

Strumenti di studio data-driven per la mobilità toscana

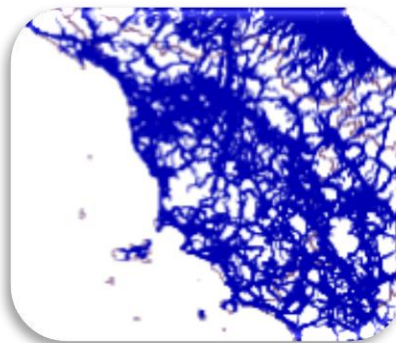
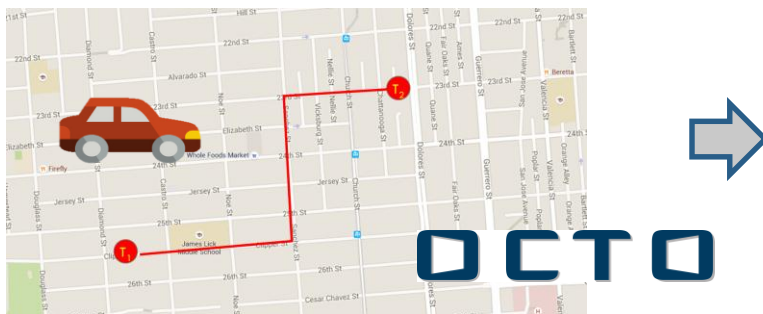
Mirco Nanni

mirco.nanni@isti.cnr.it



Sorgenti informative

- Tracce GPS di mobilità veicolare



Feb-Mar 2014

Tempo

150K

Numero di veicoli

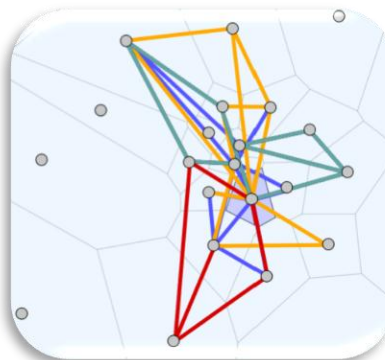
12 mln

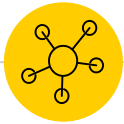
Viaggi

- Dati di telefonia (CDR)



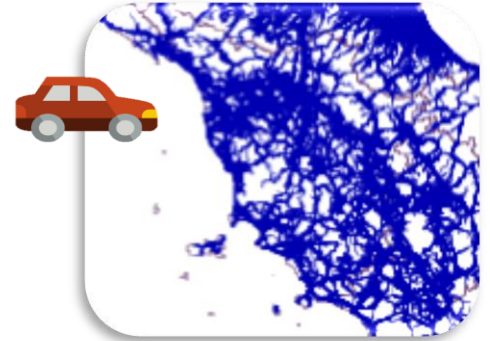
	Time start	Cell start	Cell end	Duration
10294595	"2014-02-20 14:24:58"	"PI010U2"	"PI010U1"	48
10294595	"2014-02-20 18:50:22"	"PI002G1"	"PI010U2"	78
10294595	"2014-02-21 09:19:51"	"PI080G1"	"PI016G1"	357





1. MFAD

Mobility Functional Area Discovery





Problem definition

- **Q1:** I dati di mobilità possono suggerirci quali aree funzionali esistono sul territorio?
- **Q2:** Possono aiutarci a comprendere quale tipo di struttura (monocentrica vs. policentrica) hanno?



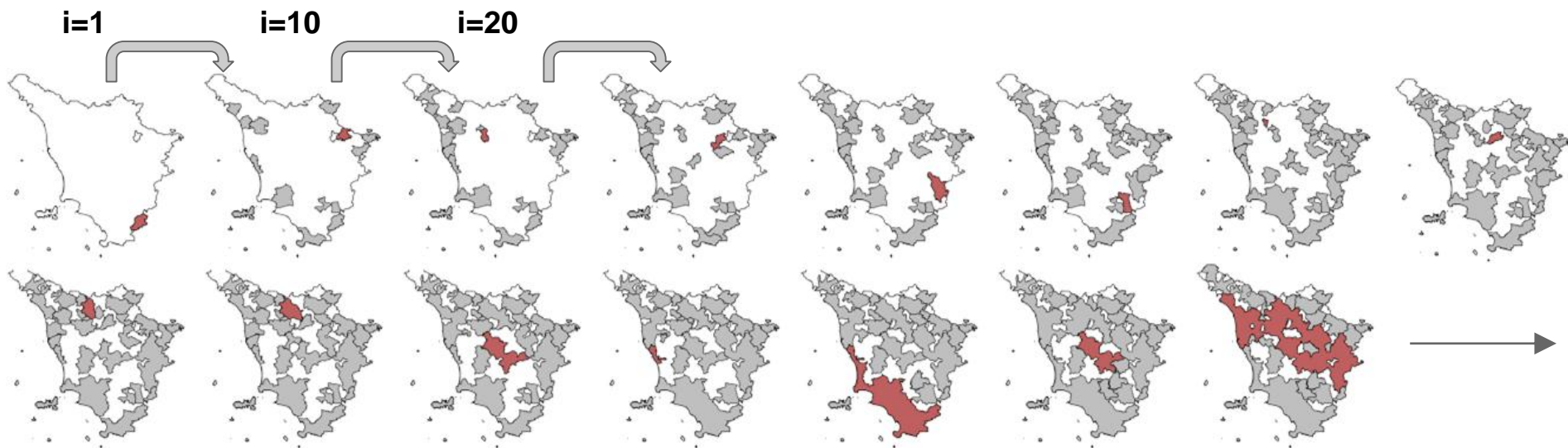
Proposta

- ◉ **Identificare i confini urbani “naturali” che emergono dai dati**
- Criterio: i comuni di un’area hanno molto più traffico interno (all’area) che da/verso l’esterno

- ◉ **Studiare la struttura interna di ogni area**
- Criterio: identificare i comuni cardine che definiscono l’area



Proposta: algoritmo agglomerativo



$$localQ(a, b, G) = \frac{F(a, b) + F(b, a)}{\sum_{(x, y) \in E \wedge \{a, b\} \cap \{x, y\} \neq \emptyset} F(x, y)}$$

