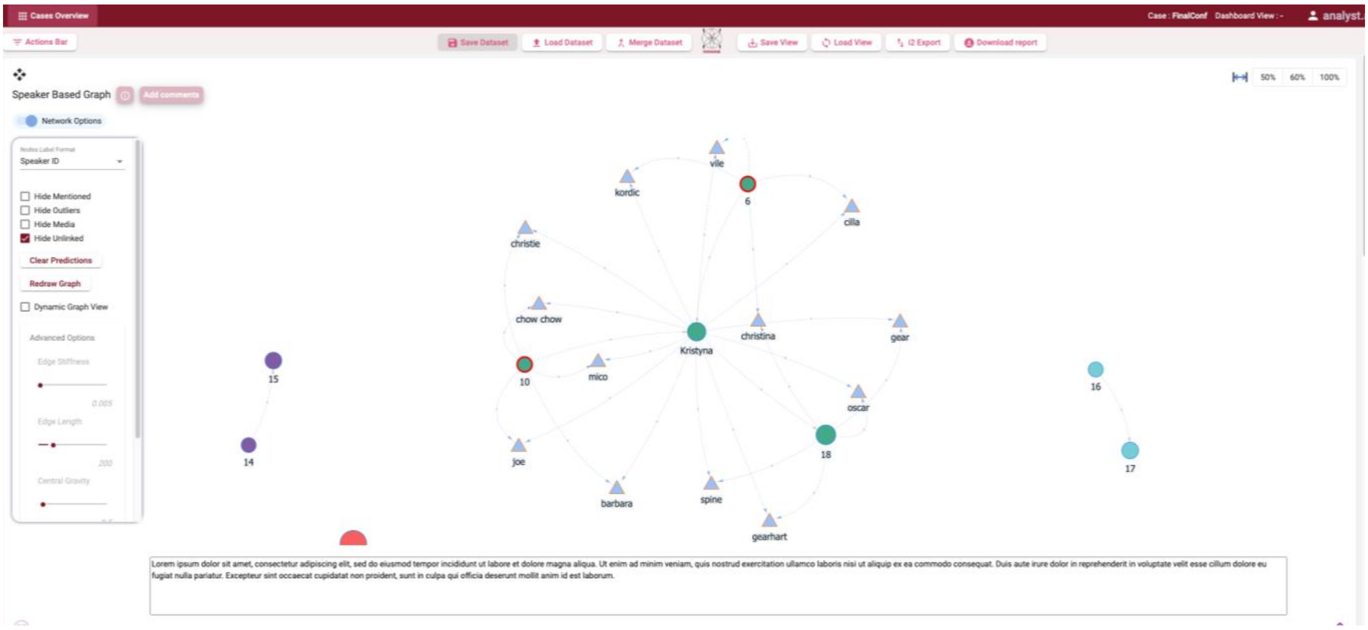


Figure 29: Detail of speaker-based graph where “cs07F_T” is predicted to interact with nodes “10” and “12”

Finally, the user can add personal notes to the social network graph, e.g., describe hypotheses about key suspects, a system requirement that is related to SR-CA29-AnnotateOnSocialNetworks. Figure 30 provides a screenshot of an exemplary note, which will be also part of the report (see Section 5.1.13).



Nodes Label Format: Speaker ID

☐ Hide Mentioned
☐ Hide Outliers
☐ Hide Media
☑ Hide Unlinked

☐ Dynamic Graph View

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Figure 30: Adding user annotations on the speaker-based network

Note that the following system requirements were eventually considered of lower priority and are not currently supported:

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

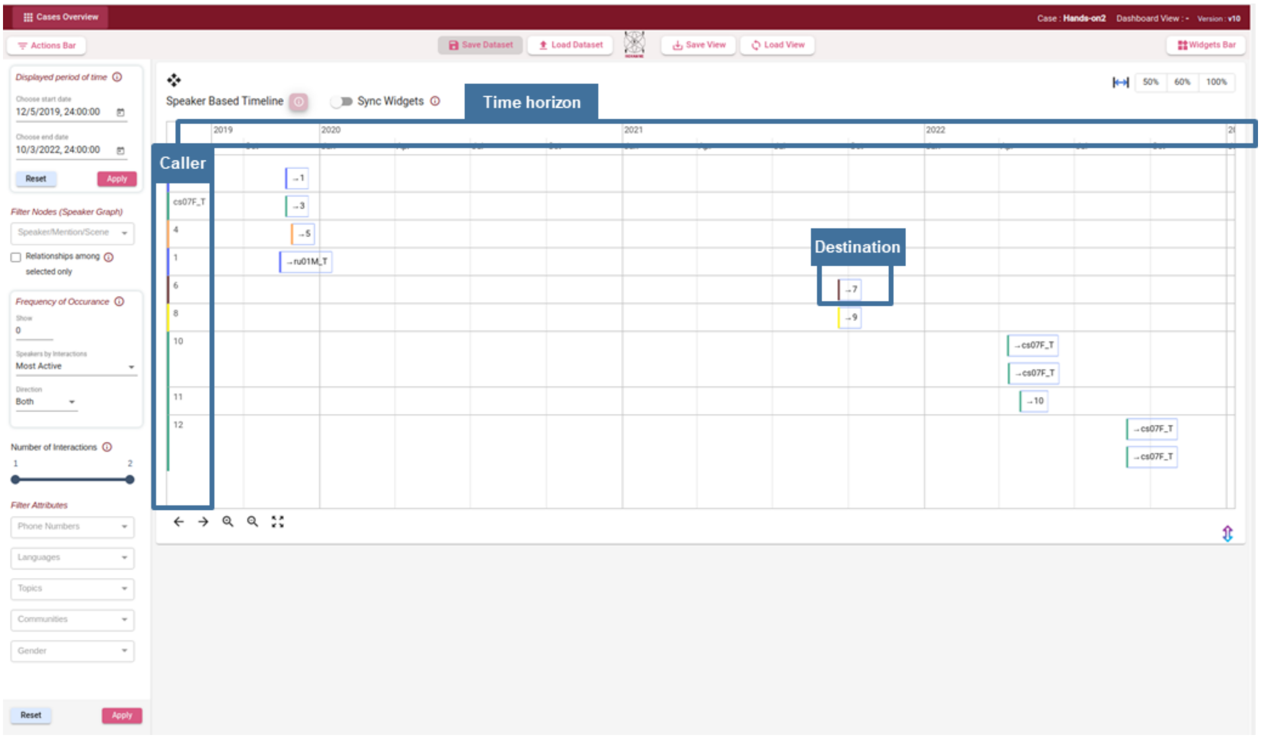
- SR-CA33-AddEvents
- SR-CA34-UpdateEntityProperties
- SR-CA35-EntityPropertiesUpdateRejection
- SR-CA36-UpdateEventProperties
- SR-CA37-EnventPropertiesUpdateRejection

5.1.5 Timeline Analysis

The Timeline Analysis widget provides the ability to “travel back in time” and identify groups of events that occurred close together in time, in order to compare the current situation with older events. The following figure presents (part of) the Timeline Analysis widget for a particular case.

In order to show such an “events” chart to the user, the visualisation engine needs to implement the following “case analysis” system requirements (which are explained in Section 3.5):

- SR-CA01-ProcessedEventsRequest, so that all processed events related to the working case are retrieved from ROXANNE backend.
- SR-CA03-PresentEventsEvolution, so that all events containing timestamp information are presented along the horizontal axis. These events (for the time being conversations) are grouped by the originating party (i.e., caller) that appears on the vertical axis. The destination party (callee) appears in the main part of the widget.
- SR-CA04-ShowEventOverview, so that details about the calls (such as languages and topics identified) appears on mouse hover.


Case : Hands-on2 Dashboard View : - Version : v10

Displayed period of time

Choose start date: 12/5/2019, 24:00:00

Choose end date: 10/3/2022, 24:00:00

Frequency of Occurance

Show: 0

Speakers by Interactions: Most Active

Direction: Both

Number of Interactions: 1 – 2

Figure 31: Screenshot of the Timeline Analysis widget (with annotations for better readability)

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

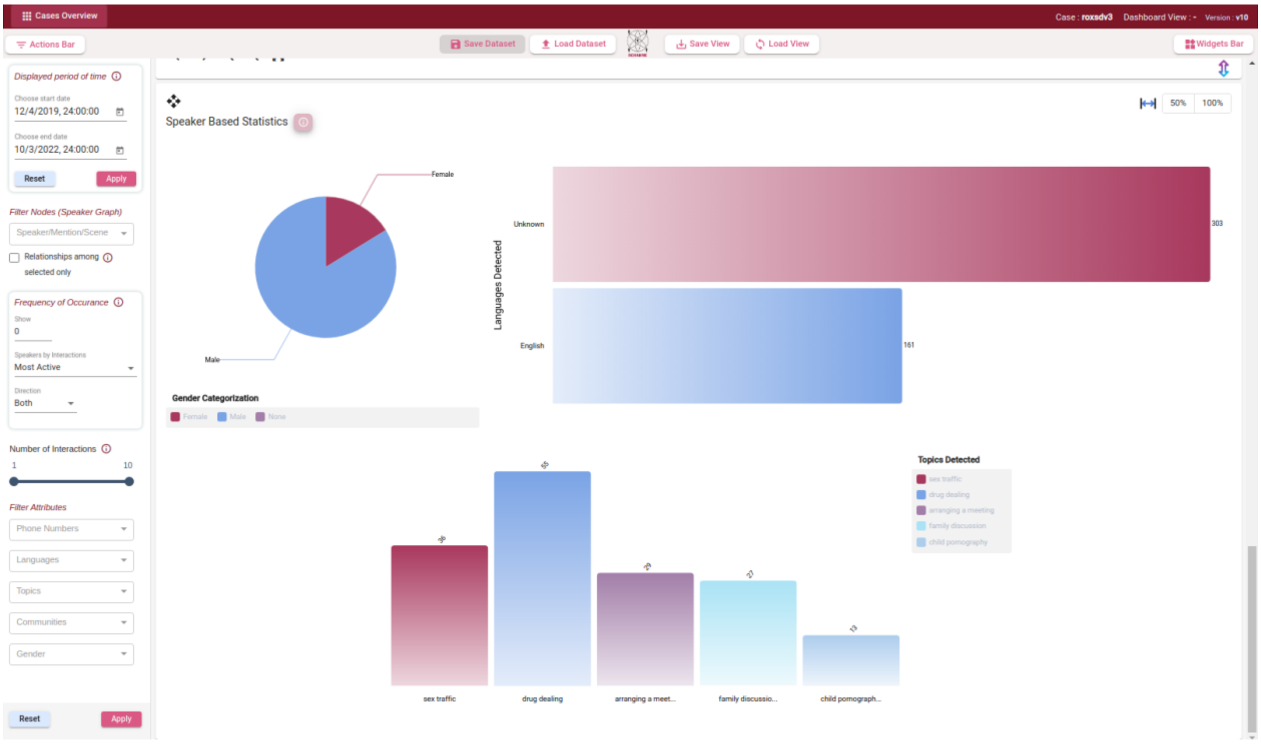
5.1.6 Statistical Analysis

The Statistical Analysis widget provides the ability to view data distributions based on categories of data using bar charts, pie charts, histograms, etc.

The following figure presents the Statistical Analysis widget for a sample case. It presents the distributions of different entity properties (namely the gender and the languages spoken) that ROXANNE components recognised from raw audio files, as well as all topics that were identified (e.g., locations, persons).

The following "case analysis" system requirements (which are explained in Section 3.5) are relevant to this widget:

- SR-CA01-ProcessedEventsRequest, so that all processed events related to the working case are retrieved from ROXANNE backend.
- SR-CA23-ShowStatisticalResults, so that statistical charts for different data attributes are presented.



Case : roxsdv3 Dashboard View : - Version : v10

Displayed period of time
Choose start date: 12/4/2019, 24:00:00
Choose end date: 10/3/2022, 24:00:00

Filter Nodes (Speaker Graph): Speaker/Mention/Scene
□ Relationships among selected only

Frequency of Occurance
Show: 0
Speakers by Interactions: Most Active
Direction: Both

Number of Interactions: 1 – 10

Figure 32: Screenshot of the Statistical Analysis widget

5.1.7 Details

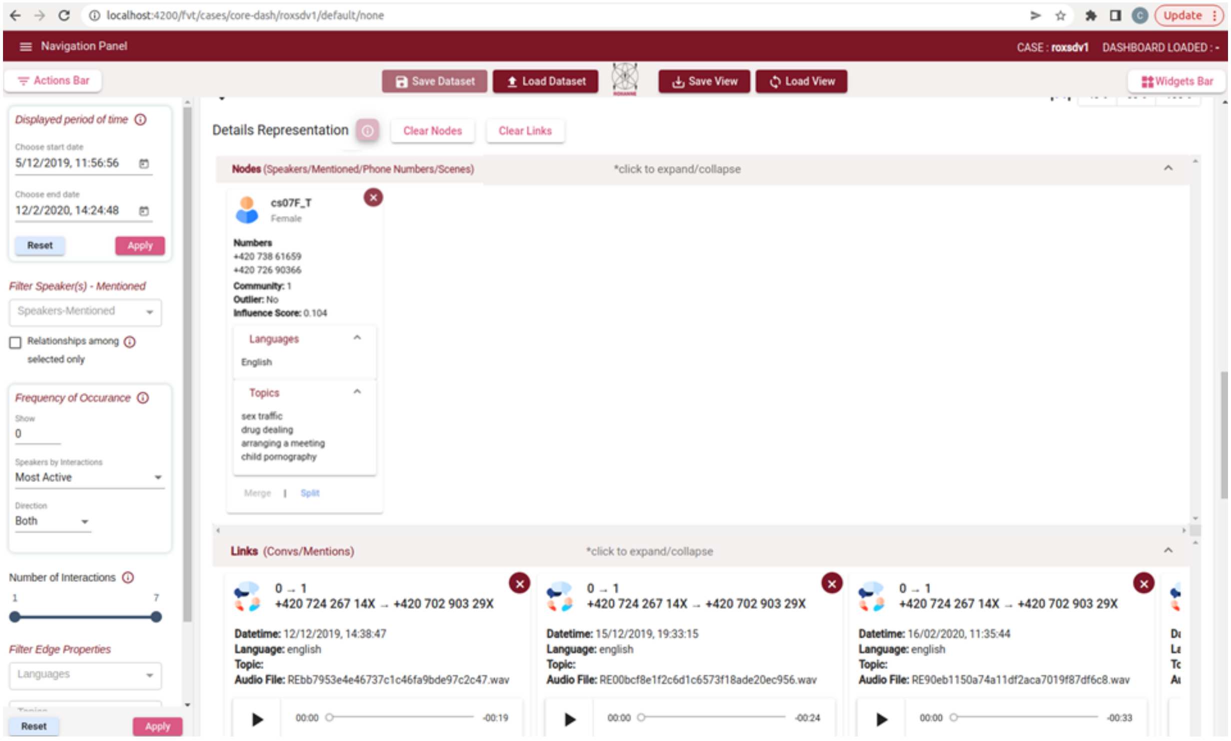
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

The Details widget is used to present information about selected entities and/or events, as well as to edit their properties.

The following figure presents the Details widget for a particular entity (called "cs07F_T") and some calls of interest.

When images and videos are ingested in Autocrime, the detail widget also provides the following lists:

- **Enrolled Faces** (for a speaker node): list of face pictures enrolled for this speaker. Clicking on the file name opens the corresponding picture. The first picture in this list is used to represent the speaker node in the Speaker-based Network.
- **Enrolled Scenes** (for a media node): list of scene pictures enrolled for this node.
- **Media list**: list of images or videos manually or automatically associated to this node.



