Association for Computing Machinery

1st ACM Summer School in Europe

ACM Digita

Con

HOME ABOUT PARTICIPATE COURSES SPEAKERS DAILY SCHEDULE VENUE SOCIAL ACTIVITIES EXPLORE ATHENS

13-19 July 2017, Athens, Greece

1st ACM Europe Summer School | Data Science

Data ethics and privacy-preserving analytics

Dino Pedreschi

KDD Lab

UNIVERSITÀ DI PISA

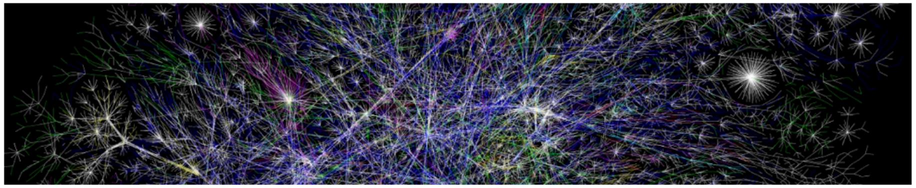
Università di Pisa & ISTI-CNR

ISTITUTO DI SCIENZA E TECNOLOGIE DELL'INFORMAZIONE "A. FAEDO"

http://kdd.isti.cnr.it

Home Research Projects People Resources Blog

Mobility Data Mining for Science of Cities



Social Network Analysis and Visual Analytics

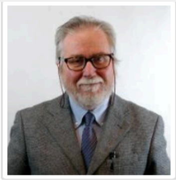
Ethical Data Mining

Analytical Platforms and Infrastructures for Social Mining

KDD Lab

http://kdd.isti.cnr.it/people

Dino Pedreschi
Full Professor
pedre@di.unipi.it



Franco Turini
Full Professor
turini@di.unipi.it

Fosca Giannotti
Director of Research
fosca.giannotti@isti.cnr.it

Salvatore Ruggieri
Full Professor
ruggieri@di.unipi.it

Anna Monreale
Assistant Professor
annam@di.unipi.it

Alina Sîrbu
Assistant Professor
alina.sirbu@unipi.it

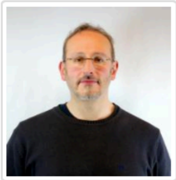
Barbara Furletti
Researcher
barbara.furletti@isti.cnr.it

Mirco Nanni
Researcher
mirco.nanni@isti.cnr.it

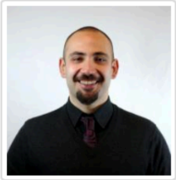
Salvatore Rinzivillo
Researcher
rinzivillo@isti.cnr.it

Roberto Trasarti
Researcher
roberto.trasarti@isti.cnr.it

Paolo Cintia
Post Doc
paolo.cintia@isti.cnr.it



Valerio Grossi
Post Doc
vgrossi@di.unipi.it



Luca Pappalardo
Post Doc
luca.pappalardo@isti.cnr.it

Giulio Rossetti
Post Doc
giulio.rossetti@isti.cnr.it

Alessandro Baroni
Ph.D. Student
baroni@di.unipi.it

Lorenzo Gabrielli
Ph.D. Student
lorenzo.gabrielli@isti.cnr.it

Riccardo Guidotti
Ph.D. Student
riccardo.guidotti@isti.cnr.it

Ioanna Miliou
Ph.D. Student
ioanna.miliou@for.unipi.it

Letizia Milli
Ph.D. Student
letizia.milli@isti.cnr.it

Laura Pollacci
Ph.D. Student
laura.pollacci@di.unipi.it

Francesca Pratesi
Ph.D. Student
francesca.pratesi@isti.cnr.it

Farzad Vaziri
Ph.D. Student
vaziri@di.unipi.it

Viola Bachini
Collaborator
viola.bachini@isti.cnr.it

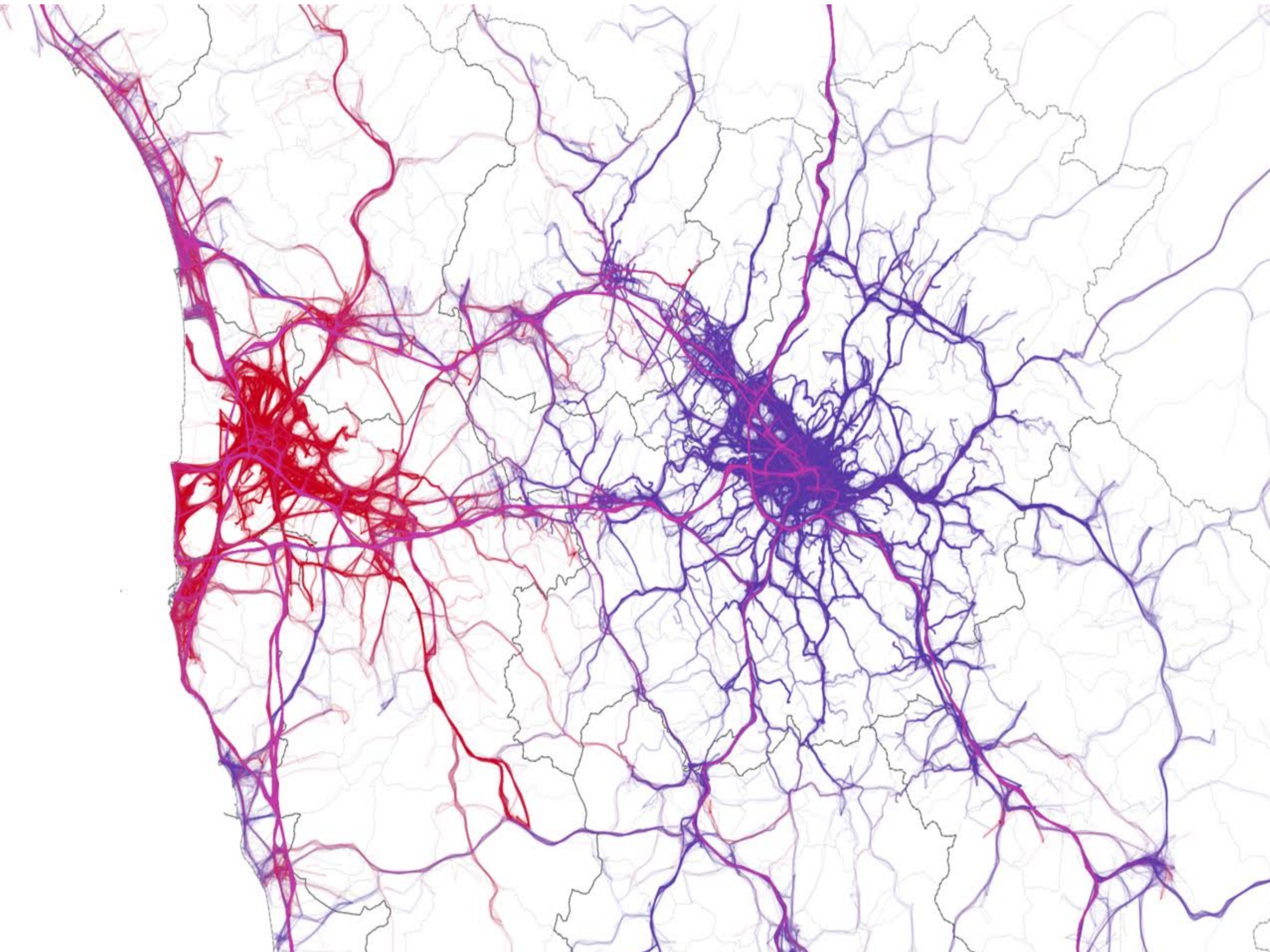
Daniele Fadda
Collaborator
daniele.fadda@isti.cnr.it

