PhD in Information Technology and Electrical Engineering

Università degli Studi di Napoli Federico II

PhD Student: Giovanni Cozzolino

XXXI Cycle

Training and Research Activities Report – Third Year

Tutor: **Antonino Mazzeo**
co-Tutor: **Flora Amato**

1. Information

- Giovanni Cozzolino, Master's Degree in Computer Engineering, in 2013 from the University of Naples Federico II.
- XXXI Cycle – ITEE
- DIETI Grant
- Tutor: Antonino Mazzeo – co-Tutor: Flora Amato

2. Study and Training activities

Courses:

- *Secure systems design*; prof. Valentina Casola
- *CLA Cambridge English: C1 Advanced (CAE)*; prof. John Crockett

Seminars

Optimal content distribution and Multi-resource allocation in software defined virtual CDNs	Claudio Sterle, A.M. Tulino	11/12/2017
Large scale integrative bioinformatics and systems biology in cancer genomics	M. Ciccarelli	18/01/2018
SeeQC-eu, Hypres quantum engineering company in Europe	O. Mukhanov, A. Kirichenko	30/01/2018
HL-LHC Upgrade, Hlumi	I.B. Alonso	11/05/2018
IBM Q: building the first universal quantum computers for business and science	F. Mattei, N. Said	16/05/2018
Promoting sparsity with Krylov iterative solvers and hierarchical Bayesian models: the L2 magic	D. Calvetti	22/05/2018
How does MathWorks accelerate the pace of engineering and science	F. Alderisio	01/06/2018
The Napoli Federico II IEEE student branch	S. Marrone	17/07/2018
Wearable Systems: design and implementation challenges	M. Ghassemian	17/09/2018

Activities schedule:

	Credits year 1								Credits year 2								Credits year 3									
		1	2	3	4	5	6			1	2	3	4	5	6			1	2	3	4	5	6			
	Estimated	bimonth	bimonth	bimonth	bimonth	bimonth	bimonth	Summary	Estimated	bimonth	bimonth	bimonth	bimonth	bimonth	bimonth	Summary	Estimated	bimonth	bimonth	bimonth	bimonth	bimonth	bimonth	Summary	Total	Check
Modules	20		6		6		6	18	12	0	3	3	3	0	0	9	6	0	6	0	0	0	0	6	33	30-70
Seminars	5							0	10	0,8	0,2	1,8	3,5	0	1,1	7,4	5	0,4	0,6	0	1,2	0,3	0,4	2,9	10,3	10-30
Research	35	6	5	8	6	4	6	35	45	8	7	8	6	8	8	45	60	10	10	10	10	10	10	60	140	80-140
	60	6	11	8	12	4	12	53	67	8,8	10	13	13	8	9,1	61	71	10	17	10	11	10	10	69	183	180

3. Research Activities

Title

A semantic methodology for (un)structured digital evidences analysis

Studies

Research activities in the second year of Ph.D. focused on the study of technologies related to Semantic Web applied to Digital Forensics domain.

Contents of the course I attended (“*Secure systems design*”) included deep insights about methodologies, protocols and tools adopted for a secure system design. These concepts were useful to understand the security mechanism that each digital investigator have to deal with in order to guarantee that evidences are not corrupted.

Description of Research Activities

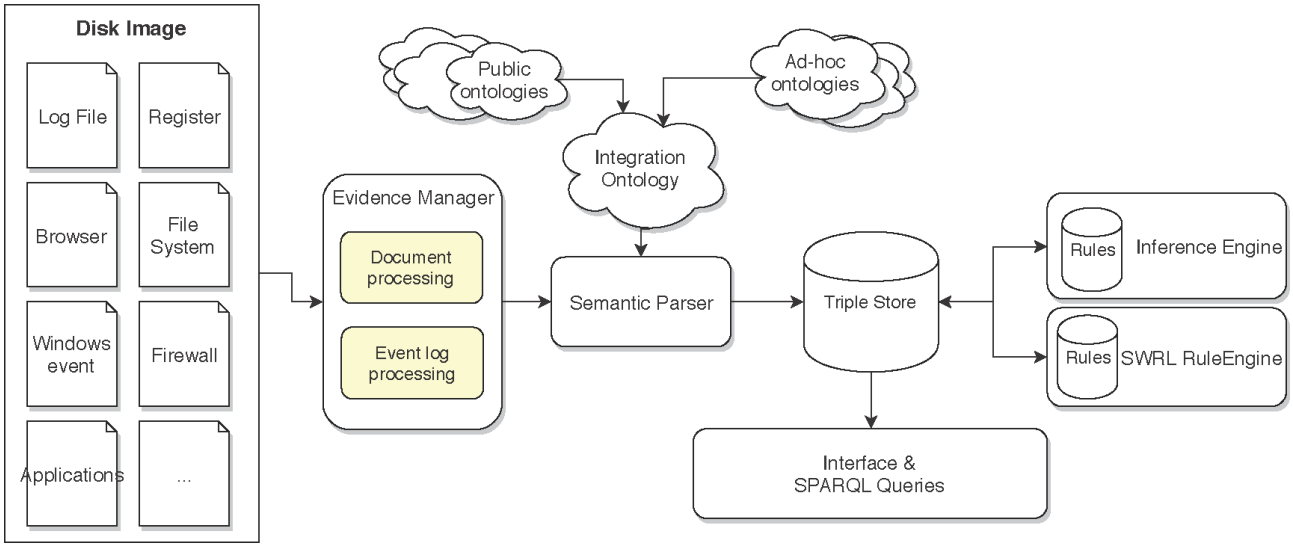
Research activities of the **first** year of this Ph.D. course had the main goal of designing a methodology, based on Semantic Web technologies, for integration, correlation and retrieval of complex information from heterogeneous data sources. The approach we proposed, which is based on semantics vocabularies, also enables the automation of some steps in the analysis of data by means of correlation.