CASE : roxsdv1 DASHBOARD LOADED : -

Displayed period of time

Choose start date: 5/12/2019, 11:56:56

Choose end date: 12/2/2020, 14:24:48

Filter Speaker(s) - Mentioned: Speakers-Mentioned

□ Relationships among selected only

Frequency of Occurance

Show: 0

Speakers by Interactions: Most Active

Direction: Both

Number of Interactions: 1 – 7

Filter Edge Properties: Languages

Nodes (Speakers/Mentioned/Phone Numbers/Scenes)

cs07F_T
Female

Numbers
+420 738 61659
+420 726 90366
Community: 1
Outlier: No
Influence Score: 0.104

Languages
English

Topics
sex traffic
drug dealing
arranging a meeting
child pornography

Links (Convs/Mentions)

0 → 1
+420 724 267 14X → +420 702 903 29X
Datetime: 12/12/2019, 14:38:47
Language: english
Topic:
Audio File: REbb7953e4e46737c1c46fa9bde97c2c47.wav
00:00 -00:19

0 → 1
+420 724 267 14X → +420 702 903 29X
Datetime: 15/12/2019, 19:33:15
Language: english
Topic:
Audio File: RE00bcf8e1f2c6d1c6573f18ade20ec956.wav
00:00 -00:24

0 → 1
+420 724 267 14X → +420 702 903 29X
Datetime: 16/02/2020, 11:35:44
Language: english
Topic:
Audio File: RE90eb1150a74a11df2aca7019f87df6c8.wav
00:00 -00:33

Figure 33: Screenshot of the Details widget

In order to realise this widget, the visualisation engine implements the following "case analysis" system requirements (for more details please see Section 3.5):

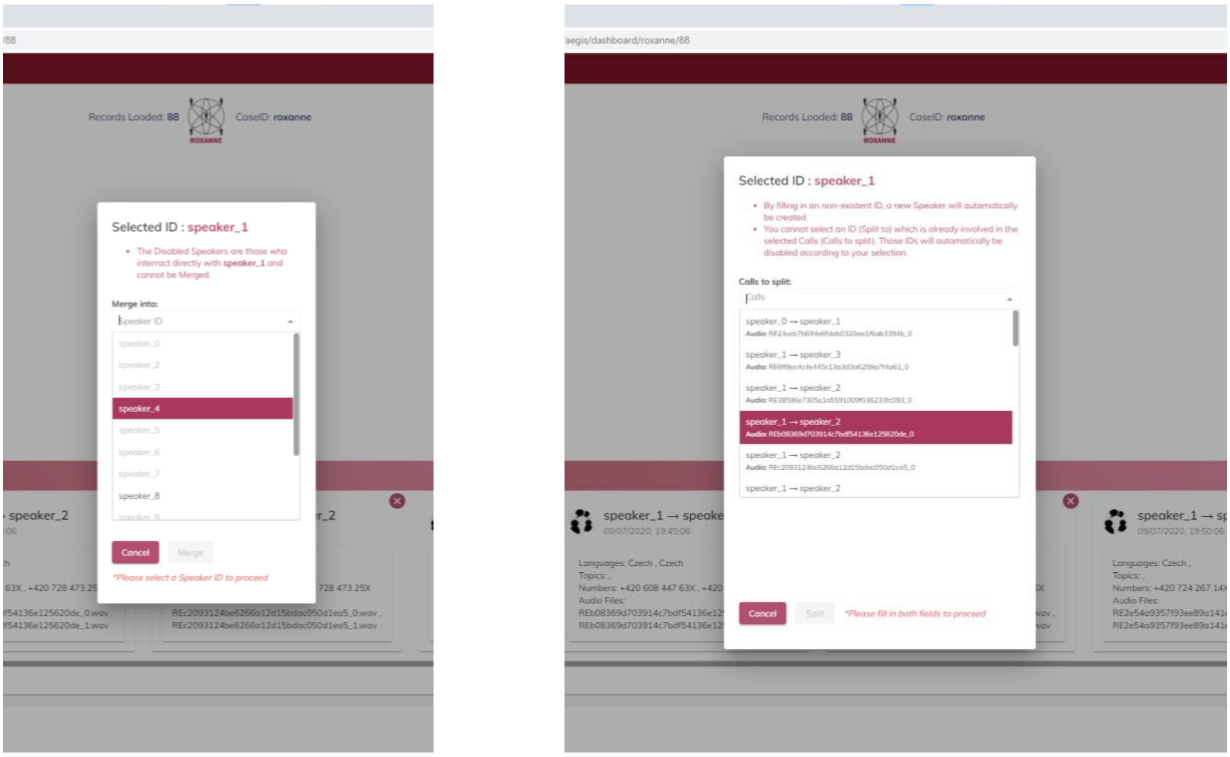
- SR-CA01-ProcessedEventsRequest, so that all processed events related to the "roxanne" case are retrieved from ROXANNE backend.
- SR-CA05-ShowEventDetails for presenting information about selected events.
- SR-CA14-ShowEntityDetails for providing details about selected entities.
- SR-CA15-TranscribedAudioPlayback which shall allow users to listen to audio files and view automatically generated transcription, where named entities such as locations, dates, etc. will be highlighted.
- SR-CA16-VideoPlayback, which shall allow users to play video files (e.g., the ones where a selected entity appears).

Note that the current version of the ROXANNE platform does <u>not</u> support free-text search on processed text (see SR-CA20-FreeTextSearch), as this system requirement was ranked of lower importance to end-users.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

The following figure shows how the following system requirements are currently supported

- SR-CA30-MergeEntities, which allows an authorised user to merge two or more entities into one.
- SR-CA31-SplitEntities, which allows an event to be attributed to another entity (compared to the one that ROXANNE modules originally inferred).
- SR-CA32-AddEntities that allows a new entity to be manually added.



Records Loaded: 88 CaseID: roxanne

Selected ID : speaker_1

- The Disabled Speakers are those who interract directly with **speaker_1** and cannot be Merged.

Merge into:

Speaker ID

speaker_0
speaker_2
speaker_3
speaker_4
speaker_5
speaker_6
speaker_7
speaker_8
speaker_9

*Please select a Speaker ID to proceed

Selected ID : speaker_1

- By filling in an non-existent ID, a new Speaker will automatically be created.
- You cannot select an ID (Split to) which is already involved in the selected Calls (Calls to split). Those IDs will automatically be disabled according to your selection.

Calls to split:

Calls

speaker_0 → speaker_1
Audio: RE24ceb7b694e6fdab0320aa1f6ab3394b_0

speaker_1 → speaker_3
Audio: RE8ff8ec4c4e445c13a3d3a6258a7f4a61_0

speaker_1 → speaker_2
Audio: RE38596e7305c1a5591009f036233fc093_0

speaker_1 → speaker_2
Audio: REb08369d703914c7bdf54136e125620de_0

speaker_1 → speaker_2
Audio: REc2093124be6266a12d15bdac050d1ea5_0

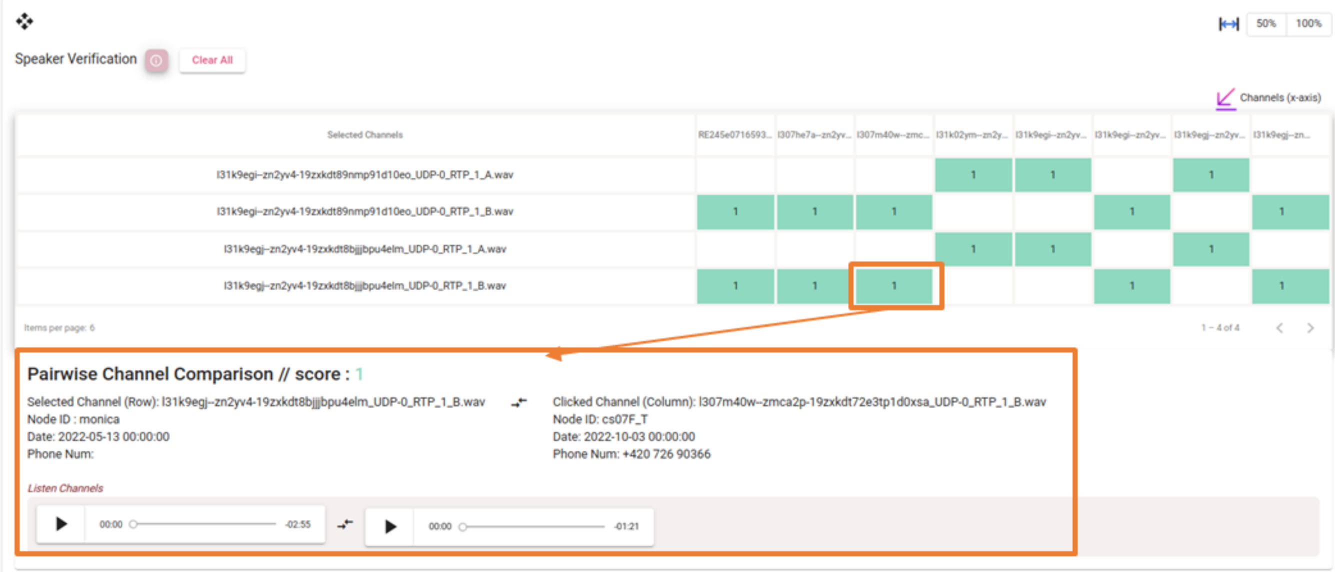
speaker_1 → speaker_2

*Please fill in both fields to proceed

Figure 34: Screenshot of Merging (left) and Splitting (right) Entities

Furthermore, the user is supported when assessing the outputs of the ROXANNE platform (see SR-CA39-SupportQualityAssessment system requirement). For example, the Speaker verification analysis widget that appears in Figure 35, compares the similarity of voiceprints on different audio channels/parts so that the user can assess the accuracy of the speaker identification process.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.



Speaker Verification

Selected Channels	RE245e0716593...	l307he7a--zn2yv...	l307m40w--zmc...	l31k02ym--zn2y...	l31k9egi--zn2yv...	l31k9egi--zn2yv...	l31k9egj--zn2yv...	l31k9egj--zn...
l31k9egi--zn2yv4-19zxkdt89nmp91d10eo_UDP-0_RTP_1_A.wav				1	1		1	
l31k9egi--zn2yv4-19zxkdt89nmp91d10eo_UDP-0_RTP_1_B.wav	1	1	1			1		1
l31k9egj--zn2yv4-19zxkdt8bjjjbpu4elm_UDP-0_RTP_1_A.wav				1	1		1	
l31k9egj--zn2yv4-19zxkdt8bjjjbpu4elm_UDP-0_RTP_1_B.wav	1	1	1			1		1

Pairwise Channel Comparison // score : 1

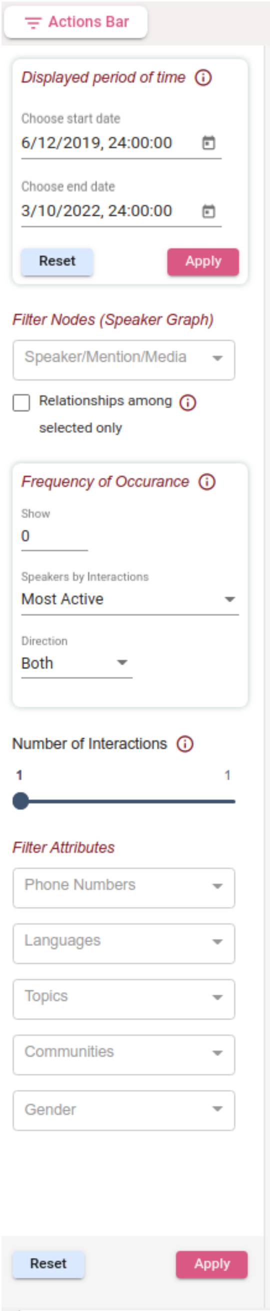
Selected Channel (Row): l31k9egj--zn2yv4-19zxkdt8bjjjbpu4elm_UDP-0_RTP_1_B.wav
Node ID : monica
Date: 2022-05-13 00:00:00
Phone Num:

Clicked Channel (Column): l307m40w--zmca2p-19zxkdt72e3tp1d0xsa_UDP-0_RTP_1_B.wav
Node ID: cs07F_T
Date: 2022-10-03 00:00:00
Phone Num: +420 726 90366

Figure 35: Assisting the user when assessing the similarity of the voiceprints for a selected pair of audio channels (rows) to the rest of the audio files (columns).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

5.1.8 Synced filters



Displayed period of time

Choose start date: 6/12/2019, 24:00:00

Choose end date: 3/10/2022, 24:00:00

Filter Nodes (Speaker Graph)

Speaker/Mention/Media

□ Relationships among selected only

Frequency of Occurance

Show: 0

Speakers by Interactions: Most Active

Direction: Both

Number of Interactions: 1 – 1

Filter Attributes

Phone Numbers

Languages

Topics

Communities

Gender

Figure 36: Screenshot of Synced Filters feature

Synced filters is a feature allowing the user to narrow down the results that appear across all visualisation widgets (e.g., social network analysis, timeline analysis, etc.).

Figure 36 shows the “Synced Filters” feature that is activated by pressing the “Actions Bar” button.

The following system requirements are relevant to this feature:

- SR-CA17-FilterEntitiesByProperty, which allows the user to filter on entities (e.g., caller, callee) and their properties (e.g., community, age, etc.).
- SR-CA18-FilterEventsByTiming, so that users can filter results based on the timing of events (e.g., phone calls).
- SR-CA21-ActivateSyncedFilters, so that the user can choose one or more filters and these to be activated when she/he is ready.
- SR-CA22-DeactivateFilters, so that the user is able to reset all active filters.

Nevertheless, filtering on the location of events (i.e., SR-CA19-FilterEventsByLocation system requirement) is not currently supported.

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

5.1.9 Custom Workspace

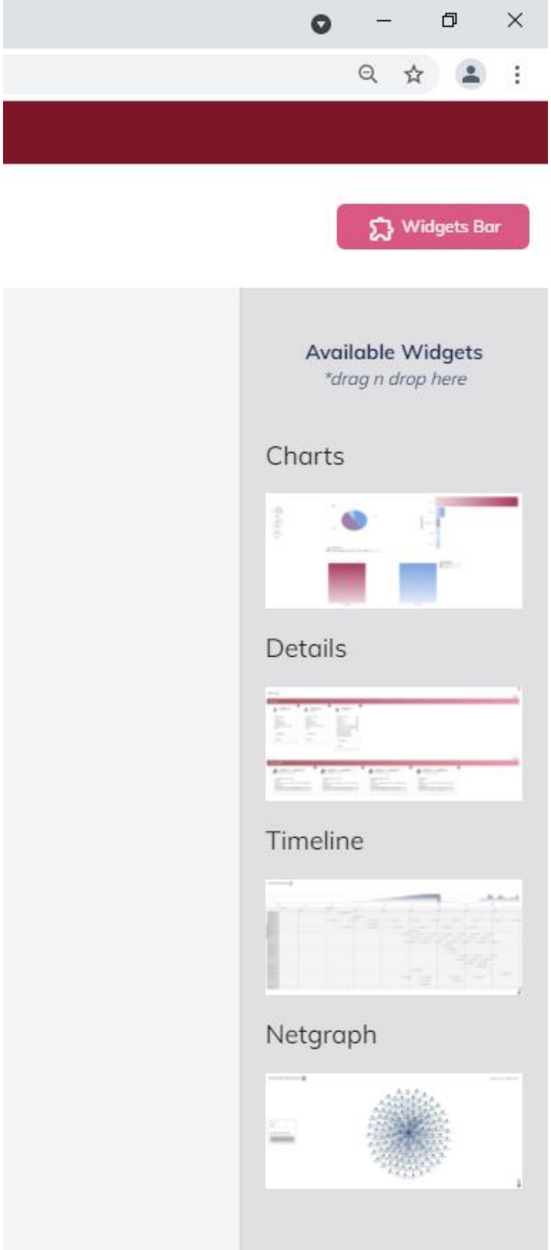


Figure 37: Synced Filters feature

The “Custom Workspace” feature allows the user to adjust the screen real estate in a convenient way.

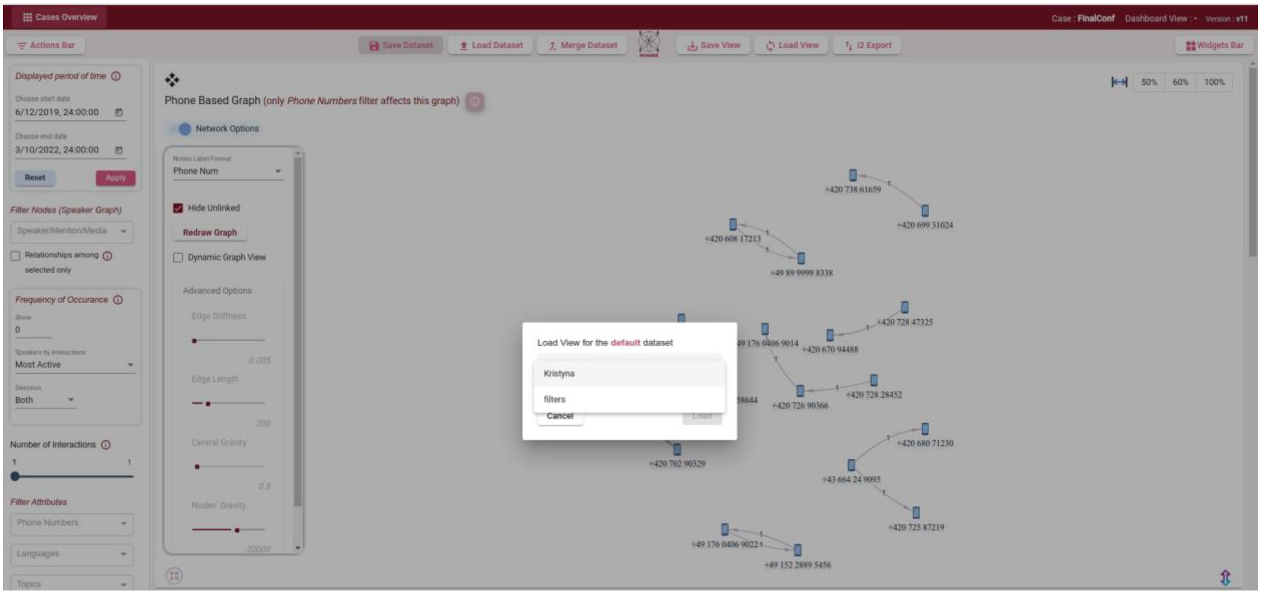
Figure 37 shows the “Custom Workspace” feature that is activated by pressing the “Widgets Bar” button.

The following system requirements are relevant to this feature:

- SR-CA24-ChooseVisibleCharts, so that the user can select the widgets / features that are visible. This could be done either for optimising usage of screen real estate or improving responsiveness.
- SR-CA25-ResizeVisibleCharts, so that users can resize charts according to the information present, user preferences/needs and technical constraints (e.g., screen resolution).
- SR-CA26-RelocateVisibleCharts, which allows charts to be relocated for customising screen real estate.

