



# PhD in INGEGNERIA DELL'INFORMAZIONE / INFORMATION TECHNOLOGY - 39th cycle

**Research Area n. 1 - Computer Science and Engineering**

**PNRR 117 Research Field: KNOWLEDGE INFUSION IN MACHINE LEARNING APPROACHES  
FOR DATA INTEGRATION AND EXPLAINABILITY IN LARGE INDUSTRIES**

**Monthly net income of PhDscholarship (max 36 months)**

**€ 1400.0**

In case of a change of the welfare rates during the three-year period, the amount could be modified.

## Context of the research activity

**Motivation and objectives of the research  
in this field**

The Italian National Plan for Recovery and Resilience emphasizes the development of new added-value services and business models based on satellite technologies, targeting different industries and a broad set of application domains. In an aerospace market that requires innovative product development in a challenging time-to-market, reaching agility to define models and long-term scenarios is mandatory. In the current scenario, modern industries have a large amount of data that needs to be analyzed and interpreted effectively. In the environment of Hardware Engineering for The Italian National Plan for Recovery and Resilience emphasizes the development of new added-value services and business models based on satellite technologies, targeting different industries and a broad set of application domains. In an aerospace market that requires innovative product development in a challenging time-to-market, reaching agility to define models and long-term scenarios is mandatory. In the current scenario, modern industries have a large amount of data that needs to be analyzed and interpreted effectively. In the environment of Hardware Engineering for Satellite Industries services, either directed to institutional or business organizations, analytics solutions have been traditionally developed through data collection processes embedded into



	<p>information systems and business intelligence applications. Despite ensuring traceability and quality requirements, this approach does not allow for catching the opportunity to quickly and automatically find correlations between data and kick off advanced analytics practices. One of the root causes is the extremely high variability of customer products, which includes many different requirements and innovative technologies, making it difficult to standardize the hardware in a model. In future Constellations Programs, Interplanetary Missions, and Planetary Defense missions, Thales Alenia Space (TAS) expects advanced analytics technologies to play a vital role in improving efficiency and reducing time-to-market. Consider that Satellite electronic equipment typically comprises several standard modules (like DCDC Converter, Transmitter/Receiver modules, etc..) and custom modules specific to the mission (like Digital modules, Processors, and Firmware based payload). Currently, TAS rely on different project and portfolio management applications, such as Primavera P6, Microsoft Project, Jira, and Kantree, which help it manage and monitor its projects' progress. However, these applications generate a vast amount of data that is often stored in silos and can be challenging to integrate and analyze. Moreover, TAS perceives the need to also integrate planning and monitoring data together with real-time data streams from Production Environment through Industrial IoT, and Verification &amp; Validation tests. The ultimate goal is to provide users with updated analyses and easily consultable insights.</p>
<p><b>Methods and techniques that will be developed and used to carry out the research</b></p>	<p>The main objective of this Ph.D. program is to develop and validate models, methods, and implementations that inject semantic knowledge (in the form of knowledge graphs and/or semi-structured knowledge extracted from product documentation and user-generated content) into machine learning processes and pipelines that integrate historical data from different project &amp; portfolio management applications and data streams from the Production Environment through Industrial IoT, and Verification &amp; Validation tests. This Ph.D. program will cover the following phases towards its final the research</p>



objectives:

- To review the literature on knowledge extraction and injection, project & portfolio management data structures, IoT data streams, Verification & Validation test reports, data integration, and knowledge graphs both from a historical and a streaming perspective.

