

## INDUSTRIAL PhD PROGRAMME PROFILE

**PhD Programme:** *"Big Data and Artificial Intelligence"*

**Cycle:** 38th

**Presumed start date of Programme:** 1 December 2022

**Duration:** 3 years, of which:

- ❖ 18 months of study and research at a company facility (spread over the three years of the programme);
- ❖ 6 months of study and research abroad.

**Coordinator:** Prof. Barbara Martini - SSD ING-INF/05

**Administrative Headquarters:** Piazza Mattei, 10 - Rome (Italy) 00186

**Scientific Disciplinary Fields:** FIS/01, ICAR/04, ICAR/05, ICAR/08, ING-IND 17, ING-IND/16, ING-IND/35, ING-INF/03, ING-INF/05, IUS 07, IUS/01, IUS/02, IUS/09, IUS/10, MED-01, M-GGR/02, SECS-P/07, SECS-P/10, SECS-S/03, SPS/08

**Programme webpage:** [www.unimercuratorum.it/ricerca/dottorati-di-ricerca/dottorato-industriale-in-big-data-ed-intelligenza-artificiale](http://www.unimercuratorum.it/ricerca/dottorati-di-ricerca/dottorato-industriale-in-big-data-ed-intelligenza-artificiale)

**Curricula:** YES

1. *Big data management for digital transition:* The term Data Management is extremely broad and encompasses not just technological but also organisational, process and compliance aspects. Data Management enables the valorisation of a strategic asset, the data capital, by putting it at the service of the organisation's development objectives. In order to seize new business opportunities in areas such as Industry 5.0, Economics and Marketing, Smart Cities and Smart Health, Data Science infrastructure, tools and skills are needed. Data-driven enterprises are those that consider data management not as a technical factor, but as a strategic business pillar. To be data-driven means to be guided by numbers, to have a data-based approach, to make informed decisions based on objective facts, not personal sensations. Transformation towards a data-driven company cannot therefore take place by way of technology alone, but through a change management process capable of introducing the culture of data at all company levels. Aside from this, representing the complexity of phenomena is also the new challenge for governance in an area. Each event can be measured against the environment surrounding the event, linking different heterogeneous sources. The focus is not just on what happens within a given space, within a precise location, but the concept of place becomes an element of mediation between various levels of information, which enables us to govern phenomena at the level of the urban space.
2. *Artificial Intelligence for Industry 5.0 and the Circular Economy:* The curriculum aims to seize the opportunities arising from the development of Artificial Intelligence in industrial settings, in order to pursue a new concept of society in which not only Artificial Intelligence, as well as technologies, such as IoT, robots, cybersecurity, cloud and high-speed networks, are actively employed to pursue an economic advantage, but above all to improve human living conditions and the environment. The curriculum aims to train experts who can contribute to the advancement of scientific knowledge through the integrated use of technological research methodologies on the one hand and statistical and socio-economic research on the other. Students are introduced to a training pathway that allows them to build up integrated knowledge and a "complex" vision of the ecosystem of applications, technologies and

## Annex 1.A

methodological solutions, capable of tackling problems according to a systemic and multi-disciplinary approach. In particular, it aims to train experts in Data Science and Artificial Intelligence and their practical applications in the areas of Industry 5.0 and the economy based on the principles of the Circular Economy, as a regenerative economic model. Combined training in data science and machine learning models, on the one hand, and advanced technologies, on the other, will be central to enabling the design and implementation of this model.

**Total available positions: 16** of which

- ❖ **2 positions** with scholarships funded by the University;
- ❖ **5 positions** with scholarships funded by Ministerial Decree no. 352/2022 - Curriculum 1:
  - **1 position** with a scholarship co-funded by *Angelini Holding SPA*;
  - **1 position** with a scholarship co-funded by *Borsa Merci Telematica Italiana SCPA*;
  - **1 position** with a scholarship co-funded by *Centro Studi delle Camere di Commercio Guglielmo Tagliacarne SRL*;
  - **1 position** with a scholarship co-funded by *InfoCamere SCPA*;
  - **1 position** with a scholarship co-funded by *PricewaterhouseCoopers Business Services SRL*;
- ❖ **5 positions** with scholarships funded by Ministerial Decree no. 352/2022 - Curriculum 2:
  - **1 position** with a scholarship co-funded by *Alilauro SPA*;
  - **1 position** with a scholarship co-funded by *Almaviva SPA*;
  - **1 position** with a scholarship co-funded by *Conerobus SPA*;
  - **1 position** with a scholarship co-financed by *Magnaghi Aeronautica SPA*;
  - **1 position** with a scholarship co-funded by *Romeo Gestioni SPA*;
- ❖ **4 positions** without scholarships.