Furthermore, the user can save the configuration of dashboard (including the active widgets and active filters) and load it in the future, so that the analysis can be easily resumed. These features are related to SR-CA27-SaveDashboardConfiguration and SR-CA28-LoadDashboardConfiguration system requirements. Figure 38 presents a screenshot of the ROXANNE platform, where the user loads a previously saved dashboard configuration.

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.



Displayed period of time

Choose start date: 6/12/2019, 24:00:00

Choose end date: 3/10/2022, 24:00:00

Filter Nodes (Speaker Graph): Speaker/Mention/Media

☐ Relationships among selected only

Frequency of Occurance

Show: 0

Speakers by Interactions: Most Active

Direction: Both

Number of Interactions: 1 – 1

Filter Attributes: Phone Numbers, Languages, Topics

Phone Based Graph (only *Phone Numbers* filter affects this graph)

Nodes Label Format: Phone Num

☑ Hide Unlinked

☐ Dynamic Graph View

Advanced Options: Edge Stiffness 0.005; Edge Length 200; Central Gravity 0.5; Nodes' Gravity -20000

Load View for the default dataset

Kristyna

filters

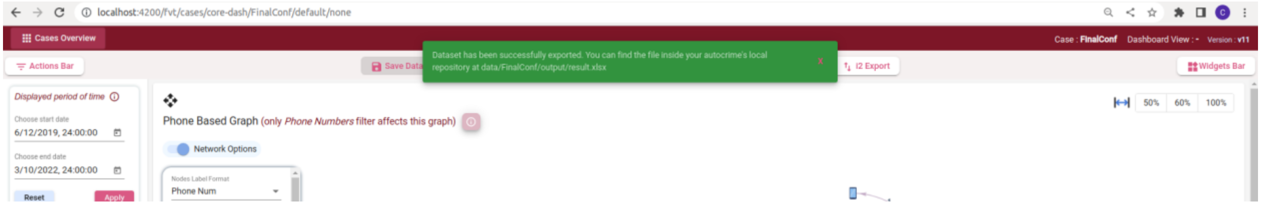
Cancel   Load

Figure 38: Loading a previously saved dashboard view

Saving and loading dashboard views can be also useful when the user wants to make a side-by-side comparison of charts that are produced a) from the same case but using two different sets of filters (see SR-CA40-CompareDifferentInferences system requirement) or b) from different cases. As the ROXANNE platform is a web application, the user can initiate a second instance of the user interface on a browser, navigate to the same case and choose to apply different dashboard views.

5.1.10 Data export

The user can also export the analysis results to suitable formats that can be subsequently imported to other investigative tools. This is related to SR-CA44-ExportResultsUsingPopularFormats system requirement and Figure 39 shows the confirmation message for successfully generating such a file after clicking on the "I2 Export" button on the top ribbon.



localhost:4200/fvt/cases/core-dash/FinalConf/default/none

Dataset has been successfully exported. You can find the file inside your autocrime's local repository at data/FinalConf/output/result.xlsx

Displayed period of time

Choose start date: 6/12/2019, 24:00:00

Choose end date: 3/10/2022, 24:00:00

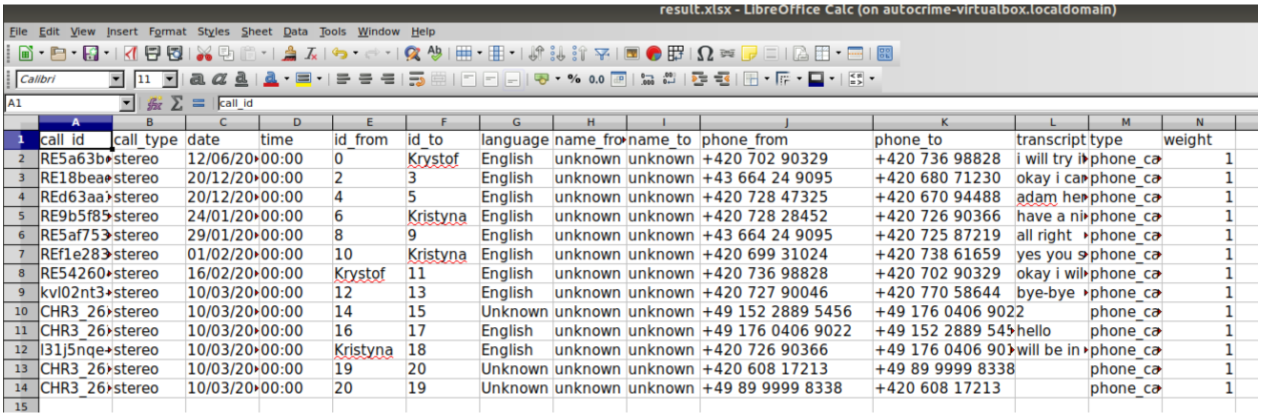
Phone Based Graph (only *Phone Numbers* filter affects this graph)

Nodes Label Format: Phone Num

Figure 39: A screenshot of the confirmation message notifying the user that a snapshot of the ROXANNE outputs has been exported to a format that can be used with I2 Analyst Notebook.

In particular, it takes as input the social network description and prepares a file that is compatible with the i2 Analyst Notebook. I2 Analyst's Notebook can import data from spreadsheet files, Extensible Markup Language (XML) files, and from any text file in which each line describes a similar but discrete piece of information. The ROXANNE platform generates a spreadsheet file like the one that appears in Figure 40.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.



result.xlsx - LibreOffice Calc (on autocrime-virtualbox.localdomain)
File Edit View Insert Format Styles Sheet Data Tools Window Help
Calibri
11
A1
call_id

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	call_id	call_type	date	time	id_from	id_to	language	name_fro	name_to	phone_from	phone_to	transcript	type	weight
2	RE5a63b	stereo	12/06/20	00:00	0	Krystof	English	unknown	unknown	+420 702 90329	+420 736 98828	i will try i	phone_ca	1
3	RE18bea	stereo	20/12/20	00:00	2	3	English	unknown	unknown	+43 664 24 9095	+420 680 71230	okay i ca	phone_ca	1
4	REd63aa	stereo	20/12/20	00:00	4	5	English	unknown	unknown	+420 728 47325	+420 670 94488	adam he	phone_ca	1
5	RE9b5f85	stereo	24/01/20	00:00	6	Kristyna	English	unknown	unknown	+420 728 28452	+420 726 90366	have a ni	phone_ca	1
6	RE5af753	stereo	29/01/20	00:00	8	9	English	unknown	unknown	+43 664 24 9095	+420 725 87219	all right	phone_ca	1
7	REf1e283	stereo	01/02/20	00:00	10	Kristyna	English	unknown	unknown	+420 699 31024	+420 738 61659	yes you s	phone_ca	1
8	RE54260	stereo	16/02/20	00:00	Krystof	11	English	unknown	unknown	+420 736 98828	+420 702 90329	okay i wil	phone_ca	1
9	kvl02nt3	stereo	10/03/20	00:00	12	13	English	unknown	unknown	+420 727 90046	+420 770 58644	bye-bye	phone_ca	1
10	CHR3_26	stereo	10/03/20	00:00	14	15	Unknown	unknown	unknown	+49 152 2889 5456	+49 176 0406 9022		phone_ca	1
11	CHR3_26	stereo	10/03/20	00:00	16	17	English	unknown	unknown	+49 176 0406 9022	+49 152 2889 545	hello	phone_ca	1
12	l31j5nqe	stereo	10/03/20	00:00	Kristyna	18	English	unknown	unknown	+420 726 90366	+49 176 0406 90	will be in	phone_ca	1
13	CHR3_26	stereo	10/03/20	00:00	19	20	Unknown	unknown	unknown	+420 608 17213	+49 89 9999 8338		phone_ca	1
14	CHR3_26	stereo	10/03/20	00:00	20	19	Unknown	unknown	unknown	+49 89 9999 8338	+420 608 17213		phone_ca	1
15														

Figure 40: Screenshot of the file exported using the I2 Analyst Notebook format

For most formats supported by I2 Analyst Notebook, the user needs to create an import specification to describe how the data is interpreted as entities and links on the chart surface. After creating a new specification for the imported data, the user can proceed with importing the data from the spreadsheet that was generated by the ROXANNE platform. Subsequently, the user can follow a step-by-step process to map the imported data to the internal data schema of I2 Analyst Notebook (e.g., in Figure 41 we use the network of telephone calls).

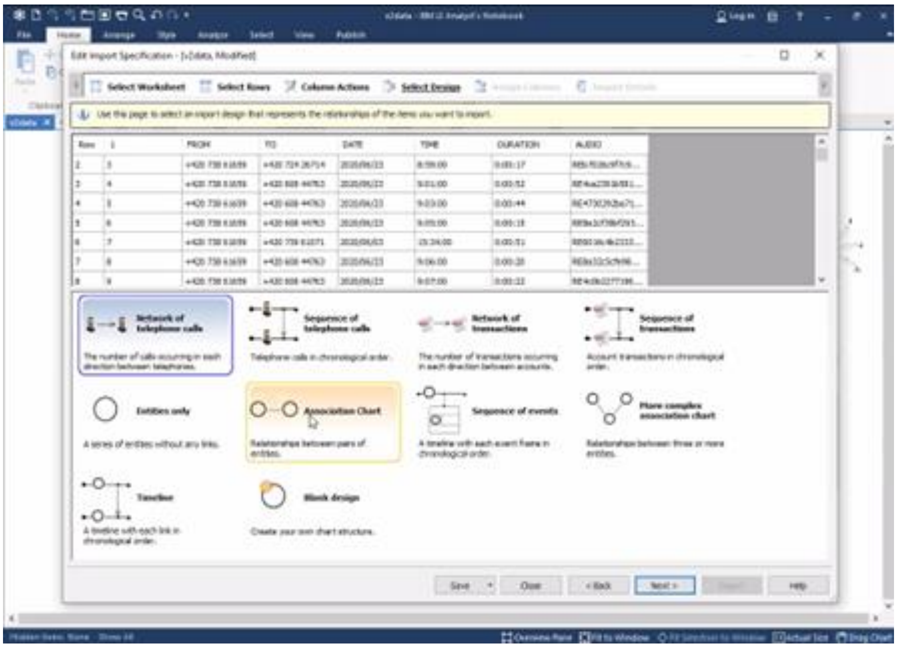


Figure 41: Mapping the data exported from the ROXANNE platform into the internal data schema of I2 Analyst Notebook

Eventually, the user can run the import process and visualise the ROXANNE outputs using the I2 Analyst Notebook widgets (e.g., Figure 42 presents a network of telephone calls).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

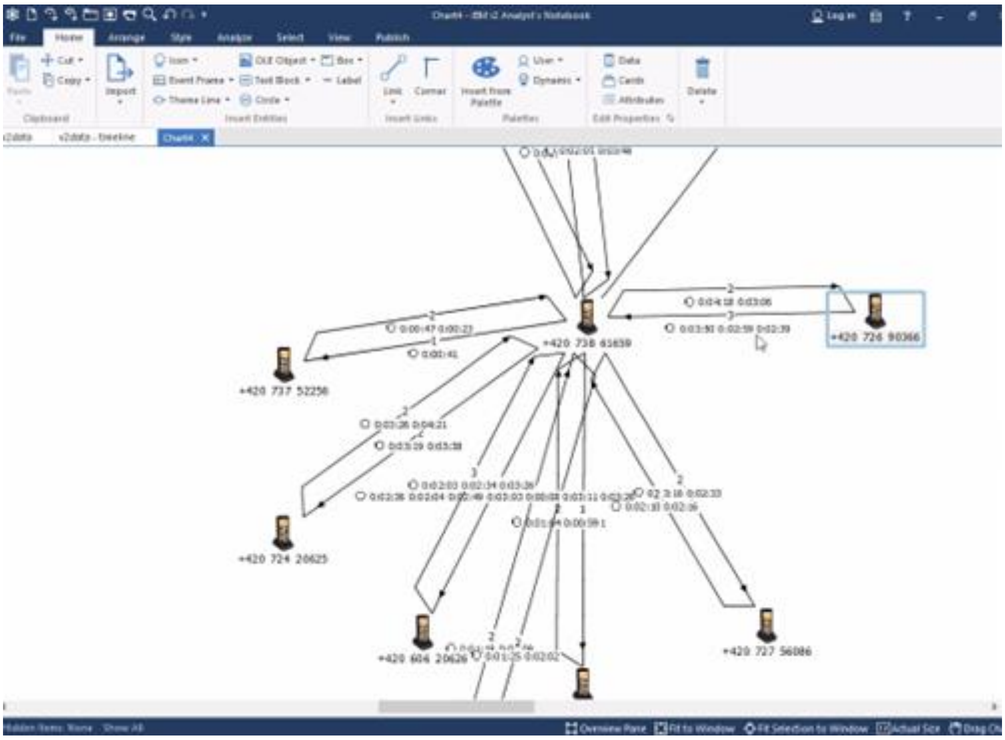
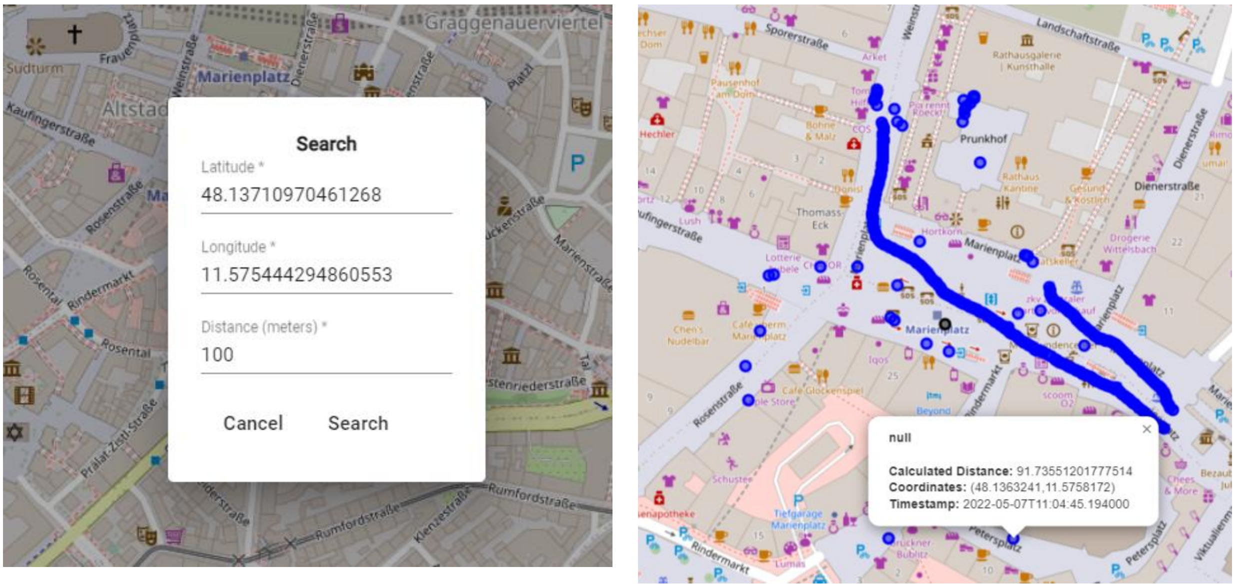


Figure 42: Visualising the ROXANNE outputs using the I2 Analyst Notebook

5.1.11 Geospatial analysis

A stand-alone, experimental web application has been created that allows the user to see all available events (nodes) found around a specific location, by providing the coordinates (longitude and latitude) of the location and the desired radius from the specified location.



Search

Latitude *
48.13710970461268

Longitude *
11.575444294860553

Distance (meters) *
100

null

Calculated Distance: 91.73551201777514
Coordinates: (48.1363241,11.5758172)
Timestamp: 2022-05-07T11:04:45.194000

Figure 43: Searching for events near a specific location (left) and viewing the events on a map (right)

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

Although this feature is currently not integrated with the rest ROXANNE platform, it is intended to show the location of the events that include geographic information (based on the SR-CA06-PresentEventsLocation system requirement).

5.1.12 Cross-case analysis

It is typical for a Law Enforcement Practitioner to check one or more suspects are common across a set of cases. This feature is related to SR-CA41-AssociateDifferentCases system requirement and Figure 44 presents how the ROXANNE platform visualises the entities (and their relationships) present in two different cases using a single social network. Note that the entities are grouped in several categories (e.g., case 1 only, case 2 only, common entities that were suggested by the user, etc.).