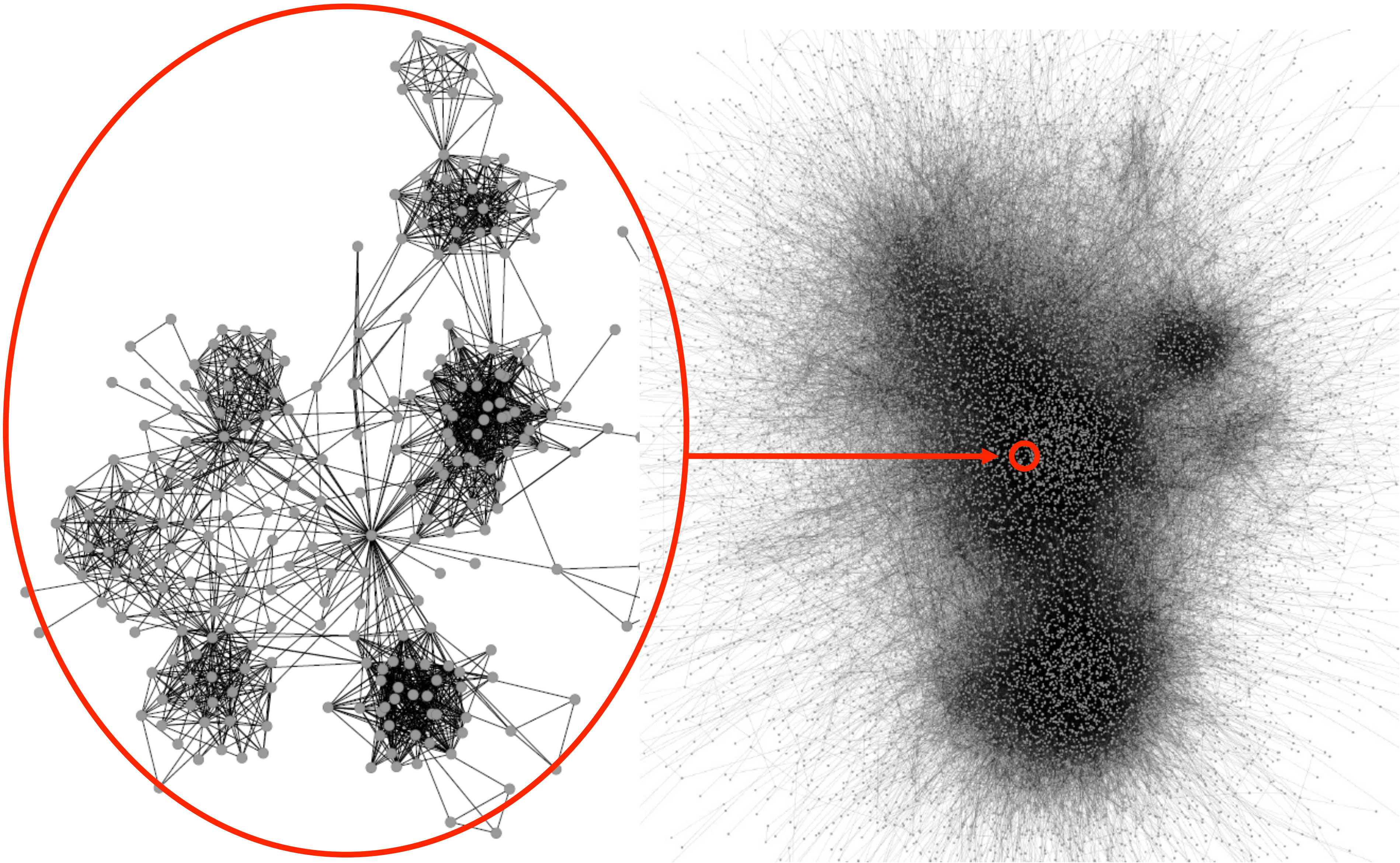
Knowledge Discovery & Data Mining Lab
http://kdd.isti.cnr.it

Wiki of the course

- http://didawiki.di.unipi.it/doku.php/wma/acm-athens-july2017

- Special thanks to
  - Anna Monreale, University of Pisa

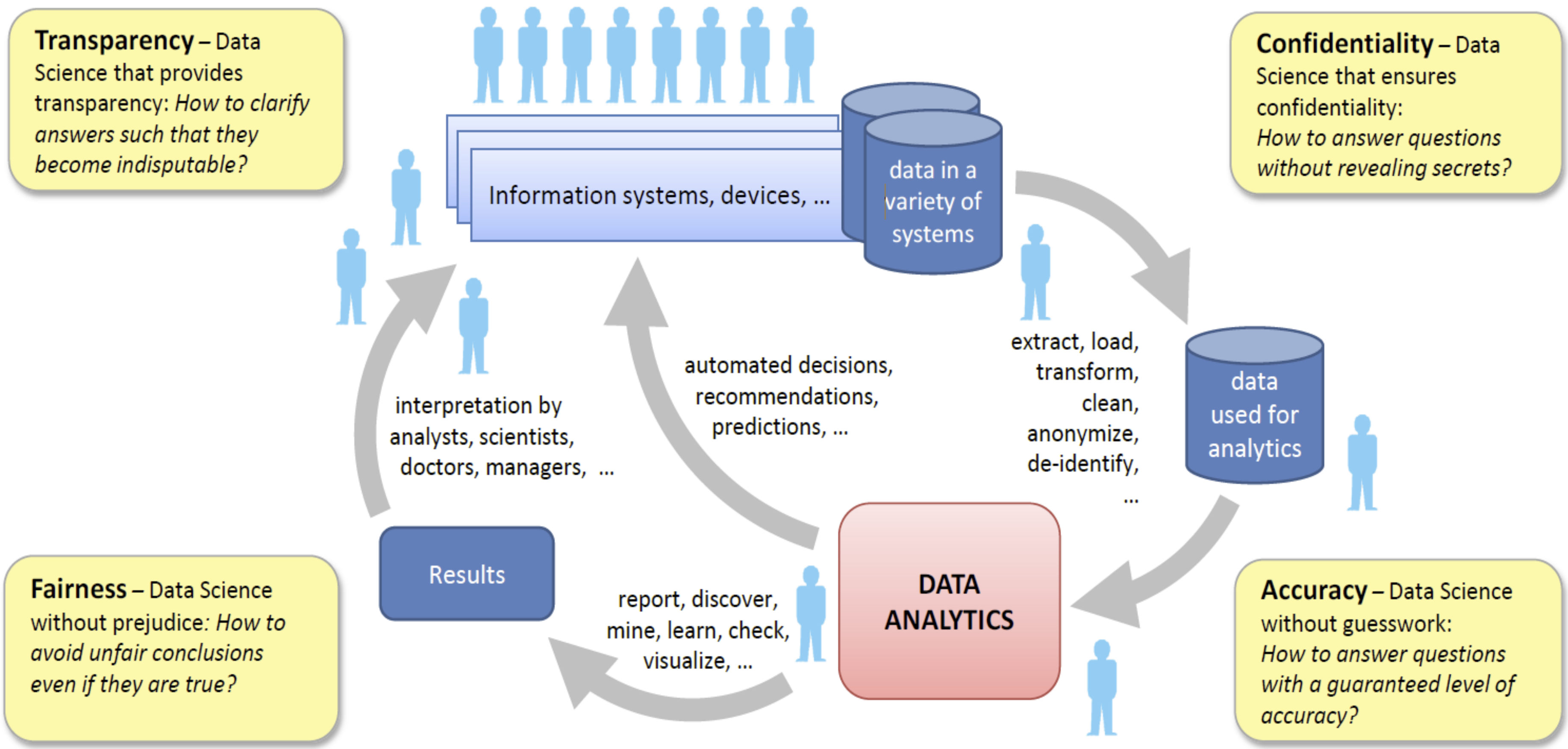




- Data science created unprecedented opportunities but also new **risks**.
- Data Science techniques might expose sensitive traits of individuals and invade their **privacy**,
- this information could be used to **discriminate** people based on their presumed characteristics, or profiles.