Red

Comunità/aree create nell'ultima iterazione

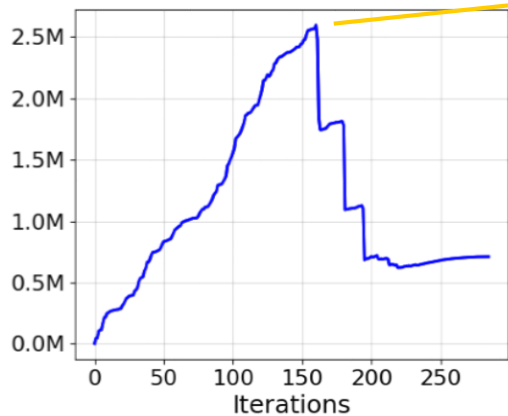
Grey

Comunità/aree create nelle precedenti iterazioni

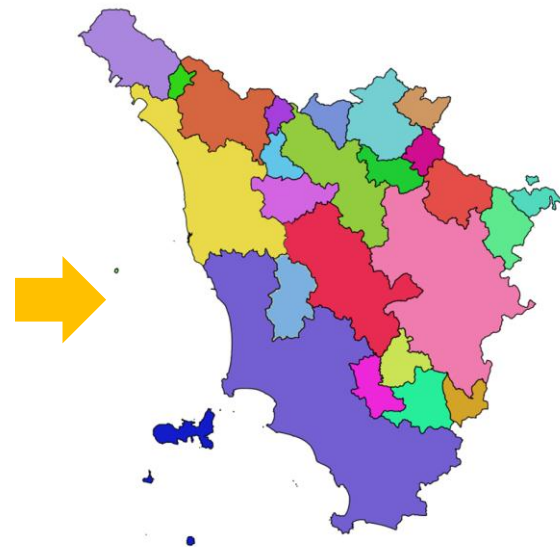


Città policentrica: **RISULTATI**

$$globalQ(G) = \sum_{(i,j) \in E} F(i,j) - F(i \rightarrow) * \frac{F(\rightarrow j)}{K}$$



25 Comunità scoperte



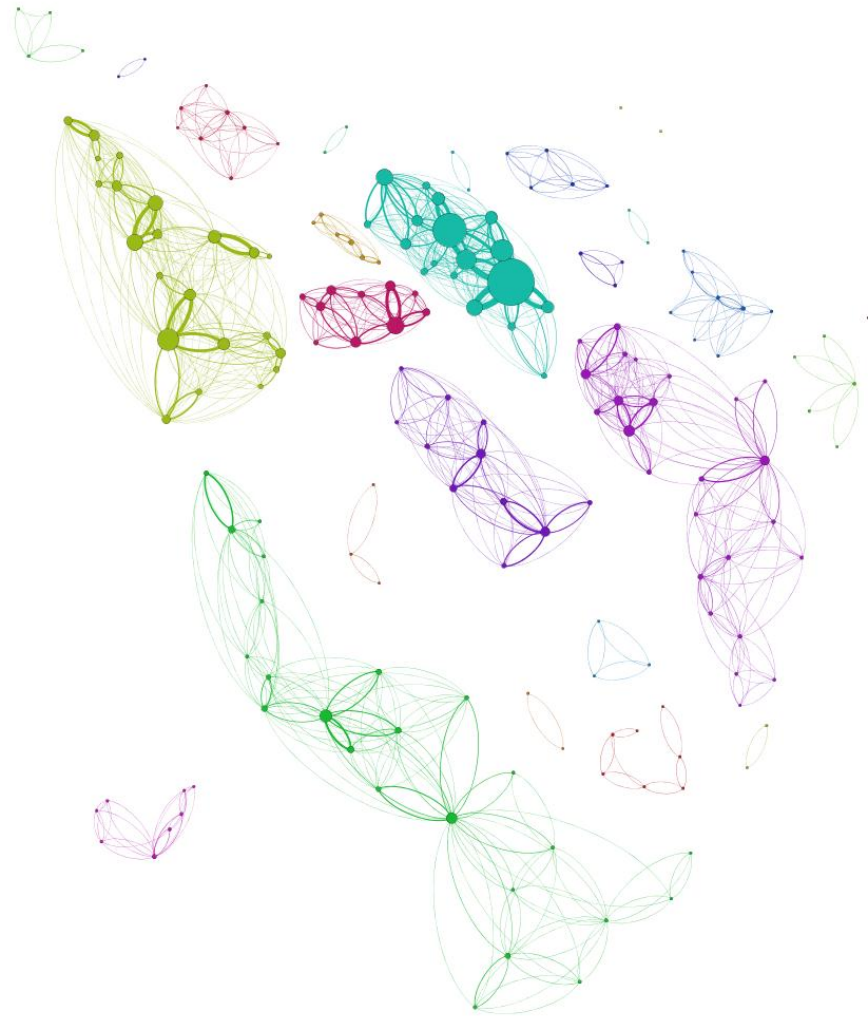
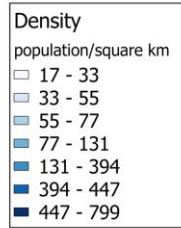
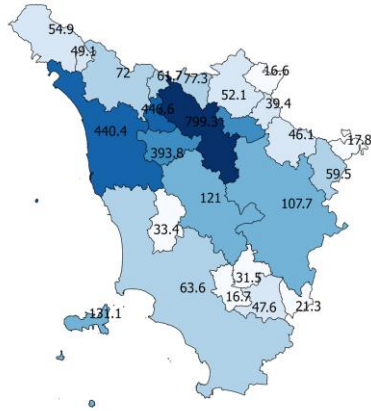
Saturazione