The following Figure shows the basic steps of the methodology:

Research activities during the **second** year of this Ph.D. course included the definition and the implementation of a system architecture able to enact the aforementioned methodology. We

provided many use cases too, in order to test and validate the methodology and to evaluate its scalability and adaptability to different scenarios.

In particular, I mainly focused **third** year on Forensic Investigation. The goal of Digital Forensics is not only the gathering, management and analysis of data stored on digital devices, but, above all, it requires interpretation of evidences. Correlation of information is very important in forensics analysis, because it is the only way to allow for the contextualization of digital evidences, promoting them as clues.



System Architecture consists of an ontology and five modules:

1. Evidence Manager: it loads binary content of digital evidences, identifying the type of given source and verifying its integrity by using hashing;
2. Semantic Parser: it generates an OWL representation of knowledge extracted from digital evidence; it instantiates the ontology that resumes concepts both from public and custom domains;
3. Inference Engine: it performs automated reasoning, according to the OWL specifications, and reflecting domain expertise reasoning.
4. SWRL Rule Engine: it uses SWRL rules in order to correlate different individuals or to establish relationships among individuals belonging to different ontologies but representing similar concepts.
5. SPARQL Queries: it is responsible for accepting SPARQL queries from users and for retrieving results by using a SPARQL query engine.

During the **third** year of the course mainly I refined the definition of my methodology and the system architecture, enhancing the capabilities of Evidence Manager module, with the implementation of:

- A full-text Document Processing component, which is responsible to process textual document and to extract structured information in a common RDF format, able to be integrated in a knowledge base, through semantic tools;
- A Log Processing Module, able to parse and analyse logs coming from different sources (like operating system, firewall, applications, etc.) and to represent extracted information in the same common format to be integrated in the knowledge base.

Collaborations

- Founded European Project CREA (Conflict Resolution with Equitative Algorithms). Justice Programme, Grant Agreement number: 766463 —CREA —JUST-AG-2016/JUST-AG-2016-05 Coordinated by University of Naples "Federico II".
- Conferenza dei Rettori delle Università Italiane, Ministero della Giustizia, Dipartimento dell'organizzazione giudiziaria, del personale e dei servizi – Direzione Generale per i sistemi informativi automatizzati (DGSIA): "*La gestione del servizio di formazione qualificata, ricerca applicata e certificazione di professionalità a seguito della riorganizzazione del Ministero della Giustizia per il tramite dei sistemi ICT, su strumenti e funzionalità del Processo Civile e Penale telematico, nell'ambito della riduzione dei tempi della giustizia, su profili di sicurezza dei sistemi informativi in uso presso il Ministero della Giustizia e gli uffici Giudiziari*".
    - Applied research on tools, functionality and security protocols of Ministry of Justice information systems.
    - Document management, encrypted full text indexing, semantic searches
    - System architecture, virtualization, authentication and authorization protocol

4. Products

Publications:

- *A MAS model for reaching goals in critical systems* - Amato, F., Cozzolino, G., Mazzeo, A., Moscato, F. - Smart Innovation, Systems and Technologies
- *Data mining in social network* - Amato, F., Cozzolino, G., Moscato, F., (...), Picariello, A., Sperli, G. - Smart Innovation, Systems and Technologies
- *Using semantic tools to represent data extracted from mobile devices* - Cozzolino, G. - Proceedings - 2018 IEEE 19th International Conference on Information Reuse and Integration for Data Science, IRI 2018
- *Detect and correlate information system events through verbose logging messages analysis* - Amato, F., Cozzolino, G., Mazzeo, A., Moscato, F. - Computing
- *Using multilayer perceptron in computer security to improve intrusion detection* - Amato, F., Cozzolino, G., Mazzeo, A., Vivenzio, E. - Smart Innovation, Systems and Technologies

5. Conferences and Seminars

- The 10-th International Conference on Intelligent Networking and Collaborative Systems (INCoS-2018) - Comenius University in Bratislava, Slovakia September 5 - 7, 2018
  - Presentation of the work: Semantic Analysis of Social Data Streams
  - Presentation of the work: Using semantic tools to represent data extracted from mobile devices
- The 21st International Conference on Network-Based Information Systems (NBiS-2018)) - Comenius University in Bratislava, Slovakia September 5 - 7, 2018
- The 10th International Symposium on Cyberspace Safety and Security (CSS 2018) – October 29-31, 2018 – Amalfi – Italy
  - Presentation of the work: An advanced methodology to analyse data stored on mobile devices

6. Tutorship

- Sistemi informativi
  - Type: Seminar
  - Subject: BPMN and Bonita BPM
  - Hours: 6