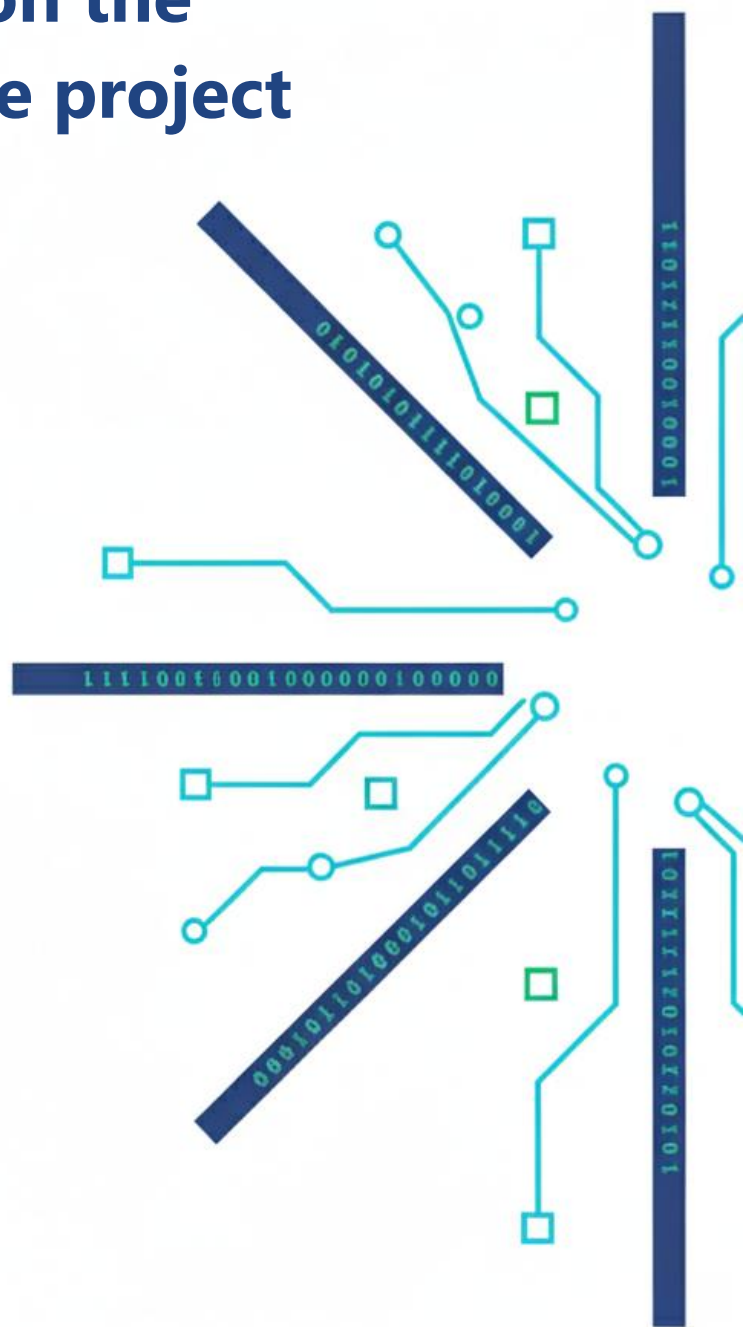


D06 – Final report on the achievements of the project



The project Automatization of digital forensics and incident response (ADFIR) funded by the European Union – Next GenerationEU through the Recovery and Resilience Plan of the Slovak Republic under project No. č. 09I05-03-V02-00079.

Contents

Contents	1
Introduction	4
1 PhD students whose research focuses on issues related to the ADFIR project.....	6
1.1 Definition of the indicator.....	6
1.2 Analysis of digital traces using machine learning methods (Eva Marková)	6
1.3 Electronic Evidence (Miroslav Srholec)	6
2 Joint projects supported by other schemes	8
2.1 Definition of the indicator.....	8
2.2 Overview of projects	8
2.3 UPJŠ Cyber Security Competence Centre (17R05-04-V01-00007)	8
2.4 Legal and Technical Aspects of Cyber Security Situational Awareness (APVV-23-0137)	10
2.5 Legal, technical and forensic aspects of cyber security with an element of artificial intelligence (APVV-25-0162)	10
2.6 AI-driven cyber incident response technology (IstroSec s. r. o.).....	12
2.7 Research and development of advanced artificial intelligence solutions for the detection of cyber threats and defence against sophisticated attacks (401101C360).....	12
3 Research publications disseminating the project's results	14
3.1 Definition of the indicator.....	14
3.2 List of publications	14
3.2.1 Article 1.....	14
3.2.2 Article 2.....	14
3.2.3 Article 3.....	15
3.2.4 Article 4.....	15
3.2.5 Article 5.....	16
4 National and international cooperation in the academic sphere with both the public and private sectors	17
4.1 Definition of the indicator.....	17
4.2 National collaborations	17
4.3 International collaborations.....	18
5 Activities aimed at disseminating results to potential partners and customers.....	19

5.1	Definition of the indicator	19
5.2	ADFIR Project Closing Conference	19
5.3	Digital Forensic Analysis – AI Security (Košice)	19
5.4	Digital Forensic Analysis – AI Security (Rožňava)	20
5.5	Workshop – Digital forensics in criminal proceedings: practice, pitfalls and evidence	20
5.6	Workshop – Data Analysis: From Theory to Practice: Analysis of Digital Traces	21
5.7	Securing digital traces and their usability as electronic evidence in criminal proceedings	21
5.8	Poster presentation at the DFRWS EU 2026 conference	22
5.9	AI Community Meetup	23
6	Published datasets and databases	24
6.1	Definition of the indicator	24
6.2	Overview of outputs	24
6.3	D14 — Synthetic dataset	24
6.4	D22 — Behavioural Patterns Database	25
7	Published methodologies	27
7.1	Definition of the indicator	27
7.2	Overview of outputs	27
7.3	D01 — Methodology for contractual relationships	27
7.4	D03 — Methodology for the processing of digital traces	28
7.5	D23 — Methodology for creating a case timeline	29
8	Published methods and models	31
8.1	Definition of the indicator	31
8.2	Overview of outputs	31
8.3	D02 — Method for anonymising digital traces	32
8.4	D12 — Method for the automated collection of digital traces	33
8.5	D13 — Method for creating a dataset	34
8.6	D15 — A model for extracting digital traces into a matrix representation	34
8.7	D16 — Model for extracting digital traces into a graph representation	35
8.8	D17 — Model for aggregating and linking digital traces	36
8.9	D18 — Model for enriching digital traces	37
8.10	D19.1 — Method for identifying digital traces: clustering and large language models	38
8.11	D19.2 — Method for identifying digital traces: anomaly detection	39

8.12	<i>D20.1 — Method for identifying relationships: factorisation of formal context</i>	40
8.13	<i>D20.2 — Method for identifying relationships: significant time windows</i>	41
8.14	<i>D21 — Method for determining the timeframe and phases of an attack</i>	42
9	Other published project outputs	43
9.1	<i>D24 — Extended report on forensic analysis</i>	43

Introduction

The project **Automatization of digital forensics and incident response** (hereinafter referred to as “**ADFIR**”) is funded by the **European Union – Next GenerationEU through the Recovery and Resilience Plan of the Slovak Republic** under project No. č. 09I05-03-V02-00079. This project addresses one of the key challenges in cybersecurity and information security – how to process the massive volume of digital evidence generated during cybersecurity incidents or forensic investigations. Currently, this process is highly demanding in terms of human resources and time. Therefore, automation using machine learning methods can significantly **improve the quality of digital forensic analysis** and reduce the time required to perform it. Overall, this enables security teams to respond more effectively to cyber threats. Main benefits of this project are:

- **Accelerated Resolution of Cybersecurity Incidents.** The project ADFIR introduces automated approaches to collecting, processing, and analyzing digital traces. As a result, security teams can identify the causes of incidents more quickly and adopt effective measures to address them.
- **Reduced Workload for Forensic Analysts.** Routine and time-consuming tasks involved in processing digital traces will be replaced by automated methods. This will allow analysts to focus on more complex cases and strategic decision-making.
- **Higher Quality and Consistency of Outputs.** The use of unified methodologies and tools ensures that the processed digital traces will be more accurate, consistent, and easily verifiable. This significantly reduces the risk of errors caused by human factors.
- **Potential Use in Criminal Proceedings.** The project outputs will be developed in compliance with legal requirements and standards, allowing the digital traces to be accepted as relevant evidence for investigations and court proceedings.