Responsible Data Science

http://www.responsibledatascience.org/



Lecture roadmap

- The new General Data Protection Regulation – GDPR
- Privacy-by-design and risk assessment
- Personal data analytics & the new deal on data
- Transparency of machine learning algorithms

The research challenges of the new General Data Protection Regulation

EU Legislation for protection of personal data

- **European directives:**
  - Data protection directive (95/46/EC)
  - ePrivacy directive (2002/58/EC) and its revision (2009/136/EC)
  - **new EU Regulation (Proposed: 25 Jan 2012, Published: 4 May 2016, into force: May 2017)**

- **Opinions by EU Article 29 Data Protection Working Party**

EU: Personal Data

- **Personal data** is defined as any information relating to an identity or **identifiable** natural person.

- An **identifiable person** is one who can be identified, **directly or indirectly**, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity.

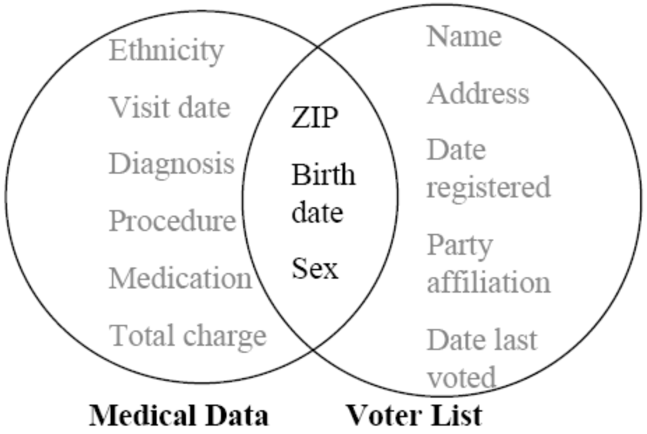
Anonymity according to 1995/46/EC

- The principles of protection must apply to any information concerning an identified or identifiable person;
- To determine whether a person is identifiable, account should be taken of **all the means likely reasonably to be used** either by the controller or by any other person to identify the said person
- **The principles of protection shall not apply to data rendered anonymous in such a way that the data subject is no longer identifiable**

Re-identification of Massachussetts' governor

- Sweeney managed to re-identify the medical record of the governor of Massachussetts
  - MA collects and publishes sanitized medical data for state employees (microdata) left circle
  - voter registration list of MA (publicly available data) right circle

- looking for governor's record
- join the tables:
  - 6 people had his birth date
  - 3 were men
  - 1 in his zipcode



Latanya Sweeney: k-Anonymity: A Model for Protecting Privacy. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 10(5): 557-570 (2002)

AOL Search History Release (2006)

A Face Is Exposed for AOL Searcher No. 4417749

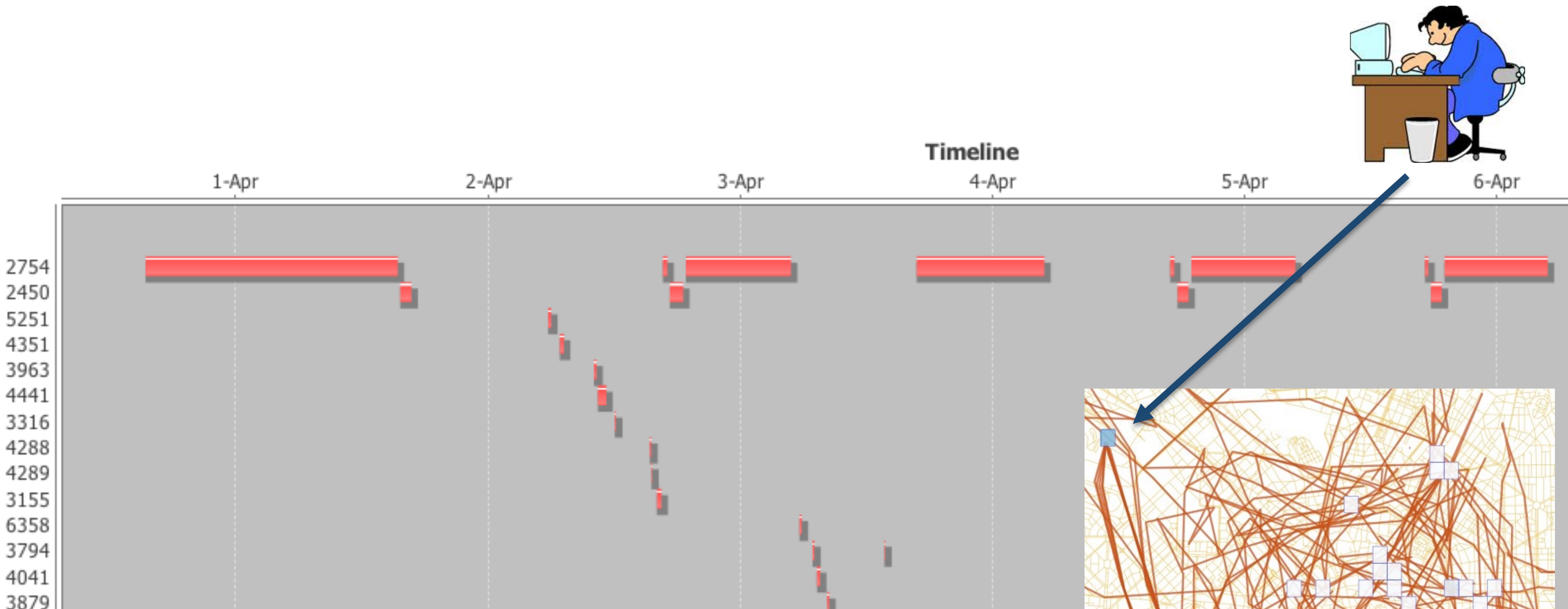
By MICHAEL BARBARO and TOM ZELLER Jr. The New York Times
Published: August 9, 2006

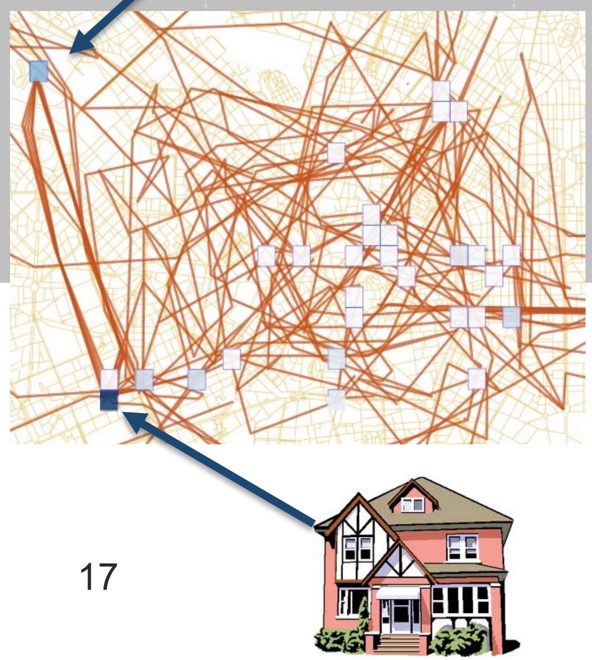
Buried in a list of 20 million Web search queries collected by AOL and recently released on the Internet is user No. 4417749. The number was assigned by the company to protect the searcher's anonymity, but it was not much of a shield.

