INSPECTr Project
Intelligence Network & Secure Platform for Evidence Correlation and Transfer

Quarterly Newsletter: Seventh Edition

Edition: November 2022

Intelligence Network & Secure Platform for Evidence Correlation and Transfer

Principal Objectives Brief Summary

To develop a shared intelligence platform and a novel process for gathering, analysing, prioritising, and presenting key data to help in the prediction, detection, and management of crime in support of multiple agencies at local, national, and international level. This data will originate from the outputs of free and commercial digital forensic tools complemented by online resource gathering. The final developed platform will be freely available to all Law Enforcement Agencies (LEAs).

INSPECTr Newsletter Seventh Edition

Welcome to the INSPECTr project newsletter, a guide to our latest work and news. In this, our seventh edition, we will provide updates on our last quarter activities, information about meetings and events attended, our upcoming events, recent dissemination activities and a blog about the continuing platform development and Living Labs Experimentation.

BLOG

Update on INSPECTr Platform Development During Living Labs Experimentation Phases 3.5, 4, and 5.

Living Lab experimentation continues to be an integral part of INSPECTr platform development. There has been a great deal of highly focussed activity during the most recent project quarter where the

INSPECTr development team have been working intensively through several phases of the Living Lab Experimentation schedule to develop, troubleshoot and provide solutions to the ever-evolving INSPECTr platform.

Living Labs Experimentation - Phase 3.5

During Living Lab 3.5 (LL3.5) INSPECTr project LEA participants were invited to practice several tasks remotely using an updated version of the INSPECTr framework on their individual server nodes via a secured VPN connection. Participants were asked to create and link a new user account to perform a series of practical tasks, including the following:

- Populating tables with Tokens, Countries, and Crime Types
- Checking available INSPECTr analysers (gadgets) and related data types
- Searching existing artefacts and running INSPECTr analysers (to run reports)
- Examining the INSPECTr analyser generated reports
- Selecting items in the report related to a mocked use case (Terror, CSAM, or Fraud) and creating new artefacts as appropriate.

As well as testing the platform, the LEAs were tasked with providing live feedback regarding issues that they encountered as they worked through the testing process. This feedback was also provided more formally via feedback surveys. Both feedback formats served to inform the development team of the technical capabilities of the platform and to provide suggested improvements to help the development team best tailor the platform for LEA end-users.

Examples of feedback presented by LEA testers were syncing issues across a subset of components, artifact creation within the Case Management System (CMS), report visualisation and its integration across components, and running of analysers on certain nodes. These issues received attention by the INSPECTr development team as they occurred during the Living Lab and served to provide further direction for subsequent platform software development.

INSPECTr Technical Team Meeting Athens 24th- 28th October 2022
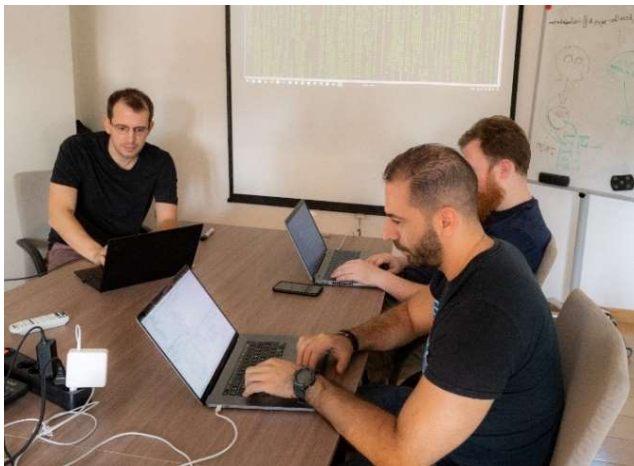
An INSPECTr developer team meeting was held in Athens during October to focus on the ongoing technical development of the platform. The meeting commenced with some big-picture planning with a review undertaken of issues being experienced in the platform and on functionality that was yet to be implemented.

Pressing issues related to forthcoming Living Labs were considered and following this review any issues that could be addressed within the available time were fixed and the platform tested to ensure everything was working as expected. In particular, some of the issues examined and fixed pertained to the running of analysers, creating artefacts using reports of analysers run previously and viewing images and downloading them and account management and population of Elastic Search indices in charge of this.

Issues regarding the Case Management System [CMS] were also discussed following feedback received from LEA testers during Living Lab 3.5, a Living Lab that was held concurrently with this technical

meeting. Discussions were also held regarding widgets and Siren including issues in the sorting and pagination of widgets and switching the language used in widgets' codebase from javascript to typescript. Discussions also included SIREN adding functionality to offer NLP results within its platform. Towards this, the NLP toolbox would be developed to offer a new set of endpoints, so that models' results could be returned.