This document (Output D06 – Final report on the achievements of the project) details the fulfilment of the ADFIR project’s measurable indicators. For each indicator, it sets out its definition, the method used to set the target value, the requirements for demonstrating fulfilment, and the specific results achieved in the project. It serves as a basis for monitoring reports, as well as an overview for project researchers.

A summary of performance is set out in Table 1. Details for each indicator are provided in the individual chapters.

Expected results and impact of project activities	Expected number	Actual performance
PhD students whose research focuses on issues related to this project	2	2
Joint projects supported by other schemes	2	5

Expected results and impact of project activities	Expected number	Actual performance
Research publications disseminating the project's results	4	5
National collaborations within the academic sphere and with both the public and private sectors	2	3
International collaboration in the academic sphere and with the public and private sectors	2	3
Activities aimed at disseminating results to potential partners/clients	3	7
Published datasets and databases	2	2
Published methodologies	3	3
Published methods and models	10	12

Table 1 – Expected results and impact of project activities

Chapters 6 to 8 of this report contain a detailed description of the individual published datasets and databases, methodologies, methods and models, including a summary of their content, the technologies used, the results achieved and their links to specific project objectives. For each output, the title of the document in which the output is described in detail is also provided.

1 PhD students whose research focuses on issues related to the ADFIR project

Expected number	Actual performance	Status
2	2	Achieved

1.1 Definition of the indicator

This indicator was estimated on the basis of current and planned PhD positions, the topics of which reflect the issues defined in the ADFIR project. The indicator tracks the development of research capacity in the field of automated digital forensic analysis.

1.2 Analysis of digital traces using machine learning methods (Eva Marková)

Data	Value
Thesis topic	Analysis of digital traces using machine learning methods
Student	RNDr. Eva Marková
Department	Institute of Informatics, Faculty of Natural Sciences, UPJŠ in Košice
Supervisor	doc. JUDr. RNDr. Pavol Sokol, PhD. et PhD.
Supporting documentation	Eva_Marková_thesis_assignment_SK_EN

Link to the project: the topic directly addresses the core content of work packages KPB3 and KPB4 – the representation of digital traces and their analysis using machine learning methods. The PhD student is a co-author of a project publication presented at the ARES 2025 conference.

1.3 Electronic Evidence (Miroslav Srholec)

Data	Value
Thesis topic	Electronic evidence
Student	JUDr. Miroslav Srholec, PhD.
Affiliation	Pavol Jozef Šafárik University in Košice

Data	Value
Supervisor	prof. JUDr. Sergej Romža, PhD.
Supporting documentation	Miroslav_Srholec_thesis_assignment_SK_EN https://www.upjs.sk/app/uploads/sites/11/2025/07/oznam_Srholec.pdf

Link to the project: the topic covers the legal aspect of the project addressed in work package KPB1, namely the admissibility and evidential value of the outputs of automated forensic analysis in criminal proceedings.

2 Joint projects supported by other schemes

Expected number	Actual performance	Status
2	5	Exceeded

2.1 Definition of the indicator

The estimate was based on ongoing and planned project proposals relating to the issues addressed in the ADFIR project.

2.2 Overview of projects

No.	Project	Principal Investigator	Status
1	Cyber Security Competence Centre at Pavol Jozef Šafárik University in Košice (17R05-04-V01-00007)	UPJŠ	Completed
2	Legal and technical aspects of situational awareness in cyber security (APVV-23-0137)	UPJŠ	Completed
3	Legal, technical and forensic aspects of cyber security incorporating artificial intelligence (APVV-25-0162)	UPJŠ	Submitted
4	AI-driven cyber incident response technology (09I05-03-V03-00005)	IstroSec s. r. o.	Completed
5	Research and development of advanced artificial intelligence solutions for the detection of cyber threats and defence against sophisticated attacks (401101C360)	IstroSec s. r. o.	Completed

2.3 UPJŠ Cyber Security Competence Centre (17R05-04-V01-00007)

Data	Value
Project code	17R05-04-V01-00007
Scheme	Slovak Republic Recovery and Resilience Plan, Component 17 – Digital Slovakia

Data	Value
Beneficiary	Pavol Jozef Šafárik University in Košice
Project duration	January 2025 – March 2026
Total eligible expenditure	EUR 1,585,417.19 including VAT
Location	Pavol Jozef Šafárik University in Košice and its specialist departments

Objective and summary: The aim of the project is to establish a Cyber Security Competence Centre at Pavol Jozef Šafárik University in Košice (KC KB UPJŠ). The centre carries out both compulsory and optional activities focused on education, research and expert work in the fields of information and cyber security, data protection, cybercrime and protection against disinformation. The KC KB UPJŠ also engages in international cooperation with academic partners from abroad and provides consultancy services to assist with the preparation and submission of projects in the field of cyber security. The establishment of the KC KB UPJŠ reflects a number of issues currently identified in the field of information and cyber security — the need to raise security awareness amongst relevant stakeholders, particularly public sector employees and students at universities and secondary schools, to train and educate new specialists, to investigate cyber threats and identify appropriate responses to them, and to enhance operational security in the public sector through the provision of expert services by the CSIRT team. The project's main outputs include a curriculum for the Master's degree programme in Applied Informatics with a specialisation in cyber security and related new or modified modules; training programmes for various target groups of public administration staff; enhanced cooperation with secondary schools through the activities of CyberTeams; recommendations for the public sector in various areas of cyber and information security, campaigns aimed at raising security awareness, scientific publications and new research collaborations, as well as the upgrading of equipment and training for the university's CSIRT team, with the possibility of providing expert support to accredited CSIRT teams in the Slovak Republic.

Link to the ADFIR project: The Centre of Excellence provides the institutional and human resources infrastructure for research in the fields of cyber and information security and digital forensic analysis at UPJŠ, and overlaps with the research team involved in the ADFIR project. The university's CSIRT team and the Competence Centre's educational activities also provide a natural channel for transferring the results of the ADFIR project into incident response practice and teaching.

2.4 Legal and Technical Aspects of Cyber Security Situational Awareness (APVV-23-0137)

Data	Value
Project number	APVV-23-0137
Scheme	Agency for the Support of Research and Development — general call
Lead organisation	Pavol Jozef Šafárik University in Košice
Link	https://portal.ccvapp.upjs.sk/project-view/2810
Status	In progress

Objective and summary: Situational awareness in cyber security, which encompasses the collection and storage of security data, its analysis and sharing, as well as its prediction, presents numerous legal and technical challenges that need to be examined in synergy with one another. The research team, comprising experts from selected fields of law and technical disciplines, aims to analyse the current legal framework relating to situational awareness with regard to the collection, storage and sharing of security data, as well as an analysis of the legal framework regarding the use of specific machine learning methods for the analysis of security data and the prediction of situational awareness. The project's unique outputs will be recommendations for obligated entities under cybersecurity legislation, a technical approach to the collection of security data taking legal aspects into account, and the design and implementation of machine learning methods for identifying relevant security data and for predicting situational awareness.

Link to the ADFIR project: The project shares a methodological core with the ADFIR project – the integration of legal analysis with data analysis and machine learning applied to security data. Whilst ADFIR focuses on digital traces and post-incident forensic analysis, this project addresses situational awareness before and during an incident. There is also an overlap in the research teams of both projects.

2.5 Legal, technical and forensic aspects of cyber security with an element of artificial intelligence (APVV-25-0162)

Data	Value
Project number	APVV-25-0162
Scheme	Agency for the Support of Research and Development

Data	Value
Principal Investigator	prof. JUDr. Jozef Čentéš, DrSc.
Research Department	050509 – Criminal Law
Implementation	2026–2030
Status	Approved project proposal

Abstract and objective: The development of artificial intelligence is fundamentally changing the nature of cyber security risks and incidents. Artificial intelligence systems are becoming an integral part of critical processes in public administration, the economy and critical infrastructure, giving rise to new types of cyber security incidents. These incidents are characterised by a high degree of automation, adaptability and the ability to generate specific digital artefacts that differ fundamentally from traditional electronic evidence. However, current legal, methodological and forensic frameworks only partially reflect these specific characteristics. The aim of the project is to create a practical legal-technical and criminalistic-forensic framework for the identification, analysis and resolution of cyber security incidents involving artificial intelligence. The project focuses on the systematic identification and classification of incidents or malfunctions in artificial intelligence systems, including cyber security incidents involving artificial intelligence ('AI incidents'), as well as on proposing a methodology for working with digital traces generated by or influenced by the operation of artificial intelligence systems ('AI traces'). The project is conceived as applied interdisciplinary research that bridges criminal law, cyber security and digital forensics. It utilises doctrinal and comparative legal analysis, criminalistic and forensic methods, and empirical research carried out in collaboration with the Cyber Security Competence and Certification Centre and CSIRT units. The expected outputs will take the form of methodological, analytical and advisory documents directly applicable in practice, particularly in providing methodological support to public administration bodies, resolving cyber incidents and cooperating with law enforcement agencies. The project will contribute to increasing the state's cyber resilience, strengthening legal certainty and improving the integration of legal and technical tools in addressing threats associated with artificial intelligence.