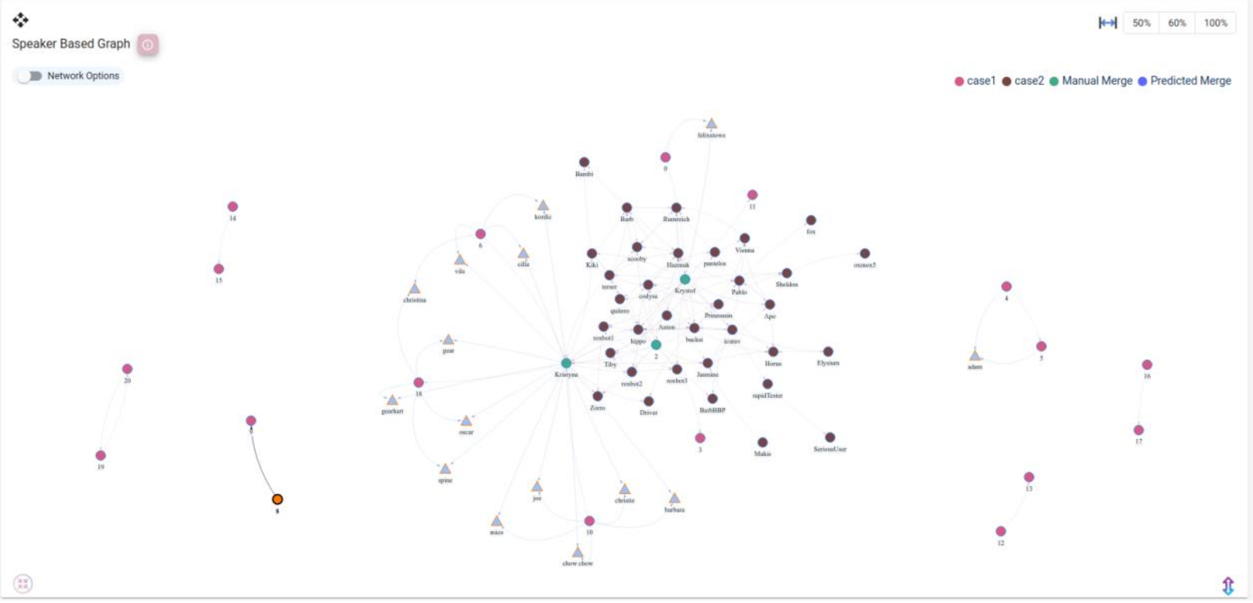
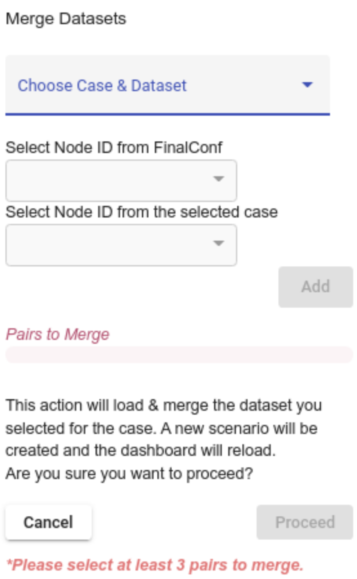


Figure 44: A screenshot of the social network generated when two cases are merged

When it comes to the last group of entities, Figure 45 shows how a user would suggest that some entities are common across the two cases.

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.



Merge Datasets

Choose Case & Dataset

Select Node ID from FinalConf

Select Node ID from the selected case

Add

Pairs to Merge

This action will load & merge the dataset you selected for the case. A new scenario will be created and the dashboard will reload.
Are you sure you want to proceed?

Cancel

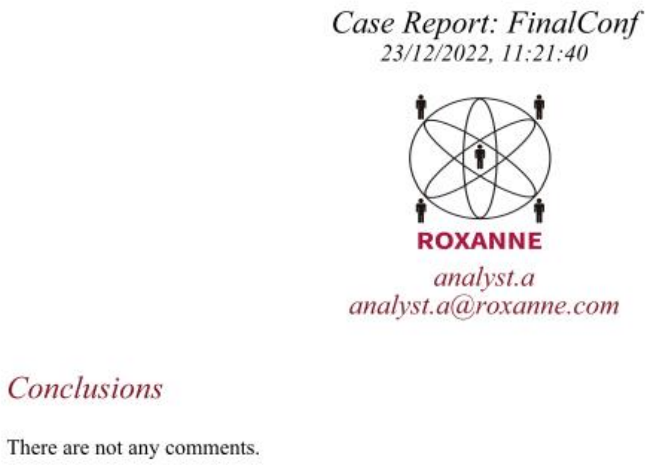
Proceed

**Please select at least 3 pairs to merge.*

Figure 45: A screenshot of the inputs expected by the user when merging two cases

5.1.13 Reporting

The ROXANNE platform also supports the automated generation of the analysis results as a report that can be forwarded to a colleague (see Figure 46 and Figure 47 for sample screenshots). Note that the "Conclusions" section is populated with user's notes on the Speaker-based graph. The reporting feature is related to the SR-CA45-GenerateCaseAnalysisReport system requirement.



Case Report: FinalConf
23/12/2022, 11:21:40

ROXANNE

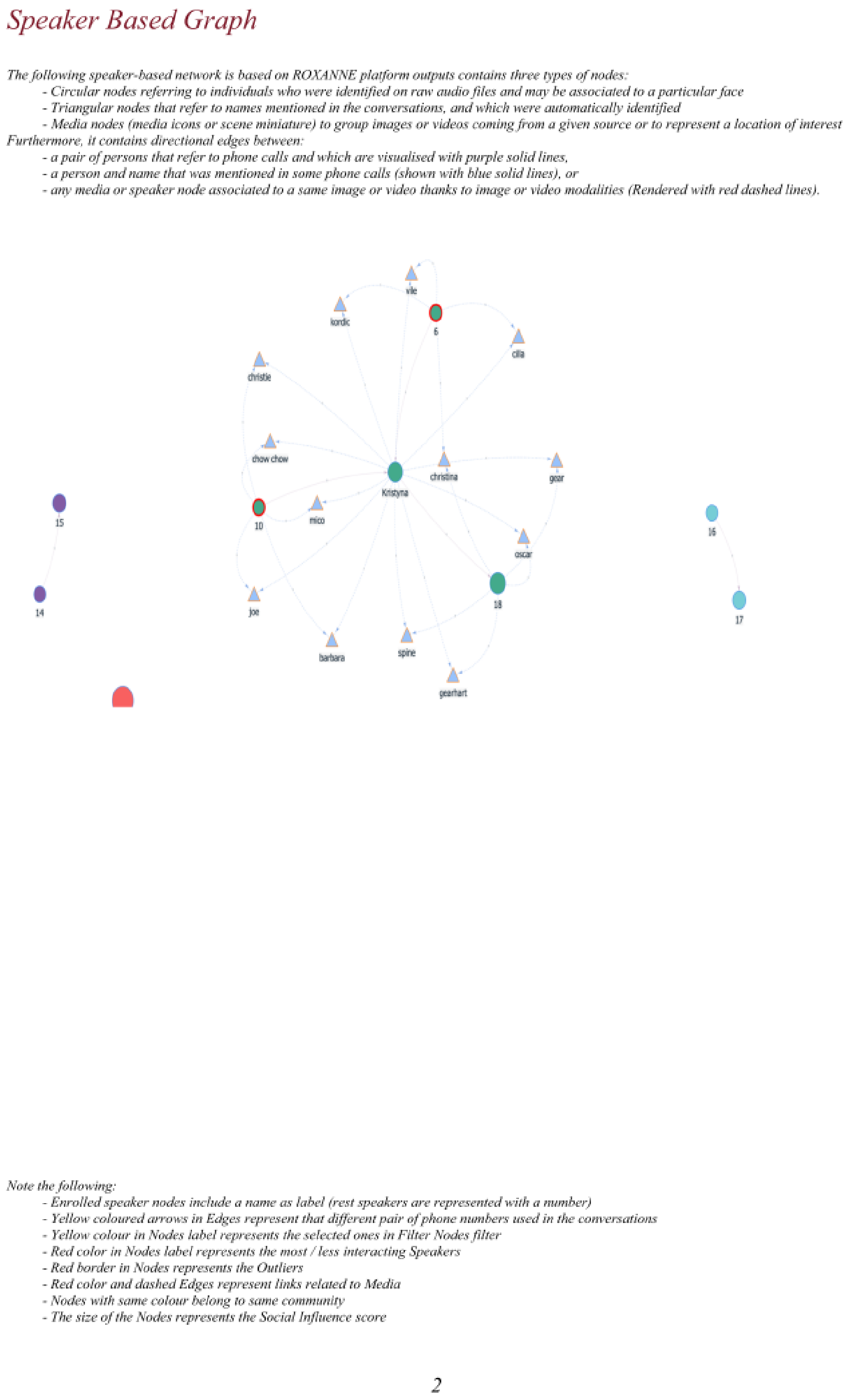
analyst.a
analyst.a@roxanne.com

Conclusions

There are not any comments.

Figure 46: The first page of a sample report generated by the ROXANNE platform

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.



Speaker Based Graph

The following speaker-based network is based on ROXANNE platform outputs contains three types of nodes:

- *Circular nodes referring to individuals who were identified on raw audio files and may be associated to a particular face*
- *Triangular nodes that refer to names mentioned in the conversations, and which were automatically identified*
- *Media nodes (media icons or scene miniature) to group images or videos coming from a given source or to represent a location of interest*

Furthermore, it contains directional edges between:

- *a pair of persons that refer to phone calls and which are visualised with purple solid lines,*
- *a person and name that was mentioned in some phone calls (shown with blue solid lines), or*
- *any media or speaker node associated to a same image or video thanks to image or video modalities (Rendered with red dashed lines).*

Note the following:

- *Enrolled speaker nodes include a name as label (rest speakers are represented with a number)*
- *Yellow coloured arrows in Edges represent that different pair of phone numbers used in the conversations*
- *Yellow colour in Nodes label represents the selected ones in Filter Nodes filter*
- *Red color in Nodes label represents the most / less interacting Speakers*
- *Red border in Nodes represents the Outliers*
- *Red color and dashed Edges represent links related to Media*
- *Nodes with same colour belong to same community*
- *The size of the Nodes represents the Social Influence score*

2

Figure 47: A sample report generated by the ROXANNE platform containing information about the speaker-based network

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

5.1.14 Platform management

In order to be able to run the ROXANNE platform on any computer, it can be installed in two ways:

- Natively (suitable for Linux Ubuntu and macOS)
- as a virtual machine hosted in Windows environments,

realising the SR-PM01-NativeInstallation and SR-PM02-VirtualMachine system requirements.

The ROXANNE platform can be deployed on any host type (e.g., server, desktop, laptop) depending on the number of users and hardware specifications (see Figure 48).

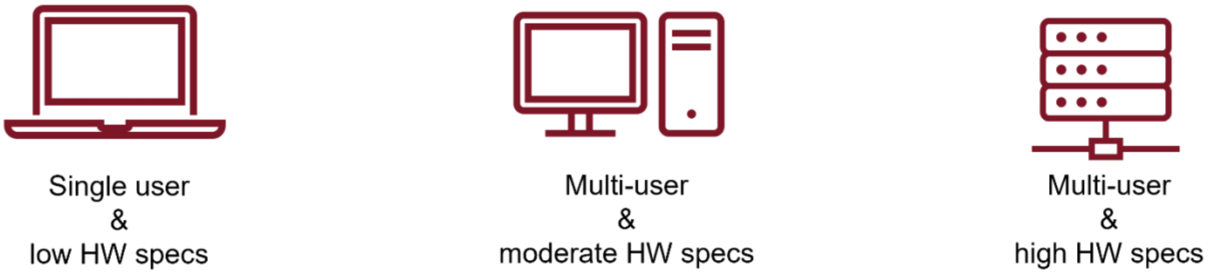


Figure 48: Candidate deployment models of the ROXANNE/Autocrime platform

Figure 49 presents the desktop of the ROXANNE platform in the case of a Virtual Machine.

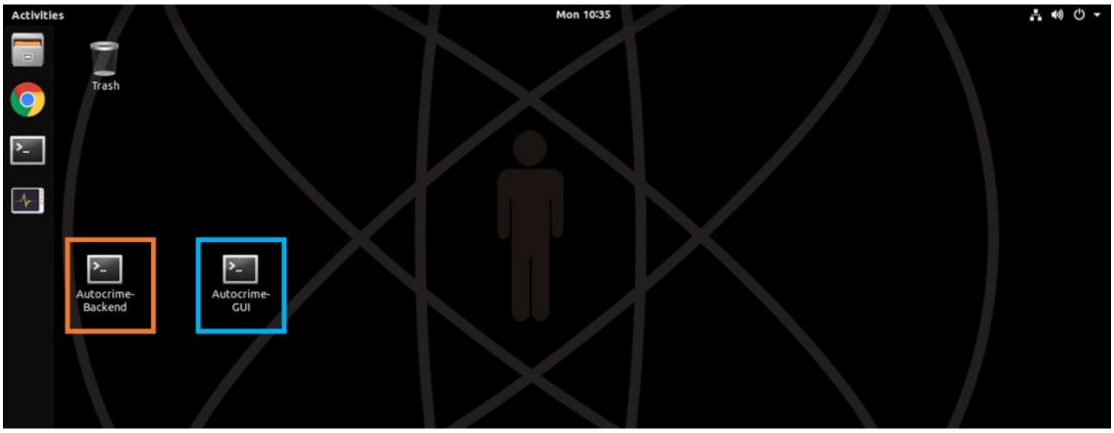
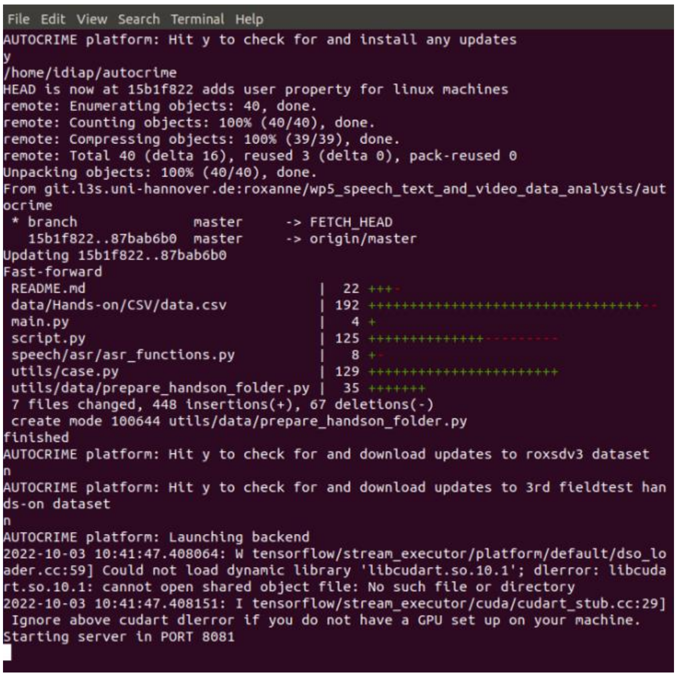


Figure 49: A screenshot of the Autocrime desktop in case of Virtual Machine instance

The user can choose whether the ROXANNE platform will be updated during start up (see SR-PM03-UpdatedPlatform system requirement). As soon as the user double clicks on the shortcut “Autocrime Backend” (i.e., the shortcut marked with an orange box of Figure 49), s/he will have to confirm if any updates to the platform should be downloaded and installed. Furthermore, the user can download newly available datasets from the default data repository (see SR-PM04-UpdatedDatasetsUsedByPlatform system requirement).



```
AUTOCRIME platform: Hit y to check for and install any updates
y
/home/idiap/autocrime
HEAD is now at 15b1f822 adds user property for linux machines
remote: Enumerating objects: 40, done.
remote: Counting objects: 100% (40/40), done.
remote: Compressing objects: 100% (39/39), done.
remote: Total 40 (delta 16), reused 3 (delta 0), pack-reused 0
Unpacking objects: 100% (40/40), done.
From git.l3s.uni-hannover.de:roxanne/wp5_speech_text_and_video_data_analysis/aut
ocrime
 * branch              master     -> FETCH_HEAD
   15b1f822..87bab6b0  master     -> origin/master
Updating 15b1f822..87bab6b0
Fast-forward
 README.md                           |  22 +++-
 data/Hands-on/CSV/data.csv          | 192 +++++++++++++++++++++++++++++++++--
 main.py                             |   4 +
 script.py                           | 125 ++++++++++++++---------
 speech/asr/asr_functions.py         |   8 +-
 utils/case.py                       | 129 +++++++++++++++++++++++
 utils/data/prepare_handson_folder.py |  35 +++++++
 7 files changed, 448 insertions(+), 67 deletions(-)
 create mode 100644 utils/data/prepare_handson_folder.py
finished
AUTOCRIME platform: Hit y to check for and download updates to roxsdv3 dataset
n
AUTOCRIME platform: Hit y to check for and download updates to 3rd fieldtest han
ds-on dataset
n
AUTOCRIME platform: Launching backend
2022-10-03 10:41:47.408064: W tensorflow/stream_executor/platform/default/dso_lo
ader.cc:59] Could not load dynamic library 'libcudart.so.10.1'; dlerror: libcuda
rt.so.10.1: cannot open shared object file: No such file or directory
2022-10-03 10:41:47.408151: I tensorflow/stream_executor/cuda/cudart_stub.cc:29]
 Ignore above cudart dlerror if you do not have a GPU set up on your machine.
Starting server in PORT 8081
```

Figure 50: A screenshot of exemplary console outputs when starting the ROXANNE platform's backend

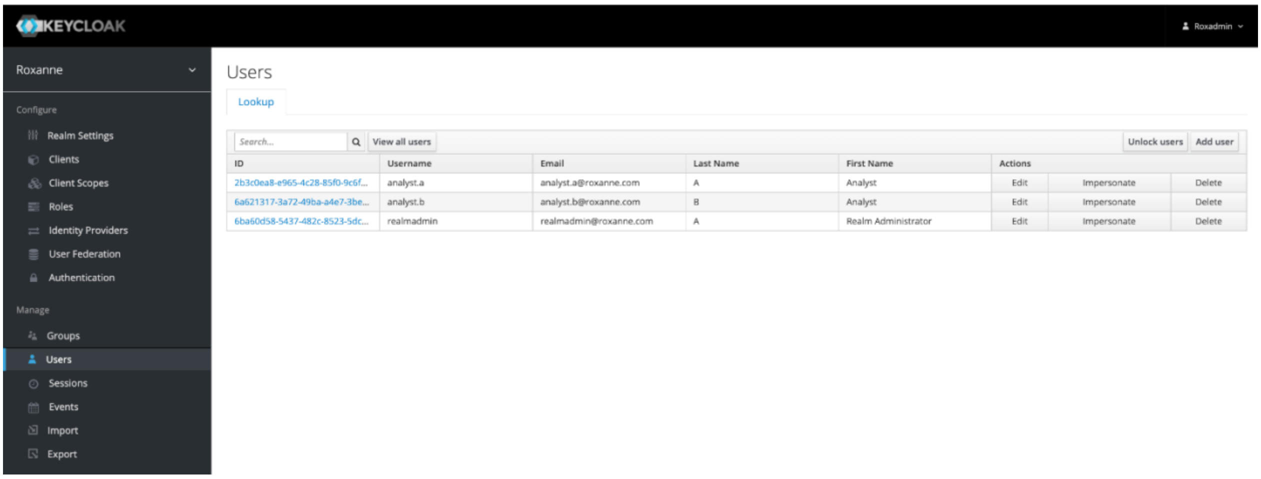
5.1.15 User management

Figure 51 presents the Users management dashboard, as offered by the Keycloak instance of the ROXANNE platform. An entity responsible for managing users' privileges (e.g., an administrator), can utilise the user interface of the Keycloak instance for the following activities:

- Register a new user (related to system requirements SR-UM01-NewUserAccountRequest and SR-UM02-NewUserAccountRegistration)
- See details about a particular user (in response to system requirements SR-UM05-UserDetailsRequest and SR-UM06-UserDetailsView)
- Update details of a user (e.g., manage users' privileges) according to system requirements SR-UM07-UserDetailsUpdateRequest, SR-UM08-UserDetailsUpdateConfirmation, SR-CM11-LimitedAccessGrantingRequest and SR-CM13-FullAccessGrantingRequest. In particular, the entity responsible for administering users' privileges can grant/revoke rights by supplying the required information based on the following JSON data schema (see [9] for more details).

```
{
    "caseA": ["create", "read", "update"],
    "caseB": ["read", "update"],
    "caseC": ["read"]
}
```

- Delete existing users in response to SR-UM09-UserDeleteRequest and SR-UM10-UserDeleteConfirmation.