The security of the platform was discussed from various aspects. In most cases it was decided that due to time restrictions and, in an effort, to push further with the development of other functionality, issues would be examined at a later point in the project.

A discussion was also held regarding the envisioned usage of the AI tools in the future, how the CASE standard could be used to describe their results, and how all this data could be later harnessed from the Knowledge Graph stored in Neo4j.

On the closing day of the meeting the PubSub component of the platform was discussed and specifically:

- The usage of the eCodex framework (Domibus, Domibus Connector, and Domibus Connector Client components) and the need to deploy this on a Windows machine.
- The existence of Linux containers for the Domibus and Domibus Connector components, which were provided to us (ILS) late in the project unfortunately, requiring us to proceed with the Windows machine solution due to time restrictions. However, it is certain that in any future development of this component, a move to a Linux-compatible solution will be achievable and beneficial.
- The development of the Information Request Management Engine (IRME) component which will act as a UI-enabled "middleware" between the CMS and the PubSub functionality, allowing users to request information and view results of previous requests, as well as allow administrators full control over the requests and thus enhanced security of the information transmitted.

The above summary of work illustrates the extensive work being undertaken in the technical development of the platform and how LEA user testing in the Living Labs environment is helping to inform the direction of the evolving platform.

Living Labs Experimentation - Phase 4

For Living Lab 4 (LL4), LEA participants travelled to CCI, University College Dublin from November 7th to November 11th 2022 where they reviewed and tested the latest version of the INSPECTr framework locally in Dublin. Accessing their individual INSPECTr server nodes, LEAs created admin and user accounts to test the functionality and usability of the Case Management System (CMS) as well as the SIREN digital visualisation environment on selected mocked use case artefacts. LL4 tasks included:

- Creating new admin and user accounts in the CMS
- Ingesting mocked data sets, and running INSPECTr analysers on them
- Examining the INSPECTr analyser generated reports for investigative purposes
- Running the SIREN digital visualisation component of the INSPECTr platform

As usual during a Living Lab, as well as testing the platform, LEAs were asked to provide live feedback regarding issues that they encountered as they worked through the testing process. This provides the software development team with priority bug reporting as well as suggested usability improvements for LEA end-users.

Examples of issues presented by LEA testers were issues finding investigative links in the use case mocked data, digital visualisation of evidence in the SIREN component, and the running of analysers on some nodes. These issues received attention by the INSPECTr development team as they were presented during the Living Lab and continue to provide direction for subsequent software upgrades and usability improvements for end-users.

LL4 LEA and Technical Team Participants in University College Dublin 7th – 11th November 2022

Living Labs Experimentation – Phase 5 – Live Demonstration by GN
INSPECTr partner, GN (French Gendarmerie), presented an online live demonstration to the project's LEA partners of some of the AI platform tools. One of those is the Gendarmerie's crime forecasting tool. The tool uses past criminal reports' data like geolocation, date and type of crime to predict future tension zones and displays them using heatmaps. This allows LEAs to distribute resources according to the zones of tension and better manage them in time and space. GN also presented tools such as speech recognition, OCR and image annotator tools.

Further Opportunities for INSPECTr Dissemination and Cross-Project Learning and Collaboration

1)   iLEAD 6th – 8th September 2022. Presentation by INSPECTr Consortium member French Gendarmerie (Judicial Police of the National Gendarmerie (PJGN)

The i-LEAD project is funded by the European Union's Horizon2020 programme; a research fund that, until Horizon Europe, was the EU's flagship initiative. It is the first H2020 project led by a law enforcement agency, the Central Unit of the Netherlands Police. The project brings together a diverse

group of partners from law enforcement, industry, and academia to promote innovation through dialogue and ongoing discussions. The discussions take place primarily through a series of practitioner workshops led by the UK Home Office. The meetings provide a forum for practitioners to discuss gaps, needs, and requirements in five distinct crime groups: Frontline Policing, Cybercrime, Cross-Border crime, Crime and Intelligence, and Forensics. The outcomes of these workshops enable a group of partners to search the market for appealing solutions and innovations produced by industry and research to meet the needs identified. The industry and research event is one of the actions that helps to present these solutions to practitioners and provide a space for providers to receive feedback. Over 100 people attended a fruitful and intense meeting between the 6th and 8th of September, where representatives from 15 countries interacted and participated actively. Practitioners from 44 LEAs were among those in attendance.