No. 4417749 conducted hundreds of searches over a three-month period on topics ranging from "numb fingers" to "60 single men" to "dog that urinates on

Name: Thelma Arnold
Age: 62
Widow
Residence: Lilburn, GA

De-identified User Trajectory





- Discovering persons living in that home and working in that company we can identify the user

EU Article 29 Data Protection Working Party: Opinion 05/2014 on Anonymization techiniques

- Provides recommendations to handle these techniques by taking account of **the residual risk** of identification inherent in each of them

- Identifies the following attacks:
  - **Singling out** an individual in a dataset
  - **Linking** two records within a dataset (or between two separate datasets)
  - **Inferring** any information in such dataset

Opinion 05/2014: Techniques

- Anonymity by randomization
- Anonynity by generalization
- Differential-privacy
- l-diversity
- t-closeness
- Pseudonymisation

The GDPR Regulation

- Will be applied on 25 May 2018 and will take the form of a Regulation
- Introduces important novelties
  - New Obligations
  - New Rights
  - **New vision on the research context**

New Elements in the EU GDPR

- New Obligations for Data Processors
- GDPR Outside EU
- Accountability Principle
- Privacy by Design
- Principle of Transparency
- Data Portability
- Right of Oblivion
- Profiling
- The right of explanation
- Research Data & GDPR

Obligations for Data Processors

- Introduces a set of novelties with respect to the previous Directives especially concerning the obligations of the Data Processor
  - Which processes personal data on behalf of the controller

- Data Processors have direct obligations for the first time (Articles 28-37)

- It must maintain a written record of all processing activities carried out on behalf of each controller
  - The documentation has to include details about any processing activity, about any transfer to third countries and description of the technical and organizational security measures.

outside the EU States?

- It tries to catch data controllers and data processors outside the EU

- The basic idea is that a non-EU company which is targeting EU consumers will be subject to the GDPR

- This aspect is particularly interesting in the context of Cloud Infrastructures since its nature does not assure that data will stay in EU
  - data transfer necessary for various reasons*: reduction of costs, redundancy and reliability, backup operations, performances*

Accountability Principle

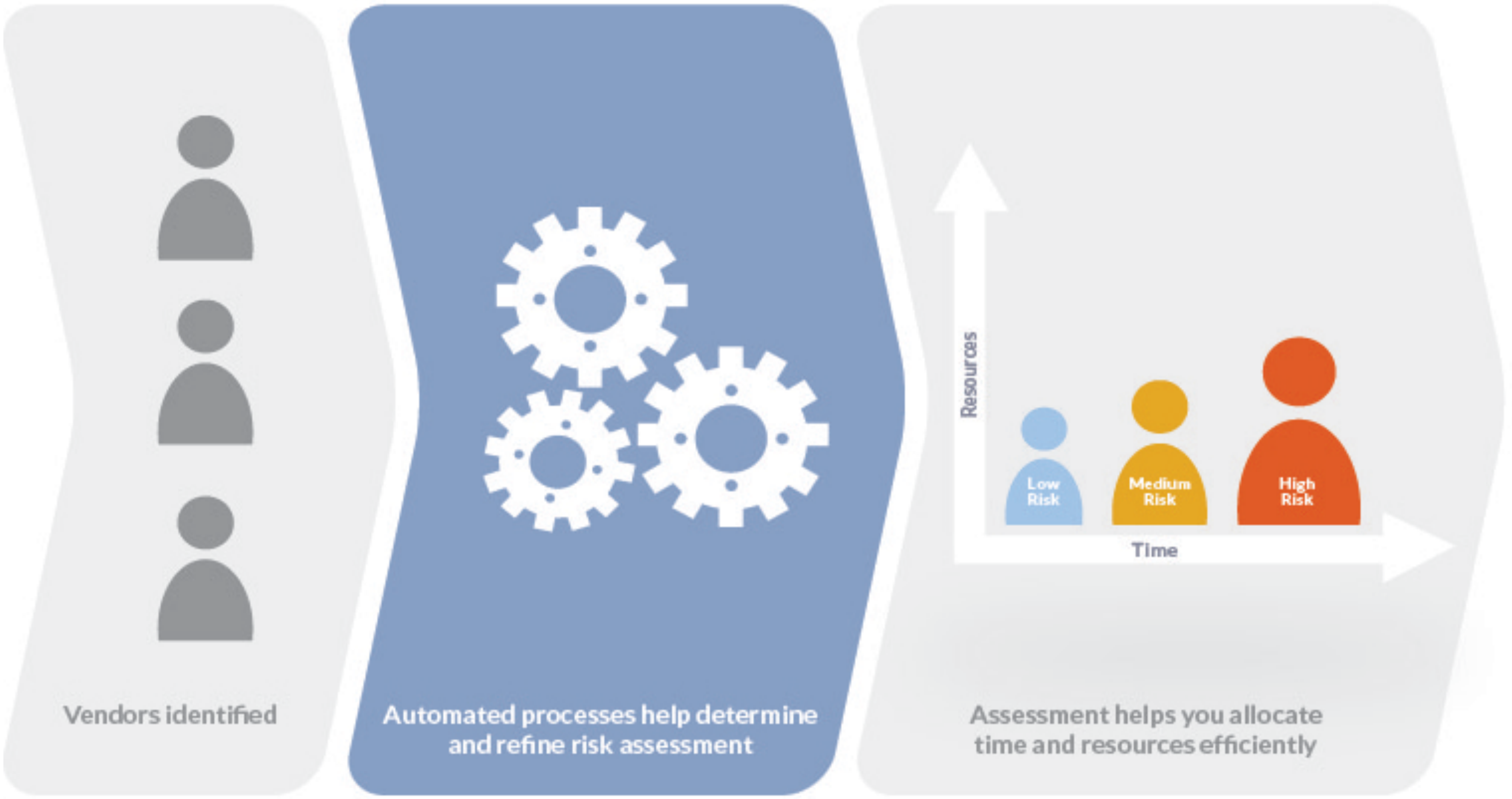
- Data controllers have to show how they comply with the rules
  - E.g. by documenting the decisions taken about a processing activity (Article 5(2)).
- To demonstrate compliance the Data Controller shall:
  - Implement technical and organizational measures ensuring and demonstrating compliance;
  - Maintain relevant documentation on processing activities;
  - Conduct a **data protection impact assessment for more risky processing;**
  - Implement **data protection by design and by default**, e.g. data minimization, pseudonymization; transparency; creating and improving security features.

Privacy by Design

- Data Controllers and Data Processor must implement appropriate security measures and data protection by design and by default.
- What does it mean appropriate?
  - The appropriate measures depend on different factors: level of sensitivity of the data, the evaluation of the **risks associate to individuals**, etc.
- Data Controllers and Processors have to also test regularly the effectiveness of any security measures adopted