Assegnazione comuni isolati



Sub network

24 communities

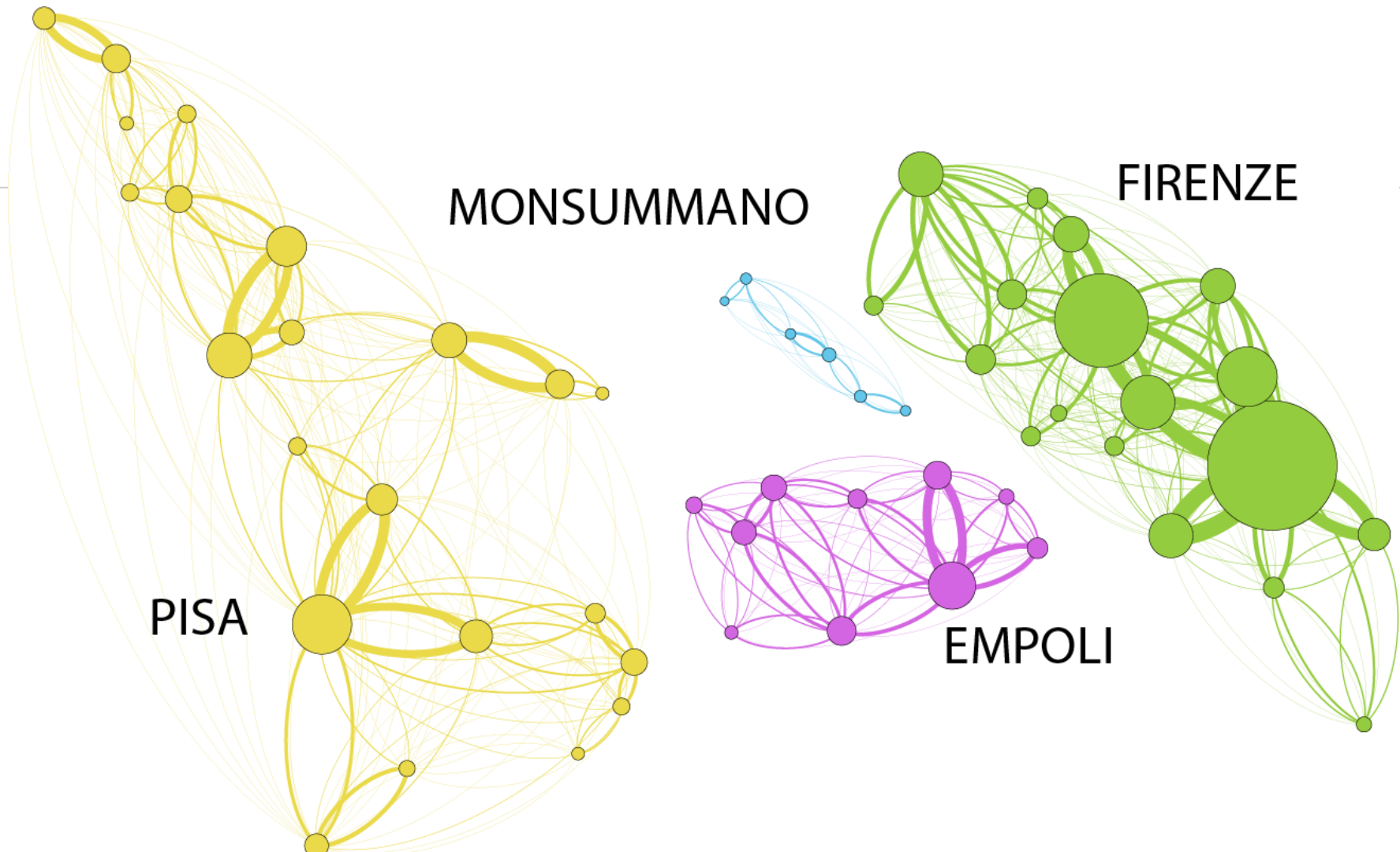


MONSUMMANO

FIRENZE

PISA

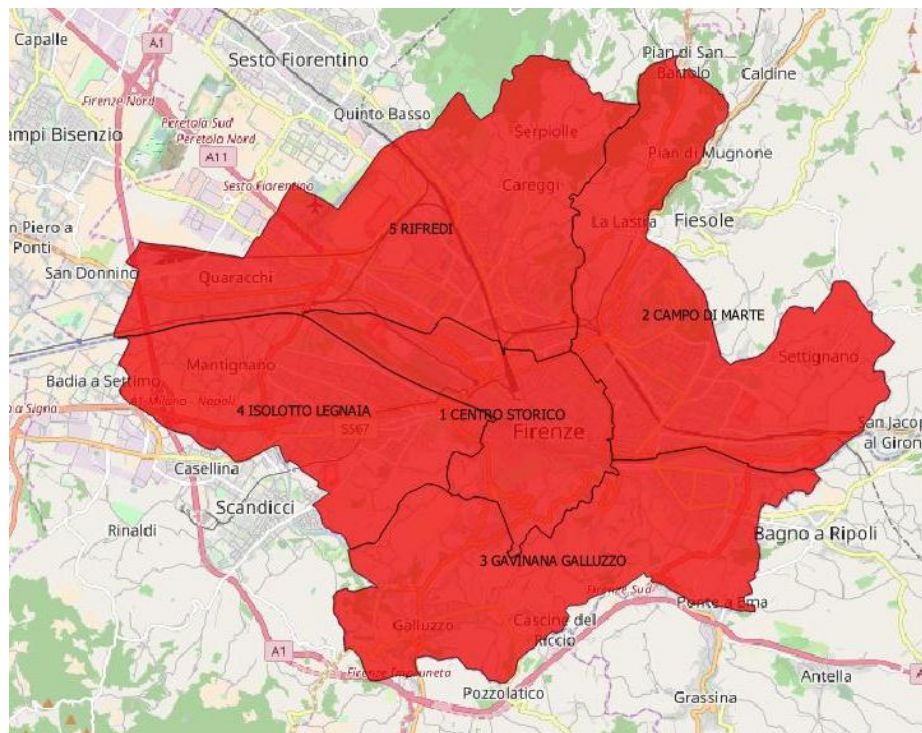
EMPOLI

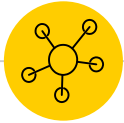




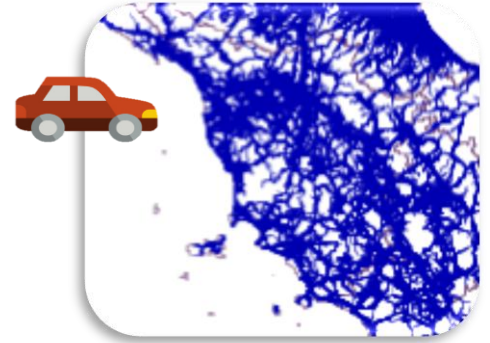
Focus: i quartieri di Firenze

Quartieri	Traiettorie	Incidenza (%)
Rifredi	212.651	31,9
Campo di Marte	126.046	18,9
Centro	103.041	15,6
Gavinana	88.825	13,4
Isolotto	134.171	20,2





2. Analisi di attrattori





Analisi degli attrattori: aeroporti



**Analizzare l'influenza dei grandi attrattori sulla
mobilità dei territori circostanti**

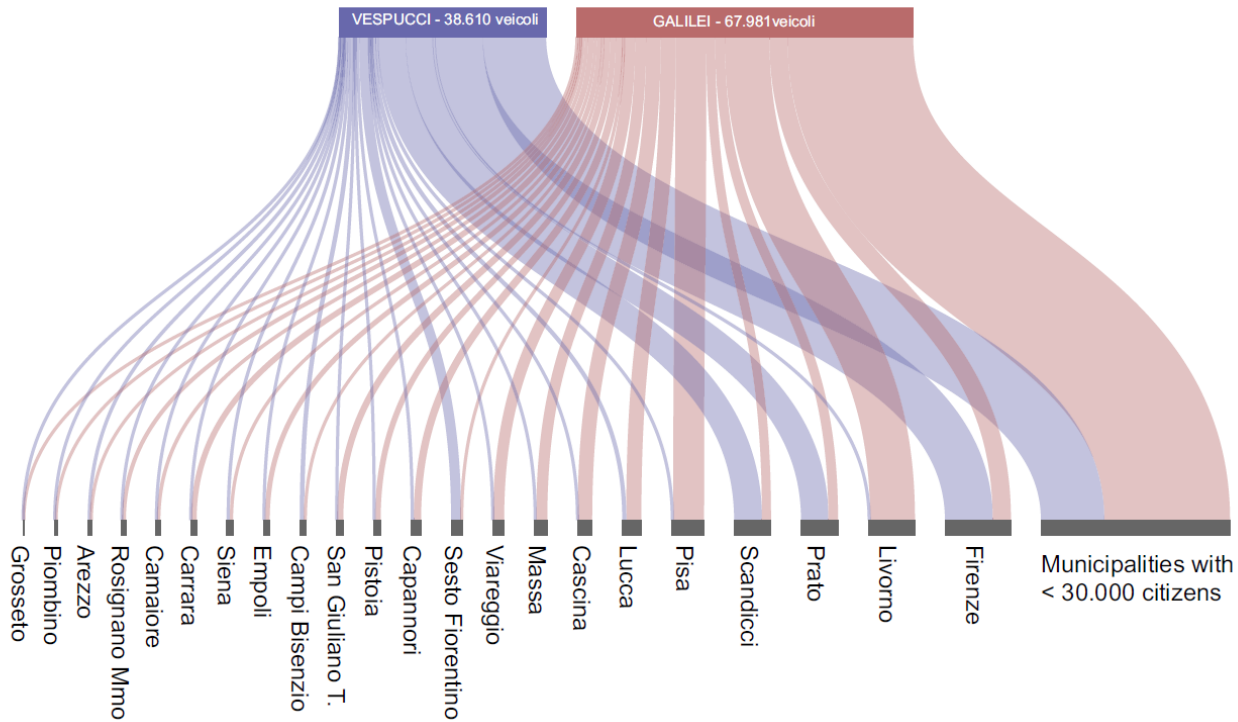
CASO STUDIO:

gli aeroporti di Firenze e Pisa e la propensione dei
residenti toscani all'uso delle due infrastrutture.