ID	Username	Email	Last Name	First Name	Actions		
2b3c0ea8-e965-4c28-85f0-9c6f...	analyst.a	analyst.a@roxanne.com	A	Analyst	Edit	Impersonate	Delete
6a621317-3a72-49ba-a4e7-3be...	analyst.b	analyst.b@roxanne.com	B	Analyst	Edit	Impersonate	Delete
6ba60d58-5437-482c-8523-5dc...	realmadmin	realmadmin@roxanne.com	A	Realm Administrator	Edit	Impersonate	Delete

Figure 51: A screenshot of the User Management page provided by the Keycloak instance of the ROXANNE platform

Note that seeing a history of events that were logged (according to system requirements SR-CM15-LoggedEventsRequest and SR-CM17-LoggedEventsListing) is not currently supported.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

6. Conclusions

This deliverable provided a detailed description of the ROXANNE platform, which aims to support law enforcement practitioners in several phases of the investigation process. In particular, the ROXANNE platform supports with:

- **Collation/Processing** phase by processing and combining data of various modalities (e.g., audio recordings from wiretapped phone calls and video files from seized devices) for automating the most time-consuming tasks during evidence extraction. Furthermore, the ability to manage cases (e.g., creation, comparison and administration) and configure what processing should take place are key for efficiency.
- **Analysis** phase by automatically producing information on the entities (e.g., persons engaged in a phone call and/or appearing on a video footage) and their importance, the topics that are discussed, the languages used, etc. The user is offered with a wide range of visualisation schemes and filters to choose from so that different inferences can be formulated and tested. Furthermore, the user can assess the platform outputs and, whenever needed, to provide expert knowledge back to the platform.
- **Dissemination** phase by exporting key outputs in other commercial platforms that can be shared with other colleagues and generating reports.

The ROXANNE platform consists of four (4) families of technologies that are closely integrated: a) Speech-related technologies that include Speaker Recognition and Automated Speech Recognition, as well as Voice Activity Detection, Voiceprint extraction and Language identification, b) Text-processing technologies (namely Named-Entity Recognition, Topic Detection, Mention Network and Relation Extraction), c) Video-processing technologies (in particular face characterization and scene, object characterization) and d) Network analysis methods that refer to Community Detection, Social Influence Analysis, Link Prediction, Cross-Network Merging and Outlier Detection. The LEA practitioner uses the Graphical User Interface to configure the processing that should take place, to provide any intelligence already available, to explore the automated results and, if and when needed, to provide expert knowledge back to the platform by validating and updating the outputs.

By building upon the thorough analysis of current Law Enforcement Agents' (LEAs) practices, pain points and feedback received during the three field test demonstrations, this report documented a set of 37 (thirty-seven) user requirements which can be grouped into 5 (five) broad categories, namely platform management, user management, case management, evidence processing and case analysis. Furthermore, in terms of their importance, these user requirements were characterised (in decreasing order) as Must-have, Should-have or Could-have.

Figure 52 presents the number of supported and not-supported user requirements by priority type, where we see that the ROXANNE platform offers features that enables all high-priority (i.e., Must-have) requirements, as well as the majority of the rest requirements (see Table 7 in Annex B for a detailed overview). These features include the following: Cases Dashboard, Social Network Analysis, Timeline Analysis, Statistical Analysis, Details, Synced filters, Custom Workspace, Cross-case analysis, Reporting, Platform management and User management. For each of those ROXANNE platform features, a short description is provided along with selected screenshots that demonstrate key functionalities offered.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635.

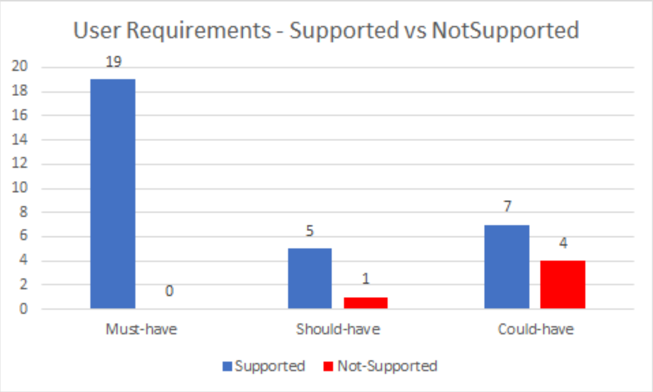


Figure 52: A screenshot of the User Management page provided by the Keycloak instance of the ROXANNE platform

References

[1] “Criminal Intelligence for Front Line Law Enforcement”, The United Nations Office on Drugs and Crime, 2010, available online from https://www.unodc.org/documents/organized-crime/Law-Enforcement/Criminal_Intelligence_for_Front_Line_Law_Enforcement.pdf

[2] D2.3 “End User Requirements”, ROXANNE consortium, public report, September 2020, available online from https://roxanne-euproject.org/results

[3] D8.4 “First Field test report and recommendations”, ROXANNE consortium, public report, October 2020, available online from https://roxanne-euproject.org/results

[4] D8.9: 2nd field test report and recommendations, ROXANNE consortium, confidential report

[5] D8.6: “Final field test report and recommendations” , ROXANNE consortium, confidential report

[6] D7.3: “Data visualization and exploratory analysis V1”, ROXANNE consortium, public report, June 2021, available online from https://roxanne-euproject.org/results

[7] D7.4: ”Description of ROXANNE platform V1, data export”, ROXANNE consortium, public report, December 2022, available online from https://roxanne-euproject.org/results

[8] D6.1 “Preliminary report on network analysis”, ROXANNE consortium, public report, June 2020, available online from https://roxanne-euproject.org/results

[9] D4.4: “ROXANNE case management and data preprocessing”, ROXANNE consortium, public report, June 2021, available online from https://roxanne-euproject.org/results

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

Annex A - State of the art on visualisation and exploratory analysis tools

In this section we focus on the most popular visualisation and exploratory analysis tools that law enforcement practitioners who participated in the global user requirements survey currently use (or plan to use in the near future). The following chart shows that Cellebrite, I2 Analyst's notebook and Forensic Toolkit (FTK) are most in use by the respondents.

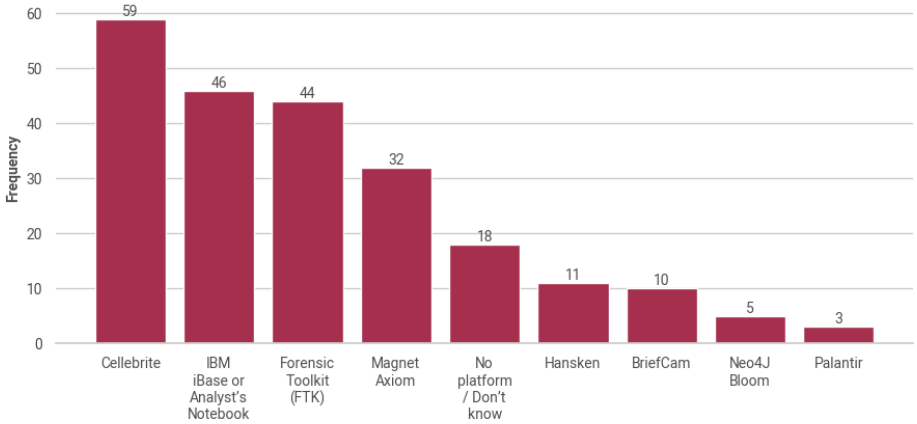


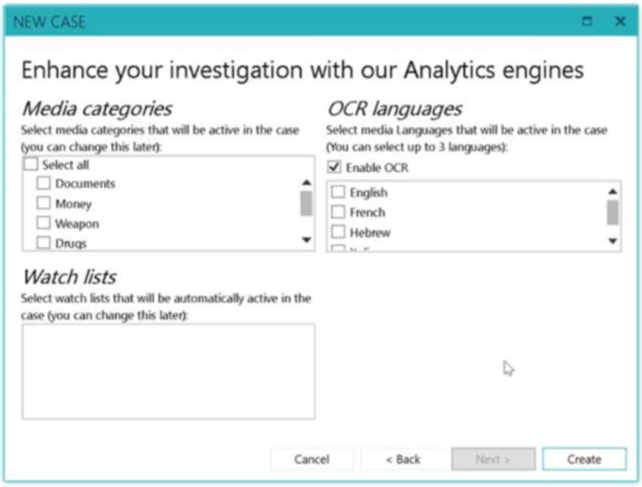
Figure 53: Most popular investigation platforms used by respondents to ROXANNE questionnaire on user requirements

4.1 Cellebrite Pathfinder

Cellebrite Pathfinder (formerly known as Cellebrite Analytics) is an investigative analytics solution that:

- Supports case management
- allows digital investigators to automate processing of structured (e.g., XML, CSV, and unstructured digital data, including multimedia files from cell phones, computers and metadata (such as CDRs).
- performs recognition and categorization of case-specific topics using AI and machine learning in order to:
  - extract entities such as person names, locations from text in multiple languages.
  - identify objects (e.g., drugs, money, guns) in images and videos

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.



NEW CASE

Enhance your investigation with our Analytics engines

Media categories

Select media categories that will be active in the case (you can change this later):

□ Select all
□ Documents
□ Money
□ Weapon
□ Drugs

OCR languages

Select media Languages that will be active in the case (You can select up to 3 languages):

☑ Enable OCR

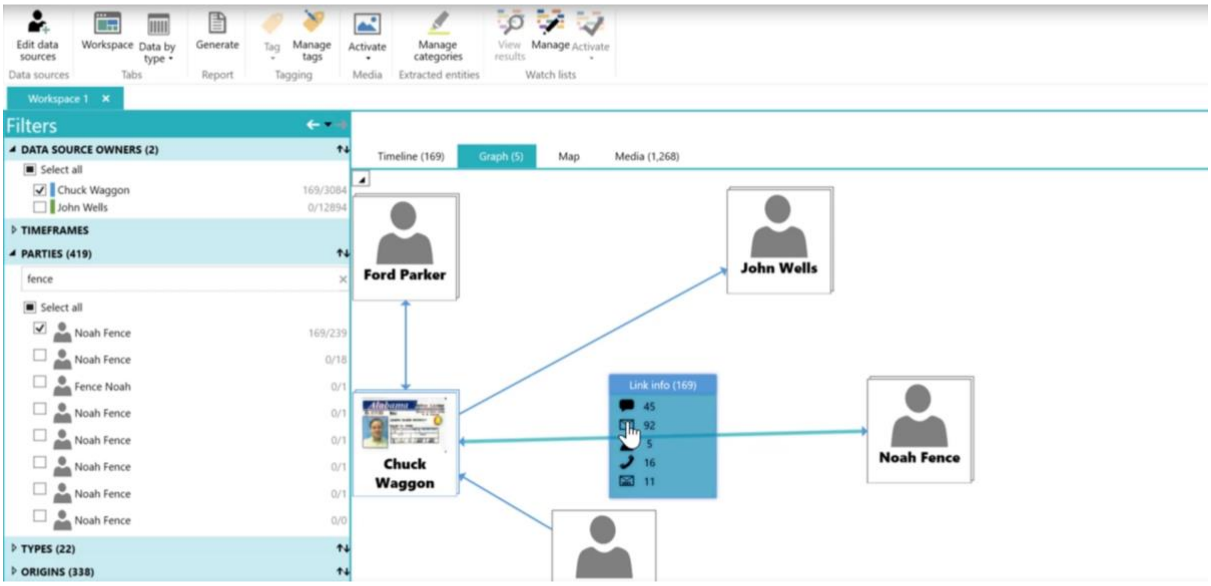
□ English
□ French
□ Hebrew

Watch lists

Select watch lists that will be automatically active in the case (you can change this later):

Figure 54: Screenshot of Cellebrite Pathfinder for enabling analytics engines that identify key entities and names in text or objects in images and videos

- provides analysts with features, including advanced filtering, as well as builds visual narratives (e.g., criminal network graphs and timeline analysis), in order to identify new insights and connections during the investigation.



Filters

DATA SOURCE OWNERS (2)

■ Select all
☑ Chuck Waggon 169/3084
□ John Wells 0/12894

TIMEFRAMES

PARTIES (419)

fence

■ Select all
☑ Noah Fence 169/239
□ Noah Fence 0/18
□ Fence Noah 0/1
□ Noah Fence 0/1
□ Noah Fence 0/1
□ Noah Fence 0/1
□ Noah Fence 0/1
□ Noah Fence 0/0

TYPES (22)

ORIGINS (338)

Figure 55: Screenshot of Cellebrite Pathfinder for manually (instead of automatically) identifying individuals

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

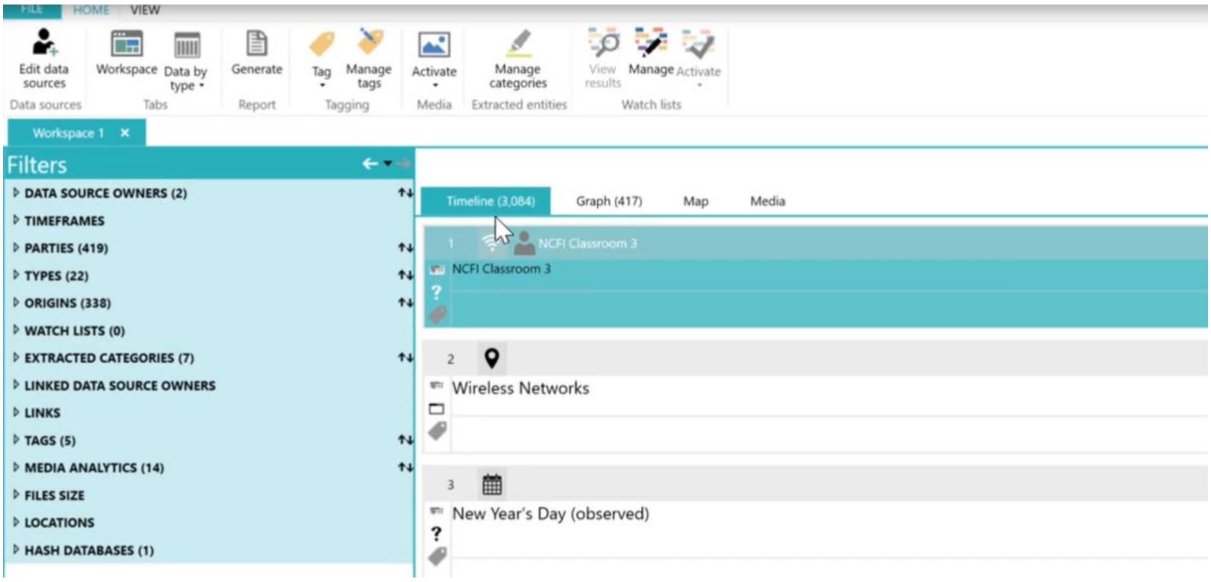


Figure 56: Screenshot of Cellebrite Pathfinder showing the set of filters that can be applied on processed data

Furthermore, Cellebrite Pathfinder:

- supports collaboration across departments, systems as well as external agencies by sharing case data and data tags, offering virtual dashboards and generating reports.
- manages user accounts, controls access to the collected evidence.
- offers high availability via backup and recovery capabilities and performance monitoring.

