



Social Mining &
Big Data Ecosystem
H2020 - www.sobigdata.eu
September 2015- August 2019

Roberto Trasarti
roberto.trasarti@isti.cnr.it



Objective

A **Multidisciplinary European Infrastructure for Big Data and Social Data Mining** providing an integrated ecosystem for **ethically sensitive scientific discoveries** and advanced applications of social data mining on the various dimensions of social life, as recorded by “big data”.



The Consortium

Italy

United Kingdom

Germany

Estonia

Finland

Switzerland

Nederlands



The
University
Of
Sheffield.



UNIVERSITÀ DI PISA



Fraunhofer
FHR



TARTU ÜLIKOOOL



INSTITUTE
FOR ADVANCED
STUDIES
LUCCA



Leibniz
Universität
Hannover

KING'S
College
LONDON



SCUOLA
NORMALE
SUPERIORE



Aalto University

ETH *Zürich*



TU Delft Delft
University of
Technology

Existing national RI's to be integrated



general architecture
for text engineering



Fraunhofer
Toscana TECH 2017 IGD



Research Center

nervousnet



Stakeholders

- **Big data analysts and social informatics researchers**, who want to enhance their algorithms to deal with social data, gain multi-disciplinary research skills, harmonise existing data and analytics infrastructures, and engage other research communities in the development of these key enabling technologies for the future digital economy and society;
- **Economists, social science and humanities researchers, journalists, policy and law makers**, who have to analyse the avalanche of (big) social data, in order to gain insight and actionable knowledge.
- **Researchers in related communities**, who would like to use the algorithms, the analytical competences and data infrastructure;
- **Industrial innovators & startupper**s, who would like to create rapid proof-of-concepts of data-driven innovative ideas and services;
- **The public as a whole**, who would like to understand their role in the production, consumption and value-creating of social data.



Exploratories

1

City of Citizens

Smart City, Mobility Data Analysis, Transportation system, Multi-modal planning, Car-Pooling, etc.

2

Societal Debate

Social Media analysis, newspaper articles, social debate, text analysis, sentiment analysis, etc.

3

Well being & Economy

Economic indicators, health, diffusion model, cost of life, crisis prediction, etc.

4

Migration Studies

Migration flows, sentiment analysis, complex network analysis, etc.



Accesses to SoBigData infrastructure

[Exploratories](#)[Access](#)[Training](#)[Dissemination](#)[Consortium](#)[EU Project](#)[Blog](#)

VIRTUAL ACCESS



Enter the Virtual Research Infrastructure and start your journey

[READ MORE](#)