**Admission qualifications:** “*laurea magistrale*” (master's degree) or the corresponding “*laurea specialistica*” (specialised degrees) or the corresponding “*laurea vecchio ordinamento*” (ordinary degree awarded according to the regulations in force prior to Ministerial Decree 509/99) or the corresponding equivalent qualifications.

**Admission procedure:** Admission will be based on an evaluation of the applicant’s qualifications and research project and the results of an oral test. During the oral test, the research project submitted by the applicant at the time of application will also be discussed and knowledge of English as a foreign language tested. Evaluation will take place in accordance with Article 6 of the call for applications.

**Procedure and schedule of tests:** will be published on the programme webpage:  
[www.unimercatorum.it/ricerca/dottorati-di-ricerca/dottorato-industriale-in-big-data-ed-intelligenza-artificiale](http://www.unimercatorum.it/ricerca/dottorati-di-ricerca/dottorato-industriale-in-big-data-ed-intelligenza-artificiale)

**Useful contacts:** [dottorati@unimercatorum.it](mailto:dottorati@unimercatorum.it)

## Scientific and educational objective of the enterprises co-funding the scholarships, pursuant to DM 352/2022 - Curriculum 1

### **Enterprise:** *Angelini Holding SPA*

DATA DRIVEN DECISION MAKING IN HUMAN RESOURCE MANAGEMENT: Following the explosion of web technologies, enterprises are adopting new strategies for dealing with their employees. On the one hand enterprises are interested in capturing the needs and expectations of employees, on the other they are engaged in introducing new technologies, in-house, which can improve the processes of communication, engagement, attraction and retention of talent. The ultimate goal is to improve job satisfaction among human resources, which inevitably leads to improved individual and company-wide performance. The analysis of Big Data derived from human resources management systems, such as performance management, learning & development management, reward & compensation, such as employees' questions on company forums/chatrooms, information exchanges and communications via Facebook, Twitter, Telegram or WhatsApp, are indicators of the technological and communicative evolution of enterprises. The objective of this scientific and training project is to build models with and integrate the data sets of HR practices, which today are kept isolated, through links and correlations that can be usefully used in responding to the great challenge of rethinking the human resources management model. Through the analysis of Big Data, using already known algorithms, such as textual analysis, or implementing new ones, the project aims to quantify the correlation between performance and engagement, performance and learning attitude, attraction and welfare programmes, retention and total reward models.

### **Enterprise:** *Borsa Merci Telematica Italiana SCPA*

The PhD Programme in Big Data (Data Science) and Artificial Intelligence trains researchers with a scientific background both in the field of data science (Data Science), aimed at deepening methodological aspects, such as statistical and computational data analysis, big data management and machine learning algorithms, and oriented towards a theoretical and applicative knowledge of artificial intelligence, with particular reference to computer science and applied mathematics, capable of combining strong analytical skills and algorithmic developments. Big Data is everywhere and influences the strategic decisions of enterprises. The term "Big Data" literally refers to the vast amount of data and information that is collected and managed on a daily basis by enterprises or entities. Rather than on the magnitude of this data, however, the focus is on how it can be used, i.e. how to analyse it in order to mine important information for all those enterprises dealing with statistics and market analysis. Big Data, in fact, would be valueless if it were not possible to analyse it to extract information: machines, compared to humans, have the ability to analyse a large amount of information in a very short time.

In other words, simply having access to large data sets is not sufficient to produce a result. The convergence between big data and artificial intelligence is therefore also inevitable to make business processes more agile, faster and smarter. The production process of big data is yet to be fully understood and concerns not only the disciplinary boundaries between statistics, computer science, social sciences and economics for the creation of an interdisciplinary data science. What is needed is appropriate statistical literacy, which must also extend to include data from the new communication tools, for a correct interpretation of the data. Indeed, the analysis of Big data must take into account the quality of the data in order to avoid erroneous and distorted representations of reality. The programme will provide cross-sectoral skills to be applied to different contexts from an interdisciplinary perspective, with the aim of creating experts who can contribute to the advancement of new scientific knowledge for both developing methods and models for interpreting data and providing answers to new frontier areas of research in contemporary science and society.

