



**PhD in Information Technology and Electrical Engineering**  
Università degli Studi di Napoli Federico II

**PhD Student: Simona De Vivo**

---

Cycle: XXXVII

**Training and Research Activities Report**

**Year: First**

Simona De Vivo

**Tutor: Prof. Domenico Cotroneo**

**Date: December 13, 2022**

# Training and Research Activities Report

PhD in Information Technology and Electrical Engineering

Cycle: XXXVII

Author: Simona De Vivo

## 1. Information:

- **PhD student:** Simona De Vivo
- **DR number:** 996112
- **Date of birth:** 22/07/1997
- **Master Science degree:** Master's degree in Computer Engineering (cum laude), University of Naples Federico II
- **Doctoral Cycle:** XXXVII
- **Scholarship type:** PhD student grant (Grant Type: MUR PON)
- **Tutor:** Prof. Domenico Cotroneo

## 2. Study and training activities:

Activity	Type <sup>1</sup>	Hours	Credits	Dates	Organizer	Certificate <sup>2</sup>
Vivi la nostra Digital Innovation	Seminar	2	0,4	01/02/2022	Prof. Domenico Cotroneo	Y
Rails Mid-Term Workshop	Seminar	5	1	25/02/2022	Prof.ssa Valeria Vittorini, Zhiyuan Lin	Y
Project Vāc: Can a Text-to-Speech Engine Generate Human Sentiments?	Seminar	1	0,2	28/02/2022	Dip. Fisica, "Ettore Pancini" - DIETI, Unina	Y
Virtualization technologies and their applications	Ad hoc Course	23	5	January 17;21;24; 28; February 01; 04; 07; 11; 14; 18; March 04 2022	Dr. Luigi De Simone, DIETI	Y
Explainable Natural Language Inference	Seminar	1h 30 min	0,3	13/04/2022	Prof. Francesco Cutugno, DIETI, Unina	Y
An Introduction to Deep Learning for Natural Language,	Seminar	1	0,2	13/04/2022	Prof. Francesco Cutugno, DIETI, Unina	Y
On using simple optimization techniques for tuning of UAVs	Seminar	2	0,4	27/04/2022	Dr. Fabio Ruggiero - DIETI - Unina	Y

# Training and Research Activities Report

PhD in Information Technology and Electrical Engineering

Cycle: XXXVII

Author: Simona De Vivo

Software Product Line Management and Software Versioning in SADAS	Seminar	2	0,4	27/04/2022	Prof. Sergio Di Martino, DIETI - Unina	Y
Using Delays for control	Seminar	1	0,2	21/04/2022	Prof. Stefania Santini - DIETI - Unina	Y
Laboratorio di MATLAB e SIMULINK per l'Ingegneria Elettrica	Tutorship	5	5	21/03/2022; 28/03/2022		Y
Statistical data analysis for science and engineering research	Ad hoc Course	12	4	22-24-29-31/03 and 05-07/04, 2022	Prof. Roberto Pietrantuono - DIETI	Y
Scientific Programming and Visualization with Python	Ad hoc Course	8	2	May, 2022	Prof. Alessio Botta - DIETI	Y
Accelerated Deep Learning via Efficient, Compressed and Managed Communication	Seminar	1	0,2	03/05/2022	Prof. A. Pescapè, DIETI - Unina	Y
AR for remote use of measurement instrumentation	Seminar	2	0,4	24/05/2022	Prof.ssa Annalisa Licardo	Y
Fixed Wireless Access	Seminar	6	1,2	17/05/2022	Prof. Antonia Maria Tulino	Y
Population and medical genomics applications to human traits and diseases	Seminar	1	0,2	29/04/2022	Computational and Quantitative Biology PhD	Y
QoE management in 5G networks	Seminar	2	0,4	08/06/2022	Prof.ssa Antonia Maria Tulino	Y
Reference standards for next generation sequencing assays on cytological samples: A worldwide ring trial study	Seminar	1	0,2	12/05/2022	Prof. Pasquale Arpaia	Y