4.2 i2 Analyst's Notebook

The Analyst's Notebook is a popular family of tools for law enforcement agencies who analyse increasing volumes of data, visualise insights in order to discover networks, patterns and trends of criminal activity and communicate intelligence with their colleagues in order to identify, predict, prevent and disrupt criminal, terrorist and fraudulent activities.

i2 Analyst’s Notebook is a single-user, standalone desktop application with multidimensional visual analysis capabilities so that analysts can uncover hidden connections and patterns in data. Key Features include:

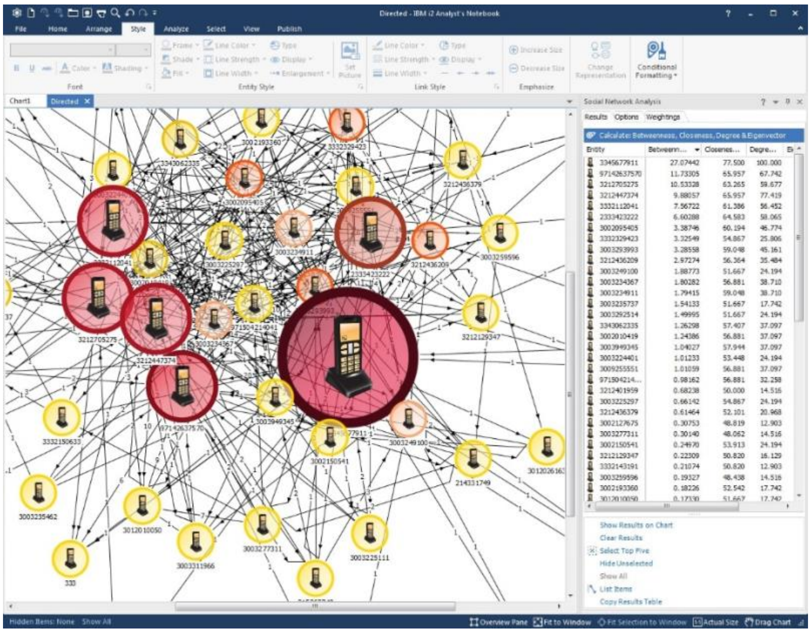
- Data extraction from text files and structured data, especially spreadsheets of data and connections.
- Support for geospatial data.
- Social network analysis for examining relationships between entities, such as people and organizations, found in input data and illustrating how they are interconnected via several attributes. Furthermore, it allows an analyst to find hidden connections, clusters of entities, and includes some basic functions for social influence analysis.
- Timeline analysis for presenting a sequence of events during a selected time window.
- Statistical views for looking how data is distributed via bar charts and histograms.
- Ability to import data from spreadsheet files, XML files, and from any text file in which each line describes a similar but discrete piece of information.

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

The i2 Analyst's Notebook can be combined with other tools for additional features. For example, it can be combined with the i2 iBase repository for gaining workgroup-level database that can serve as back-end, or the i2 Information Exchange for Analysis (iXa) that provides external search capabilities. Furthermore, analysts can use the "i2 Text Chart" for manual entity and event extraction based on ontologies. With respect to collaboration, analysts can share insights with colleagues who do not have access to i2 Analyst's Notebook via the i2 Chart Reader.

At the same time, there are several additional configurations:

- i2 Analyst's Notebook Premium which also provides local, desktop data repository
- and i2 Enterprise Insight Analysis (EIA)



Calculate: Betweenness, Closeness, Degree & Eigenvector

Entity	Betweenn...	Closenes...	Degre...
3345677911	27.07442	77.500	100.000
97142637570	11.73305	65.957	67.742
3212705275	10.53328	63.265	59.677
3212447374	9.88057	65.957	77.419
3332112041	7.56722	61.386	56.452
2333423222	6.60288	64.583	58.065
3002095405	3.38746	60.194	46.774
3332329423	3.32549	54.867	25.806
3003293993	3.28558	59.048	45.161
3212436209	2.97274	56.364	35.484
3003249100	1.88773	51.667	24.194
3003234367	1.80282	56.881	38.710
3003234911	1.79415	59.048	38.710
3003235737	1.54133	51.667	17.742
3003292514	1.49995	51.667	24.194
3343062335	1.26298	57.407	37.097
3002010419	1.24386	56.881	37.097
3003949345	1.04027	57.944	37.097
3003224401	1.01233	53.448	24.194
3009255551	1.01059	56.881	37.097
971504214...	0.98162	56.881	32.258
3212401959	0.68238	50.000	14.516
3003225297	0.66142	54.867	24.194
3212436379	0.61464	52.101	20.968
3002127675	0.30753	48.819	12.903
3003277311	0.30140	48.062	14.516
3002150541	0.24970	53.913	24.194
3212129347	0.22309	50.820	16.129
3332143191	0.21074	50.820	12.903
3003259596	0.19327	48.438	14.516
3002193360	0.18226	52.542	17.742
3012010050	0.17330	51.667	17.742

Figure 57: Screenshot of i2 Analyst Notebook showing a social network graph[24]

4.3 Forensic Toolkit

Forensic Toolkit (FTK) is another popular commercial product targeting investigators who want to perform digital forensic investigations and analyse those finding in order to get a clearer picture of events, relationships, and patterns.

FTK retrieves user files from captured electronic devices (e.g., mobile phones, computers etc) and data from operating system and applications, such as email and chat communications and stores this data in a full-featured database. User files can be viewed in a near-native format, while data belonging to each case can be exported for offline review. Premium versions of FTK also exist with additional features. For example,

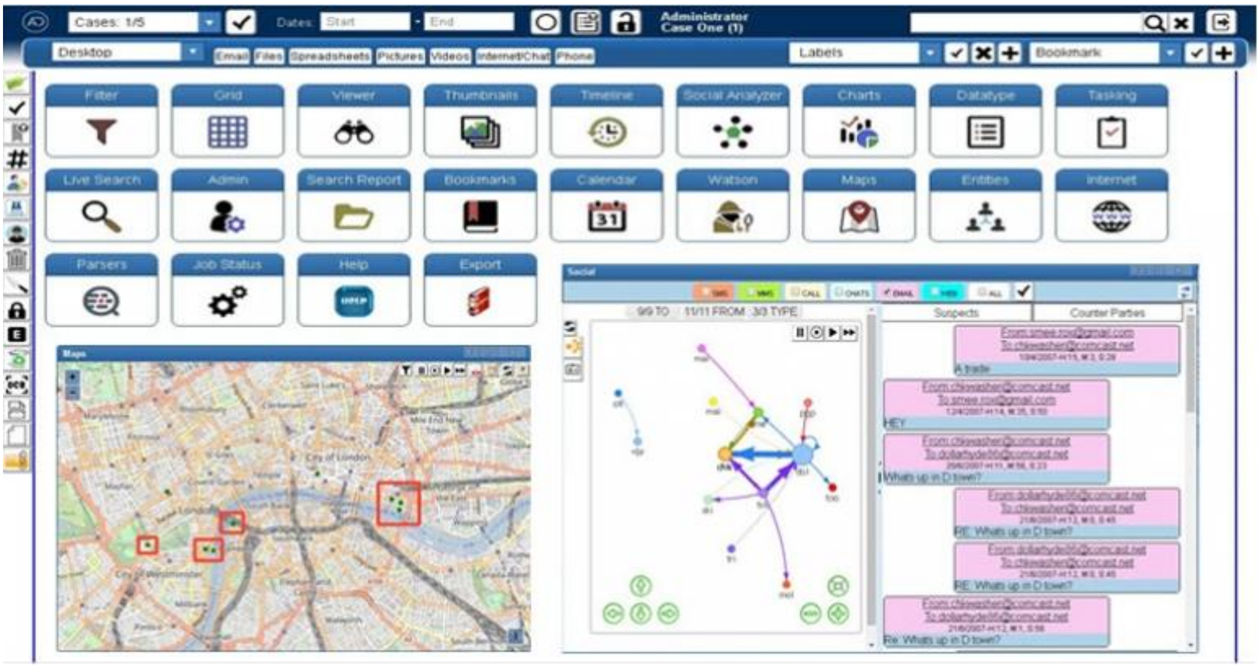
- FTK Central support remote collection of data from other institutional and external sources like Microsoft Office 365 and Slack, allows repetitive tasks to be automated.

[24] https://www.ibm.com/products/i2-analysts-notebook

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

- Labs offers collaborative features such as the ability for administrators to grant full edit rights or view-only access to other colleagues for each separate case, audit log and evidence tracking to improve chain of custody documentation, automated email notifications when evidence collection or processing is complete.

FTK offers advanced search functionality, such as detection of similar faces and objects inside the case or across cases, as well as tagging and labeling interesting findings by using AI techniques. The user can use the case dashboard for viewing the status of processing jobs, data in timeline charts, pie charts and maps as well as the relationships between entities. Nevertheless, entities are recognised based on identifiers found in data, such as email address, rather than finetuned based on the context and content of data.



From smee.roo@gmail.com
To chiwasher@comcast.net
[illegible]
A trade

From chiwasher@comcast.net
To smee.roo@gmail.com
[illegible]
HEY

From chiwasher@comcast.net
To dollarhyde86@comcast.net
[illegible]
Whats up in D town?

From dollarhyde86@comcast.net
To chiwasher@comcast.net
[illegible]
RE: Whats up in D town?

From dollarhyde86@comcast.net
To chiwasher@comcast.net
[illegible]
RE: Whats up in D town?

From chiwasher@comcast.net
To dollarhyde86@comcast.net
[illegible]
Re: Whats up in D town?

Figure 58: Screenshot of Forensic Toolkit workspace

4.4 Magnet Axiom

Axiom is a digital investigation and analysis platform like Forensic Toolkit. It can be used for recovering digital evidence from several user types of user devices, ingesting user-requested archive files, as well as warrant returns from user-generated content providers. In particular, it can be used for acquiring digital evidence, such as pictures, chat messages, and browser history, from smartphones, computers, cloud services like Google, Apple, Facebook, third-party services such as Cellebrite and Oxygen and other encrypted applications (e.g., Snapchat). Then, Axiom uses built-in machine learning algorithms and other plug-ins to flag inappropriate files (e.g., those containing nudity, weapons, drugs) and allows the user to report and adjust any errors identified for improving the defensibility and reproducibility of examinations.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

Furthermore, Axiom employs visualisation techniques to automatically show relationships between artifacts, files, and people. The figure below shows where different files came from, who they are connected to, and where they were stored.

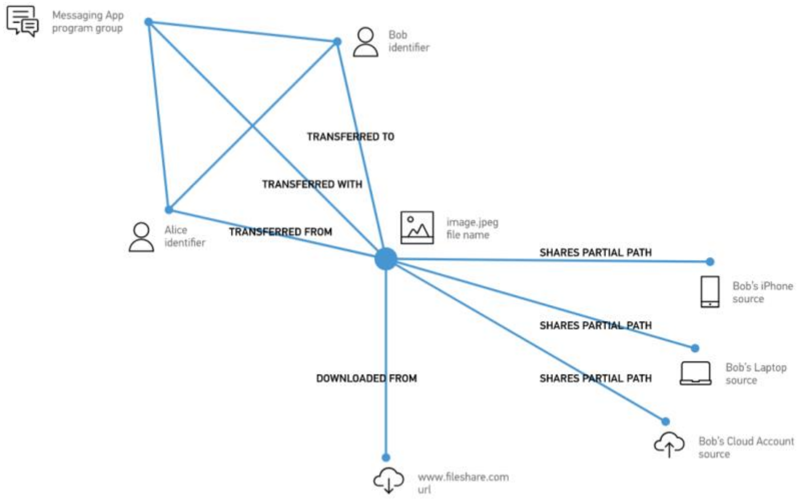
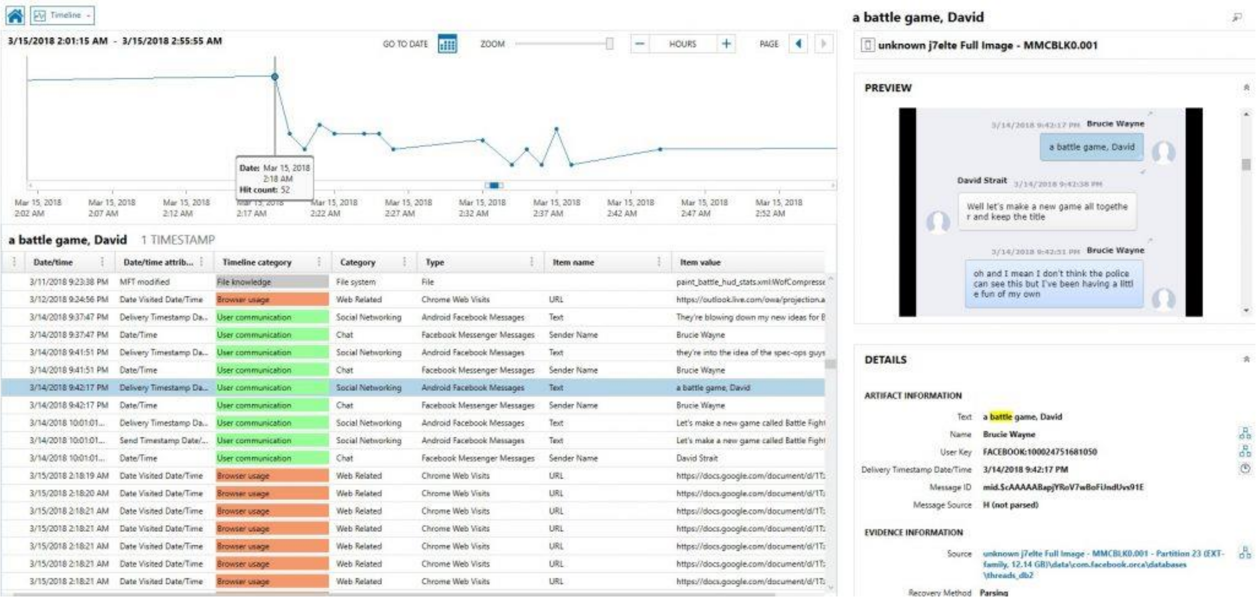


Figure 59: Screenshot of Magnet Axiom showing a social network graph[25]

Similarly, Axiom can present evidence on a timeline for reconstructing how an incident occurred. The analyst can view all timestamps recovered from devices' files and file system details on the top left pane of the following figure, while the bottom left window includes all the timestamps in a tabular format categorized into activity types (e.g., Deleted File, Physical Location, File/Folder Opening, Account Usage, User Communication and Social Activity). Once the user has selected a particular entry, a preview of the evidence will appear on the top right corner (if available) along with further details on the bottom right.



3/15/2018 2:01:15 AM - 3/15/2018 2:55:55 AM

Date: Mar 15, 2018 2:18 AM
Hit count: 52

a battle game, David 1 TIMESTAMP

Date/time	Date/time attrib...	Timeline category	Category	Type	Item name	Item value
3/11/2018 9:23:38 PM	MFT modified	File knowledge	File system	File		paint_battle_hud_stats.xml:WofCompresse
3/12/2018 9:24:56 PM	Date Visited Date/Time	Browser usage	Web Related	Chrome Web Visits	URL	https://outlook.live.com/owa/projection.a
3/14/2018 9:37:47 PM	Delivery Timestamp Da...	User communication	Social Networking	Android Facebook Messages	Text	They're blowing down my new ideas for B
3/14/2018 9:37:47 PM	Date/Time	User communication	Chat	Facebook Messenger Messages	Sender Name	Brucie Wayne
3/14/2018 9:41:51 PM	Delivery Timestamp Da...	User communication	Social Networking	Android Facebook Messages	Text	they're into the idea of the spec-ops guys
3/14/2018 9:41:51 PM	Date/Time	User communication	Chat	Facebook Messenger Messages	Sender Name	Brucie Wayne
3/14/2018 9:42:17 PM	Delivery Timestamp Da...	User communication	Social Networking	Android Facebook Messages	Text	a battle game, David
3/14/2018 9:42:17 PM	Date/Time	User communication	Chat	Facebook Messenger Messages	Sender Name	Brucie Wayne
3/14/2018 10:01:01...	Delivery Timestamp Da...	User communication	Social Networking	Android Facebook Messages	Text	Let's make a new game called Battle Fight
3/14/2018 10:01:01...	Send Timestamp Date/...	User communication	Social Networking	Android Facebook Messages	Text	Let's make a new game called Battle Fight
3/14/2018 10:01:01...	Date/Time	User communication	Chat	Facebook Messenger Messages	Sender Name	David Strait
3/15/2018 2:18:19 AM	Date Visited Date/Time	Browser usage	Web Related	Chrome Web Visits	URL	https://docs.google.com/document/d/1T:
3/15/2018 2:18:20 AM	Date Visited Date/Time	Browser usage	Web Related	Chrome Web Visits	URL	https://docs.google.com/document/d/1T:
3/15/2018 2:18:21 AM	Date Visited Date/Time	Browser usage	Web Related	Chrome Web Visits	URL	https://docs.google.com/document/d/1T:
3/15/2018 2:18:21 AM	Date Visited Date/Time	Browser usage	Web Related	Chrome Web Visits	URL	https://docs.google.com/document/d/1T:
3/15/2018 2:18:21 AM	Date Visited Date/Time	Browser usage	Web Related	Chrome Web Visits	URL	https://docs.google.com/document/d/1T:
3/15/2018 2:18:21 AM	Date Visited Date/Time	Browser usage	Web Related	Chrome Web Visits	URL	https://docs.google.com/document/d/1T:
3/15/2018 2:18:21 AM	Date Visited Date/Time	Browser usage	Web Related	Chrome Web Visits	URL	https://docs.google.com/document/d/1T:

a battle game, David

unknown j7elte Full Image - MMCBLK0.001

3/14/2018 9:42:17 PM Brucie Wayne
a battle game, David

David Strait 3/14/2018 9:42:38 PM
Well let's make a new game all togethe r and keep the title

3/14/2018 9:42:51 PM Brucie Wayne
oh and I mean I don't think the police can see this but I've been having a littl e fun of my own

ARTIFACT INFORMATION

Text	a battle game, David
Name	Brucie Wayne
User Key	FACEBOOK:100024751681050
Delivery Timestamp Date/Time	3/14/2018 9:42:17 PM
Message ID	mid.$cAAAAABapjYRoV7wBoFiJndUvs91E
Message Source	H (not parsed)

EVIDENCE INFORMATION

Source	unknown j7elte Full Image - MMCBLK0.001 - Partition 23 (EXT-family, 12.14 GB)\data\com.facebook.orca\databases\threads_db2
Recovery Method	Parsing

[25] https://www.magnetforensics.com/products/magnet-axiom/

Figure 60: Screenshot of Magnet Axiom showing a timeline graph[26]

Finally, the user can generate a report that can be customised to the case or the intended audience, by choosing the template to be used and the case-related data that should be included.