The availability of large quantities of data can foster the convergence of different disciplines, for the development of models and algorithms capable of explaining the complexity of social, economic, biological, natural, technological and cultural phenomena in greater depth. In particular, the use of Big Data – combined

## Annex 1.A

with the inherent algorithms of Artificial Intelligence – enables a range of applications in the following fields: Economy and Society; Industry 5.0; Advanced Digital Platforms; Data Driven Science; Spatial Big Data; Circular Economy.

**Enterprise:** *Centro Studi delle Camere di Commercio Guglielmo Tagliacarne s.r.l.*

**ANALYSIS OF NEW BUSINESS STRATEGIES THROUGH BIG DATA:** Following the explosion of web technologies, enterprises are adopting new communication strategies with their customers and suppliers. On the one hand, enterprises are engaged in capturing the needs and expectations of consumers, and on the other they are committed to introducing new technologies, in-house, to improve their production processes and products to better capture consumer needs. The evolution of web-based business-to-consumer (B-to-C) and business-to-business (B-to-B) communications is the result of this widespread use of new technologies. The analysis of Big Data derived from web-based communications, such as consumer queries on Google or other search engines, information exchanges and communications via Facebook, Twitter, Telegram or WhatsApp, point to the technological and communication-based evolution of businesses. The objective of this scientific and educational project is to quantify the number of enterprises that adopt these new tools by determining their purpose. The project aims, through the analysis of Big Data, using already known algorithms, such as text analysis, or implementing new ones, to quantify the number of enterprises present on the web with their own site or with the simple use of the best-known communication apps, differentiating them according to the main purpose of the communication (input for new products, product returns, product improvement, assistance, etc.).

**Enterprise:** *InfoCamere SCPA*

**ANALYSIS OF NEW STRATEGIES FOR AUTOMATING BACK OFFICE AND FRONT OFFICE PROCESSES THROUGH BIG DATA AND ARTIFICIAL INTELLIGENCE**

The evolution of web, social and mobile technologies in recent years has irreversibly conditioned the management of business processes, not only in the production of material goods, thanks also to the drive of Industry 4.0, but in the services sector as well. In particular, repetitive activities, such as the handling of standard paperwork by public authorities or answering FAQs from users, have already been widely automated thanks to machine learning and data analysis technologies. Innovative start-ups and research projects have helped introduce, in both large enterprises and SMEs, virtuous experiences of Robotic Process Automation (RPA) and Hyperautomation, i.e. the robotization of back- and front-office processes, transforming operations traditionally carried out by human operators into robot-guided or mixed-operation processes. These solutions have become especially important during the Covid-19 pandemic by ensuring operational continuity at a time when operators were required to observe physical distancing rules.

The objective of this scientific and training project is to identify and analyse real experiences, both in the public and private sectors, in which Big Data analysis and the use of artificial intelligence have been key factors in the reconfiguration of back office and front office processes (e.g. in customer care), as well as to study the extent of the resulting digital transformation, in terms of its effects and benefits on service quality and user satisfaction (reduction of errors, improvement of time to market, cost reduction, efficiency, operational continuity, organisational resilience, etc.).

**Enterprise:** *PricewaterhouseCoopers Business Services SRL*

**DEEP LEARNING FOR SIGNAL AND VIDEO PROCESSING AND POSSIBLE ON EDGE APPLICATIONS.**

The analysis of video sources represents a major challenge in today's society. Suffice it to think of the dynamics introduced by the Covid-19 pandemic, and the possibility of being able to check compliance with mask-wearing or social distancing mandates without direct human intervention in vast areas through video