TRANSNATIONAL ACCESS

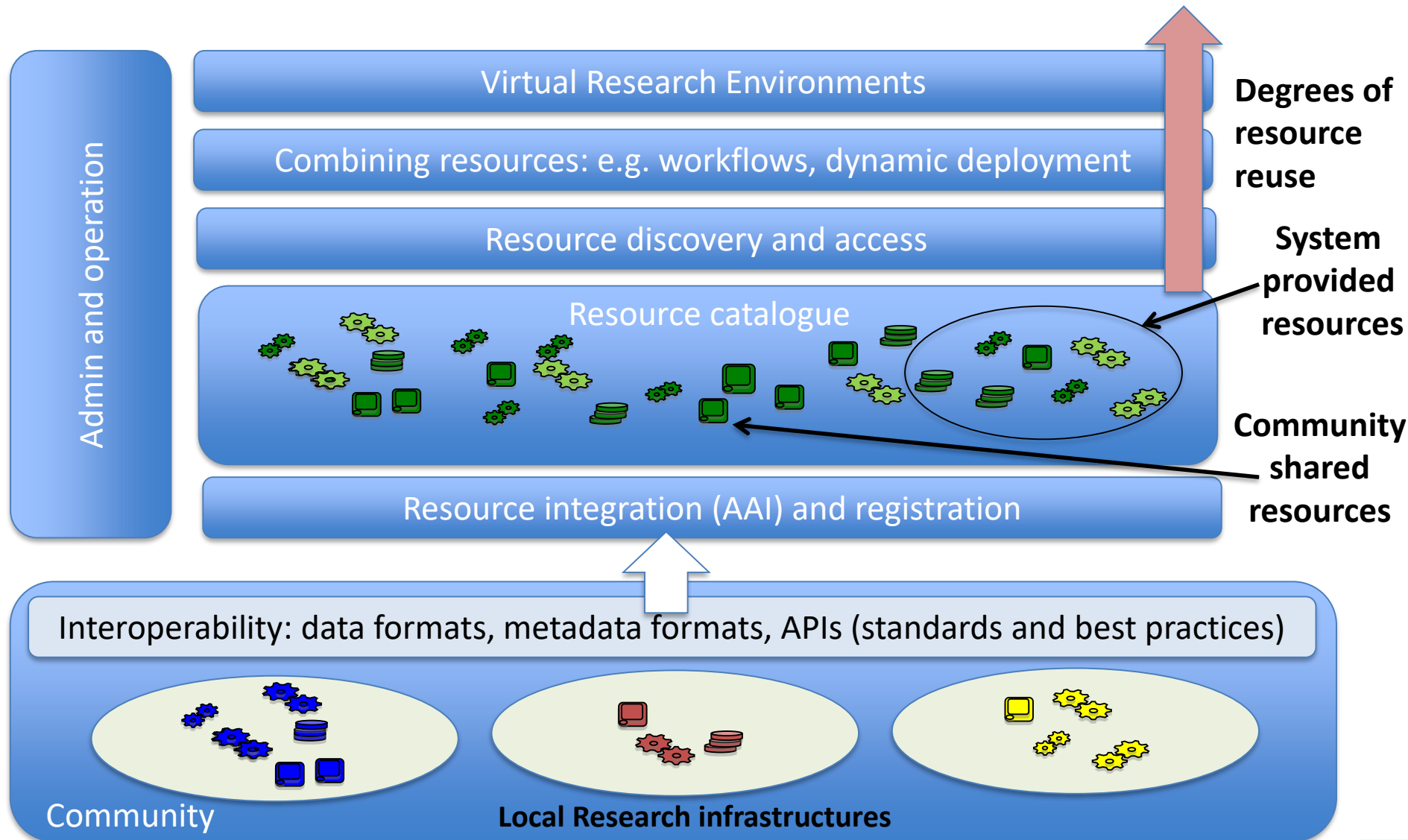


Come to visit us to make your research on big data. See Transnational Calls open.

[READ MORE](#)



E-Infrastructure Architecture





SoBigData.eu VREs



[Exploratories](#) [Access](#) [Training](#) [Dissemination](#) [Consortium](#) [EU Project](#) [Blog](#) [Q](#)



Go to the **SoBigData LAB** to perform your experiments



Go to VRE **Exploratories** to see the thematic environments of SoBigData



Go to the **Catalogue** to have a list of our methods and datasets



Go to SoBigData **Virtual Research Environments - VREs**



Exploratory VRE

SoBigData
City of Citizens

City of Citizens Administration Members Catalogue Story 1: Investigating City Mobility

Home / Groups / City Of Citizens

SoBigData Products Activity Stream About

Investigating City Mobility

The idea of the story is to produce a comprehensive set of analyses able to produce an overview of the city and the people living in it. In particular the city will be described by a set of basic and complex statistics such as: Incoming and outgoing traffic, different access points, distribution in space and time of the traffic, systematic vs occasional traffic, distribution of the radius of gyration and distribution of different types of users in the city. Those statistics will be generated on different cities but also on partitions of city area according to the usage of it. A predictive tool will be used to forecast the traffic 20 minute in advance.

Select one city on the map to visualize the available services and applications

Urban Mobility Atlas
An overview of mobility of a city by means of a set of visual indicators

Trip Builder

A tool to generate personalized tours of the city. Useful for the users and city managers who want to build tourist guides according to specific preferences.

Car Pooling

Analyzing Users behaviour allow this tool to produce recommendation for a car pooling service as well as indicating the potential impact of the service in a city

Carpooling Network Analysis

Analyzing carpooling network it is possible to discover the existence of drivers/passengers micro-communities

Mobility Profiles

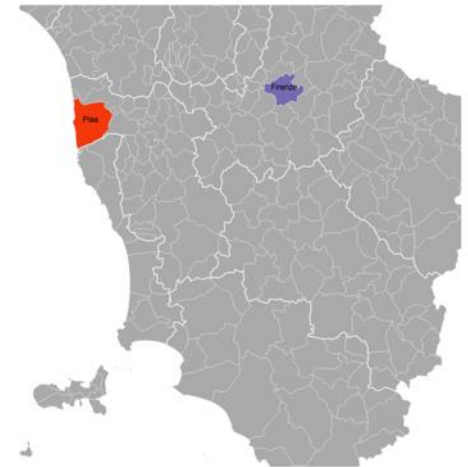
Understanding the systematic behavior of the user is the key to enable proactive services and analyze the traffic of a city using a new perspective.

Exploration of time use

A tool to estimate the detect significant personal places and identify them as home, work or place of social activities based on temporal patterns of a person's presence in these places.