Link to the ADFIR project: The project proposal builds directly on the results of the ADFIR project and extends them to include aspects of artificial intelligence in forensic analysis. Whilst ADFIR developed methods for the automated analysis of digital traces and a legal framework for their use in criminal proceedings, the follow-up project focuses on a new category of traces generated by the operation of artificial intelligence systems and on incidents in which these systems are involved.

2.6 AI-driven cyber incident response technology (IstroSec s. r. o.)

Data	Value
Project number	09I05-03-V03-00005
Beneficiary	IstroSec s. r. o.
Scheme	Recovery and Resilience Plan of the Slovak Republic, Investment 5
Implementation period	November 2024 – March 2026
Eligible expenditure	EUR 1,181,314.51, of which EUR 792,613.07 comes from the mechanism
Location	Vajnorská 100/B, 831 04 Bratislava 3
Reference	https://istrosec.com/sk/eu-projects/ai-driven-cyber-incident-response-technology/

Objective: The aim of the project is to research and develop advanced artificial intelligence-driven technology for the detection and resolution of cyber incidents, with a view to automating processes and enabling self-learning in artificial intelligence models. The project focuses on the research and development of an advanced tool for threat analysis and predictive analytics, and represents a significant step forward in the field of cyber security innovation.

Link to the ADFIR project: This project addresses the automation of responses to cyber incidents using artificial intelligence methods, which corresponds in substance to work packages KPB2 and KPB4 of the ADFIR project, namely the collection and analysis of digital traces. The partner IstroSec can thus directly apply the results of the ADFIR project to the development of a commercial incident response tool.

2.7 Research and development of advanced artificial intelligence solutions for the detection of cyber threats and defence against sophisticated attacks (401101C360)

Data	Value
Project code	401101C360

Data	Value
Call	SVaV – Strategic Research and Development through Partnerships (PSK-MSVVM-020-2024-DV-EFRR)
Scheme	Slovakia Programme
Beneficiary	University of Žilina in Žilina
Partners	IstroSec s. r. o., Syteli s. r. o., Technical University of Košice
Implementation period	April 2026 – October 2029
Eligible expenditure for IstroSec	EUR 3,470,896.45, of which a non-repayable grant of EUR 2,776,717.16
Location	Žilina
Project website	https://www.cybersecurityresearch.sk
Link	https://istrosec.com/sk/eu-projects/vyskum-a-vyvoj-ai-na-detekciu-kybernetickych-hrozieb-a-obranu-proti-utokom/

Objective: The aim of the project is to research advanced detection methods that integrate artificial intelligence and behavioural models, thereby significantly advancing the current state of the art. The project focuses on mapping current threats, researching methods for analysing attacker behaviour and testing security measures under realistic conditions.

Link to the ADFIR project: Research into methods for analysing attacker behaviour directly builds on work package KPB4 of the ADFIR project, particularly the identification of behavioural patterns and the behavioural pattern database (deliverable D22). The project also extends the collaboration of partner IstroSec to include other Slovak research institutions.

3 Research publications disseminating the project's results

Expected number	Actual performance	Status
4	5	Exceeded

3.1 Definition of the indicator

The estimate was based on the number of previous scientific articles published by key members of the research team. The indicator includes peer-reviewed publications that disseminate the project's results within the professional community. The publications constitute output D25 – Peer-reviewed scientific articles.

3.2 List of publications

3.2.1 Article 1

Hennelová, Z., Marková, E., & Sokol, P. (2025). **The Impact of Anti-forensic Techniques on Data-Driven Digital Forensics: Anomaly Detection Case Study**. In International Conference on Availability, Reliability and Security (ARES 2025), Lecture Notes in Computer Science (pp. 131–148). Cham: Springer Nature Switzerland.

Link to the article: https://link.springer.com/chapter/10.1007/978-3-032-00635-6_8

Abstract: Digital forensic analysis is of paramount importance in responding to cyber threats and incidents, as it enables investigators to identify, collect and analyse digital evidence. As attacks become increasingly sophisticated, attackers are using anti-forensic techniques to evade detection and hinder forensic investigations. These techniques seek to remove, obscure or manipulate digital traces, thereby reducing the effectiveness of forensic analysis. Despite growing interest in anti-forensic techniques, only limited effort has so far been devoted to systematically assessing their actual impact. This paper examines the effects of selected anti-forensic techniques — the deletion of files and records, the concealment of data in alternative data streams, encryption, the creation of false evidence and the manipulation of timestamps — on two key aspects: the integrity and availability of digital traces, and the accuracy of algorithmic anomaly detection in data-driven forensic analysis.

3.2.2 Article 2

Hennelová, Z., Sokol, P., Bajtoš, T., & Krišáková, S. P. (2025). **Data-Driven Digital Forensics: Anomaly Detection in Mozilla Firefox**. In ITAT 2025: Information Technologies — Applications and Theory (pp. 404–413). CEUR Workshop Proceedings, Vol. 4092.

Link to the article: <https://ceur-ws.org/Vol-4092/paper33.pdf>

Abstract: The use of web browsers generates a large amount of data that can serve as valuable forensic artefacts in the investigation of cyber incidents. These artefacts contain information about visited pages, downloaded files and timestamps of user activity, which help to identify the attack vector. Existing tools allow this data to be visualised; however, its analysis is time-consuming and heavily dependent on the analyst's experience. This paper utilises a data-driven digital forensic analysis methodology to automate the identification of suspicious events using machine learning algorithms. Data from the Mozilla Firefox browser is processed into a unified timeline, against which anomalies are detected. The results show that the ECOD and COPOD methods represent a suitable compromise between accuracy and the volume of output data.

3.2.3 Article 3

Antoni, Ľ., Sokol, P., Krišáková, S. P., Kotlárová, D., Krídlo, O., & Krajčí, S. (2025). **Formal Concept Analysis and Attribute Dependencies of a NIST Data Leakage Case**. In ITAT 2025: Information Technologies — Applications and Theory (pp. 358–368). CEUR Workshop Proceedings, Vol. 4092.

Link to the article: <https://ceur-ws.org/Vol-4092/paper29.pdf>

Abstract: Security incidents involving unauthorised access and data leaks remain one of the most critical challenges in contemporary cyber security and require transparent and interpretable methods of behaviour analysis. This paper presents the application of formal conceptual analysis and association rule mining to the NIST Data Leakage Case EVT dataset, which simulates a security incident involving potential data exfiltration. The aim is to uncover and interpret logical dependencies between binary audit attributes and time series extracted from system logs. The authors construct a conceptual lattice comprising 26 formal concepts over a dataset containing 10,306 rows and 11 attributes, derive a canonical base of attribute implications with 100 per cent reliability, and extend it with approximate association rules. The results highlight typical access patterns as well as anomalous combinations involving weekend and night-time activity.

3.2.4 Article 4

Bača, M., Krišáková, S. P., & Sokol, P. (2026). **Identification of Relevant Digital Forensic Traces Using Time Window Aggregation and Hierarchical Clustering**. In International Conference on Computational Intelligence in Security for Information Systems (CISIS 2026), Communications in Computer and Information Science (pp. 224–236). Cham: Springer Nature Switzerland.

Abstract: Digital forensic investigations are increasingly faced with the challenge of analysing large volumes of heterogeneous data, whilst the identification of relevant digital traces remains a complex and resource-intensive task. This paper proposes a methodology for the systematic identification of relevant forensic artefacts based on aggregation into time windows and hierarchical clustering. The proposed approach introduces quantitative metrics, such as attack rate and attack lift, to assess the relevance of clusters, whilst emphasising

interpretability through the identification of key attributes that characterise individual clusters. The methodology was validated on several publicly available datasets; the results show that the approach effectively supports the identification of incident-related behaviour whilst maintaining interpretability.

