# PhD in Computer and Control Engineering

## Research Title: CityScience: Data science for smart cities

<table>
<thead>
<tr>
<th>Funded by</th>
<th>Comitato ICT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Supervisor</strong></td>
<td>Silvia Chiusano - <a href="mailto:silvia.chiusano@polito.it">silvia.chiusano@polito.it</a></td>
</tr>
<tr>
<td><strong>Contact</strong></td>
<td><a href="http://dbdmg.polito.it/">http://dbdmg.polito.it/</a></td>
</tr>
</tbody>
</table>

### Context of the research activity

A smart city is an urban environment in which social and environmental resources are combined to make cities more livable by increasing urban prosperity and competitiveness. From a data science perspective, smart data emerging from smart cities give rise to a lot of challenges that constitute a new inter-disciplinary field of research.

On the one side, the establishment of innovative technologies related to mobile or wearable computing and smart city infrastructure led to the continuous massive generation of heterogeneous data. On the other side, due the search for better-informed decisions by urban stakeholders, there is a natural tendency to collect, process, and analyse these data by transforming them into information and actionable insights.

Urban computing thus entails the acquisition, integration, and analysis of big and heterogeneous data collections generated by a diversity of sources in urban spaces to profile the different facets and issues of the city environment.

### Objectives

The objective of the PhD is the study and development of proper solutions for the collection, management, and analysis of huge volumes of heterogeneous urban data, to finally extract useful insights for increasing the efficiency,
accessibility and functionality of offered services, and finally the well-being of citizens.

The analysis of urban data collections to gain actionable insights results in various folds of challenges in the context of data analytics, that will be addressed during the PhD, such as:

- Definition of cross-domain data fusion methods. To fully characterize the urban environment heterogeneous data sources should considered and integrated. For example, to predict the air quality in a city area different data sources, consisting of traffic, meteorology, Point Of Interests (POIs), road networks, should be considered. Since these data are usually collected with different spatial and temporal granularities, suitable data fusion techniques should be devised to support the data integration phase, and provide a spatio-temporal alignment of collected data.

- Adoption of no relational data models. The storage of heterogeneous urban data collections requires the use of alternative data representations to the relational model such as NoSQL databases (e.g., MongoDB).

- Adoption of technologies and design of algorithms for big data analytics. The capability of smart cities to generate and collect data of public interest has increased at an unprecedented rate, to such an extent that data can rapidly scales towards Urban Big Data. Consequently, the analysis of this massive volume of data demands distributed processing algorithms and technologies (e.g., Apache Spark).

- Design of machine learning algorithms to deal with spatial and spatio-temporal data. Urban data is usually characterized by spatio-temporal coordinates describing when and where data has been acquired. Spatio-temporal data has unique properties, consisting of spatial distance, spatial hierarchy, temporal smoothness, period and trend, which entails the design of suitable data analytics methods.

Different smart city scenarios will be considered as reference case studies such as urban mobility, citizen-centric contexts, and healthy city. More urban scenarios will be considered with the aim of exploring the different facets of urban data and evaluating how the proposed solutions perform on different data collections. The objectives of the research activity consist in identifying the peculiar characteristics and challenges of each considered application domain and devise novel solutions for the management and analysis of urban data for each domain.
| Skills and competencies for the development of the activity | The research activity involves multidisciplinary knowledge and skills such as database, machine learning and advanced programming. Knowledge of data analytics, machine learning methods is desirable. |