## Annex 1.A

surveillance. The analysis of video recordings of enclosed spaces (construction sites, trains, commercial buildings) can ensure compliance with safety regulations (for example, through the recognition of personal protective equipment on authorised personnel), spot the occurrence of anomalies (such as crossing through a danger zone at a construction site) and the implementation of quality controls (such as the observation of semi-finished products in a production chain). Furthermore, the automatic classification of online video content has applications in recommendation engines and marketing strategies, moderating sensitive content and creating enriched archives in enterprise environments. Today, Artificial Intelligence can automate all these activities, with the possibility of introducing a competitive advantage for enterprises in terms of reducing costs and human resources required for such lower value-added activities as continuous monitoring, video surveillance and compliance with security regulations, resulting in the investment of resources in more complex and challenging objectives. Furthermore, its application in the marketing context and to online content can significantly enrich the information assets associated with unstructured information and enable commercial strategies with a considerable impact. Last but not least, technological evolution enables real-time and on-edge applications of these solutions. In particular, on-edge inference refers to the possibility of executing these algorithms directly on the instruments acquiring the data (cameras, camcorders). Such a possibility significantly enhances interest in these applications, due to their simple execution and frugal architecture. The scientific objective is the study of Artificial Intelligence algorithms that represent the state of the art in Computer Vision for the aforementioned applications, with a special focus on on-edge inference.

## Scientific and educational objective of the enterprises co-funding the scholarships, pursuant to DM 352/2022 - Curriculum 2

### **Enterprise:** *Alilauro SPA*

#### BIG DATA ANALYSIS FOR RECONSTRUCTING MOBILITY PATTERNS

The PhD Programme in Big Data (Data Science) and Artificial Intelligence trains researchers with a scientific background both in the field of data science (Data Science), aimed at deepening methodological aspects, such as statistical and computational data analysis, big data management and machine learning algorithms, and oriented towards a theoretical and applicative knowledge of artificial intelligence, with particular reference to computer science and applied mathematics, capable of combining strong analytical skills and algorithmic developments. Big Data is everywhere and influences the strategic decisions of enterprises. The term "Big Data" literally refers to the vast amount of data and information that is collected and managed on a daily basis by enterprises or entities. Rather than on the magnitude of this data, however, the focus is on how it can be used, i.e. how to analyse it in order to mine important information for all those enterprises dealing with statistics and market analysis. Big Data, in fact, would be valueless if it were not possible to analyse it to extract information: machines, compared to humans, have the ability to analyse a large amount of information in a very short time.

In other words, simply having access to large data sets is not sufficient to produce a result. The convergence between big data and artificial intelligence is therefore also inevitable to make business processes more agile, faster and smarter. The production process of big data is yet to be fully understood and concerns not only the disciplinary boundaries between statistics, computer science, social sciences and economics for the creation of an interdisciplinary data science. What is needed is appropriate statistical literacy, which must also extend to include data from the new communication tools, for a correct interpretation of the data. Indeed, the analysis of Big data must take into account the quality of the data in order to avoid erroneous and distorted representations of reality. The programme will provide cross-sectoral skills to be applied to different contexts from an interdisciplinary perspective, with the aim of creating experts who can contribute to the advancement of new scientific knowledge for both developing methods and models for interpreting data and providing answers to new frontier areas of research in contemporary science and society.

The availability of large quantities of data can foster the convergence of different disciplines, for the development of models and algorithms capable of explaining the complexity of social, economic, biological, natural, technological and cultural phenomena in greater depth. In particular, the use of Big Data - combined with the inherent algorithms of Artificial Intelligence - enables a range of applications in the following fields: Economy and Society; Industry 5.0; Advanced Digital Platforms; Data Driven Science; Spatial Big Data; Circular Economy.

### **Enterprise:** *Almaviva SPA*

Big Data is everywhere and influences the strategic decisions of enterprises. The term "Big Data" literally refers to the vast amount of data and information that is collected and managed on a daily basis by enterprises or entities. With the explosion of web technologies, enterprises are adopting new strategies for communicating with customers and suppliers. On the one hand, enterprises are engaged in capturing consumer needs and expectations, and on the other they are committed to introducing new technologies, in-house, that can improve production processes and products to better capture consumer needs. It is therefore interesting to explore how this data can be analysed, in relation to topics such as intelligence, safety, connectivity and reliability for buses, metropolitan railways, trains and ships. The aim of this scientific and educational project is to integrate new technologies, such as 5G, Artificial Intelligence and interconnection with the IoT world, into the study of this data. The proposed innovations further target the needs of operators, drivers and passengers, addressing new mobility trends such as big data analysis to support planning, predictive diagnostics and increasingly adaptive service regulation. Special attention is paid to safety, which is increasingly perceived as a key component of service reliability.