# Training and Research Activities Report

PhD in Information Technology and Electrical Engineering

Cycle: XXXVII

Author: Simona De Vivo

Switched differential algebraic equations: jumps and impulses	Seminar	1	0,2	01/06/2022	Prof. Raffaele Iervolino, DIETI, Unina	Y
Thermoacustics for renewable energies	Seminar	1	0,2	01/06/2022	Osvaldo Danisi	Y
Vine robots: design challenges and unique opportunities	Seminar	1	0,2	31/05/2022	Dr. Mario Selvaggio – DIETI, Unina	Y
Wireless collaborative intelligent with goal-oriented communications	Seminar	2	0,4	10/06/2022	Prof.ssa Antonia Maria Tulino	Y
Imprenditorialità Accademica	Ad hoc Course	10	4	26/05 - 14/06 2022	Prof. Pierluigi Rippa, DIETI	Y
ARTISAN Summer School (Role and effects of ARTificial Intelligence in Secure ApplicatioNs)	Doctoral School	32	6	4-7 July 2022	Université Grenoble Alpes LCIS, 50	Y
Introduction to Intellectual Property Management	Seminar	2	0,4	19/07/2022	Prof.ssa Antonia Maria Tulino	Y
Machine Learning for Science and Engineering Research	Ad hoc Course	20	5	20-21-22-23-24-27-28-29-30/06 - 01/07 2022	DIETI ITEE PhD	Y
DataWeek	Course	8	1	18-19-20-21/10/2022	International tech academy, entirely online	Y
Privacy-Preserving Machine Learning	Seminar	2	0,4	14/10/2022	Proff. Simon Pietro Romano, Roberto Natella (DIETI, UNINA)	Y

# Training and Research Activities Report

PhD in Information Technology and Electrical Engineering

Cycle: XXXVII

Author: Simona De Vivo

Data Security	Master Degree in Data Science – Course		6	1 <sup>st</sup> semester 2022-2023	Università degli Studi di Napoli Federico II	Y
Critical Data Visualization	Master Degree in Data Science – Course		6	1st semester 2022-2023	Università degli Studi di Napoli Federico II	Y
Connecting the dots: Investigating an APT	Seminar	2	0,4	11/11/2022	Proff. S.P. Romano, R. Natella – DIETI	Y
Privacy and Data Protection	Seminar	2	0,4	22/11/2022	Proff. S.P. Romano, R. Natella – DIETI	Y
Cybercrime and Information warfare: national and international actors	Seminar	2	0,4	18/11/2022	Proff. S.P. Romano, R. Natella – DIETI	Y
Digital Forensics	Seminar	2	0,4	06/12/2022	Proff. S.P. Romano, R. Natella – DIETI	Y
From Cyber Situational Awareness to Adaptive Cyber Defense: Leveling the Cyber Playing Field	Seminar	2	0,4	13/12/2022	Prof. Giancarlo Sperli	Y

1) Courses, Seminar, Doctoral School, Research, Tutorship

2) Choose: Y or N

## 2.1. Study and training activities - credits earned

	Courses	Seminars	Research	Tutorship	Total
Bimonth 1	12	2	6		20
Bimonth 2		1,6	6		7.6
Bimonth 3	5	1,5	6	5	17.5
Bimonth 4	6	3,6	6		15.6
Bimonth 5	10	0,4	6		16.4
Bimonth 6	6	0,4	6		12.4
<b>Total</b>	<b>39</b>	<b>9,5</b>	<b>36</b>	<b>5</b>	<b>89.5</b>
<b>Expected</b>	<b>30 - 70</b>	<b>10 - 30</b>	<b>80 - 140</b>	<b>0 – 4.8</b>	

### 3. Research activity:

**Title:** Leveraging Green AI Techniques to Perform Anomaly Detection in Complex Systems

#### Description and Study

##### Green AI

*Artificial intelligence* (AI) leverages computers and machines to mimic the problem-solving and decision-making capabilities of the human mind. AI includes the sub-fields of machine learning and deep learning. These disciplines are comprised of AI algorithms that typically make predictions or classifications based on input data. The field of AI has reported remarkable progress on a broad range of capabilities including object recognition, game playing, speech recognition, and machine translation. Much of this progress has been achieved by increasingly large and computationally intensive deep learning models.