PRIVACY RISK ASSESSMENT IN BIG DATA ANALYTICS AND USER-CENTRIC DATA ECOSYSTEMS

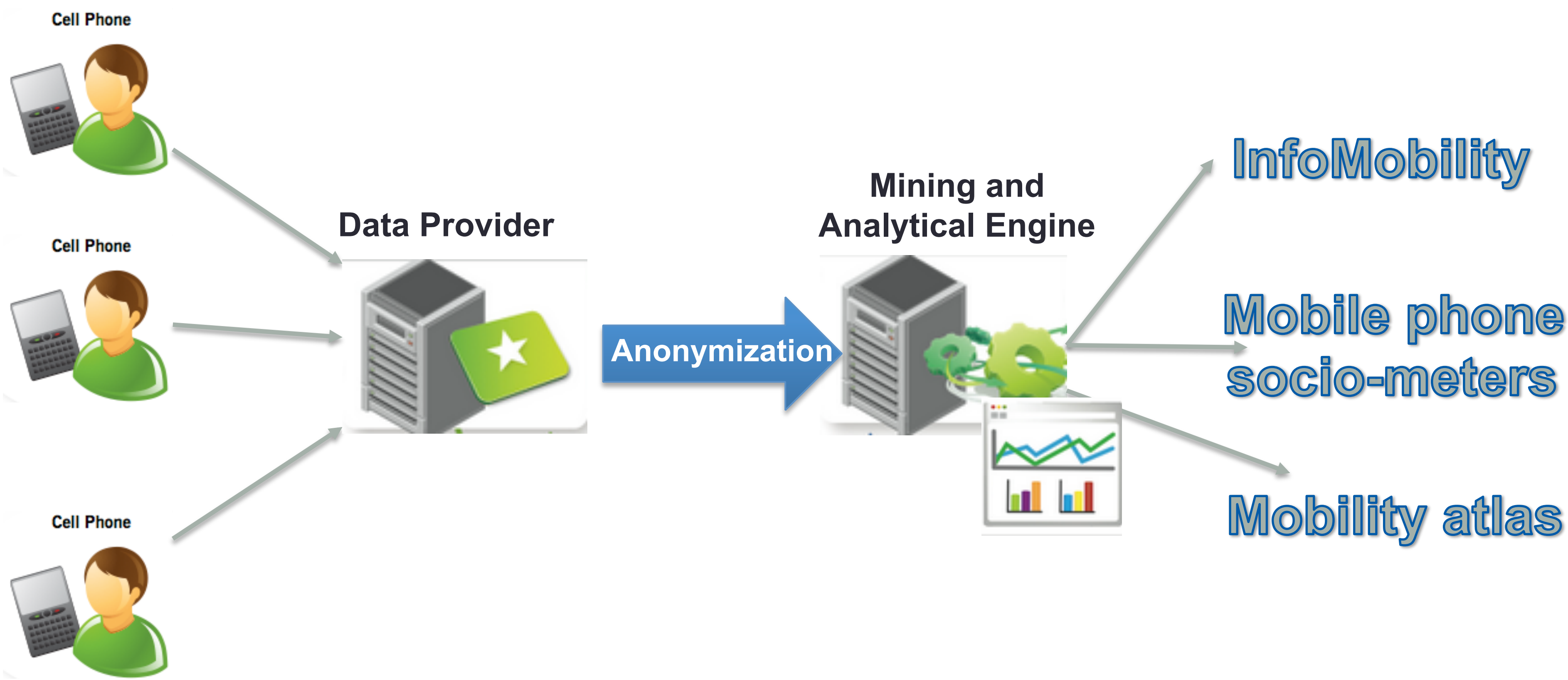
Privacy Risk Assessment



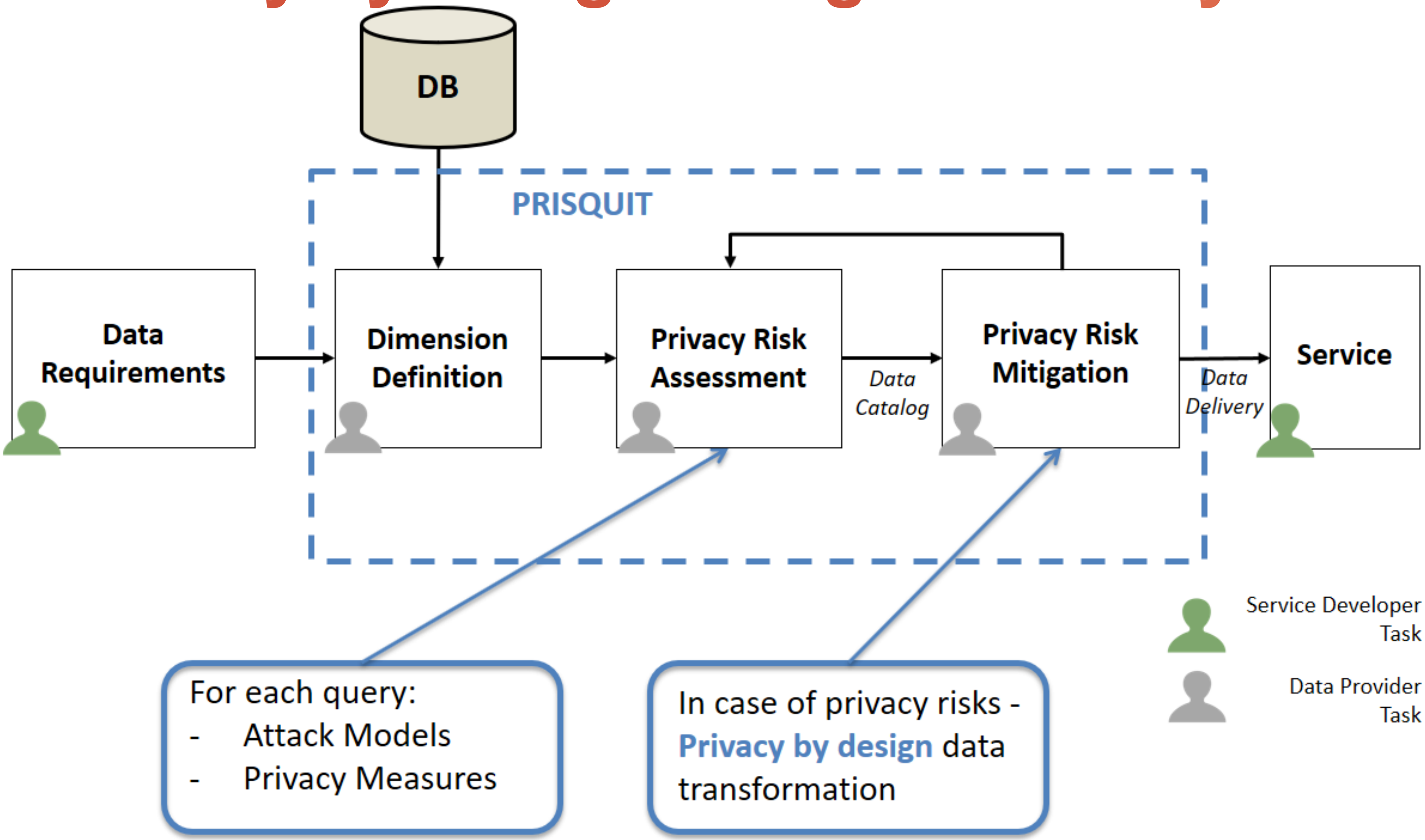
Privacy by Design in Big Data Analytics

- Design frameworks
  - to counter the threats of privacy violation
  - without obstructing the knowledge discovery opportunities of data analysis
- Trade-off between privacy quantification and data utility