3.2.5 Article 5

Kordík, M. (2026). **Outline of Principles for the Analysis of Open-Source Intelligence (OSINT) as Evidence in Criminal Proceedings**. *Acta Facultatis Iuridicae Universitatis Comenianae*, 45(2), pp. 3–16.

Link to the article: <https://afi.flaw.uniba.sk/index.php/AFI/en/article/view/1478>

Abstract: This article examines the possibility of using open-source intelligence analysis as evidence in criminal proceedings and seeks to define minimum standards for its use so that the requirements of the Code of Criminal Procedure are met, both in terms of positive principles and those defined negatively. The paper addresses the terminology used in this field and analyses the forthcoming amendment to the rules on the seizure of electronic evidence, which, however, does not bring about any relevant change from this perspective. In its argumentation, it draws on the case law of the International Criminal Court and provides an interpretation of the provisions of the Budapest Convention regarding the possibility of sharing OSINT analysis as one of the fundamental sources for the sharing of digital evidence in Europe. The author distinguishes between negative principles — the prohibition on overcoming resistance and the prohibition on infiltrating closed groups — and positive principles of necessity, identification, objectivity and the absence of conflicts of interest, the supportive function and the security of OSINT procedures.

4 National and international cooperation in the academic sphere with both the public and private sectors

Type of cooperation	Expected number	Actual performance	Status
National collaborations	2	3	Exceeded
International collaborations	2	3	Exceeded

4.1 Definition of the indicator

The estimate was based on existing collaborations and the potential for future collaborations with entities that have a natural interest in the project's results. For existing partnerships, contracts are listed with a reference to the Central Register of Contracts, together with a justification of the connection. In the case of potential new collaborations, Letters of Interest are listed.

4.2 National collaborations

Organisation	Form	Link
DISIG, a. s.	Cooperation Agreement	https://www.crz.gov.sk/zmluva/7377499/
ESET, spol. s r. o.	Cooperation Agreement	https://www.crz.gov.sk/zmluva/8052347/
Binary Confidence s. r. o.	Letter of intent to cooperate	Letter_of_intent_Binary_confidence

DISIG, a. s. is a Slovak company operating in the field of trust services and information security. The aim of the collaboration is the research and application of machine learning in the identification of relevant records within SIEM systems. The collaboration directly follows on from work package KPB4, namely the identification of relevant digital traces using machine learning methods.

ESET, spol. s r. o. is a global company operating in the field of cyber security, headquartered in Slovakia. The aim of the collaboration is to utilise the scientific expertise of the Faculty of Natural Sciences at UPJŠ in research into the integration of data analysis and machine learning into the activities of SOC teams. The collaboration links the project's results to the operations of security monitoring centres.

Binary Confidence s. r. o. is a Slovak company providing cyber security incident monitoring and response services. As the operator of a security monitoring centre, it is a natural user of the project's results in the field of automated forensic analysis.

4.3 International collaborations

Organisation	Country	Form
Axenta a. s.	Czech Republic	Letter_of_intent_Axenta_CZ
The Science, Entrepreneurship and Technologies University (SET)	Ukraine	Letter_of_Intent_SET
University of Málaga	Spain	Letter_of_intent_University_Malaga

Axenta a. s. is a Czech company specialising in cyber security and the operation of security monitoring centres. This collaboration enables the verification of the transferability of the project's results to the Czech market and to the practical work of commercial incident response teams.

The Science, Entrepreneurship and Technologies University (SET) is a Ukrainian higher education institution. This collaboration brings to the project the perspective of an environment characterised by a high intensity of cyber-attacks and broadens the international reach of the project's results.

The University of Málaga is a Spanish public university with a research background in the field of cyber security and formal methods of data analysis. This collaboration complements, in particular, the methods of formal conceptual analysis applied in work package KPB4.

5 Activities aimed at disseminating results to potential partners and customers

Expected number	Actual performance	Status
3	7	Exceeded

5.1 Definition of the indicator

The estimate was based on the planned activities for disseminating results within the project, namely conferences and workshops. These activities fulfil the outputs of work package KPB5, namely D26 — Workshop for students and D27 — Workshop for potential stakeholders. A total of seven activities were carried out, compared to the three planned, with the target groups comprising secondary school pupils, university students of computer science and law, lawyers, and experts from academia, public administration and security companies.

5.2 ADFIR Project Closing Conference

Data	Value
Date	24 June 2026
Venue	Amphitheatre, 2nd floor of the old building, Faculty of Law, Comenius University in Bratislava
Partners	UPJŠ, EISI, IstroSec
Number of participants	31
Type of participants	professionals

Link: <https://adfir.upjs.sk/2026/06/17/zaverecna-konferencia-projektu-adfir/>

Abstract: The closing conference was dedicated to the presentation of the results of the ADFIR project

and to an expert discussion on the automation of digital forensic analysis, its legal limitations and potential applications beyond criminal proceedings. All three project partners were involved in the organisation, and the event's focus brought together the technical, research and legal aspects of the project. The conference was organised beyond the scope of deliverable D27 and is reported as a separate dissemination activity.

5.3 Digital Forensic Analysis – AI Security (Košice)

Data	Value
Date	3 November 2025
Venue	Faculty of Natural Sciences, UPJŠ in Košice, Park Angelinum 9, Košice
Supervisor / Lecturer	Mgr. Sophia Petra Krišáková
Number of participants	17
Type of participants	secondary school pupils

Abstract: An educational activity for secondary school pupils introducing the basics of digital forensic analysis and the security aspects of artificial intelligence. The activity fulfils the project's objective of raising awareness of the topic amongst future professionals.

5.4 Digital Forensic Analysis – AI Security (Rožňava)

Data	Value
Date	23 February 2026
Venue	Secondary Technical School, Hviezdoslavova 5, Rožňava
Tutor / Lecturer	Mgr. Sophia Petra Krišáková
Number of participants	37
Type of participants	secondary school pupils

Abstract: A repeat of the educational activity at a technical secondary school with a broader range of participants. The activity extends the project's geographical reach beyond the regional capital.

5.5 Workshop – Digital forensics in criminal proceedings: practice, pitfalls and evidence

Data	Value
Date	4 May 2026

Data	Value
Venue	Faculty of Natural Sciences, UPJŠ in Košice, Jesenná 5, Košice
Supervisor / Lecturer	Ing. Jaroslav Oster
Number of participants	20
Type of participants	university students (computer science and law)

Abstract: An interdisciplinary workshop bringing together students of computer science and law, focusing on the practical challenges of securing and interpreting digital traces and their admissibility as evidence in criminal proceedings.

5.6 Workshop – Data Analysis: From Theory to Practice: Analysis of Digital Traces

Data	Value
Date	13 November 2025
Venue	Faculty of Natural Sciences, UPJŠ in Košice, Jesenná 5, Košice
Supervisor / Lecturer	Mgr. Sophia Petra Krišáková
Number of participants	approximately 10
Type of participants	university students (computer science)

Abstract: A practical workshop for computer science students focusing on the process of moving from the theoretical foundations of data analysis to its application to real-world digital traces.

5.7 Securing digital traces and their usability as electronic evidence in criminal proceedings

Data	Value
Date	24 June 2026, 10:00 – 12:00
Venue	Slovak Bar Association building, Kolárska 6148/4, Bratislava

Data	Value
Tutor / Lecturer	Assoc. Prof. JUDr. Marek Kordík, LL.M., PhD., JUDr. Petra Dražová, Assoc. Prof. RNDr. JUDr. Pavol Sokol, PhD. and PhD.
Number of participants	20
Type of participants	lawyers

Abstract: A professional event for lawyers organised in collaboration with the Slovak Bar Association. It focused on the preservation of digital traces and the conditions for their admissibility as electronic evidence. The event directly fulfils the project's objective of promoting the acceptance of a new methodological approach to the analysis of digital evidence by the legal profession and, as one of three key activities for potential users of the results, fulfils the project's reference indicator.