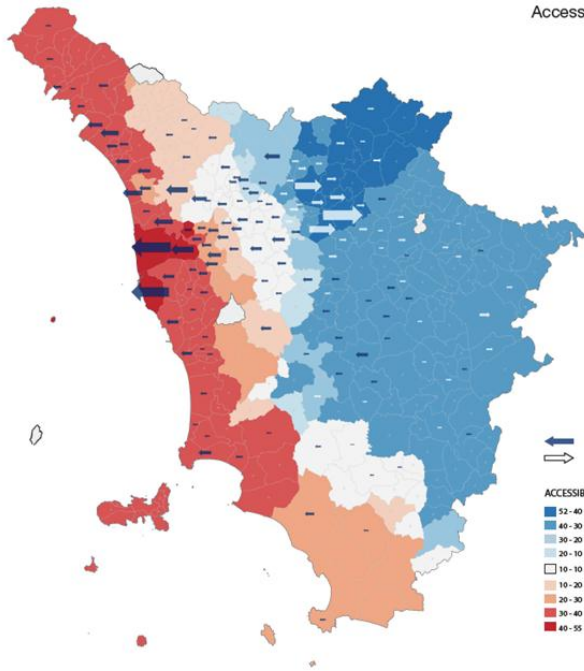




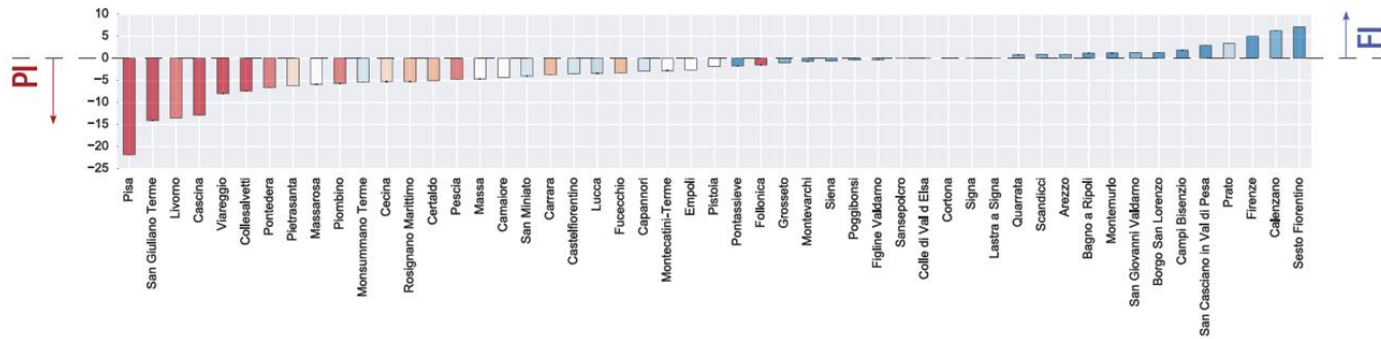
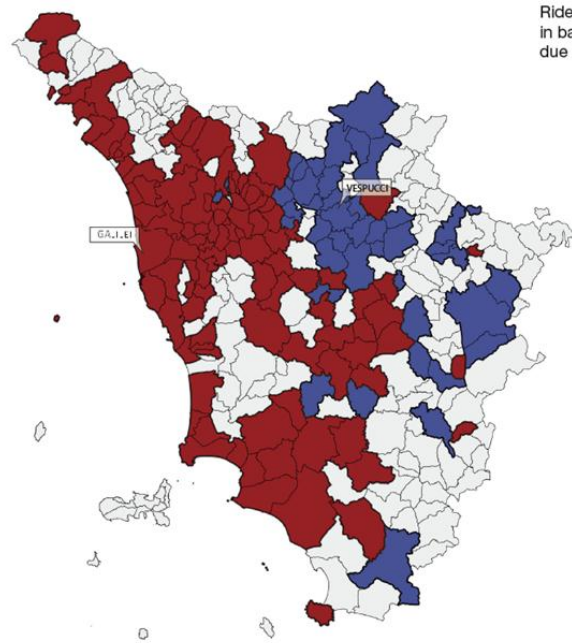
Analisi degli attrattori: aeroporti



Accessibilità vs preferenza



Ridefinizione del territorio in base all'attrazione delle due infrastrutture



Modelli di investimento vs. attrattività

Modeling Investments and Attractiveness on Tuscan Airports.
Ioanna Miliou, Diana Knipf, Salvatore Rinzivillo, Seven R. Bishop, Dino Pedreschi.

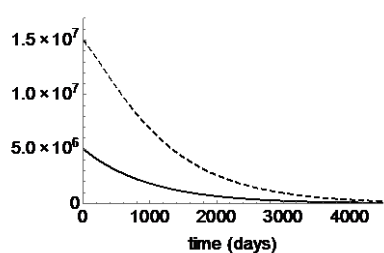
An intertwined system based on investment and attractiveness

$$\frac{d}{dt}A = s(mF - (k + e)A), \quad A \rightarrow \text{Attractiveness of airport}$$

$$\frac{d}{dt}F = -rF + re \frac{bA}{1 + bhA}. \quad F \rightarrow \text{Number of passengers served}$$

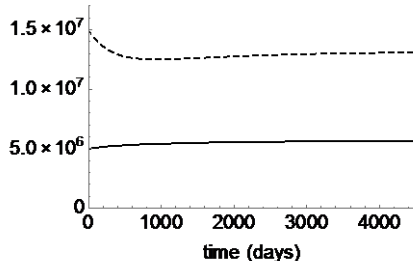
Simple case: non spatial model

No extra investments ($e=0$)



(a) $e = 0$

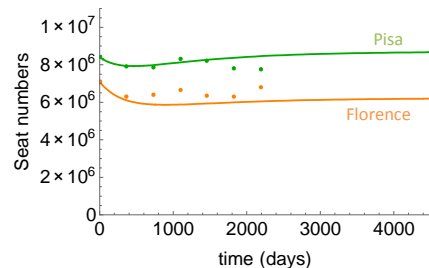
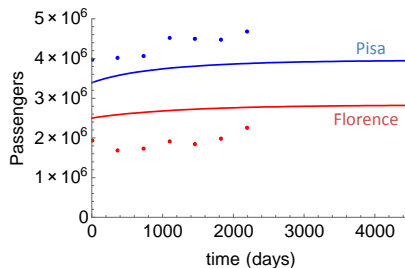
Structural investments ($e=0.05$)