**Enterprise:** *Conerobus SPA*

Digital skills are a key strategic asset for social and economic growth. Following the explosion of web technologies, businesses are adopting new strategies for communicating with customers and suppliers, the population is becoming increasingly digitally aware, public administrations and businesses are being prompted to improve the services they provide, and the education system is engaged in organising itself to cover the needs of digital skills development. In fact, on the one hand, enterprises are engaged in capturing consumer needs and expectations, and on the other they are committed to introducing new technologies, in-house, that can improve production processes and products to better capture consumer needs. It is therefore interesting to explore how topics such as 5G, Artificial Intelligence and interconnection with the IoT world can be analysed using new technologies and algorithms based on Big Data.

The Conerobus project is part of this broader framework, through Big Data analysis and the use of intelligent tools, through its commitment to improving transport safety, connectivity and reliability for buses, metropolitan railways and trains. In fact, the proposed innovations further target the needs of operators, drivers and passengers, addressing new mobility trends (such as Big Data analysis) to support planning, predictive diagnostics and increasingly adaptive service regulation. There is a special focus on safety, which is increasingly perceived as a key component of service reliability. Satellite localisation, voice and data communication between control centre and vehicle, integration and management of all on-board systems, passenger counting and occupancy assessment, as well as information display management, are also features that make public transport a truly smart on-board control centre.

**Enterprise:** *Magnaghi Aeronautica SPA*

THE VALUE OF DIVERSITY

Research and analysis of materiality related to diversity and inclusion within AI systems; implementation of functional strategies for the development and dissemination of inclusive cultures with a view to engagement and retention. So-called algorithm "biases" in AI can harm people, especially minorities, which is why various solutions to remedy the problem are also being studied. Solving the algorithm bias, i.e. the lack of fairness that emerges from the output of an IT system, is one of the challenges that researchers and institutions are trying to solve at the moment, to prevent the spread of artificial intelligence from broadening existing areas of discrimination and inequality, especially to the detriment of ethnic minorities, women, and the poor. The lack of fairness described in the bias inherent in the algorithm comes in various forms but can be summarised as the discrimination of a group of people on the basis of a specific categorical distinction. This bias can take various "forms", such as: racial discrimination, age discrimination, gender discrimination, etc. In seeking a solution to this problem, a step forward would be the inclusion of the concept of diversity in the initial stages of any process or project within Artificial Intelligence.

**Enterprise:** *Romeo Gestioni SPA*

APPLICATION OF AN ARTIFICIAL INTELLIGENCE SYSTEM TO REAL ESTATE USER COMMUNICATIONS.

Romeo Gestioni S.p.A. manages a Multi-channel Contact Centre, through its Operations Centre, into which flow - *inter alia* - email reports from thousands of users of managed real estate assets, as well as communications from clients, other entities (municipalities, ASL, local agency branches, etc.), public safety bodies, third parties, etc.. The communications in question are analysed by a team of professionals, thoroughly analysed and recorded in the Romeo Gestioni S.p.A. Information System, linking them to the management data in the company's Oracle database. The need to interpret communications in a non-linear and structural



## Annex 1.A

manner according to sender, described phenomena, references to addresses and building structures that are not perfectly declined, etc., constitutes a major limitation in the automation of the activities required to enter and record incoming reports to the Operations Centre. Through the creation of learning models, based on the communications received to date and relating to the Management of the Available Property Assets of the INPS, aims to create an IA system capable of analysing incoming communications and automatically records intervention requests, contextualising them within the framework of the property assets, their nature and the relative sender. In addition to the actual IA system, the use of pattern and text recognition systems is envisaged for mining content and documents from email communications, which may also encompass several levels of content enveloping (as in the case of digitally signed documentation in CADES format). The IA System will be required to provide a mechanism for operator-driven verification and validation of the proposed classifications and a self-learning module according to the corrections indicated by operators. The project ultimately aims to improve and improve the timeliness of recording of the operations transmitted to an economic operator engaged in the management of real estate assets distributed among and used by a broad spectrum of subjects.