COMPUTER AND CONTROL ENGINEERING

DAUIN - AI4CTI - Artificial Intelligence for Cyber Threat Intelligence

Funded By	Dipartimento DAUIN
Supervisor	MELLIA MARCO - marco.mellia@polito.it
Contact	GARZA PAOLO - paolo.garza@polito.it
Context of the research activity	As digital reliance grows, cyber fraud is surging, with costs projected to hit $13.8T by 2028. Social engineering attacks exploit multimedia and fake news, bypassing outdated security tools. AI is key to countering these threats, using advanced algorithms for scalable, adaptive threat detection. The candidate will develop AI-driven cybersecurity solutions, leveraging multimodal analysis to detect malicious content, despite limited ground truth data, enabling on-device protection and integration.
	Nowadays, we rely on digital services to stay informed, organise our work, manage our finances, and more. Numbers in hand, 63,1% of the global population accesses the web daily for work, social media, and any service. With this, cyber fraud and attacks are proliferating. With the explosion of social networks and instant messaging, attack vectors multiply, making social engineering attacks based on counterfeit multimedia and fake news an everyday threat. Artificial Intelligence is the only means to counter these ever-growing threats. Testified by its success in Natural Language Processing and Computer Vision applications, AI allows us to design algorithms and systems capable of identifying new threats promptly, with great scalability, automatically adapting to modifications. In this study, we study and develop ground-breaking AI-based technologies to counter social engineering attacks: a scalable and clever data collection plan feeds a graph-based data ocean, on top of which AI models – specifically designed to extract multimodal features from any internet content – will pave the road to highly-specialised downstream tasks ultimately designed to detect malicious content and prevent data loss. Research objectives: The candidate will develop AI-based solutions to counter cyber threats, focusing on the automatic detection of phishing attacks across multiple vectors, including email, websites, and messaging applications. The project will be based on three key pillars: * data collection and aggregation: Crawl the web and the dark web, in a scalable and cost-effective way, and discover and explore online groups in

Objectives	messaging applications such as Telegram or WhatsApp and Online Social Media Net-works like Instagram or TikTok. * data storage and indexing: develop an innovative graph-based data structure that allows for simplifying the query process to support the integration with AI-based algorithms that typically need to process data during training. Given that state-of-the-art graph-based platforms are still in their infancy, the candidate will contribute to new solutions specifically tailored to the web security scenario. * AI algorithms: The candidate will focus on the development of a foundation model specifically engineered for cybersecurity. This will be a cornerstone that will streamline and open applications to several use cases. Unlike Large Language Models or Computer Vision models that address a single specific domain, the model will be multi-modal in nature, given the mix of text, images, videos, languages, etc. that are found on the web. Research work plan: We foresee three phases: * During the first year, the candidate will review the state of the art, and focus on the data collection, storage and indexing platforms * During the second year, the candidate will focus on the development of AI solutions, leveraging the data collected and aggregating CTI outlets to obtain labelled data to train algorithms. These algorithms will work initially on separate domains, like text and images. * During the third year, the candidate will deep dive into AI approaches, fine-tuning the models to vertical applications like phishing detection and malicious profiles found on social media networks. Here, the models will be multimodal in nature, able to analyse images and text at the same time, References: - Boffa, M., Valentim, R. V., Vassio, L., Giordano, D., Drago, I., Mellia, M., & Houidi, Z. B. (2023). LogPrécis: Unleashing Language Models for Automated Shell Log Analysis, Computers & Security, Volume 141,2024, - Boffa, M., Milan, G., Vassio, L., Drago, I., Mellia, M., & Houidi, Z. B. (2022, June). Towards nlp-based pro-cessing of honeypot logs. EuroS&PW - Valentim, R., Drago, I., Mellia, M., Cerutti, F., X-squatter: AI Multilingual Generation of Cross-Language Sound-squatting. ACM Transactions on Privacy and Security - Valentim, R., Drago, I., Mellia, M., Cerutti, F., Lost in Translation: AI-based Generator of Cross-Language Sound-squatting, EuroS&PW, 2023

Skills and competencies for the development of the activity	- Good programming skills (such as Python, Torch, Spark) - Solid Machine Learning knowledge - Knowledge of NLP and LLM - Fundamentals of Networking and Computer Security