(b) $e = 0.05$

Attractiveness is proportional to the cost of operating the airport (k) and the extra investments (e)

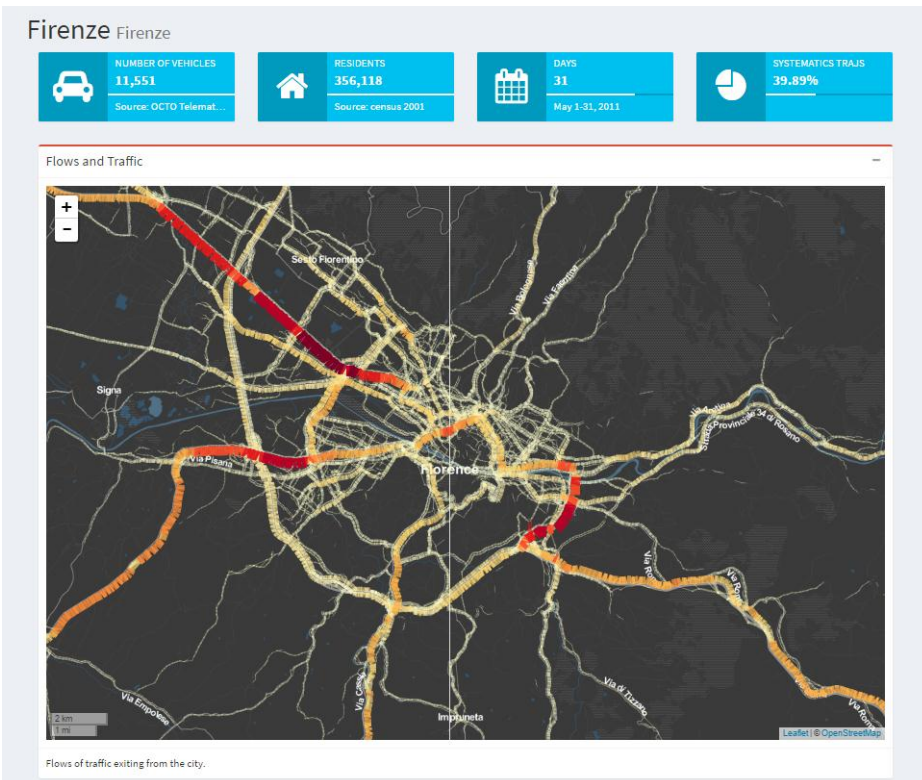
Spatial model: two airports, two populations



The two airports reach an equilibrium: neither of the two is overwhelming the other



Urban Mobility Atlas (UMA)

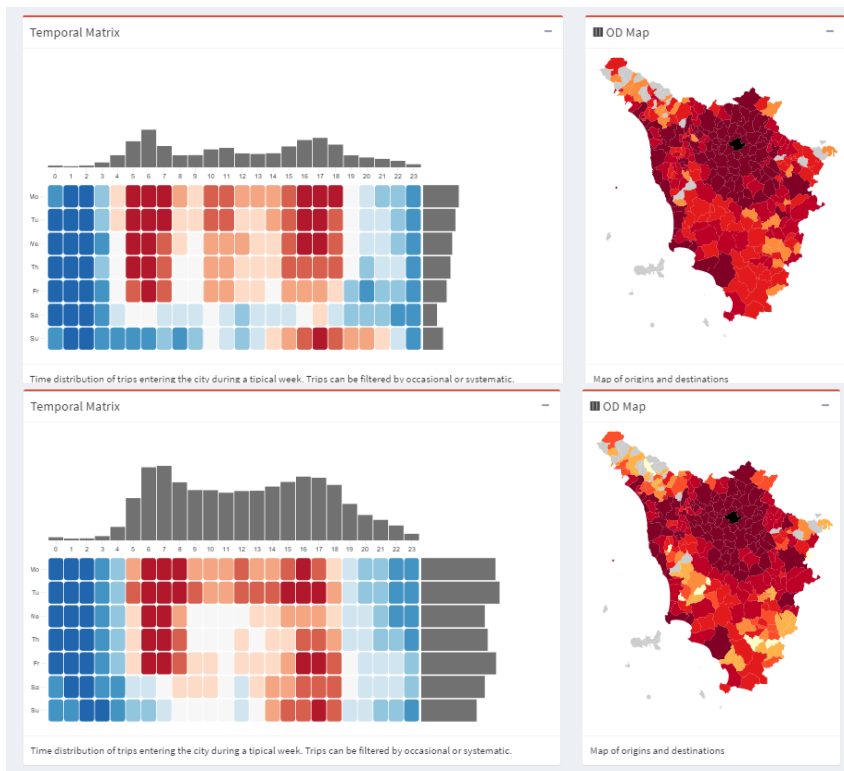


<http://kdd.isti.cnr.it/uma2/>

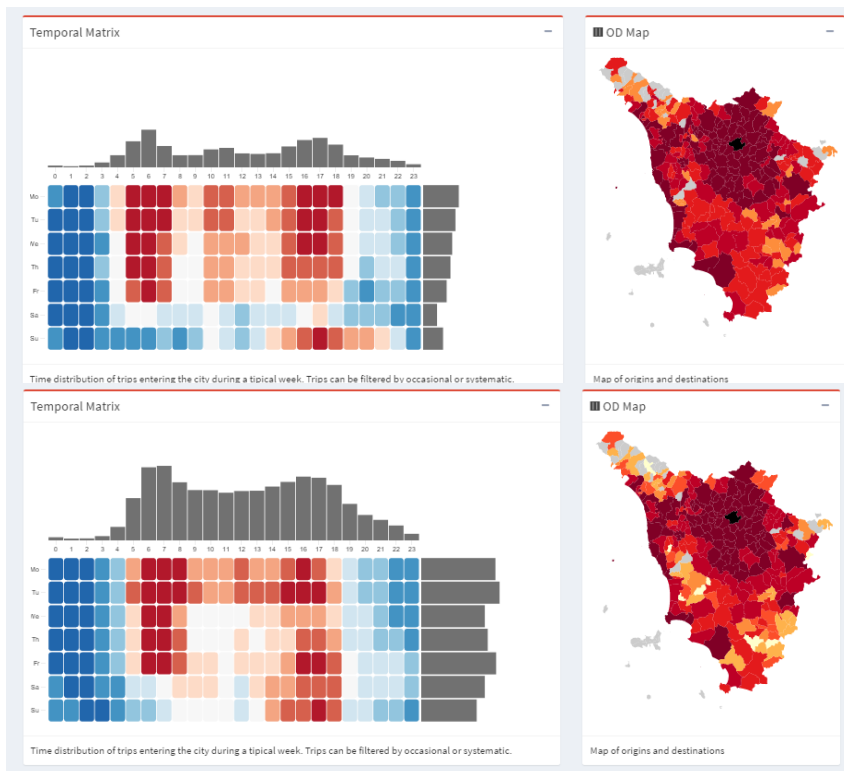


Urban Mobility Atlas (UMA)