5.8 Poster presentation at the DFRWS EU 2026 conference

Data	Value
Date	23–27 March 2026
Venue	Linköping University (LiU)
Supervisor / Lecturer	Sophia Petra Krišáková, Pavol Sokol, Martin Bača, Radovan Slíž, Matej Mikuláš
Number of participants	approximately 200 conference participants
Type of participants	experts from academia and public administration, representatives of security companies

Poster title: Lifecycle of Automated Digital Forensic Analysis Using Machine Learning.

Abstract: The poster presented a comprehensive lifecycle of automated digital forensic analysis using machine learning, as covered by the ADFIR project across work packages KPB2 to KPB5. The event provided an opportunity to present the project's results to the international digital forensic analysis community.

5.9 AI Community Meetup

Data	Value
Date	18 June 2026, 8.30 – 10.30
Venue	TECHNICOM University Science Park, Němcovej 5, Košice
Tutor / Lecturer	Košice IT Valley in collaboration with UPJŠ
Number of participants	26
Type of participants	IT professionals, security experts, researchers, students

Link: <https://www.kosiceitvalley.sk/events/ai-community-meetup/>

Abstract: The second meeting of the artificial intelligence community in Košice, dedicated to the intersection of artificial intelligence, machine learning and cyber security. The programme focused on how artificial intelligence assists in incident detection, the automation of security processes and digital forensic analysis. The event featured a presentation of the ADFIR project and its research and innovation outcomes in the field of automated response to cyber incidents. The event was aimed at IT professionals, security experts, researchers and students, and its objective was also to foster connections between the academic and corporate sectors. Attendance was free of charge, subject to compulsory registration due to limited capacity.

6 Published datasets and databases

Expected number	Actual performance	Status
2	2	Achieved

6.1 Definition of the indicator

The estimate was based on the number of datasets and databases that need to be prepared to achieve the project's objectives. The indicator covers outputs of the 'dataset' and 'database' types intended for publication. These are two outputs covering two distinct layers of the project's data chain: input forensic data for training and comparing methods (D14) and a knowledge base of attacker behaviour patterns obtained through analysis (D22).

6.2 Overview of outputs

Number	Output	Work package	Responsible partner	Access
D14	Synthetic dataset	KPB2	UPJŠ	Public
D22	Database of behavioural patterns	KPB4	IstroSec	Public

6.3 D14 — Synthetic dataset

Data	Value
Output number	D14
Work package / task	KPB2, task KPB2.3
Lead partner	Pavol Jozef Šafárik University in Košice
Type	dataset
Access and dissemination	public — publication in open repositories
Document with a detailed description	ADFIR-D14-KPB2-03-Synthetic-dataset (EN)

Summary: The output is a synthetic forensic dataset created for the purpose of comparing methods of digital trace analysis. The research team based their work on the observation that,

in the field of digital forensic analysis, there is no dataset suitable for all research purposes, and that existing datasets are either not shared or lack documentation and a formal description of their structure. The dataset was therefore created primarily from datasets used in Capture The Flag competitions and has been appropriately supplemented with realistic data to depict scenarios that may occur in real-world security incidents. The output includes full documentation of the dataset's construction process, a description of the artefacts included, and a definition of its limitations — the research team explicitly states that such a dataset may not cover all situations that may arise in a real-world environment. The dataset covers Windows operating system artefacts, including the NTFS file system, and is organised in a structure suitable for subsequent processing in the KPB3 working package.

Link to project objectives: The output fulfils Specific Objective 1 of the project — the identification of potential sources of forensic artefacts, their collection and the identification of a suitable representation. The dataset forms the input data basis for representation models in the KPB3 work package and for analytical methods in the KPB4 work package.

6.4 D22 — Behavioural Patterns Database

Data	Value
Output number	D22
Work package / task	KPB4, task KPB4.4
Lead partner	IstroSec s. r. o.
Type	Database
Access and dissemination	public — publication in open repositories
Document with a detailed description	ADFIR-D22-KPB4-04-Database-of-Behavioural-Patterns (EN)

Summary: The output is a catalogue of detection rules, in which each rule describes one specific, observable indicator of attack activity. The rules are based on three types of telemetry: process creation events (typically Event ID 4688, Sysmon Event ID 1 or EDR telemetry), entries in Windows and ETW event logs, and registry change events, such as the creation or modification of a persistence key. The rules are deliberately static — they do not utilise machine learning or probabilistic scoring, and each has a clearly defined trigger that is evaluated as true or false. As a result, every detection is fully explainable, auditable and reproducible, and the analyst always knows exactly why the rule triggered. Each database record has five columns: the MITRE ATT&CK framework technique identifier to which the rule is mapped (for example, T1113 Screen Capture or T1547 Boot or Logon Autostart Execution), the type of telemetry and triggering event, the exact trigger condition, a short, human-

readable name for the rule, and a detailed description of the detected activity and its context. Mapping to the MITRE ATT&CK framework provides the rules with a common taxonomy, enables the measurement of coverage of attacker techniques, identifies gaps in detection, allows findings to be reported in a language that is understandable to other teams, and enables the correlation of findings with the behaviour of specific attacker groups. The document also includes examples of how to interpret the rules, as well as notes on maintaining and expanding the catalogue.

Link to project objectives: The output fulfils Specific Objective 2 of the project — the identification of applicable data analysis and machine learning methods enabling the automation of a forensic analyst's analytical work, specifically the search for behavioural patterns in data.

Assessment of the indicator: the expected value of two published datasets and databases has been fully achieved. The indicator has been met. Both outputs are intended for public access in open repositories in accordance with the principles of open science.

7 Published methodologies

Expected number	Actual performance	Status
3	3	Achieved

7.1 Definition of the indicator

The estimate was based on the number of methodologies that need to be developed to achieve the project's objectives. Unlike individual methods and models, these are comprehensive methodological frameworks that determine how a given task is approached. The first framework defines the contractual and legal environment for the provision of digital forensic analysis services; the second sets out the principles and stages of the actual processing of digital evidence; and the third outlines the procedure for creating a case timeline.

7.2 Overview of outputs

Number	Output	Work package	Responsible partner	Access
D1	Methodology for contractual relationships	KPB1	EISI	Confidential
D3	Methodology for the processing of digital traces	KPB1	EISI	Public
D23	Methodology for creating a case timeline	KPB5	UPJŠ	Public

7.3 D01 — Methodology for contractual relationships

Data	Value
Output number	D1
Work package / task	KPB1, task KPB1.1
Lead partner	European Information Society Institute
Type	Methodology
Access and dissemination	not public — internal project document archive

Data	Value
Document containing a detailed description	ADFIR-D1-KPB1-01-Methodology-of-Contractual-Relationships (EN)

Summary: This methodology establishes a uniform framework of recommendations and procedures for entering into contractual relationships when providing digital forensic analysis and cyber incident response services. Its aim is to ensure legal certainty, data protection, the preservation of the evidential value of digital evidence, and the effective resolution of security incidents. The methodology organises the conclusion of contractual relationships into two consecutive phases. The pre-contractual phase involves identifying the client's needs and objectives, conducting a preliminary analysis of their environment, verifying the provider's competence, protecting confidential information, and assessing the processing of personal data. The contract conclusion phase involves selecting a suitable contractual model, defining the subject matter of the contract and the scope of authorisations, setting out the service level agreement and response times, protecting information and personal data, handling digital evidence, liability for damage, the price, the duration and termination of the contract, dispute resolution and the handling of deliverables. The document includes two flowcharts summarising the recommended procedure for both phases. The methodology is based on the Slovak Commercial Code, the Personal Data Protection Regulation and other relevant legislation in the fields of cyber security, trade secret protection, copyright and criminal law, and is anchored in the technical standards of the ISO/IEC 27037, 27041, 27042 and 27043, and in the ISO 21043-2 standard. The arguments are supported by case law from the Supreme Administrative Court of the Slovak Republic and the regional courts.

Link to the project's objectives: The output fulfils specific objective 3 of the project — the use of automated digital forensic analysis in criminal proceedings. It establishes the contractual and legal prerequisites for the outputs of automated forensic analysis to stand up to scrutiny regarding their admissibility and probative value.