To design and develop a knowledge infusion and integration methodology specifically targeting a diverse set of sources, spanning streaming data, The main objective of this Ph.D. program is to develop and validate models, methods, and implementations that inject semantic knowledge (in the form of knowledge graphs and/or semi-structured knowledge extracted from product documentation and user-generated content) into machine learning processes and pipelines that integrate historical data from different project & portfolio management applications and data streams from the Production Environment through Industrial IoT, and Verification & Validation tests. This Ph.D. program will cover the following phases towards its final the research objectives:

- To review the literature on knowledge extraction and injection, project & portfolio management data structures, IoT data streams, Verification & Validation test reports, data integration, and knowledge graphs both from a historical and a streaming perspective.
- To design and develop a knowledge infusion and integration methodology specifically targeting a diverse set of sources, spanning streaming data, historic data, formalized knowledge bases and informal knowledge.
- To devise the best configurations of machine learning solutions that benefit of the infused knowledge to maximise the performance of the models with respect to the baseline data-driven only solutions.
- To evaluate the effectiveness and efficiency of the proposed methods through experiments, case studies, and user feedback.
- To propose guidelines for implementing industrial deployments of the proposed approaches.
- To feed the process with data available in Thales Alenia Space database.



	Alenia Space database.
<b>Educational objectives</b>	<p>The main educational objectives are:</p> <ul style="list-style-type: none"> <li>• Consolidate background on knowledge infusion, project and portfolio management, Industrial IoT, Verification and Validation, data integration, and knowledge graphs.</li> <li>• Develop theoretical frameworks for the scenario described in the objectives.</li> <li>• Develop experimental know-how in the field.</li> <li>• Integrated multidisciplinary skills aiming at supporting and implementing the solution in the Space Industry application scenario.</li> </ul> <p>Achieve soft skills particularly related to technology transfer and entrepreneurship.</p>
<b>Job opportunities</b>	<p>After completing this Ph.D. program, the candidate can pursue various job opportunities spanning multiple industries and roles.</p> <ul style="list-style-type: none"> <li>• Data Architect who is responsible for designing and maintaining an organization's data architecture, ensuring that data is stored efficiently, easily accessed, and integrated.</li> <li>• Data Integration Specialist who focuses on integrating data from different sources and ensuring that the data is accurate and consistent.</li> <li>• Data Management Consultant who advises organizations on how to manage and integrate data effectively to support their business objectives. •</li> <li>• Knowledge Management expert who is specialized in developing, maintaining and integrating knowledge graphs in broader data-centric projects in organizations.</li> <li>• Business Intelligence Analyst who is responsible for gathering and analyzing data from various sources to provide insights and support decision-making.</li> <li>• Research Scientist who develops new knowledge graph technologies and applies them to solve complex</li> </ul>



	data integration problems.
<b>Composition of the research group</b>	1 Full Professors 1 Associated Professors 1 Assistant Professors 2 PhD Students
<b>Name of the research directors</b>	Emanuele Della Valle and Marco Brambilla

<b>Contacts</b>	
emanuele.dellavalle@polimi.it +393389375810 <a href="http://emanueledellavalle.org/">http://emanueledellavalle.org/</a>	

<b>Additional support - Financial aid per PhD student per year (gross amount)</b>	
<b>Housing - Foreign Students</b>	--
<b>Housing - Out-of-town residents (more than 80Km out of Milano)</b>	--

<b>Scholarship Increase for a period abroad</b>	
<b>Amount monthly</b>	700.0 €
<b>By number of months</b>	6

<b>National Operational Program for Research and Innovation</b>	
<b>Company where the candidate will attend the stage (name and brief description)</b>	Thales Alenia Space Italia S.p.A
<b>By number of months at the company</b>	6
<b>Institution or company where the candidate will spend the period abroad (name and brief description)</b>	Thales Alenia Space Italia S.p.A Francia
<b>By number of months abroad</b>	6

<b>Additional information: educational activity, teaching assistantship, computer availability, desk availability, any other information</b>
<p>EDUCATIONAL ACTIVITIES (purchase of study books and material, including computers, funding for participation in courses, summer schools, workshops and conferences): financial aid per PhD student.</p> <p>TEACHING ASSISTANTSHIP: availability of funding in recognition of supporting teaching activities by the PhD student There are various forms of financial aid for activities of support to the teaching practice. The PhD student is encouraged to take part in these activities, within the limits allowed by the regulations.</p> <p>COMPUTER AVAILABILITY: individual use.</p>



DESK AVAILABILITY: individual use.