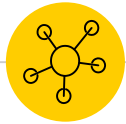
Sistematico



Occasionale



<http://kdd.isti.cnr.it/uma2/>



3. Analisi di eventi

*St. Peter's Square
(Piazza San Pietro)*



*Olympic Stadium
(Stadio Olimpico)*



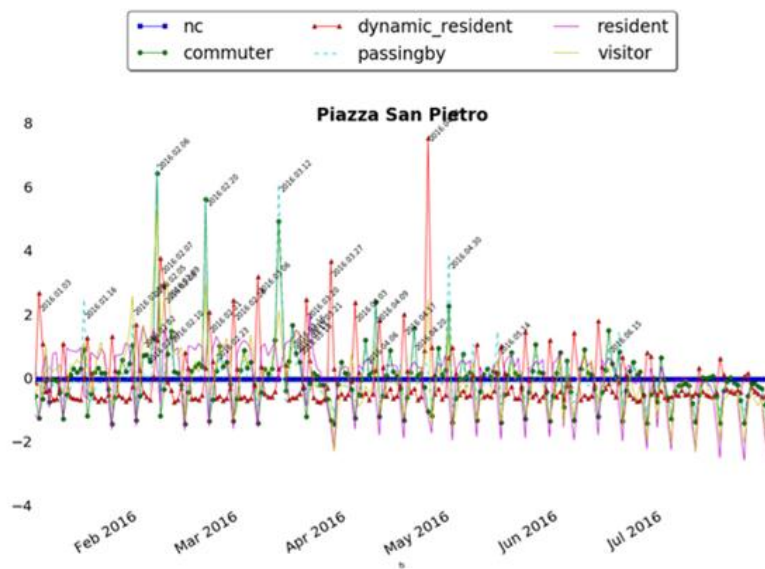
*Circus Maximus
(Circo Massimo)*





Rilevazione e misurazione eventi

Peak detection



Piazza San pietro

Misuriamo le presenze distinguendo i diversi *City User*:

- Residenti
- Pendolari / Visitatori costanti
- Turisti / Visitatori occasionali

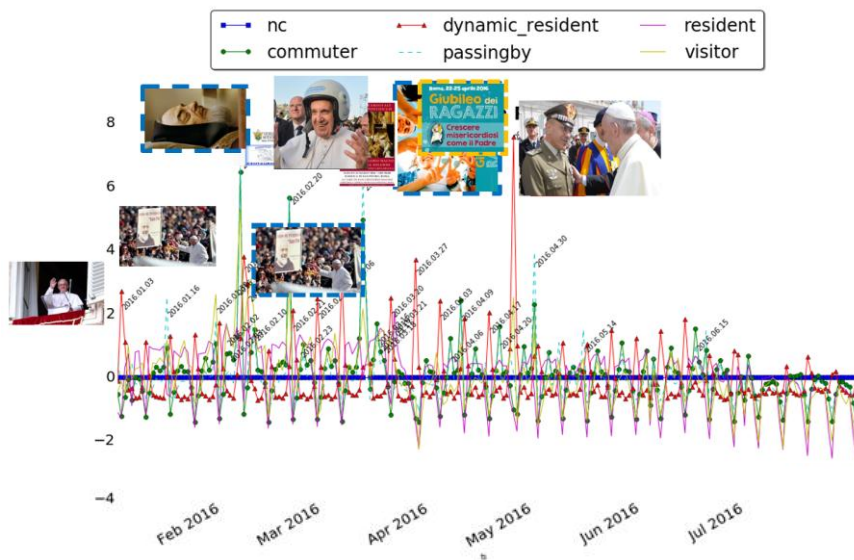
Identifichiamo presenze anomale

- in un period specific
- rispetto a presenze passate



Rilevazione e misurazione eventi

Peak detection



Piazza San pietro

Misuriamo le presenze distinguendo i diversi *City User*:

- Residenti
- Pendolari / Visitatori costanti
- Turisti / Visitatori occasionali

Identifichiamo presenze anomale

- in un period specific
- rispetto a presenze passate

Social Mining & Big Data Analytics



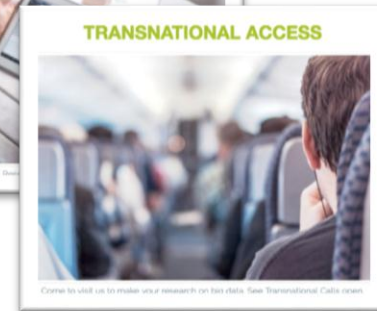
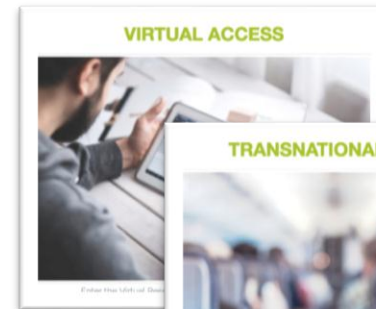
Social Mining &
Big Data Ecosystem
H2020 - www.sobigdata.eu
September 2015- August 2019



Objectives and Stakeholders

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”.

- Big data analysts and social informatics researchers
- Economists, social science and humanities researchers
- journalists, policy and law makers
- Researchers in related communities
- Industrial innovators & startups
- The public as a whole





Virtual Research Environments

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 provide 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

Tip 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 recommendations 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 observe the existence of clusters/passenger micro-communities

Mobility Profiles

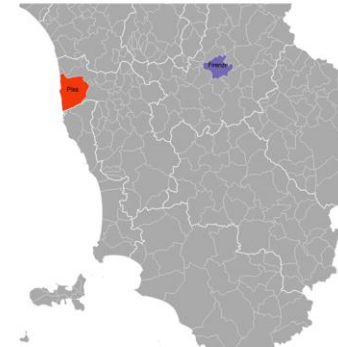
Understanding the systematic behavior of the user is the key to analyze proactive services and analyze the traffic of a city using a new perspective.

Exploration of time use

A tool to estimate the detect significant personal places and analyze the patterns of a person's presence in these places.

Trajectory Prediction

Predicting the near future of a user and analyzing the user's past movements and use the mobility profiles to create a correct mobility prediction.