7.4 D03 — Methodology for the processing of digital traces

Data	Value
Output number	D3
Work package / task	KPB1, task KPB1.3
Lead partner	European Information Society Institute
Type	Methodology
Access and dissemination	public — publication on the project website

Data	Value
Document with a detailed description	ADFIR-D3-KPB1-03-Methodology-for-processing-digital-traces (EN)
Availability	https://adfir.upjs.sk/

Summary: The methodology defines the principles and stages of processing digital traces so that the requirements for their integrity, authenticity, admissibility and relevance are met. The document first sets out the basic principles and then formulates five principles for the collection, preservation and interpretation of digital evidence: identification of the necessary and appropriate scope of investigation; acquisition of a sufficient quantity of relevant information; documentation and chaining of evidence, together with transparency, verifiability and credibility of the procedure; impartiality and objectivity; and, finally, risk analysis, tactics and professionalism. The second part of the methodology elaborates on the five stages of digital analysis — identification of digital information, data collection, data preservation, verification of data and procedures, and analysis and interpretation of data. The methodology thus provides a process framework that links the technical implementation of forensic analysis with the requirements of the legal system and with the technical standards of the ISO/IEC 27037 to 27043 series.

Link to project objectives: This output fulfils specific objective 3 of the project. The methodology serves as a binding process framework for all technical outputs of the project, as it specifies the conditions under which a digital trace processed using automated methods remains admissible as evidence.

7.5 D23 — Methodology for creating a case timeline

Data	Value
Output number	D23
Work package / task	KPB5, task KPB5.1
Lead partner	Pavol Jozef Šafárik University in Košice
Type	Methodology
Access and dissemination	public — publication in open repositories
Document with a detailed description	ADFIR-D23-KPB5-01-Methodology-for-creating-a-case-timeline (EN)

Summary: This methodology describes the data processing workflow for the case timeline and the ETL Timeline tool, which was developed for this purpose. The tool integrates the outputs of two independent analytical workflows into a single environment for joint interpretation: it links standardised metadata records generated from the outputs of Eric Zimmerman’s tools (EZ Tools) with analytical results derived from data processed by the Plaso tool. Events extracted from forensic artefacts are thus displayed in a shared temporal context alongside analysis results, such as identified clusters, anomalies, attack phases, and relationships between events. The document describes in detail the tool’s dependencies and prerequisites, mandatory and optional input data, how to run the tool, background data processing and data input. A separate section is devoted to the graphical user interface and its functionalities — displaying the timeline graph, enriching the timeline with outputs from further analyses, working with text data, and filtering events. The concluding chapter outlines the tool’s forensic significance: the integrated view eliminates the need to manually link results from multiple tools and enables the analyst to identify temporal and logical relationships between individual events more effectively.

Link to project objectives: The output fulfils specific objective 2 of the project — the automation of a forensic analyst’s analytical work, specifically the identification of the time frame of a security incident and its individual phases.

Assessment of the indicator: the expected value of three published methodologies was achieved in full. The indicator has been met. Furthermore, the methodological framework for the analysis of open-source information as evidence in criminal proceedings was also published in the form of a peer-reviewed scientific article (Chapter 3).

8 Published methods and models

Expected number	Actual performance	Status
10	12	Exceeded

8.1 Definition of the indicator

The estimate was based on the number of methods and models that need to be developed to achieve the project's objectives. The indicator covers outputs of the 'method' and 'model' types across work packages KPB1 to KPB4. All are intended for publication, either in open repositories or on the project website.

Outputs D19 and D20 were delivered as pairs of separate outputs, namely D19.1 and D19.2, and D20.1 and D20.2 respectively. Each addresses a different methodological approach and is presented as a separate document. The total number of published methods and models therefore reached twelve, compared to the ten originally planned.

8.2 Overview of outputs

Number	Output	KPB	Type	Responsible
D02	Method for anonymising digital traces	KPB1	Method	EISI
D12	Method for the automated collection of digital traces	KPB2	Method	IstroSec
D13	Method for creating a dataset	KPB2	Method	IstroSec
D15	Model for extracting digital traces into a matrix representation	KPB3	Model	UPJŠ
D16	Model for extracting digital traces into a graph representation	KPB3	Model	UPJŠ
D17	Model for aggregating and linking digital traces	KPB3	Model	UPJŠ
D18	Model for enriching digital traces	KPB3	Model	IstroSec
D19.1	Method for identifying digital traces — clustering and large language models	KPB4	Method	UPJŠ
D19.2	Method for identifying digital traces — anomaly detection	KPB4	Method	UPJŠ

Number	Output	KPB	Type	Responsible
D20.1	Method for identifying relationships — factorisation of formal context	KPB4	Method	UPJŠ
D20.2	Method for identifying relationships — significant time windows	KPB4	Method	UPJŠ
D21	Method for determining the time frame and phases of an attack	KPB4	Method	UPJŠ

8.3 D02 — Method for anonymising digital traces

Data	Value
Output number	D02
Work package / task	KPB1, task KPB1.2
Lead partner	European Information Society Institute
Type	Method
Access and dissemination	public — publication on the project website
Document with a detailed description	ADFIR-D2-KPB1-02-Method-for-anonymising-digital-traces (EN)
Availability	https://adfir.upjs.sk/

Summary: The method defines a procedure for anonymising digital traces so that they can be used for research purposes and for training machine learning models without breaching the legal framework for the protection of personal data. The procedure consists of four sequential steps. The first is the inventory and classification of data, during which data is categorised as personal or non-personal, and other types of protected information, such as trade secrets, are specifically identified. The second step is to assess compliance with the legal requirements for the anonymisation of personal data, i.e. to verify when anonymised data ceases to be personal data within the meaning of the General Data Protection Regulation. The third step is the selection of anonymisation methods; the methodology distinguishes between two basic groups — randomisation and generalisation — and sets out the conditions for their appropriate use. The fourth step is to check the result, i.e. to verify that the anonymised dataset does not allow for the re-identification of the data subjects. The concluding section focuses on the use of anonymised digital traces for future analysis, including an explanation of the exception for processing for the purposes of scientific research.

Link to the project's objectives: This output fulfils specific objective 3 of the project and is also a prerequisite for work packages KPB2 to KPB4, as it enables work with real-world digital traces whilst ensuring the protection of personal data.

8.4 D12 — Method for the automated collection of digital traces

Data	Value
Output number	D12
Work package / task	KPB2, task KPB2.1
Lead partner	IstroSec s. r. o.
Type	Method
Access and dissemination	public — publication in open repositories
Document with a detailed description	ADFIR-D12-KPB2-01-Method-for-the-automated-collection-of-digital-traces (EN)

Summary: The method defines a robust and repeatable procedure for the automated collection of forensic artefacts, which ensures the forensic integrity and consistency of the resulting datasets. The document systematically maps the main categories of Windows operating system forensic artefacts and their sources, and classifies them into groups according to their evidential value: artefacts relating to programme execution and their presence on the system (Amcache, Background Activity Monitor, MUICache, Prefetch, RecentApps, RecentFileCache, RunMRU, Shimcache, System Resource Usage Monitor, UserAssist and Windows Activities), persistence artefacts (the DNS Hosts file, registry persistence, scheduled tasks, services, Shim databases, driver files and the WMI repository) and user activity artefacts. The method employs a hybrid approach that combines data collection from simulated attacks in a controlled environment with anonymised artefacts from real-world scenarios. For the collection and processing itself, it proposes a suite of technological solutions combining established industry tools with IstroSec's in-house tools, whilst explicitly defining the use of Windows operating system forensic artefact parsers for the rapid triage and extraction of standard artefacts.

Link to project objectives: The output fulfils Specific Objective 1 of the project — the identification of potential sources of forensic artefacts and their collection whilst preserving the integrity and authenticity of digital traces.

8.5 D13 — Method for creating a dataset

Data	Value
Output number	D13
Work package / task	KPB2, task KPB2.2
Lead partner	IstroSec s. r. o.
Type	Method
Access and dissemination	public — publication in open repositories
Document with a detailed description	ADFIR-D13-KPB2-02-Method-for-creating-a-dataset (EN)

Summary: The method addresses one of the main challenges in the field, namely the lack of high-quality forensic datasets that meet the requirements for the availability of comprehensive case studies and a sufficient number of relevant digital artefacts. The document first distinguishes between three basic types of datasets suitable for machine learning in digital forensic analysis — real-world datasets obtained directly from the resolution of actual security incidents, datasets created from simulated attacks and attacker techniques in a controlled environment, and databases derived from Capture The Flag competitions — and compares their advantages and limitations. It then presents two related methods. The first describes the creation of forensic data sets based on attacker techniques and is demonstrated using the APT-19 group scenario, distinguishing between the preparation phase and the execution phase of the attack; it focuses specifically on the detection of fileless threats and .NET-based threats, the exploitation of trusted paths, and the camouflaging and circumvention of modern protection mechanisms. The second method describes the creation of forensic datasets from materials from Capture The Flag competitions, including the identification of potential data sources.