INSPECTr is a H2020-EU funded project. As part of the INSPECTr project, the Judicial Police of the National Gendarmerie (PJGN) is working on the development of an AI toolkit that groups various services. During the iLEAD event two solutions were presented being those of Stylometry and NLP annotator.

The automatic stylometry tool could be most useful when the user has a large amount of data which makes it almost impossible to analyse using traditional stylometry. This is particularly true for online data such as web discussion forums (twitter, reddit, darknet marketplaces, etc …) with millions of posts every day. In darknet marketplaces for example, a user with criminal intentions can create multiple accounts to convey one idea or sell a product and other accounts to reinforce the value of his/her opinion or product. This information may help investigators understand the target avatar dynamics and establish different investigative tactics. During an investigation, it could also be useful to know that the suspect doesn't only sell drugs but is also involved in child exploitation. This might open up new leads that wouldn't be considered if the two accounts were not linked by an automatic tool.

We proposed a method and a framework for authorship identification in web forums. The suggested model is a combination of both linguistic features and embeddings to identify specific writing styles. The model is based on cosine similarity and yield an average accuracy of up to 99.4% when using 200 forum comments or posts per author. The framework also shows specific similarities in writing styles whether they are lexical or syntactic.

For the NLP annotator, today, in order to train Named Entity Recognition (NER), Semantic Relation Extraction (SRE) or Text Classification (TC) models, it is necessary to build a labeled dataset with entities, relations and topics of interest. However, manual labelling is a time-consuming task. To facilitate this process, we presented a semi-automatic & multilingual text annotation tool. The platform is based on a Web Interface that sends requests to a REST Server that automatically identifies and tags entities and relations between them from the input text. The human annotator is then asked to correct the model prediction instead of annotating the text from scratch. The key idea is to change the task from a manual annotation to a manual review, making the annotation process much faster and more pleasant.

2)   French Gendarmerie AI Platform: GN winning the first place in Europol Innovation Award!

The artificial intelligence tools platform was created to simplify the work of police officers and provide them access to state-of-the-art AI tools. To avoid any black box effect, it was fully developed in-house by the French Gendarmerie. The platform integrates over 10 tools, each of which have been designed to meet the needs of law enforcement agencies. Many of these tools are already available through

Europol. A typical use case of the platform would be to combine multiple tools for the purpose of tracking offenders online, who may be active in multiple dark web forums. The platform was chosen as the most innovative project by Europol.

3) CERIS Annual Event 2022 - 27th-28th September 2022 – Fighting Crime and Terrorism/Resilient Infrastructure

DG HOME organised a workshop in Brussels to bring together law enforcement practitioners, local authorities, infrastructure operators, policy makers and researchers to discuss the lessons drawn from FCT and INFRA projects, reach recommendations, and determine their most effective implementation strategies in the future research and innovation activities. The aim of this workshop was to facilitate the sharing of experiences and research findings in order to identify current achievements, possible synergies, and strengthen collaboration on a European level. This event was attended by Dr Ray Genoe, Coordinator of the INSPECTr Project.

4) International Conference Innovation in Law Enforcement Prague, 2nd – 4th November 2022

INSPECTr partner the French Gendarmerie, as finalists of the Europol's Innovation Award, were invited to attend the International Conference Innovation in Law Enforcement and present their work. The AI platform was presented, and the presented tools raised a great deal of interest from the conference participants. Europol also manifests high level of interest in incorporating even more of the platform tools into Europol's tool repository.

5) Workshop at the European Public Prosecutor's Office (EPPO) in Luxembourg on 10th November 2022

This workshop aimed to discuss the data management strategy for their Case Information Management System (CIMS). INSPECTR project partner Fabrizio Turchi of CNR attended this workshop.

EPPO is primarily focused on their CIMS to support cross-border investigations. They automatically extract relevant data from a variety of documents (most of them financial), either not structured or poorly structured, and move them in a model exploited by the CIMS to find cross-references between different investigative cases. Therefore, they are interested in better understanding the UCO/CASE ontology and verify if part of the derived model can be integrated in their model. CASE (Cyber-investigation Analysis Standard Expression) was also designed for this purpose, and they should map their data to determine coverage and gaps. Furthermore, they are interested in INSPECTr because it has been developing a real platform for cross-border investigations relying on the UCO/CASE model.

Questions were asked about the backward compatibility of UCO/CASE. In other words, EPPO wanted to know if the ontology is capable to handle "hot" extensions of the data model without interfering with the applications based on the old versions. Fabrizio answered that the current 1.0.0 version can maintain this compatibility because it's highly probable that new versions will extend the previous models so the applications should continue to be valid without interruption even if they won't use the additional features until they won't be updated.

