ICT PHD

Research project for a PhD curriculum in ICT – Computer Engineering and Science

**Tutor**: Laura Po

**(\*) Italian [Industrial] Co-tutor:**

 **(\*\*) Foreign Co-tutor:**

**Proposed Title of the research:**

*Big Data Analytics in a Sustainable Smart City*

**Keywords: (5)**

Big Data Analysis and Integration

Sensor Data Management and Outlier Detection

HPC technologies

Geographical and Visual Data Exploration

**Research objectives: --(max 10 rows)**

This proposal is focused on the management and analysis of huge amount of heterogeneous urban data streams.

This proposal aims to couple the efforts from several large research areas (semantics, statistics and data integration) for developing techniques and tools for integrating, analyzing and correlating traffic data and air quality data.

These techniques will be oriented to the development of a methodology for the creation of a large dataset from private and public datasets, real-time sensor data, available on the web or within the urban context.

This includes the study of techniques for

* Data series management
* Data profiling
* Dimension reduction for sensor data
* Data integration
* Data cleaning
* Data analysis
* Statistical techniques to identify outliers and missing data

The dataset will be complemented with a visualization engine allowing people to retrieve geographical data of interest from it.

**Proposed research activity -- (max 10 rows)**

The research activity is carried out in the TRAFAIR project, a project funded by European Commission (CEF TELECOM) that aims at defining a standard set of metadata able to represent urban air quality maps, at providing real-time estimation on air pollution in the city on a urban scale, development of a service for prediction of urban air quality based on weather forecast and traffic flows, at realising an open dataset describing the urban air quality maps and the prediction maps in 6 European cities of different size.

The project is led by UNIMORE and involved other 8 partners: the City of Modena, the Regional Company LepidaSpA, the University of Florence, the Tuscany Region, the University of Santiago de Compostela, the City of Santiago de Compostela, the Supercomputing Center of Galicia (CESGA), the University of Zaragoza.

The main activities can include:

• Survey and Analysis of the available/relevant data (sources of pollution, weather forecast, traffic data etc.)

• Analysis and implementation of techniques for the population of a unified Smart City Dataset

• Urban sensor data analysis, management, and querying

• Analysis of the existing best practices, standards, and services for releasing air quality data, metadata definition and dataset population for air quality data

• Analysis and use of HPC technologies to store the different source of data and to elaborate statistics and simulations on scalable datasets

• Development of policies and tools for the effective monitoring of the current situation of pollutants and traffic flows within the city

• Development of a tool for the effective display of urban air pollution maps

In particular, these activities will be implemented in the urban context of the city of Modena and compared to the other 5 urban areas.

**Supporting research projects (and Department)**

"TRAFAIR - Understanding Traffic Flows to Improve Air quality" project (2017-EU-IA-0167) funded under the CEF Telecom Call 2017-3 on Public Open Data by the European Commission

**Possible connections with research groups, companies, universities.**

Università degli Studi di Firenze (prof. Paolo Nesi)

Universidade de Santiago de Compostela (Prof. José Ramón Ríos Viqueira );

Universidad de Zaragoza (Spain – Prof. Raquel Trillo Lado )

Fundación Centro Tecnológico de Supercomputación de Galicia, CESGA (Spain- Dr. Ignacio López Cabido)

LepidaSpA (Regional company)

 (\*) optional

(\*\*) optional/to be completed on the second year