Privacy-by-Design in Big Data Analytics



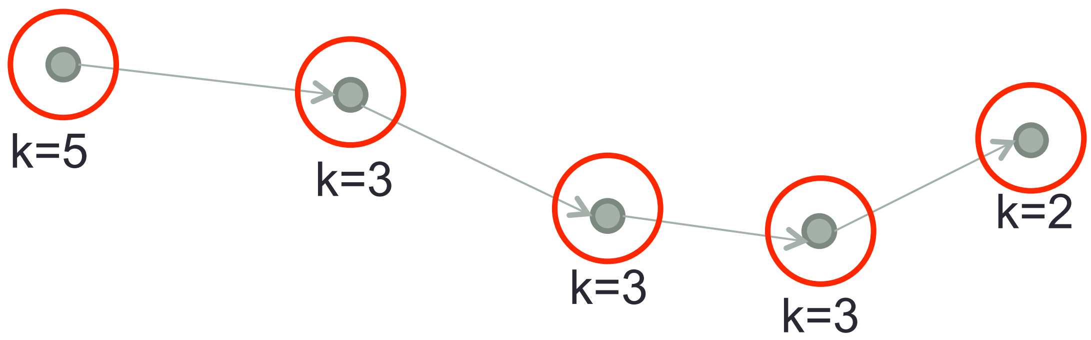
Privacy-by-Design in Big Data Analytics



Privacy risk measures

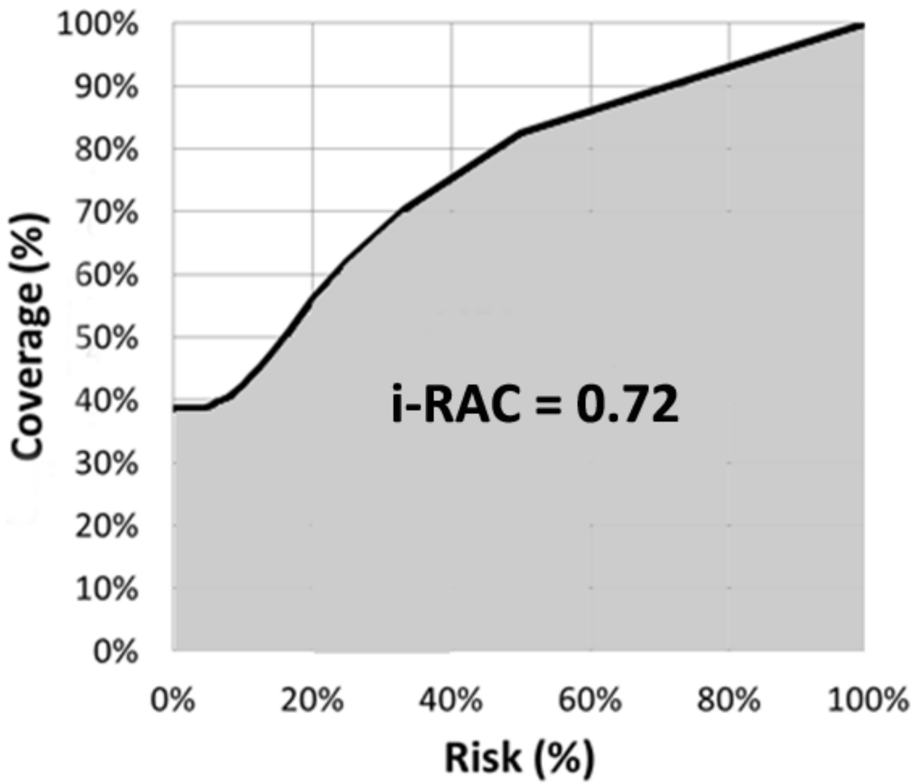
Probability of re-identification denotes the probability to correctly associate a record to a unique identity, *given* a BK

Risk of re-identification is the maximum probability of re-identification *given* a set of BK



Risk and Coverage (RaC) curve

- A diagram of coverage (% of data preserved) at varying values of risk
- Concept has analogies with ROC curves.
- Each curve can be summarized by a single measure, e.g. AUC (area under the curve) – the closer to 1, the better



RAC_U → for each risk value, quantifies the percentage of users in U having that risk

RAC_D → for each risk value, quantifies the data in D covered by only users having at most that risk

The approach

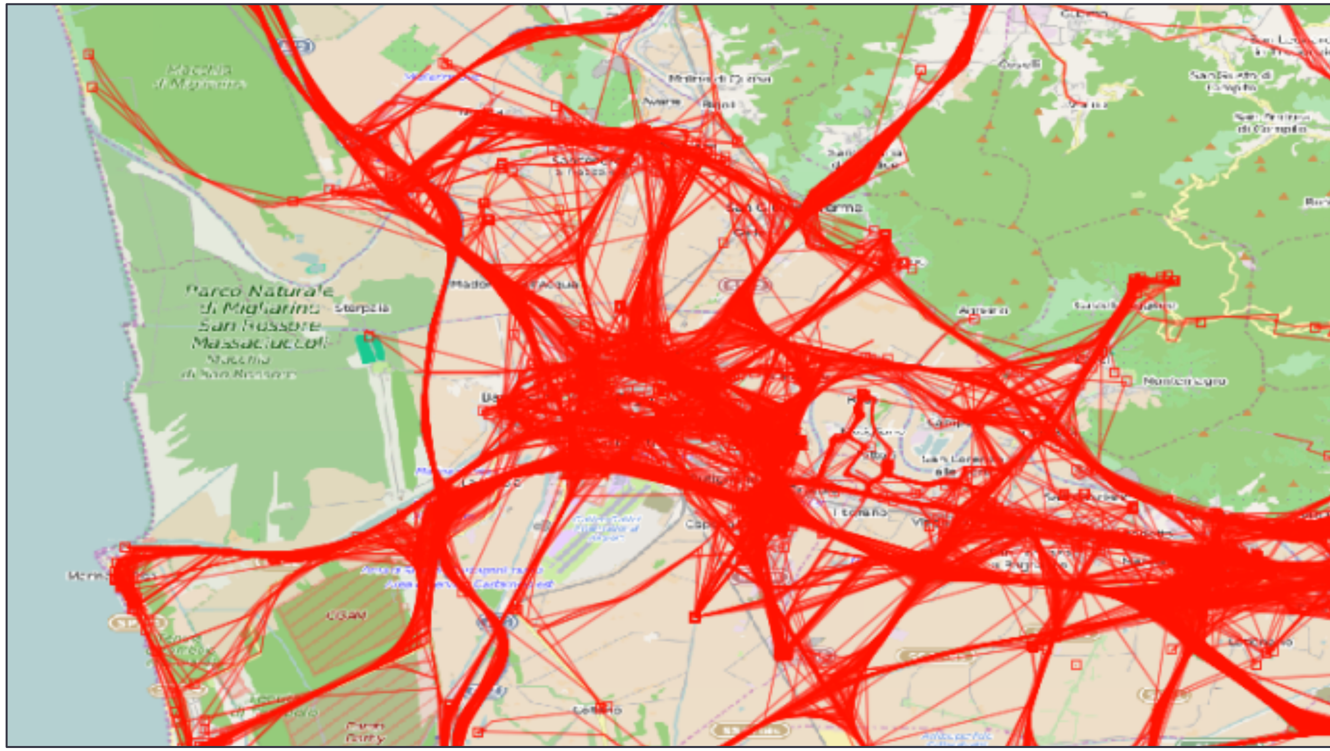
Generalize from exemplary set of services (data, query, requirements, BK, risk)

Key issue: the language of BK – how to specifies the set of possible attacks in a general way for mobility data.

Several kinds of data:

- **presence** (individual frequent locations)
- trajectory (individual movements)
- road segment (collective frequent links)
- profiles (individual systematic movements)
- individual call profiles (from CDR data)

Data Statistics



Area Covered: 726 Km^2

Number of trajectories: 247.633
Number of users: 10.355
Temporal window: 1 month

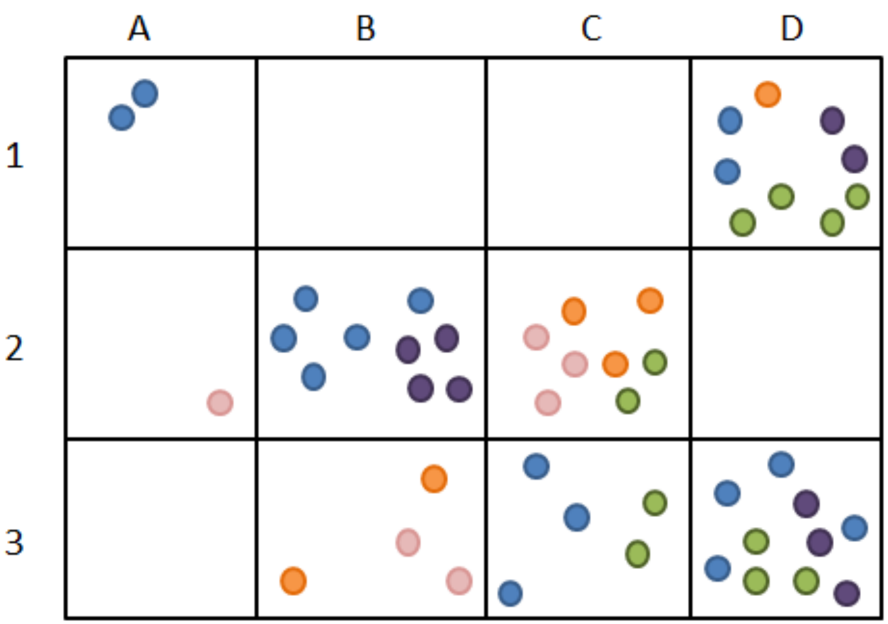
Only active users are selected: at least 7 trajectories in 1 month.