Organisations / SoBigData Catalogue / MyWay - Trajectory Prediction

MyWay - Trajectory Prediction

Product Groups Activity Stream

Followers 0

Organisation

SoBigData Catalogue

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](#)

Field	Value
Author	Trassari Roberto
Maintainer	Trassari Roberto
Version	1
Last Updated	6 dicembre 2016, 14:53 (UTC-01:00)
Created	16 novembre 2016, 10:12 (UTC-01:00)
Accessibility	Both
Accessibility Mode	Download
Basic Rights	Download
Creation Date	2016-11-01
Creator	Trassari, Roberto, roberto.trassari@iit.it
Field/Scope of use	Non-commercial research only
Owner	Trassari, Roberto, roberto.trassari@iit.it

Trajectory Prediction - Methodology and Application on Pisa users

Mobility Profiles as basic concept

User's Personal Mobility Data Store

PMS PMS PMS

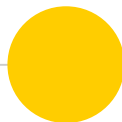
Mobility Prediction Activity Prediction Trajectory Prediction

Predicting trajectories

The test shows the accuracy of the individual position regarding only one user



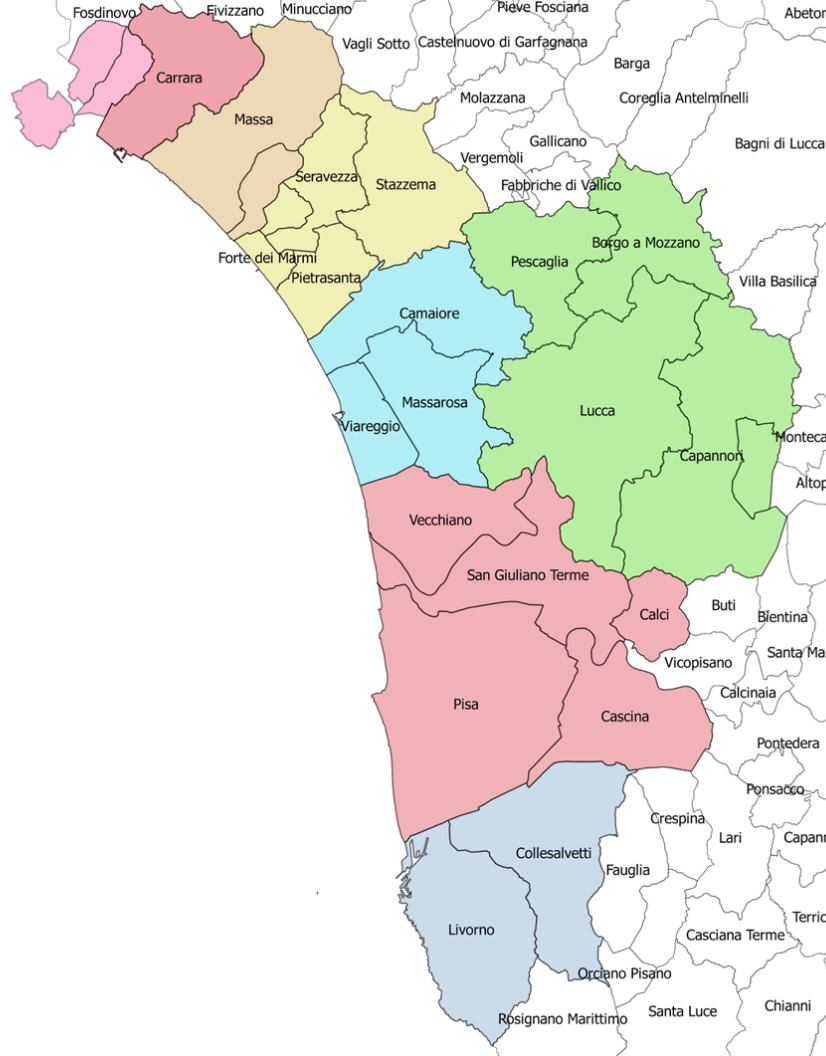
Training and Industry



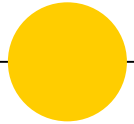


Grazie!

domande ?



BACKUP Slides





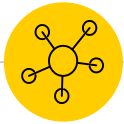
Competitors (state of the art)

Network Based

- ◉ Louvain
MODULARITY BASED
- ◉ Demon
EGO NETWORK BASED
- ◉ Infohiermap
CONDUCTANCE BASED

Cluster Based

- ◉ DBSCAN
- ◉ KMedoid



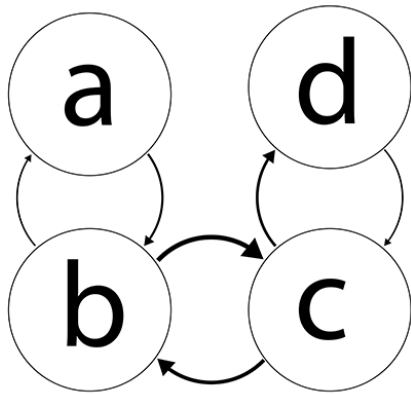
NETWORK approach

In order to *compare our approach* with the state of art we observe those measures:

internal density conductance modularity



Quality score locale: **indice di autocontenimento**



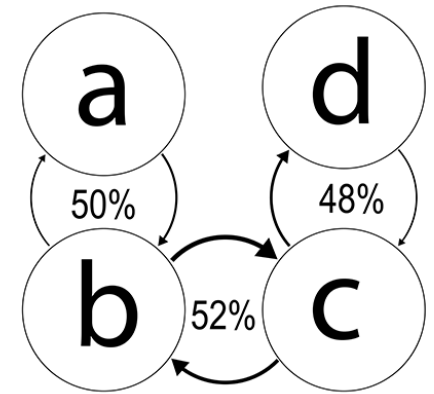
Iteration n

Local Function	Inner traffic
localQ(a,b)	50%
localQ(a,d)	0%
localQ(a,c)	0%
localQ(b,c)	52%
localQ(d,c)	48%

$$\text{Evaluation}_n(t,z) = tUz / (t^* + z^*)$$

$$aUb = 50\% \quad \mathbf{bUc = 52\%}$$

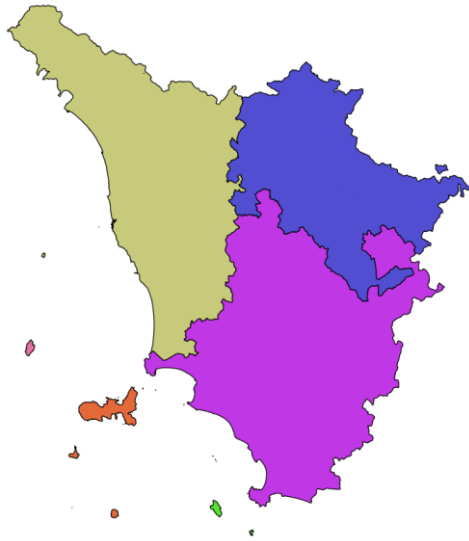
$$cUd = 48\% \quad dUe = 40\%$$



Iteration $n+1$



Network approach **LOUVAIN**

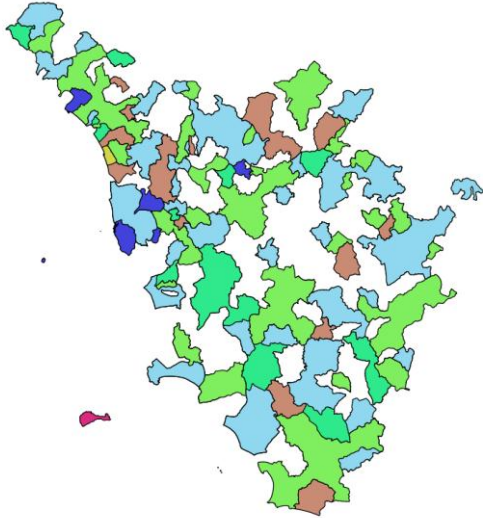


Measure min/max/avg/std	Louvain	Policentrometer
Internal Edge Density	0.15/0.32/0.21/0.07	0.27/0.75/ 0.49 /0.20
Conductance	0.014/0.58/0.38/0.27	0.014/0.97/ 0.88 /0.19
Modularity	0.16	-0.06