We distinguish between **Red AI** and **Green AI**.

Red AI is AI that focuses only on results — not on resources consumed. There are two major factors that make AIs “Red”. The first is using larger datasets so the AI has more to learn from. The second is using more complex AI models (i.e., more math) so the AI learns more effectively. These two approaches have two things in common. They both require huge amounts of resources. Often are used large amounts of data to guarantee better results, but this is a bad solution because it is often simply redundant data or not very relevant to the problem to be solved that only contribute to increasing the energy consumed during model training.

Green AI refers to AI research that yields novel results while considering the computational cost, encouraging a reduction in resources spent.

Green AI promotes approaches that have favourable performance/efficiency trade-offs. If measures of efficiency are widely accepted as important evaluation metrics for research alongside accuracy, then researchers will have the option of focusing on the efficiency of their models with positive impact on both inclusiveness and the environment.

In the view of reduce energy consumptions and improve performances, I am developing a *feature extraction* technique. Feature extraction increases the accuracy of learned models by extracting features from the input data. This phase of the general framework reduces the dimensionality of data by removing the redundant data. Of course, it increases training and inference speed. To this aim, I am leveraging natural language processing (NLP) techniques to automatically select the most important features and attributes from data.

# Training and Research Activities Report

PhD in Information Technology and Electrical Engineering

Cycle: XXXVII

Author: Simona De Vivo

---

## Anomaly Detection

In data analysis, *anomaly detection* (also known as outlier detection and sometimes as novelty detection) is generally defined as the recognition of rare elements, events, or observations that deviate significantly from most data and do not conform to a well-defined notion of normal behavior. Such examples may give rise to the suspicion of being generated by a different mechanism or appear inconsistent with the rest of that data set.

Anomaly detection finds application in many industries, including cybersecurity, medicine, computer vision, statistics, neuroscience, law enforcement, financial fraud, etc.

My research proposal aims to exploit AI, including NLP techniques, to monitor infrastructures and identify the anomalies due to cyber-attacks.

The purpose is to implement an intelligent solution for analyzing and detecting attacks by using online data processing. Infrastructure monitoring will also serve as basis for collecting information on attacks (cyber threat intelligence), which is used to understand what cyber threats target the infrastructures and are essential information for identifying and preventing such attacks.

To this goal, I am working on an approach that can satisfy the requirements of the green AI (i.e., using as few computational resources as possible), trying to maximize the efficiency of the model's training, the model's inference, and the usage of data (e.g., by using transfer learning techniques or pre-trained models).

I am also working on the possibility of implement the solution in an industry case study (e.g., railway), considering the resources constraints of embedded systems.

## 4. Research products:

In this first year, I have produced the following product.

### 4.1. Publications

#### Conference Paper

1. Liguori Pietro, Improta Cristina, **De Vivo Simona**, Natella Roberto, Cukic Bojan, & Cotroneo Domenico (2022). "Can NMT Understand Me? Towards Perturbation-based Evaluation of NMT Models for Code Generation". IEEE/ACM 1st International Workshop on Natural Language-Based Software Engineering (NLBSE), 2022.

# Training and Research Activities Report

PhD in Information Technology and Electrical Engineering

Cycle: XXXVII

Author: Simona De Vivo

---

## 5. Conferences and seminars attended

I participated the following conference:

Conference name	Place	Dates	Number of accepted papers
The 1st Intl. Workshop on Natural Language-based Software Engineering Co-located with ICSE 2022	Virtual	May 8, 2022	14

As the author, I **presented** the following paper:

- “Can NMT Understand Me? Towards Perturbation-based Evaluation of NMT Models for Code Generation”, NLBSE2022, virtual.

## 6. Activity abroad:

...

## 7. Tutorship

In this first year I carried out 5 hours of tutoring as part of the "additional activities" called "Matlab and Simulink for electrical engineering".