Number of trajectories: 235.306
Number of active users: 3.780
***Temporal window*: 1 month**

Data description

For each user, list of locations (grid cells) that the user has frequently visited (#visit>threshold)

User_id, Cell id



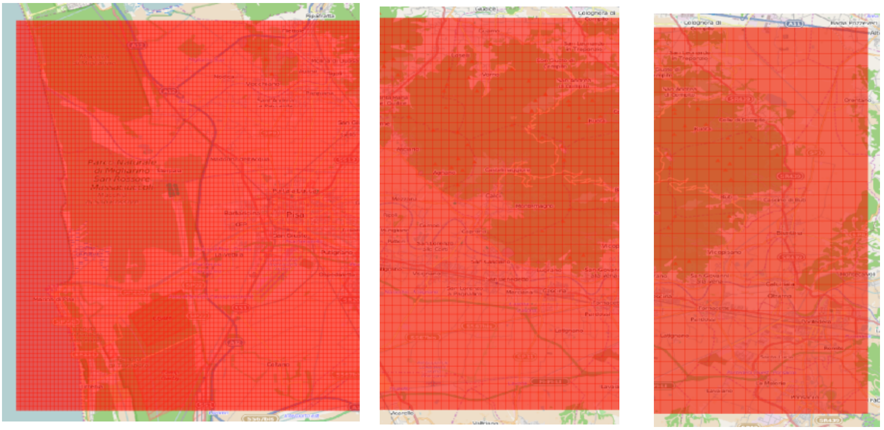
Blue: <B2,5>,<D3,4>,<C3,3>,<A1,2>,<D1,2>
Green: <D1,4>,<D3,3>,<C2,2>,<C3,2>
Orange: <C2,3>,<B3,2>
Purple: <B2,4>,<D3,3>,<D1,2>
Pink: <C2,3>,<B3,2>

Data Dimensions

Grid size: defines the granularity of the spatial information released about each user

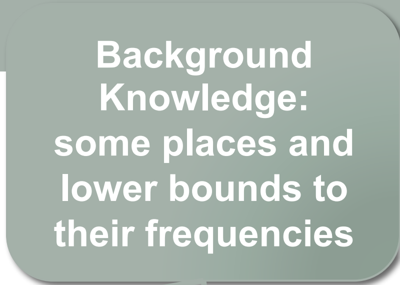
Frequency threshold: defines a filter on the data DO can distribute

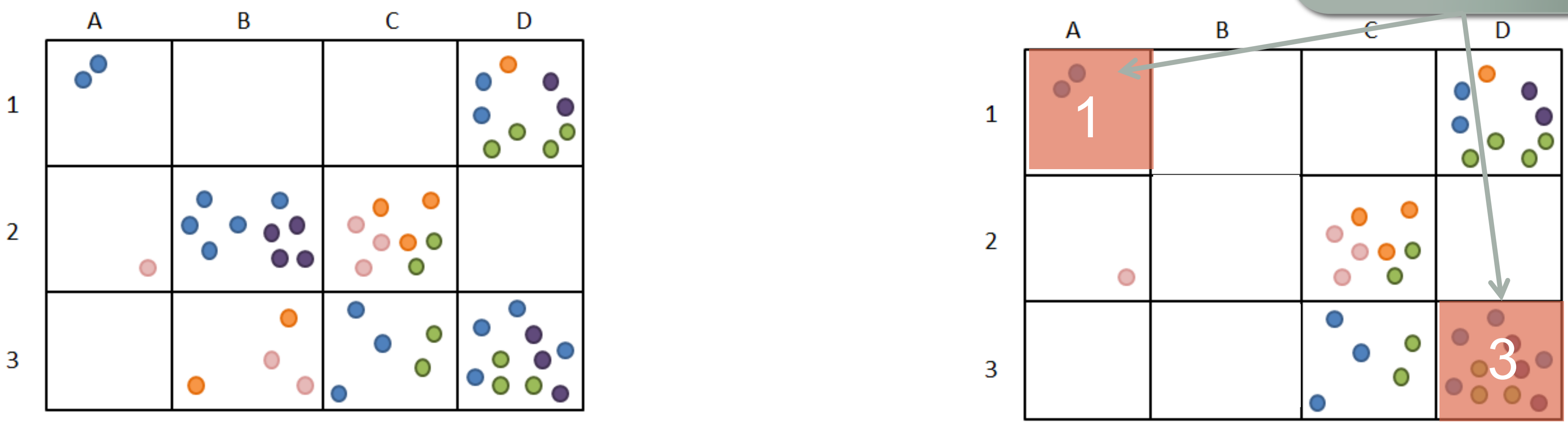
Spatial granularity used:
Grids (cell side): 250, 500 and 750 meters



Frequency threshold: 1, 4, 7, 10, 13

Attack: Casual observation





The attacker knows some location(s) with minimum frequencies

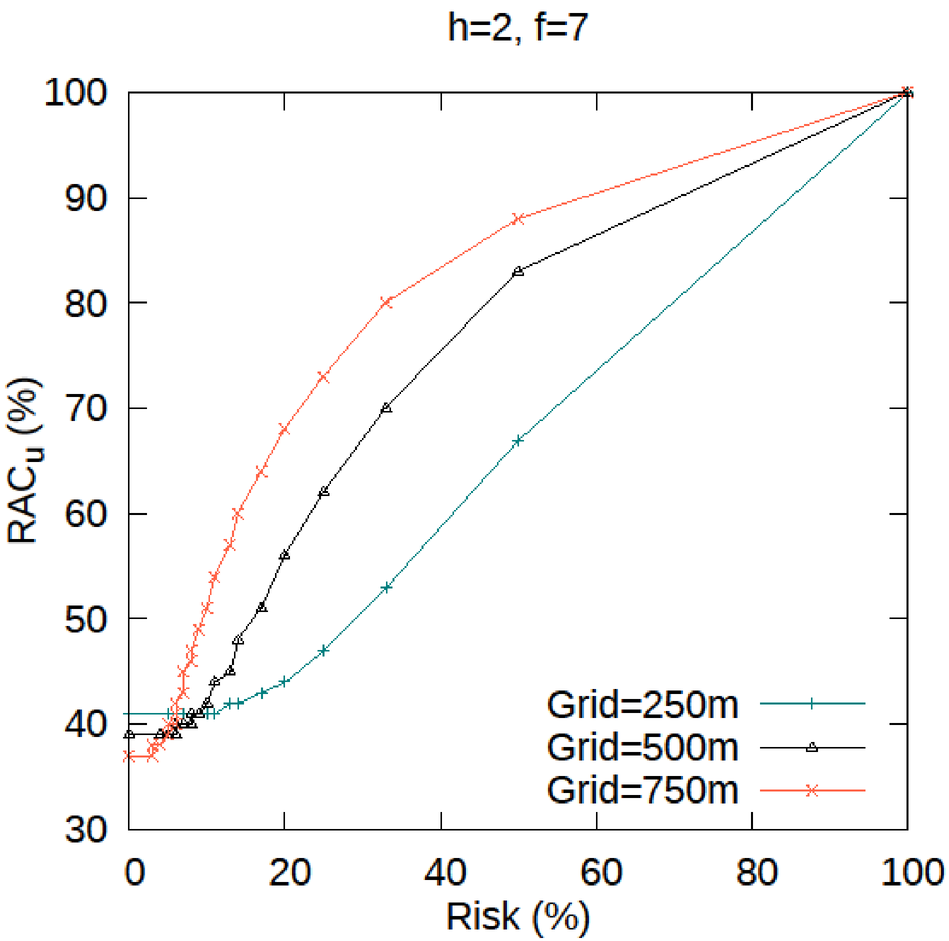
Background Knowledge Dimensions:

