Intervento realizzato da

In consideration of the determination of the Regione Piemonte – Direzione Istruzione, formazione e lavoro No. 218 of 2022, May 3 and s.m.i. which listed the higher institutions authorized to activate PhD positions in the apprenticeship format for the years 2022-2024 in the framework of a specific regional call for proposals (Apprendistato di Alta Formazione e Ricerca - Avviso Pubblico 2022-2026 per l'individuazione e la gestione dell'offerta formativa pubblica approvato con Determinazione 114 del 3/3/2022, modificato con D.D. n. 451 del 17/08/2022 e prorogato con D.D. n. 807 del 24/12/2024)

COMPUTER AND CONTROL ENGINEERING

Secure Artificial Intelligence

Company	ARUBA A.I. SRL. [P.iva/CF:13073110960]

Supervisor	CAGLIERO LUCA - luca.cagliero@polito.it

Contact	Francesco Tarasconi

Context of the research activity	The reserach project explores the role of Artificial Intelligence in optimizing IT infrastructure and online services while addressing security, privacy, and adaptability challenges. Key objectives include AI-driven predictive maintenance, anomaly detection, providing real-rime support, as well as fine-tuning domain-specific AI models, evaluating open-source vs. closed-source architectures, and developing secure, scalable AI frameworks for diverse industries. Format: The Company ARUBA AI SRL has planned for the winner of this position a collaboration within a contract of high apprenticeship according to the Italian Legislative Decree 81/2015, art. 45.

	Context: Artificial Intelligence (AI) is transforming IT infrastructure and online services by improving operational efficiency, cybersecurity, and resource optimization. However, as AI models become more sophisticated and widely adopted, new challenges arise regarding security, privacy, and computational resource management. The increasing reliance on cloud-based AI solutions and the widespread use of large language models (LLMs) necessitate a reevaluation of how AI systems are deployed and secured. A critical challenge is balancing AI performance with security and regulatory compliance. Many AI models operate on public cloud platforms, exposing sensitive data to potential breaches. Additionally, small and medium-sized enterprises (SMEs) often lack the resources to implement secure and

Objectives	customized AI solutions. Meanwhile, AI-driven software development is revolutionizing coding practices, overall reducing time-to-market, while the impact on the final quality must be evaluated closely across different applications. This research project aims to explore the role of Secure Artificial Intelligence in enhancing IT infrastructure and online services while addressing security, sustainability, and adaptability challenges. Research objectives: AI-Driven IT Infrastructure Optimization, including predictive maintenance, anomaly detection, automatic resource balancing.Development of Agentic AI or innovative approaches to integrate Generative AI in the diverse landscape of online services and tools. Exploration of new domain-specific Transformer models for industry-specific applications.Fine-tuning of pretrained large generative models on specific domains to improve accuracy and relevance for business use cases.Evaluation of the benefits and limitations of open-source vs. closed-source AI models and architectures across different industries.Developing frameworks to protect AI models from adversarial attacks and data poisoning and integrate them into the IT Infrastructure Optimization strategies.Development of scalable and modular AI frameworks to meet the needs of companies of different sizes. Tentative work plan: During the first year, the PhD candidate mainly studies and addresses the key challenges in IT infrastructures, identifying AI-based mitigation approaches, with particular attention to LLM-based agent solutions. The research envisages the adoption of autonomous agents that support the users, monitor the current situation, learn the characteristics of suboptimal usage patterns, and trigger ad hoc mitigation actions. Beyond the generative AI fundamentals, the PhD candidate will study and extend the state-of-the-art literature on anomaly detection, predictive maintenance, and optimized resource allocation. The second PhD year will be devoted to specializing Transformer- and LLM-based solutions to the use cases under study. Thanks to a more advanced knowledge of the limitations of general-purpose solutions, we envisage the development and testing of domain-specific approaches tailored to real business cases. The research will also involve the use of fine-tuning and reinforcement learning to optimize the level of AI specialization as well as requires the definition and adoption of quantitative and qualitative assessment procedures. The last year will explore the tight integration of IT infrastructure optimization modules with cybersecurity procedure and quality control modules. The aim is to develop and test comprehensive solutions for optimized IT infrastructure and online services, which leverage a global understanding of the main system and users' needs. List of possible publication venues - Conferences: ACL, EMNLP, ACM Multimedia, KDD, ACL, COLING, IEEE ICDM, ECML PKDD, ACM CIKM - Journals: IEEE TKDE, ACM TKDD, IEEE TAI, ACM TIST, IEEE/ACM TASLP, ACL TACL

Skills and competencies for the development of the activity	The candidate shall be less than 30 years old at the moment of the hiring from the company. The PhD candidate is expected to - Have the ability to critically analyze complex systems, model them and identify weaknesses; - be proficient in Python programming; - know data science fundamentals; - have a solid background on machine learning and deep learning; - have natural inclination for teamwork; - be proficient in English speaking, reading, and writing;

	- proficiency with Docker and Kubernetes software is a plus.