Trajectory Prediction

Predicting the near future of a user movement and use the mobility data collected to create a mobility profile



Home / Organisations / SoBigData Catalogue / MyWay - Trajectory Prediction

MyWay - Trajectory Prediction

Product Groups Activity Stream

Followers 0

Organisation

SoBigData Catalogue

SoBigData is the European Research Infrastructure for Big Data and Social Mining. For more details about the EU Project you can visit the Project Site: <http://www.sobigdata.eu/> [read more](#)

Additional Info

Field	Value
Author	Trasarti Roberto
Maintainer	Trasarti Roberto
Version	1
Last Updated	6 dicembre 2016, 14:53 (UTC+01:00)
Created	16 novembre 2016, 10:12 (UTC+01:00)
Accessibility	Both
AccessibilityMode	Download
Basic rights	Download
CreationDate	2016-11-01
Creator	Trasarti, Roberto, roberto.trasarti@isti.cnr.it
Field/Scope of use	Non-commercial research only
Owner	Trasarti, Roberto, roberto.trasarti@isti.cnr.it

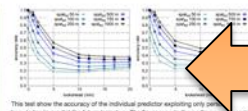
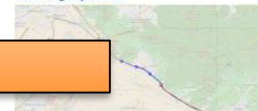
Trajectory Prediction - Methodology and Application on Pisa users

Mobility Profiles as basic concept



The resulting three predictors are shown above, for each predictor a different color is used: individual history, the individual profile and the individual predictor (red) are inside the user PMDS, while the collective predictor (blue) is outside and therefore handled by a third party that orchestrates the users' information as well as the hybrid predictor (green). This third party, usually called coordinator, has the responsibility for the storage and management of the users' profiles. In the case of the hybrid strategy the coordinator stores all the mobility profiles of the users (which are compact representations of their mobility) and resolves the queries for the predictor only in the case the individual predictor of a specific user fails. The hybrid strategy result to be the best solution.

Predicting trajectories





Transnational access first call

Proposals: 22 received – 18 EU/4 non-EU- Most popular hosts: CNR (9) and USFD (8)

Selection criteria: Applicant quality, Proposed Research Quality, Originality, Feasibility and Impact on SoBigData

Period/Budget: Feb 2017 – Jul 2017 - 2,500 euro

Successful projects: 13 successful – 12 EU / 1 non-EU - 7 male and 6 female (1 non-EU)

- Secure **ethical approval** from visitor's university
- Sign **contract** with host institution
- **Introduce visitor** to research team, arrange access to data and computational infrastructure, talks, etc.



Tuscan Big Data Challenge

- An initiative devoted to **Tuscan** enterprises interested in exploring the potential of **big data** in improving the production processes, their market or their interaction with customers and suppliers.
- The Tuscan Data Challenge gives opportunity to all the Tuscan companies, that usually do not use their **data**, to enter the world of Big Data and exploit its potential.

Who is Involved

- ***SoBigData.it*** laboratory
 - **CNR Pisa**
 - **IMT Studies Lucca**
 - **Scuola Normale Superiore**
 - **University of Pisa**



Tuscan Big Data Challenge

- In **September 2016** the challenge was launched.
- In **December 2016** researchers of ***SoBigData*** selected 9 (out of 23) of the most interesting projects and started supporting for free the selected companies in the creation of a Proof of Concept.
- Starting from **February 2017** the winning companies are working together with researchers from the SoBigData and are using the methodologies and software tools in the big-data analysis platform in the laboratory itself.



Training and Industry



IL MASTER

ISCRIZIONE

STUDENTI

CONTATTI

BLOG



PARTNERS

Almaviva

CLOUD4WI

cloudera

eesa

extra
Open Solutions, Smart Integration

GENERALI

IBM

IRPET

Istat

KPMG

NAVIONICS

net7

nova

OCTO

sas
THE POWER TO KNOW

TA
SOFTWARE SOLUTION

Trust-IT Services
Communicating 4.2 to markets

coop
Unicoop Tirreno

UnipolSai
ASSICURAZIONI



Training and Industry

PhD+



The PhD+ programme, designed and fully developed by the University of Pisa, aims at fostering innovation and entrepreneurial mind-set in Master degree students, PhD students, PhDs and academics.

PhD+ is one of the best practice of training in research valorisation, innovation and entrepreneurship, also recognized by the Network of Design for Resilient Entrepreneurship, within the [ENDuRE](#) European project





Main Contacts

- **Coordinator**
 - **CNR** Fosca Giannotti -
fosca.giannotti@isti.cnr.it
- **Co-coordinators**
 - **UNIPi** University of Pisa, *Italy*
Dino Pedreschi -
dino.pedreschi@di.unipi.it
 - **UNIPi** University of Pisa, *Italy*
Paolo Ferragina -
ferragina@di.unipi.it



www.sobigdata.eu

@SoBigData (<https://twitter.com/SoBigData>)

<https://www.facebook.com/SoBigData>