DigForASP: A European Cooperation Network for Logic-based AI in Digital Forensics

Stefania Costantini (UnivAQ)    Francesca Lisi (UniBA)
Raffaele Olivieri (RaCIS)

CILC 2019, Trieste, June 20, 2019
The Action Web Site: digforasp.uca.es
COST: ’European Cooperation in Science and Technology’

COST provides networking opportunities for researchers and innovators in order to strengthen Europe’s capacity to address (interdisciplinary) scientific, technological and societal challenges.

COST implements its mission by funding excellence-driven, open and inclusive networks for peaceful purposes in all areas of science and technology.
COST Funding Scheme

Through these networks, the so-called ‘COST Actions’, COST provides funds for:
- meetings
- training schools
- short term scientific missions or other networking activities

Participants are invited to relevant meetings by each event’s responsible person.

Participants can apply for STSM, “Short Term Scientific Missions” providing a reason, a program and a budget.
The DigForASP Action

- The COST Action CA17124 DIGFORASP “DIGital FORensics: evidence Analysis via intelligent Systems and Practices” is financed under funds for “European Cooperation in Science & Technology, Horizon 2020”

- Action Activities: October 2018 - October 2022

- Stems from an idea by myself and my (former) Ph.D. student Raffaele Olivieri (officer of Italian Law Enforcement), proposal written with the aid of a small group of colleagues
Participants

Who are an Action’s participants? researchers from Universities, or other Institutions related to Research and/or Development and/or Applications, in particular:

- the original Action’s proponents, and
- other partners which join the Action later (so far, from 34 different countries) by applying to the Coordinator and to the national COST representative.

From which countries? all COST Countries, “Near Neighbour Countries” and “International Partner Countries” (the latter with no funding)
COST Countries

- International Partner Countries
- COST Near Neighbour Countries
- COST Full & Cooperating Members
DigForASP Proponents

Proponents: DigForASP has 55 proponents (of which 40% female) from 21 different countries, among which 9 EU countries, 10 ICT (“Inclusiveness Target Countries”, i.e., countries which, though external, are closely related with EU, plus Georgia and Russian Federation.)
DigForASP Management

- Coordinator: Prof. Jesus Medina Moreno, University of Cadiz, Spain
- Vice-Coordinator Prof. Stefania Costantini, University of L’Aquila, Italy
- Management Committee: two representatives for each participating country, selected by each national COST representative upon recommendation by the Coordinator.
- Science Communication Manager: Prof. Francesca Lisi, University of Bari, Italy
Action’s Subject: Digital Forensics (DF)

DF is a branch of criminalistics which deals with the identification, acquisition, preservation (according to precise regulations), analysis and presentation of the information content of computer systems, or in general of digital devices.

- Computer Forensics
- Live Forensics
- Mobile Forensics
- Database Forensics: concerns database analysis for the retrieval of data or of transaction activities and logs.
- Network & Internet Forensics
- Embedded Forensics: concerns the analysis of embedded systems
- Cloud Forensics
- Multimedia Forensics
Digital Forensics: Phases

1 Identification, i.e. retrieving, via various forms of investigation, devices that may possibly contain digital data useful for the investigation.

2 Acquisition, i.e., retrieving evidence (from storage devices or from network interception).

3 Preservation.

4 Evidence Analysis, where the evidence collected is examined and aggregated to identify possible sources of proof to be presented in Court.
Action’s Focus: Evidence Analysis

Weak points of human-based evidence analysis (despite the availability of off-the-shelf tools):

- outcomes should be verifiable with respect to the results, and to how such results are generated (now, results provided by available off-the-shelf tools which are ‘black-box’);
- all the above must be explainable to the involved parties.

Otherwise:

- undesirable uncertainty about the outcome of evidence analysis;
- different technicians can reach different conclusions, possibly leading to different judgments in court.
Evidence Analysis: aspects involved

- Timing of events and actions
- Possible causal correlations
- Contexts in which suspicious actions occurred
- Skills of the involved suspects
- Awareness of the involved suspects of committing a violation or a crime and of the degree of severity of the violation/crime

For each given case, there can be possible alternative scenarios (alternative consistent interpretations of the data)
Our answer: Artificial Intelligence and Automated (logical) Reasoning

Several methods, techniques and tools have been developed over the years with the aim to:

- extract useful knowledge from data;
- reason with uncertain/incomplete knowledge;
- perform causal and temporal reasoning;
- generate consistent scenarios compatible with a set of known facts.
The importance of Computational Logic

- Reasoning functionalities where the problem specification and the computational program are closely aligned: results be formally verified, visualised and explained.

