

# INTERNSHIP PROPOSAL

Data Science Students meet Companies 2022

November 30, 2022

# 1 PROPOSAL

---

## 1.1 WHO WE ARE

Engineering Group is the Digital Transformation Company, leader in Italy and expanding its global footprint, with around 12,000 associates and with over 60 offices.

The Engineering Group, consisting of over 20 companies in 12 countries, has been supporting the continuous evolution of companies and organizations for more than 40 years, thanks to a deep understanding of business processes in all market segments, fully leveraging the opportunities offered by advanced digital technologies and proprietary solutions.

It integrates best-of-breed market solutions, managed services, and continues to expand its expertise through M&As and partnerships with leading technology players. The Group strongly invests both in innovation, through its R&I division, and in human capital, with the internal IT & Management Academy. Engineering is a key player in the creation of digital ecosystems that bridge the gap between different markets, while developing composable solutions that ultimately foster a continuous Business transformation.

## 1.2 COE DATA & ANALYTICS

The Engineering Data & Analytics Center of Excellence consists of more than 200 professionals dedicated and specialized in all the issues that revolve around the analysis of data and ranging from management and manipulation on traditional tools (DB and ETL), to skills, enabling those capabilities of visual Data Discovery, on Self-Service BI tools, passing through the most innovative frontiers such as the management of unstructured data, linked to the Hadoop ecosystem, up to techniques and tools for complex data analysis that see the use of statistical, exploratory and predictive models, ontologies and semantic analysis, sentiment analysis, data and text mining, image processing, machine and deep learning, artificial intelligence.

## 1.3 INTERNSHIP PROPOSALS

Several internship topics are available at Advanced Analytics Group: most of them stem from current business activities, in order to get deeper insights about advanced analytical approaches. Therefore, further topics may be available, on the basis of new needs arising from future business projects.

A list of open opportunities can be find below:

- **Generation of synthetic data:** The training of deep learning models usually requires pretty large amount of data. This could be a challenge in many business scenarios: training phase may suffer of low quality data; information might be only partial. In this Internship activity the candidate will contribute to the development of a platform for the generation of synthetic data (in particular images) starting from:
  - tools to solve specific problems (e.g. parked cars)
  - diffusion models for the generation of targeted images addressed to business problems

More information about specific business cases will be given in the discussion of proposals.

**Robust classifiers:** Some business problems require the use of robust, self-extensible classification algorithms, which may combine supervised and unsupervised techniques, to face insufficient data availability and high content variability. The activity is aimed at apply specific, robust predictive algorithms, characterized by low data requirements in specific business problems. More information about specific business cases will be given in the discussion of proposals.

- **Emotion recognition:** the aim of the activity is to develop custom models for detecting the degree of emotionality expressed by a user, both from images and audio tracks.

**Graph Embeddings:** one of the most interesting challenges today is to leverage information structured on graphs to generate predictive analysis and conceptual inferences even in highly complex scenarios. The activity focuses on the exploration of some knowledge graph embeddings techniques, to solve specific business problems, in areas such as anomaly detection, inference, reasoning starting from unstructured data flows (in natural language). More information about specific business cases will be given in the discussion of proposals.

The desired outcome of Internship projects consists in either a standalone application, or a software component, integrated in our internal AI Laboratory. Code will be implemented in Python. A basic knowledge of Statistics, Machine Learning and Deep Learning principles is required. Knowledge of Microservices basics is a plus.