Drawback too few and too big communities



Network approach **DEMON**



Measure min/max/avg/std	Demon	Policentrometer
Internal Edge Density	0.12/0.50/0.28/0.18	0.27/0.75/ 0.49 /0.20
Conductance	0.37/0.90/0.50/0.17	0.014/0.97/ 0.88 /0.19
Modularity	-0.38	-0.06

Drawback too overlapping communities



Network approach **INFOHIERMAP**



Measure min/max/avg/std	INFOHIERMAP	Policentrometer
Internal Edge Density	0.09/0.50/0.18/0.10	0.27/0.75/ 0.49 /0.20
Conductance	0.90/0.98/ 0.95 /0.24	0.014/0.97/ 0.88 /0.19
Modularity	0.006	-0.06

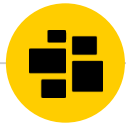
Drawback non contiguous communities



Lesson Learned

Comparison with network algorithms:

- Too few communities
- Too big
- Not contiguous



CLUSTERING approach

partitions obtained with K-medoid
and DBSCAN cluster methods

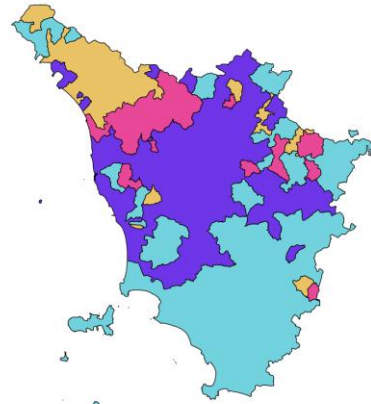


K MEDOID

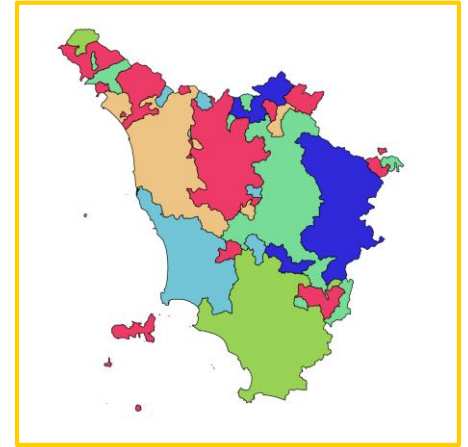
Drawback non contiguous communities



k=2



k=4

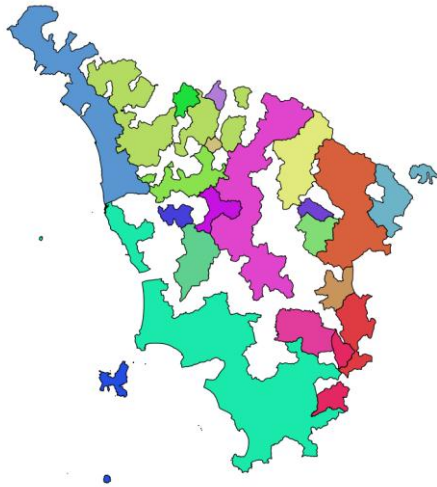


k=6

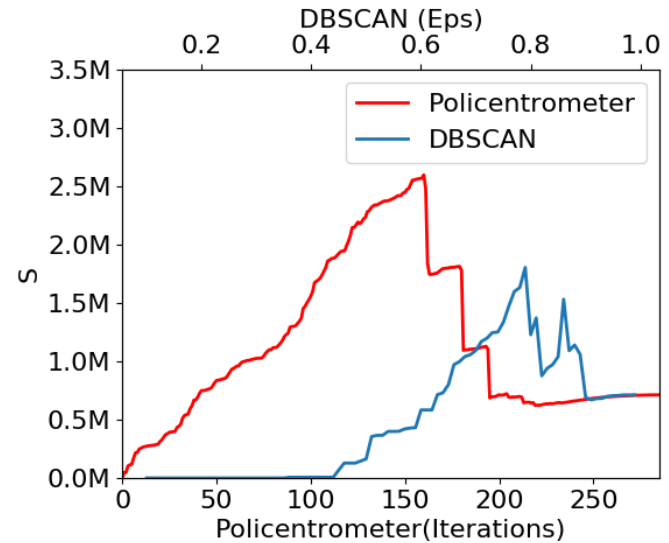


DBSCAN

Drawback: Policentrometer get a higher global score than DBSCAN



dbscan communities



optimal choice for the two methods

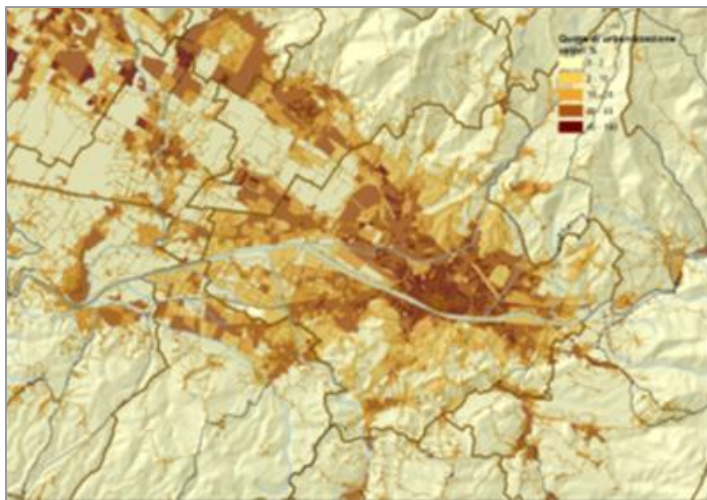


Conclusion

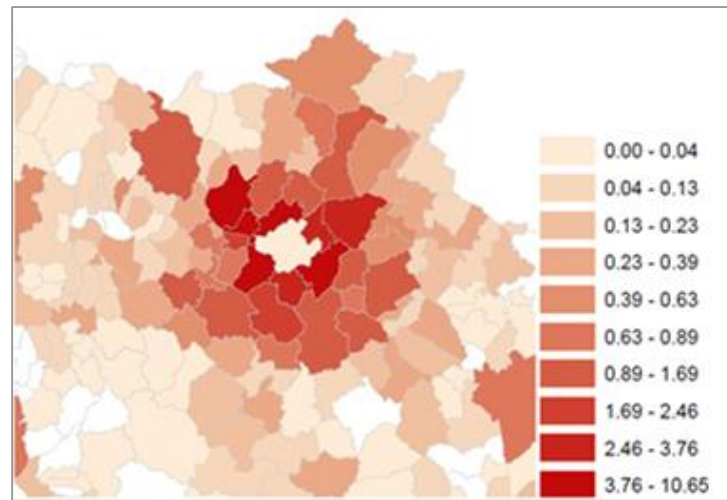
- ◉ **Flow inclusion based problem definition**
- ◉ **Ad hoc algorithm that outperforms state-of-the-art methods**
- ◉ **Preliminary evaluation of results**



Stato dell'arte



Analisi Morfologica



Analisi funzionale