4.5 Hansken

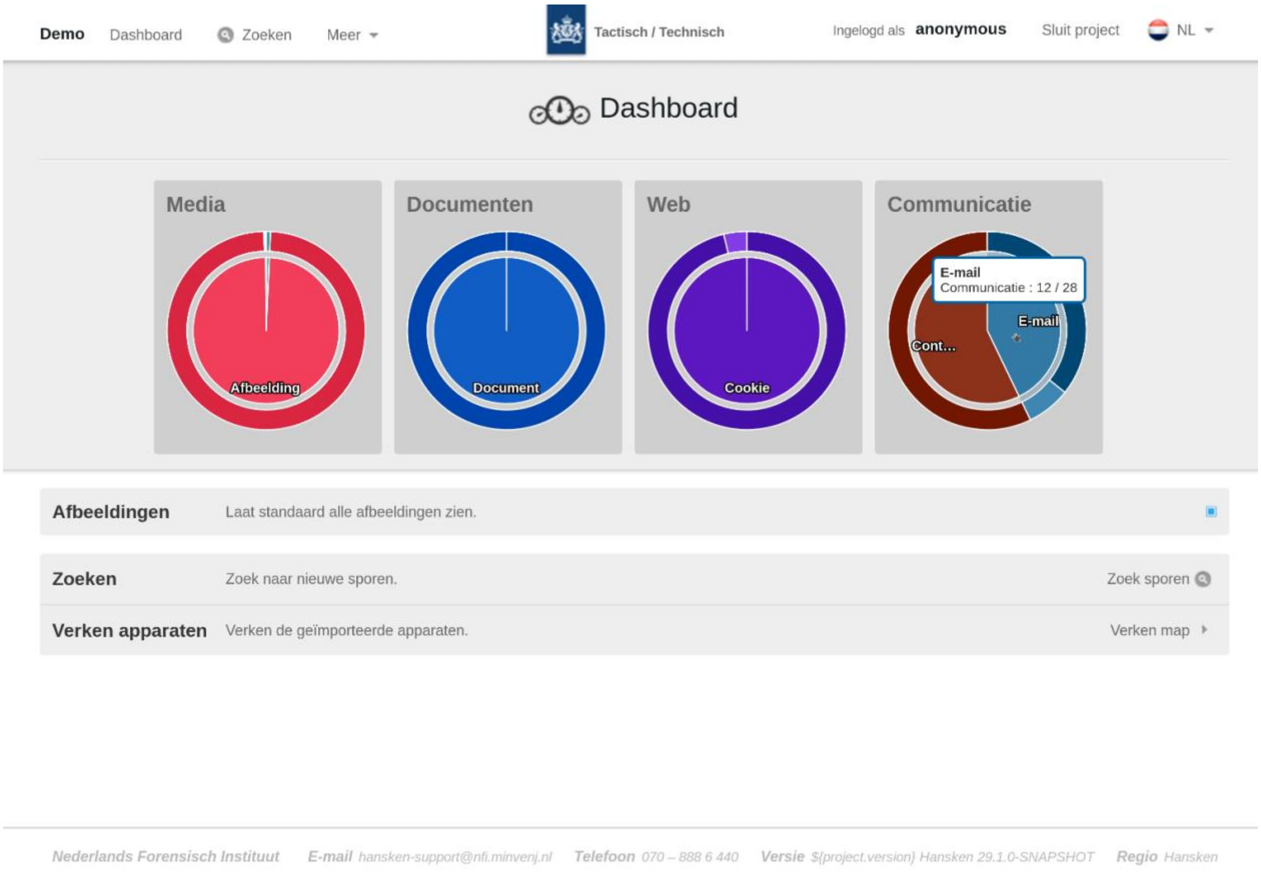
Hansken is a forensic analysis system developed at the Netherlands Forensic Institute and a successor of XIRAF. Hansken automates the collection of forensic artifacts, organizes these artifacts in a centralised database so that investigators can effectively and efficiently discover the most relevant ones for answering investigative questions.

Hansken has the following features:

- supports case management that allows digital inputs such as hard disk images to be associated with new or existing projects.
- enables user management for creating user accounts and assigning access rights for projects/cases to users.
- accepts a variety of digital evidence inputs such as files, chat logs, browser histories, processes, email, etc. to be uploaded
- allows processing workflow to be specified for different evidence inputs, i.e., the set of analysis tools to be run for automatically extracting forensic artifacts
- allows artifacts to be shared with colleagues
- allows users to search a large collection of forensic artifacts along multiple dimensions, e.g., type of evidence input used as source (such as mobile phone image, hard disk image, etc.), property of the artifact like creation date, location, etc.
- enables users to view the artifacts in multiple ways, such as opening the artifact with the appropriate application, statistical view (e.g., percentage of related pictures, videos, email messages, etc.), tabular view for providing an overview of key properties (table columns) of the related artifacts (rows), timeline view for presenting objects in chronological order according to timestamps, map view for including the location where events took place, etc.
- allows users to annotate individual object with notes
- generates predefined or custom reports

[26] https://www.magnetforensics.com/resources/using-the-new-timeline-explorer-in-magnet-axiom-3-0/

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.



Afbeeldingen	Laat standaard alle afbeeldingen zien.	
Zoeken	Zoek naar nieuwe sporen.	Zoek sporen
Verken apparaten	Verken de geïmporteerde apparaten.	Verken map

Nederlands Forensisch Instituut E-mail hansken-support@nfi.minvenj.nl Telefoon 070 – 888 6 440 Versie ${project.version} Hansken 29.1.0-SNAPSHOT Regio Hansken

Figure 61: Screenshot of Hansken showing a case dashboard[27]

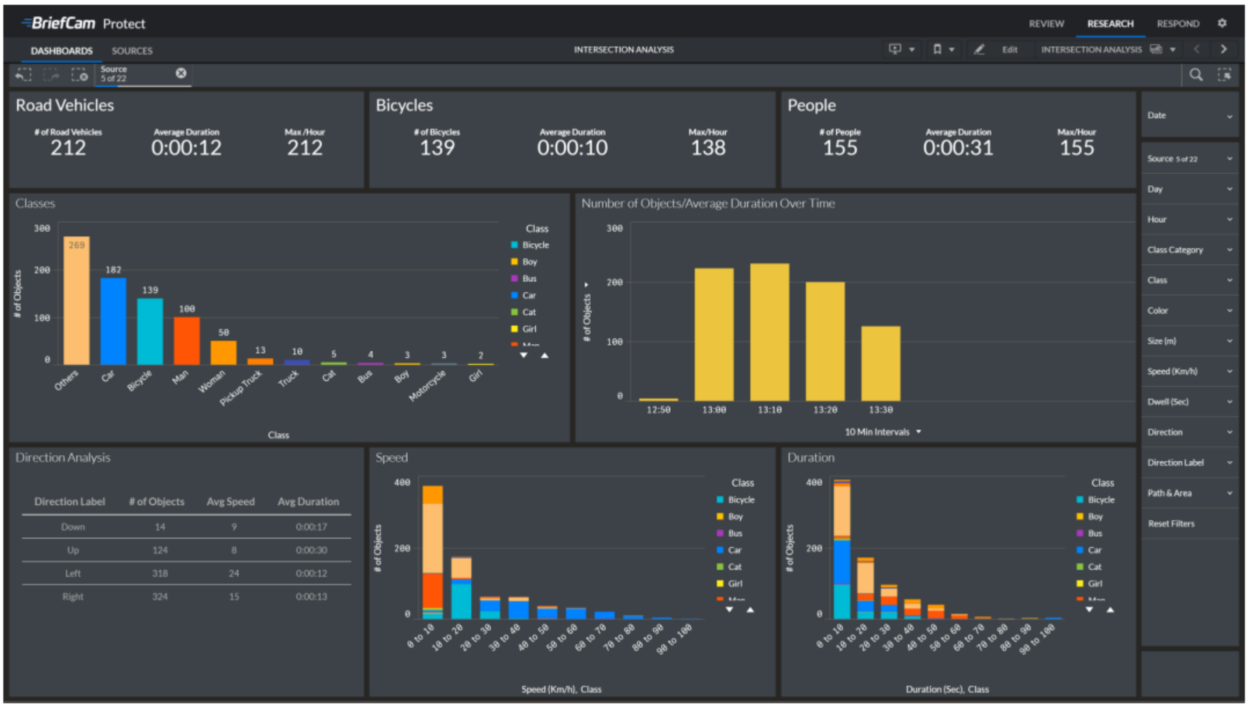
4.6 BriefCam

The BriefCam platform is an all-in-one video analytics solution, comprised of hardware and software for processing video in order to detect, track, extract and identify people, objects, their attributes and behavior. These insights are then visualised using customizable dashboards for operational decision making. BriefCam offers the following features:

- Case management, where all video assets of an investigation and related findings are organised in a single logical folder, allowing collaboration with other users to take place.
- Video synopsis creation, where a video footage is dramatically shortened by keeping a series of screenshots conveying useful information (i.e., frames showing changing scenes).
- Video analytics, such as license plate recognition, face recognition, face mask detection, proximity estimation, etc
- Real-time alerts where a notification is triggered if certain user-defined conditions (e.g., faces recognised, objects are identified) are met
- Dashboards for visualising and analysing key performance indicators related to object movement, demographic segmentations, behavior trending, hotspots and object interactions using custom and preconfigured charts.

[27] https://hansken.org/demo

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.



Road Vehicles

# of Road Vehicles	Average Duration	Max/Hour
212	0:00:12	212

Bicycles

# of Bicycles	Average Duration	Max/Hour
139	0:00:10	138

People

# of People	Average Duration	Max/Hour
155	0:00:31	155

Direction Analysis

Direction Label	# of Objects	Avg Speed	Avg Duration
Down	14	9	0:00:17
Up	124	8	0:00:30
Left	318	24	0:00:12
Right	324	15	0:00:13

Figure 62: Screenshot of Briefcam showing a case workspace

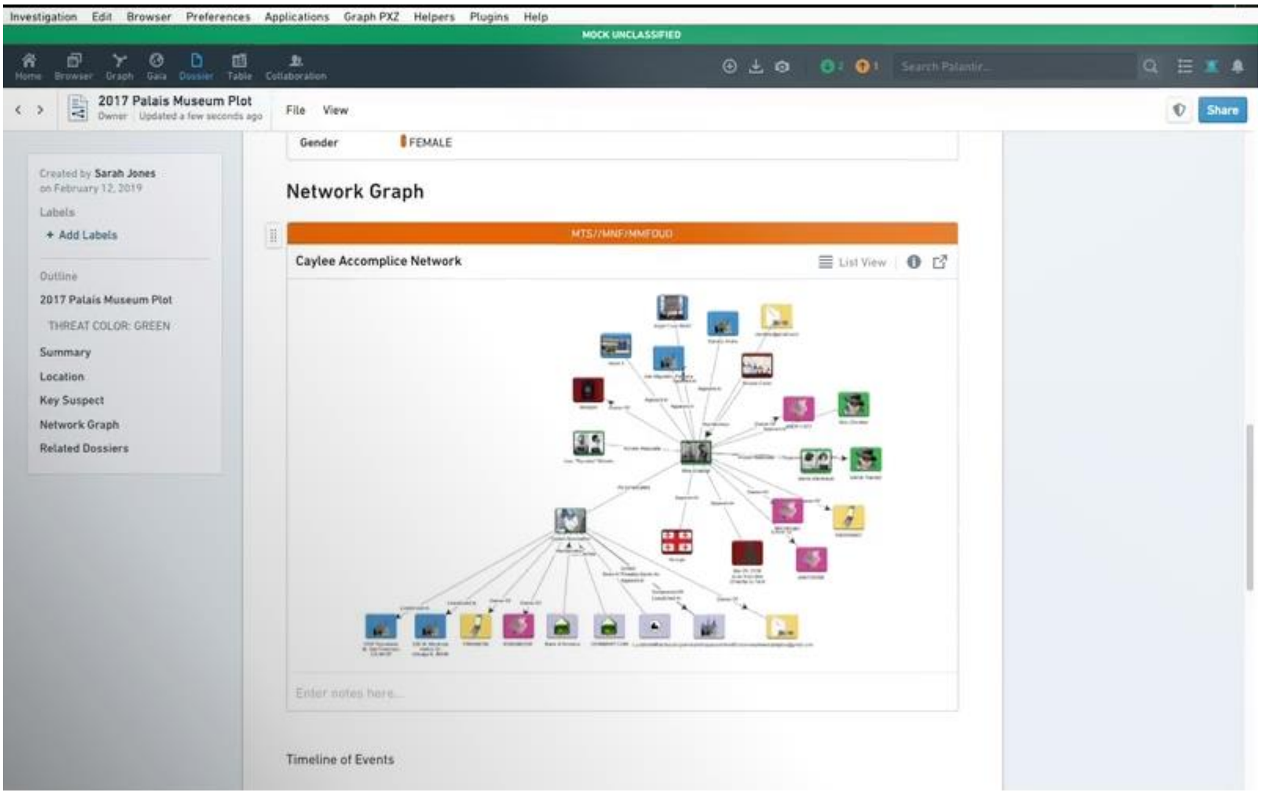
4.7 Palantir Gotham

Palantir Gotham provides a suite of tools for semantic, temporal and geospatial analysis of criminal data from multiple data sources so that law enforcement practitioners can explore hypotheses, get insights (e.g., unknown connections and patterns) and eventually share actionable intelligence with their colleagues.

Palantir Gotham has the following features:

- Support for multiple, federated and voluminous data sources spanning several cases (e.g., users can drag and drop data objects across cases).
- Enable secure and auditable collaboration between individuals, teams and organisations.
- Automated tools for investigative analysis across different data sets using artificial intelligence techniques, for example text entity extraction is supported.
- Several types of visualisation charts for different types of data and workflows (geospatial, network diagram for displaying information flows, timeline for viewing events, CDR etc.) and to be used by advanced and non-technical analysts.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.



Created by Sarah Jones on February 12, 2019

Labels

+ Add Labels

2017 Palais Museum Plot

THREAT COLOR: GREEN

Gender FEMALE

Caylee Accomplice Network

Enter notes here...

Figure 63: Screenshot of Palantir Gotham showing a social network graph[28]

- Support for filtering and search based on ontologies.
- Compilation and sharing of investigative reports internally and with external partners
- Continuous monitoring of system status and uptime

4.8 Comparing the ROXANNE platform against other commercial tools

The following table provides a high-level overview of the features currently available in the ROXANNE platform and how it stacks up against competitive solutions.