Two presentations followed, the first one being on CASE and the second one on an application being developed within a separate project for the preparation of the E-Package containing the elements of evidence expressed in CASE to facilitate the evidence exchange among Judicial Authorities in the Member States, relying on e-EDES and e-CODEX.

Regarding the INSPECTr project the main goals were illustrated being those of parsing, homogenisation and ingestion of commercial tool reports, providing data analytics (additional storage

requirements), maintaining the chain of evidence (traceability and provenance) the integrity has been maintained by blockchain, enhancing cross-border collaboration to discover linked cases, and to exchange elements of digital evidence. It was also highlighted that the standardisation is a key concept within the project and from the outset, INSPECTr opted for the open-source UCO/CASE ontology to serve as a standard for interchange, interoperability, and analysis of investigative information with a set of parsers for commercial tools having been developed to convert the XML reports generated by the tools into UCO/CASE. At the moment the tools that have received consideration have been: UFED (Cellebrite), AXIOM (Magnet Forensic), XAMN (MSAB), OXYGEN (MH Service). Finally, the data have been ingested in a Storage System based on Elastic Search for subsequent data analysis.
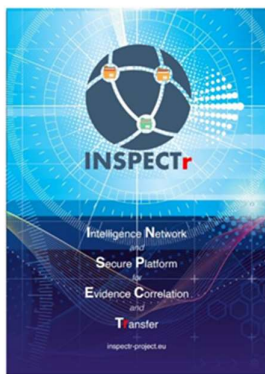
Further Opportunities for INSPECTr Dissemination

INSPECTr Website Project Resources and Project Results Pages

During the last quarter we developed a new area of the INSPECTr website in which we host our Project Resources and Project Results. The following INSPECTr Resources have been developed aiming to disseminate the INSPECTr brand and inform our audiences of the main project goals and these, together with other resources, are available on our Project Resources page. A project video is a work in progress and will be added as a Resource on completion.

Project Brochure	**Project Banner**	**Project Flyer**
For distribution to more technical audiences	For events and conference dissemination	Foldable flyer with essential project information

INSPECTr Project Public Deliverables

We are also hosting our public deliverables that have successfully completed the European Commission approval process. The following deliverables are available for viewing and download from our Project Results page. More public deliverables will be added as they come on stream.

➢ D2.1 Initial Legislative compliance relating to law-enforcement powers and evidence requirements
➢ D2.3 Reference Digital Forensics Domain Model
➢ D2.4 e-Codex infrastructure evaluation in the context of deployment in LLs
➢ D6.6 Dissemination – Period 1
➢ D8.1 INSPECTr Research ethics and data protection
➢ D8.2 First Report on Ethical Governance
➢ D8.5 Ethical, Legal and Social requirements for the INSPECTr platform and tools

Project Activities and Events between August 2022 – November 2022

- **INSPECTr Monthly Project Meetings**
- **INSPECTr Weekly Technical Meetings**
- **INSPECTr LSG Monthly Meetings**
- **Ethics Work Package Monthly Meetings**
- **EARG Ethics Advisory and Review Group**
- **INSPECTr Technical Team Meeting Athens**
- **Living Labs Experimentation Phase 3.5**
- **Living Labs Experimentation Phase 4**
- **Living Labs Experimentation Phase 5**

INSPECTr Consortium Attendance at Conferences and Workshops

iLEAD Industry and Research Days	The Hague	6th-8th September 2022
Gendarmerie AI Platform: GN winning the first place in Europol's Innovation Award for the most innovative project	The Hague	Excellence in Innovation Awards 2022
CERIS Annual Event 2022 – Fighting Crime and Terrorism/Resilient Infrastructure	Brussels	27th-28th September 2022
Innovations in Law Enforcement Conference	Prague	2nd-4th November, 2022
European Public Prosecutor's Office EPPO	Luxembourg	10th November 2022

Future Conferences, Workshops, and Events

Diffusion of Policing in the Algorithmic Society	Oslo	18th November 2022
INSPECTr Webinar Series in conjunction with CEPOL - A series of lunchtime webinars presenting the INSPECTr project.	Online Event in conjunction with CEPOL	February 2023

Closing

We look forward to updating you further in February 2023 with our final INSPECTr Newsletter. In the interim, communications from our readers are welcome and if you wish to contact us or subscribe to our Newsletter you can e-mail us directly at inspectr@ucd.ie. Further information and updates can also be found on our project website https://inspectr-project.eu/.