Link to project objectives: The output fulfils specific objective 1 of the project. The method is a direct prerequisite for the creation of dataset D14 and, at the same time, enables the reproducible expansion of the project's database with new attack scenarios.

8.6 D15 — A model for extracting digital traces into a matrix representation

Data	Value
Output number	D15
Work package / task	KPB3, task KPB3.1

Data	Value
Lead partner	Pavol Jozef Šafárik University in Košice
Type	Model
Access and dissemination	public — publication in open repositories
Document with a detailed description	ADFIR-D15-KPB3-01-Model-for-the-extraction-of-digital-traces-into-a-matrix-representation (EN)

Summary: The model transforms predominantly unstructured and text-based forensic data into a format suitable for machine learning methods and formal analytical approaches. The input data consists of data obtained using the Plaso tool, which aggregates and correlates forensic artefacts from various sources into a unified timeline, known as a ‘supertimeline’. The resulting data structure takes the form of a table comprising 17 attributes (including, amongst others, date, time, time zone, MACB-type timestamps, source and source type). The model systematically identifies relevant attributes and encodes them into binary, categorical or numerical variables in such a way as to minimise information loss. After encoding, a total of 769 attributes were produced. Due to limited computational capacity, it was not possible to process all attributes at once; therefore, the principal component analysis method was applied separately for each source type, and the individual datasets were subsequently combined into a unified representation. The resulting matrix representation consists of 178 attributes, which form a unified representation of the data across all datasets; the source artefact and a textual description are provided for each attribute.

Link to project objectives: This output fulfils Specific Objective 1 of the project — the identification of a suitable representation of forensic artefacts. It bridges the gap between manual forensic analysis and automated data processing methods and forms the basis for the subsequent aggregation and linking of digital traces in Output D17.

8.7 D16 — Model for extracting digital traces into a graph representation

Data	Value
Output number	D16
Work package / task	KPB3, task KPB3.2
Lead partner	Pavol Jozef Šafárik University in Košice
Type	Model

Data	Value
Access and dissemination	public — publication in open repositories
Document with a detailed description	ADFIR-D16-KPB3-02-Model-for-extracting-digital-traces-into-a-graph-representation (EN)

Summary: The model provides a means for the systematic transformation of forensic artefacts from disparate data sources into a unified graph structure. Digital traces are represented as a set of nodes and edges, where nodes represent, for example, devices, users, processes, files or communication entities, and edges express their mutual links, interactions or causal relationships. Unlike a tabular representation, the graph model allows not only the objects of forensic analysis themselves to be captured, but also their context and temporal relationships, which is one of the key prerequisites for the successful reconstruction of events in the field of digital forensic analysis. The output includes graph-based process and state diagrams, which provide a basis for the subsequent analysis of the incident's progression, the identification of key events and the detection of anomalous structures in the data. The model is theoretically grounded in graph theory, conceptual modelling of information systems and business process management. The document also outlines a future direction — using graph algorithms, such as subgraph search, it is possible to identify patterns representing typical, known attack schemes and either predict their development in real time or, as part of post-incident analysis, gain new insights into the attacker's behaviour.

Link to project objectives: The output fulfils Specific Objective 1 of the project and also lays the groundwork for Specific Objective 2, as the graph representation serves as input for the method of identifying relationships between digital traces in output D20.

8.8 D17 — Model for aggregating and linking digital traces

Data	Value
Output number	D17
Work package / task	KPB3, task KPB3.3
Lead partner	Pavol Jozef Šafárik University in Košice
Type	Model
Access and dissemination	public — publication in open repositories
Document with a detailed description	ADFIR-D17-KPB3-03-Model-for-the-aggregation-and-linking-of-digital-traces (EN)

Summary: The model enables the systematic processing of forensic artefacts, the reduction of the volume of records, their correlation across different sources, and their interpretation in the context of identified system or user behaviour scenarios. A key concept of the model is the meta-record — a logically coherent unit of information created by linking multiple forensic traces based on temporal, identity or contextual relationships. The process of defining meta-records involved identifying relevant behavioural scenarios, such as authentication, remote access, file manipulation or persistence, to which forensic artefacts capable of indicating the occurrence of a given event were subsequently assigned. The meta-records are then aggregated into larger analytical units with the aim of reducing the fragmentation of recorded events and creating a clearer picture of the course of the analysed activities; the first method of aggregation is grouping according to a selected time window, the size of which can be adapted to the nature of the analysed data, thereby creating time intervals representing, for example, a user session, a phase of an attack or a separate incident. A separate chapter addresses aggregation functions with regard to unknown values and distinguishes between three semantic cases — where the values 0 and 1 are equally significant, where the value 1 is more significant than 0, and where the values 0 and an unknown value are equally significant. All three cases are described by a single generalised weighted function with coefficients chosen according to the semantics of the problem, which allows for the implementation of a single modular function whose behaviour can be altered simply by adjusting the weights. The model utilises the outputs of Eric Zimmerman’s tools (EvtxECmd, MFTECmd, RECmd) and also includes a proposal for the visualisation of aggregated digital traces.

Link to project objectives: The output fulfils specific objective 1 of the project. It builds on the matrix representation from output D15 and enables a higher level of abstraction, correlation and interpretation of events.

8.9 D18 — Model for enriching digital traces

Data	Value
Output number	D18
Work package / task	KPB3, task KPB3.4
Lead partner	IstroSec s. r. o.
Type	Model
Access and dissemination	public — publication in open repositories
Document with a detailed description	ADFIR-D18-KPB3-04-Model-for-enriching-digital-traces (EN)

Summary: The model provides a comprehensive framework for systematically enriching the context of digital traces by linking artefacts to external sources, indicators of compromise and threat intelligence. The document first defines the possibilities for digitally enriching evidence and the complexity of artefact enrichment during the pre-processing phase, before systematically addressing the individual types of data that can be enriched. For each type, it provides a standard set of three pieces of information: significance and use, typical artefacts containing the data in question, and methods and tools for enriching it. Among other things, the catalogue covers API keys and tokens, cloud resource identifiers, code signatures, cryptocurrency wallet addresses, CVE vulnerability identifiers, digital certificates, DNS queries, email addresses, cryptographic file hashes and fully qualified domain names. The result is a flexible and scalable model supporting various levels of enrichment, ranging from simple attributes to complex composite information. The model enables more accurate incident assessment and resolution, faster attack attribution, and better prioritisation of activities both during incident response and forensic analysis.

Link to project objectives: The output fulfils Specific Objective 1 of the project and enhances the informative value of digital traces for the analytical methods of the KPB4 work package.

8.10 D19.1 — Method for identifying digital traces: clustering and large language models

Data	Value
Output number	D19.1
Work package / task	KPB4, task KPB4.1
Lead partner	Pavol Jozef Šafárik University in Košice
Type	Method
Access and dissemination	public — publication in open repositories
Document with a detailed description	ADFIR-D19-KPB4-01-01-Method-for-identifying-digital-traces (EN)

Summary: This paper examines two distinct approaches to identifying digital traces — clustering analysis and large language models. In the section on clustering, it introduces forensic cluster metrics, validation metrics for determining the number of clusters, and a procedure for detecting anomalies via clustering, which it subsequently applies using hierarchical clustering to pre-processed datasets, with an emphasis on the interpretability of the results. In the section on large language models, it describes data pre-processing, the analysis itself, the evaluation of performance, and the interpretability of the results. The conclusion states that the aim was not to create a universal automated system to replace the

work of an analyst, but to verify the extent to which the selected methods can support the analysis of large and complex datasets, in which the manual search for relevant events is time-consuming and practically limited. Hierarchical clustering proved to be a suitable tool for datasets in which it is not possible to define in advance the number or nature of individual clusters, which is particularly significant when working with unknown activities; large language models, on the other hand, offer the ability to interpret complex textual representations and provide contextual explanations for the identified events. Partial results were published in a peer-reviewed scientific paper on the identification of relevant forensic artefacts using aggregation into time windows and hierarchical clustering, which introduced the ‘attack rate’ and ‘attack lift’ metrics.