Table YY: Comparison of ROXANNE platform against other commercial tools

	Feature	Cellebrite Pathfinder	i2 Analyst's Notebook	Forensic Toolkit	Magnet Axiom	Hansken	Briefcam	Palantir Gotham	ROXANNE PLATFORM
Evidence files	Audio	Limited ✓	Limited ✓	Limited ✓	Limited ✓	Limited ✓	✗	Limited ✓	✓

[28] https://www.palantir.com/palantir-gotham/titan/

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

	Structured Data	✓	✓	✓	✓	✓	✗	✓	Limited ✓
	Unstructured Data (e.g., Text)	✓	✓	✓	✓	✓	✗	✓	Limited ✓
	Videos	✓	Limited ✓	✓	✓	✓	✓	✓	✓
	Images	✓	✓	✓	✓	✓	✗	✓	✓
Audio processing	Speaker Clustering and Identification	✗	✗	✗	✗	✗	✗	✗	✓
	Language Identification	✗	✗	✗	✗	✗	✗	✗	✓
	Automatic Speech Recognition	✗	✗	✗	✗	✗	✗	✗	✓
Video analysis	Face Detection	✓	Images only ✓	✓	✗	✗	✓	✓	✓
	Scene Detection	✓	Images ✓ only	✓	✗	✗	✓	✓	✓
Text processing	Named Entity Recognition	✓	✓	Keyword-based ✓	✗	✓	✗	✓	✓
	Topic Detection	✓	✗	✗	✗	✗	✗	✓	✓
Network analysis	Link Prediction	✗	✓	✗	✗	✗	✗	✗	✓

	Social Influence Analysis	✗	✓	✓	✗	✗	✗	✗	✓
	Community Detection	✗	✓	✗	✗	✗	✗	✓	✓
	Outlier Detection	✗	✗	✗	✗	✗	✗	✗	✓
Exploratory analysis	Social Network Visualization	✓	✓	✓	✓	✓	✗	✓	✓
	Timeline Analysis	✓	✓	✓	✓	✓	✓	✓	✓
	Statistical Analysis	✓	✓	✓	✓	✓	✓	✓	✓
	Geospatial Analysis	✓	✓	✓	✓	✓	✓	✓	(basic) ✓
	Details on Entities and Events	✓	✓	✓	✓	✓	✓	✓	✓
	Advanced Filters	✓	✓	✓	✓	✓	✓	✓	✓
General features	Case Management	✓	✓	✓	✓	✓	✓	✓	✓
	Reporting	✓	✓	✓	✓	✓	✓	✓	✓
	Cross Case Analysis	✓	✓	✓	✗	✓	✓	✓	✓

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

Annex B – Mapping of System Requirements to User Requirements

Table 7: Mapping of System Requirements to User Requirements and Status

System Requirement	User Requirement	Status of User Requirement	User Requirement Priority	Feature
SR-CM01-NewCaseCreationRequest	UR-CM01-CaseCreation	Supported	Must-have	Cases Dashboard
SR-CM02-NewCaseCreationConfirmation	UR-CM01-CaseCreation	Supported	Must-have	Cases Dashboard
SR-CM08-UserCasesRequest	UR-CM03-CaseOverview	Supported	Could-have	Cases Dashboard
SR-CM10-AuthorisedCasesListing	UR-CM03-CaseOverview	Supported	Could-have	Cases Dashboard
SR-CM03-CaseDescriptionUpdateRequest	UR-CM02-CaseEdit	Not-Supported	Could-have	N/A
SR-CM06-CaseDescriptionUpdateRejection	UR-CM02-CaseEdit	Not-Supported	Could-have	N/A
SR-UM03-UserAuthenticationRequest	UR-UM05-UserLogIn	Supported	Must-have	Cases Dashboard
SR-UM04-UserAuthenticationResponse	UR-UM05-UserLogIn	Supported	Must-have	Cases Dashboard
SR-EP06-MetaDataFileUploading	UR-EP03-MetadataUpload	Supported	Must-have	Cases Dashboard
SR-EP01-AudioFileUploading	UR-EP01-EvidenceUpload	Supported	Must-have	Uploading files to an existing case
SR-EP02-VideoFileUploading	UR-EP01-EvidenceUpload	Supported	Must-have	Uploading files to an existing case
SR-EP04-EvidenceFilePersistentStorage	UR-EP01-EvidenceUpload	Supported	Must-have	Uploading files to an existing case
SR-EP07-EvidenceWorkflowConfiguration	UR-EP04-ConfigureWorkflow	Supported	Must-have	Uploading files to an existing case
SR-EP03-TextFileUploading	UR-EP01-EvidenceUpload	Not-Supported	Must-have	N/A
SR-EP05-IncrementalEvidenceWarning	UR-EP02-EvidenceAlert	Not-Supported	Could-have	N/A
SR-CA07-FinetuneAlgorithmParameters	UR-CA03-EntityRelationshipsAnalysis	Supported	Must-have	Configuring an existing case
SR-CA38-AddContextSpecificKeywords	UR-CA09-CustomRepository	Supported	Must-have	Configuring an existing case
SR-EP08-RealTimeEvidenceProcessingStatusUpdate	UR-EP05-MonitorWorkflow	Supported	Could-have	Configuring an existing case
SR-CM18-AlternativeScenarios	UR-CM07-CaseScenarios	Supported	Must-have	Configuring an existing case

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

SR-CA01-ProcessedEventsRequest	UR-CA01-TemporalAnalysis	Supported	Must-have	Social Network Analysis
SR-CA01-ProcessedEventsRequest	UR-CA02-GeoSpatialAnalysis	Not-Supported	Should-have	N/A
SR-CA01-ProcessedEventsRequest	UR-CA03-EntityRelationshipsAnalysis	Supported	Must-have	Social Network Analysis
SR-CA01-ProcessedEventsRequest	UR-CA06-MultidimensionalAnalysis	Supported	Must-have	Social Network Analysis
SR-CA08-PresentSocialNetwork	UR-CA03-EntityRelationshipsAnalysis	Supported	Must-have	Social Network Analysis
SR-CA04-ShowEventOverview	UR-CA01-TemporalAnalysis	Supported	Must-have	Social Network Analysis
SR-CA04-ShowEventOverview	UR-CA06-MultidimensionalAnalysis	Supported	Must-have	Social Network Analysis
SR-CA04-ShowEventOverview	UR-CA02-GeoSpatialAnalysis	Not-Supported	Should-have	N/A
SR-CA04-ShowEventOverview	UR-CA03-EntityRelationshipsAnalysis	Supported	Must-have	Social Network Analysis
SR-CA10-AdjustSocialNetworkAppearance	UR-CA03-EntityRelationshipsAnalysis	Supported	Must-have	Social Network Analysis
SR-CA11-ChooseSocialNetworkZoomLevel	UR-CA03-EntityRelationshipsAnalysis	Supported	Must-have	Social Network Analysis
SR-CA12-SocialNetworkScrolling	UR-CA03-EntityRelationshipsAnalysis	Supported	Must-have	Social Network Analysis
SR-CA13-ShowEntityOverview	UR-CA01-TemporalAnalysis	Supported	Must-have	Social Network Analysis
SR-CA13-ShowEntityOverview	UR-CA06-MultidimensionalAnalysis	Supported	Must-have	Social Network Analysis
SR-CA13-ShowEntityOverview	UR-CA02-GeoSpatialAnalysis	Not-Supported	Should-have	N/A
SR-CA13-ShowEntityOverview	UR-CA03-EntityRelationshipsAnalysis	Supported	Must-have	Social Network Analysis
SR-CA09-ShowCommunitiesOfEntities	UR-CA03-EntityRelationshipsAnalysis	Supported	Must-have	Social Network Analysis
SR-CA42-HighlightImportantEntities	UR-CA13-AutomatedSuggestions	Supported	Could-have	Social Network Analysis

SR-CA42-HighlightImportantEntities	UR-CA03-EntityRelationshipsAnalysis	Supported	Must-have	Social Network Analysis
SR-CA43-ShowPredictedRelationships	UR-CA13-AutomatedSuggestions	Supported	Could-have	Social Network Analysis
SR-CA43-ShowPredictedRelationships	UR-CA03-EntityRelationshipsAnalysis	Supported	Must-have	Social Network Analysis
SR-CA29-AnnotateOnSocialNetworks	UR-CA07-UserAnnotation	Supported	Should-have	Social Network Analysis
SR-CA29-AnnotateOnSocialNetworks	UR-CA03-EntityRelationshipsAnalysis	Supported	Must-have	Social Network Analysis
SR-CA33-AddEvents	UR-CA08-OutputRevision	Not-Supported	Must-have	N/A
SR-CA33-AddEvents	UR-CA03-EntityRelationshipsAnalysis	Not-Supported	Must-have	N/A
SR-CA34-UpdateEntityProperties	UR-CA08-OutputRevision	Not-Supported	Must-have	N/A
SR-CA35-EntityPropertiesUpdateRejection	UR-CA08-OutputRevision	Not-Supported	Must-have	N/A
SR-CA36-UpdateEventProperties	UR-CA08-OutputRevision	Not-Supported	Must-have	N/A
SR-CA37-EnventPropertiesUpdateRejection	UR-CA08-OutputRevision	Not-Supported	Must-have	N/A
SR-CA01-ProcessedEventsRequest	UR-CA01-TemporalAnalysis	Supported	Must-have	Timeline Analysis
SR-CA01-ProcessedEventsRequest	UR-CA03-EntityRelationshipsAnalysis	Supported	Must-have	Timeline Analysis
SR-CA01-ProcessedEventsRequest	UR-CA06-MultidimensionalAnalysis	Supported	Must-have	Timeline Analysis
SR-CA01-ProcessedEventsRequest	UR-CA02-GeoSpatialAnalysis	Not-Supported	Should-have	N/A
SR-CA03-PresentEventsEvolution	UR-CA01-TemporalAnalysis	Supported	Must-have	Timeline Analysis
SR-CA04-ShowEventOverview	UR-CA01-TemporalAnalysis	Supported	Must-have	Timeline Analysis
SR-CA04-ShowEventOverview	UR-CA03-EntityRelationshipsAnalysis	Supported	Must-have	Timeline Analysis
SR-CA04-ShowEventOverview	UR-CA06-MultidimensionalAnalysis	Supported	Must-have	Timeline Analysis

SR-CA04-ShowEventOverview	UR-CA02-GeoSpatialAnalysis	Not-Supported	Should-have	N/A
SR-CA01-ProcessedEventsRequest	UR-CA02-GeoSpatialAnalysis	Not-Supported	Should-have	N/A
SR-CA01-ProcessedEventsRequest	UR-CA01-TemporalAnalysis	Supported	Must-have	Timeline Analysis
SR-CA01-ProcessedEventsRequest	UR-CA03-EntityRelationshipsAnalysis	Supported	Must-have	Timeline Analysis
SR-CA01-ProcessedEventsRequest	UR-CA06-MultidimensionalAnalysis	Supported	Must-have	Timeline Analysis
SR-CA23-ShowStatisticalResults	UR-CA05-StatisticalAnalysis	Supported	Must-have	Statistical Analysis
SR-CA01-ProcessedEventsRequest	UR-CA02-GeoSpatialAnalysis	Not-Supported	Should-have	N/A
SR-CA01-ProcessedEventsRequest	UR-CA01-TemporalAnalysis	Supported	Must-have	Details
SR-CA01-ProcessedEventsRequest	UR-CA03-EntityRelationshipsAnalysis	Supported	Must-have	Details
SR-CA01-ProcessedEventsRequest	UR-CA06-MultidimensionalAnalysis	Supported	Must-have	Details
SR-CA05-ShowEventDetails	UR-CA01-TemporalAnalysis	Supported	Must-have	Details
SR-CA05-ShowEventDetails	UR-CA02-GeoSpatialAnalysis	Not-Supported	Should-have	N/A
SR-CA05-ShowEventDetails	UR-CA03-EntityRelationshipsAnalysis	Supported	Must-have	Details
SR-CA05-ShowEventDetails	UR-CA06-MultidimensionalAnalysis	Supported	Must-have	Details
SR-CA14-ShowEntityDetails	UR-CA01-TemporalAnalysis	Supported	Must-have	Details
SR-CA14-ShowEntityDetails	UR-CA02-GeoSpatialAnalysis	Not-Supported	Should-have	N/A
SR-CA14-ShowEntityDetails	UR-CA03-EntityRelationshipsAnalysis	Supported	Must-have	Details
SR-CA14-ShowEntityDetails	UR-CA06-MultidimensionalAnalysis	Supported	Must-have	Details

SR-CA15-TranscribedAudioPlayback	UR-CA01-TemporalAnalysis	Supported	Must-have	Details
SR-CA15-TranscribedAudioPlayback	UR-CA02-GeoSpatialAnalysis	Not-Supported	Should-have	N/A
SR-CA15-TranscribedAudioPlayback	UR-CA03-EntityRelationshipsAnalysis	Supported	Must-have	Details
SR-CA15-TranscribedAudioPlayback	UR-CA06-MultidimensionalAnalysis	Supported	Must-have	Details
SR-CA16-VideoPlayback	UR-CA01-TemporalAnalysis	Supported	Must-have	Details
SR-CA16-VideoPlayback	UR-CA02-GeoSpatialAnalysis	Not-Supported	Should-have	N/A
SR-CA16-VideoPlayback	UR-CA03-EntityRelationshipsAnalysis	Supported	Must-have	Details
SR-CA16-VideoPlayback	UR-CA06-MultidimensionalAnalysis	Supported	Must-have	Details
SR-CA20-FreeTextSearch	UR-CA04-AdvancedFiltering	Not-Supported	Must-have	N/A
SR-CA30-MergeEntities	UR-CA08-OutputRevision	Supported	Must-have	Details
SR-CA30-MergeEntities	UR-CA03-EntityRelationshipsAnalysis	Supported	Must-have	Details
SR-CA31-SplitEntities	UR-CA08-OutputRevision	Supported	Must-have	Details
SR-CA31-SplitEntities	UR-CA03-EntityRelationshipsAnalysis	Supported	Must-have	Details
SR-CA32-AddEntities	UR-CA08-OutputRevision	Supported	Must-have	Details
SR-CA32-AddEntities	UR-CA03-EntityRelationshipsAnalysis	Supported	Must-have	Details
SR-CA39-SupportQualityAssessment	UR-CA10-OutputCharacterisation	Supported	Should-have	Details
SR-CA17-FilterEntitiesByProperty	UR-CA04-AdvancedFiltering	Supported	Must-have	Synced filters
SR-CA18-FilterEventsByTiming	UR-CA04-AdvancedFiltering	Supported	Must-have	Synced filters
SR-CA21-ActivateSyncedFilters	UR-CA04-AdvancedFiltering	Supported	Must-have	Synced filters

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

SR-CA22-DeactivateFilters	UR-CA04-AdvancedFiltering	Supported	Must-have	Synced filters
SR-CA19-FilterEventsByLocation	UR-CA04-AdvancedFiltering	Not-Supported	Must-have	N/A
SR-CA24-ChooseVisibleCharts	UR-CA06-MultidimensionalAnalysis	Supported	Must-have	Custom Workspace
SR-CA25-ResizeVisibleCharts	UR-CA06-MultidimensionalAnalysis	Supported	Must-have	Custom Workspace
SR-CA26-RelocateVisibleCharts	UR-CA06-MultidimensionalAnalysis	Supported	Must-have	Custom Workspace
SR-CA27-SaveDashboardConfiguration	UR-CA06-MultidimensionalAnalysis	Supported	Must-have	Custom Workspace
SR-CA28-LoadDashboardConfiguration	UR-CA06-MultidimensionalAnalysis	Supported	Must-have	Custom Workspace
SR-CA40-CompareDifferentInferences	UR-CA11-OutputComparison	Supported	Could-have	Custom Workspace
SR-CA44-ExportResultsUsingPopularFormats	UR-CA14-ExportedOutputs	Supported	Must-have	Data export
SR-CA06-PresentEventsLocation	UR-CA02-GeoSpatialAnalysis	Not-Supported	Should-have	N/A
SR-CA41-AssociateDifferentCases	UR-CA12-OutputMerging	Supported	Must-have	Cross-case analysis
SR-CA45-GenerateCaseAnalysisReport	UR-CA15-Reporting	Supported	Must-have	Reporting
SR-PM01-NativeInstallation	UR-PM01-MultiOSplatform	Supported	Must-have	Platform management
SR-PM02-VirtualMachine	UR-PM01-MultiOSplatform	Supported	Must-have	Platform management
SR-PM03-UpdatedPlatform	UR-PM02-EasyToMaintainPlatform	Supported	Must-have	Platform management
SR-PM04-UpdatedDatasetsUsedByPlatform	UR-PM02-EasyToMaintainPlatform	Supported	Must-have	Platform management
SR-UM01-NewUserAccountRequest	UR-UM01-UserAccountCreation	Supported	Must-have	User management
SR-UM02-NewUserAccountRegistration	UR-UM01-UserAccountCreation	Supported	Must-have	User management
SR-UM05-UserDetailsRequest	UR-UM02-UserAccountView	Supported	Could-have	User management

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833635. No part of this document may be used, reproduced and/or disclosed in any form or by any means without the prior written permission of the ROXANNE project partners. © 2019 – All rights reserved.

SR-UM06-UserDetailsView	UR-UM02-UserAccountView	Supported	Could-have	User management
SR-UM07-UserDetailsUpdateRequest	UR-UM03-UserAccountEdit	Supported	Could-have	User management
SR-UM08-UserDetailsUpdateConfirmation	UR-UM03-UserAccountEdit	Supported	Could-have	User management
SR-CM11-LimitedAccessGrantingRequest	UR-CM04-CaseRestrictedCollaboration	Supported	Should-have	User management
SR-CM13-FullAccessGrantingRequest	UR-CM05-CaseParallelWorkspaces	Supported	Should-have	User management
SR-UM09-UserDeleteRequest	UR-UM04-UserAccountDelete	Supported	Could-have	User management
SR-UM10-UserDeleteConfirmation	UR-UM04-UserAccountDelete	Supported	Could-have	User management
SR-CM15-LoggedEventsRequest	UR-CM06-CaseIntegrity	Not-Supported	Should-have	N/A
SR-CM17-LoggedEventsListing	UR-CM06-CaseIntegrity	Not-Supported	Should-have	N/A