- Free inference engines are available from some powerful computational logic techniques, thus allowing for fast prototyping and experiments.

Engineered tools will have to be designed and implemented in future projects possibly stemming from the DigForASP Action.
My Vision: Smart Cyber-Physical System for Digital Investigations
In a computer belonging to a suspect, the technicians found:

- a list of file names, with associated size and type;
- a set of files, with size and type, some of them with illicit contents;
- the log of a file exchange tool, reporting the names of the exchanged files.

Question: did the suspect exchange files with illicit contents?
Data Recovery & File Sharing

Filesharing

Illicit Files
Data Recovery & File Sharing

Memory

Recovered Files

Illicit Files

INDX Files

Cache Filesharing
Solution (in ASP)

1. Represent data as datalog facts.
2. Apply the well-known ‘stable marriage’ algorithm in order to try to couple files with their names; several possible scenarios can be obtained, as a name may correspond (for type and size) to more than one file.
3. Assess the plausibility of illicit file exchange, e.g., in how many scenarios such an exchange is postulated; proof element to be reported to the judge, for proper consideration in the context of the case.

Prototype implementation and experiments on realistic data (by Raffaele Olivieri in his Ph.D. Thesis)
Other developed examples (ASP)

- “Monkey and banana” for alibi verification.
- Clique identification and graph analysis in general for identifying key groups in criminal organization.
- Hidato puzzle for path verification.

References:


Aim of the Action in the short term

Synergic cooperation of experts from the Digital Forensics field, crime investigators, lawyer and experts from several areas of AI and Automated Reasoning. Why?

- DF experts alone are not even aware of the potential that European research offers for aiding them in their activities.
- Researchers alone are not familiar with the subtleties of such a challenging interdisciplinary field.

Bringing together researchers from several areas of AI and Automated Reasoning with DF experts can foster a productive exchange.
Aim of the Action in the long term

- A substantial evolution of the current paradigm of evaluation and interpretation of data in DF analysis, which might be exportable, in the future, also to other Forensic Sciences;
- A “breakthrough innovation” for the judicial system, based on the possibility of adopting intelligent, reliable and dependable decision-support systems for the reconstruction of facts.

From the socio-economical perspective, the use of automated reasoning tools will become, in the long-term, a positive benefit for all the involved stakeholders, with a twofold improvement both on efficiency and quality.
Action: Expected Results

- Explore the potential of AI and Automated Reasoning in DF.
- Cope with the technical and practical aspects but also with foundational and societal issues and with the ethical aspects involved.
- Attract Companies by effective dissemination activities.
- Become a catalyst for future specific research projects with an international and multidisciplinary composition.
Scientific Working Groups

WG1  Digital Forensics requirement analysis (Leader Dr. Raffaele Olivieri, Italy)

WG2  Research on applications of AI/Automated Reasoning to DF (Leader Prof.ssa Alessandra Mileo, Ireland)

WG3  Prototypes and Platforms (Leader Prof. Pedro Cabalar, Spain)

WG4  Benchmarks based on real cases (Leader Prof.ssa Viviana Mascardi, Italy)

WG5  Platforms integration and multi-dimensional environment (Leader Prof.ssa Esra Erdem, Turkey)

Ongoing work: Ontologies for case representation, reasoning on these ontologies by SPARQL, Datalog, ASP, ELP, etc.
How to Join a Working Group

WG’s activities now starting

Everyone who may be interested in the activity of the Action can join one or more of the WG’s: follow this link and subscribe!
https://goo.gl/forms/ni2y4T1mWd4Acs9n1
You will be funded for participation to WG Meetings

Next Meeting: WG4, Trieste (Italy) June 18-19, to join in contact the WG leader Prof. Viviana Mascardi, email viviana.mascardi@unige.it
Thank you for your attention!
Questions are welcome

Grazie per l’attenzione