Link to project objectives: This output fulfils Specific Objective 2 of the project — the identification of applicable data analysis and machine learning methods enabling the automation of a forensic analyst’s analytical work.

8.11 D19.2 — Method for identifying digital traces: anomaly detection

Data	Value
Output number	D19.2
Work package / task	KPB4, task KPB4.1
Lead partner	Pavol Jozef Šafárik University in Košice
Type	Method
Access and dissemination	public — publication in open repositories
Document with a detailed description	ADFIR-D19-KPB4-01-02-Method-for-identifying-digital-traces (EN)

Summary: This output addresses the identification of digital traces through anomaly detection using unsupervised learning methods. The theoretical section defines anomaly detection, its unsupervised implementation and three types of anomalies — point, contextual and collective — and categorises the available detection methods. The applied section describes the datasets used and their pre-processing, and experimentally compares eight methods: principal component analysis, Local Outlier Factor, Subspace Outlier Detection, Empirical-Cumulative-distribution-based Outlier Detection, Copula-Based Outlier Detection, Isolation Forest, Isolation-based Nearest Neighbour Ensemble and Lightweight On-line Detector of Anomalies. The comparison is carried out on three levels — experimental setup, evaluation of detection performance and comparison of computational complexity — and is supplemented by an interpretation of the results. Partial results were published in two peer-reviewed scientific papers from the project: in a study on anomaly detection in the Mozilla Firefox

browser, which identified the ECOD and COPOD methods as the most suitable compromise between accuracy and the volume of output data, and in a study on the impact of anti-forensic techniques on the accuracy of algorithmic anomaly detection.

Link to project objectives: This output fulfils Specific Objective 2 of the project. Anomaly detection represents a key approach to identifying relevant digital traces in an environment where annotated data is not available.

8.12 D20.1 — Method for identifying relationships: factorisation of formal context

Data	Value
Output number	D20.1
Work package / task	KPB4, task KPB4.2
Lead partner	Pavol Jozef Šafárik University in Košice
Type	Method
Access and dissemination	public — publication in open repositories
Document with a detailed description	ADFIR-D20-KPB4-01-Method-for-identifying-relationships (EN)

Summary: This paper addresses the identification of relationships in digital forensic analysis data through matrix factorisation based on formal concepts. After defining formal concept analysis, the paper progressively elaborates on the decomposition of the formal context — the factorisation of binary matrices, the interpretation of formal concepts as factors, and the set coverage problem — and presents the RSF algorithm, which combines greedy selection with the iterative generation of candidate formal concepts. The result is an accurate decomposition of the input incidence matrix, which enables the identification of groups of objects characterised by common attributes and the capture of relationships between temporal clusters of events and their properties. Factors with fewer attributes generally describe more general patterns present across a larger number of objects, whilst factors with more attributes capture more specific combinations of properties and have smaller extents. The application section distinguishes between general and specific factors, and specifically those related to attacks, and demonstrates that during analysis it is possible to focus on a selected significant attribute and track which other attributes it co-occurs with within individual factors. An important contribution of the method is the interpretability of the factors identified.

Link to the project's objectives: The output fulfils specific objective 2 of the project, namely the identification of relationships between digital traces as one of the four areas important for the work of a forensic analyst.

8.13 D20.2 — Method for identifying relationships: significant time windows

Data	Value
Output number	D20.2
Work package / task	KPB4, task KPB4.2
Lead partner	Pavol Jozef Šafárik University in Košice
Type	Method
Access and dissemination	public — publication in open repositories
Document with a detailed description	ADFIR-D20-KPB4-02-Method-for-identifying-relationships (EN)

Summary: This output elaborates on a method for identifying significant time windows in forensic data that may indicate the occurrence of non-standard or security-relevant activity. The method is based on formal conceptual analysis and attribute implications, and utilises recurring combinations of attributes specific to selected parts of the dataset. The output consists of temporal rules linking a closed attribute description of a group of objects to a specific time window, so that, in addition to localising the time interval, the method also provides its characteristics, which can be utilised in further forensic analysis. The method deliberately does not generate all valid attribute implications with a temporal attribute in the consequent, but creates only one rule with a maximum antecedent for each identified temporally homogeneous extent, thereby reducing the redundancy of the resulting rules. The output depends on the parameter ``kmax`, which represents a trade-off between the scope of the search and computational complexity: higher values may lead to the identification of additional time-specific extents, but the number of candidate sets examined increases sharply as the number of attributes rises. In the dataset used, a value of kmax equal to three was sufficient to identify significant time windows. Partial results of the formal conceptual analysis of forensic data were published in a peer-reviewed scientific paper devoted to attribute dependencies in the NIST Data Leakage Case dataset, in which a conceptual network comprising 26 formal concepts was constructed from a dataset containing 10,306 records and 11 binary attributes.

Link to project objectives: This output fulfils Specific Objective 2 of the project and also directly supports the determination of the incident timeline addressed in Output D21.

8.14 D21 — Method for determining the timeframe and phases of an attack

Data	Value
Output number	D21
Work package / task	KPB4, task KPB4.3
Lead partner	Pavol Jozef Šafárik University in Košice
Type	Method
Access and dissemination	public — publication in open repositories
Document with a detailed description	ADFIR-D21-KPB4-03-Method-for-determining-the-time-frame-and-phases-of-an-attack (EN)

Summary: The method reconstructs and analyses digital traces through time series analysis, with the aim of shifting forensic analysis from manual log review to semi-automated determination of the time frame of the start of an attack and its individual phases. Time series were generated from relevant artefacts of various types — Windows event logs, the master file table, the system resource usage monitor and the registry — across four simulated incidents. The method pursues three objectives: determining the moment of the first anomalous activity, creating an input layer for automated anomaly detection using machine learning methods, and generating a condensed attack timeline linking related events from various sources.

Link to the project's objectives: The output fulfils specific objective 2 of the project — the identification of the timeframe of a security incident or criminal offence, or, where applicable, the sub-timeframes of individual actions, as one of the four areas important to the work of a forensic analyst.

Assessment of the indicator: the expected value of ten published methods and models was achieved in full. The indicator has been met. Partial results of the methods from the KPB4 work package were also published in the form of peer-reviewed scientific papers at the international conferences ARES 2025, ITAT 2025 and CISIS 2026 (Chapter 3). Outputs intended for publication will be made available in open repositories or on the project website in accordance with the open access principle declared in the project description.

9 Other published project outputs

In addition to the outputs falling under the measurable indicators set out in Chapters 6 to 8, the project has produced a further public output in the form of a ‘report’, which, by its nature, is an applied summary of the results of the entire project and is therefore presented in a separate chapter.

9.1 D24 — Extended report on forensic analysis

Data	Value
Output number	D24
Work package / task	KPB5, task KPB5.2
Lead partner	IstroSec s. r. o.
Type	Report
Access and dissemination	public — publication in open repositories
Document with a detailed description	ADFIR-D24-KPB5-02-Extended-report-on-forensic-analysis (EN)

Summary: This output presents a template structure for an extended forensic analysis report, which integrates the results of the methods and models developed in work packages KPB2 to KPB4 into a format that can be used in practice by security teams and in legal assessments. The report opens with an executive summary and methodology, which sets out the sources of evidence, the methods used to process and parse them, and the approach to data enrichment. This is followed by information on the evidence examined, including the type of evidence and cryptographic fingerprints, and details of the device under investigation — name, operating system version, last shutdown time, IP address and network settings, installed programmes, and applicable enrichment techniques. At the heart of the document are the findings, organised by forensic categories: programme launch artefacts (Prefetch, Amcache, system resource usage monitor, Background and Desktop Activity Moderator, UserAssist, Shimcache, RecentFileCache, RunMRU and others), persistence artefacts (Run and RunOnce keys, scheduled tasks, services, startup folders, persistence via WMI and persistence in the registry), log analysis (remote access logs, Sysmon, PowerShell, Task Scheduler, application logs, firewall, DNS, proxy, VPN and alerts from EDR, SIEM and anti-malware solutions) and user activity artefacts (browser and download history, RecentDocs, recently used documents and trusted Office documents). Each category concludes with an optional summary, allowing the report to be used as a template for partial analyses as well.

Link to project objectives: The output fulfils Specific Objective 2 of the project as well as Specific Objective 4 — the dissemination of project results to the professional community,

particularly security experts and teams. The report presents the project's analytical outputs in a format that is comprehensible to both the client commissioning the analysis and law enforcement agencies.