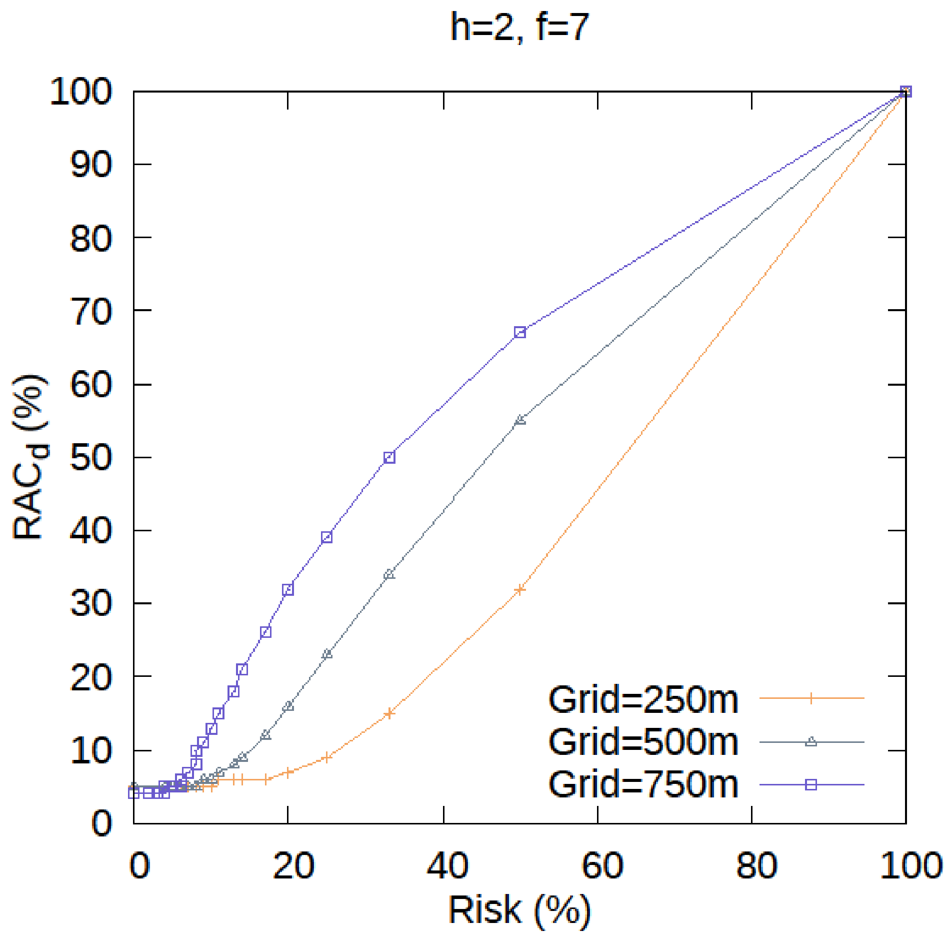
- Number of locations known (h = 1, 2, 3)
- Minimum frequency associate to the known locations (100% of original freq, 50% of original freq, only presence)

E.g., Mr. Smith was seen once in A1 and 3 times in D3

Simulation Attack Model

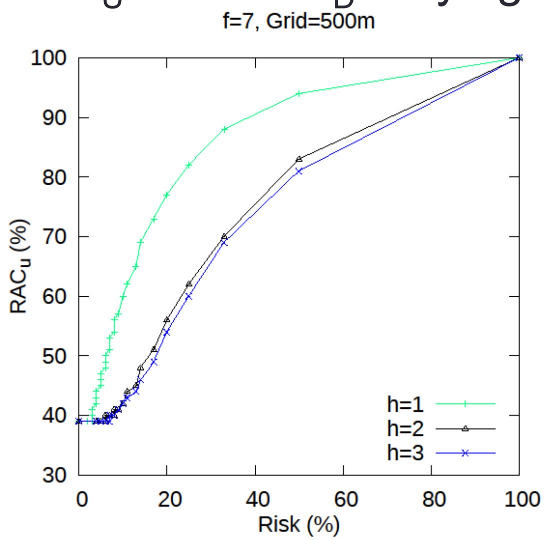
RAC_U and RAC_D varying the **grid** and fixing #location and frequency

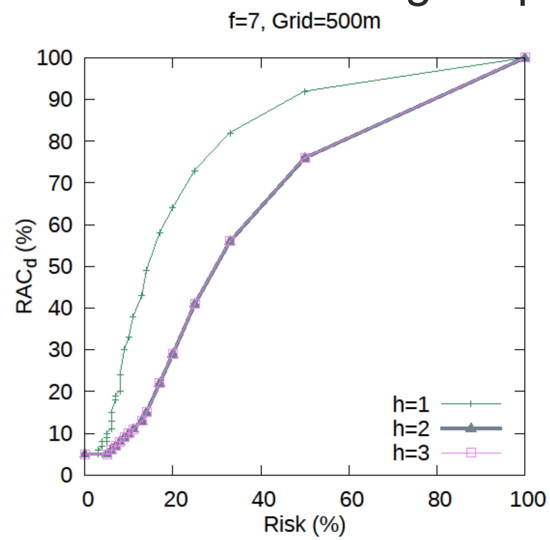




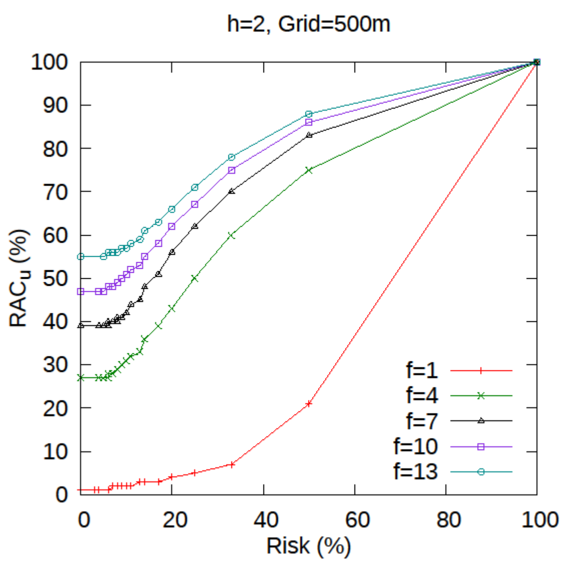
Simulation Attack Model (2)

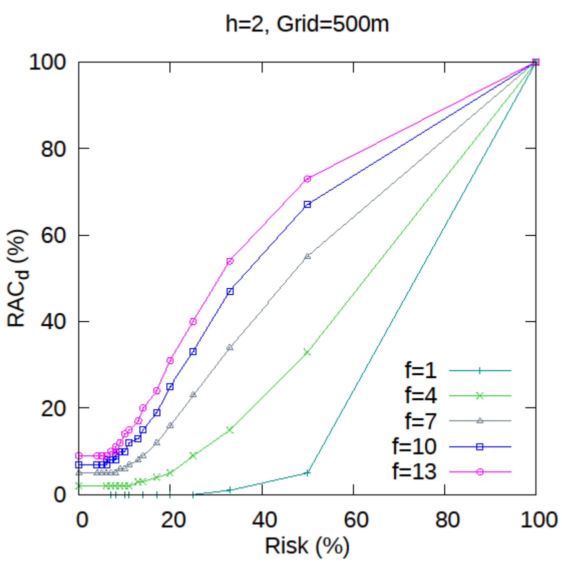
RAC$_U$ and RAC$_D$ varying the **#location** and fixing frequency and grid



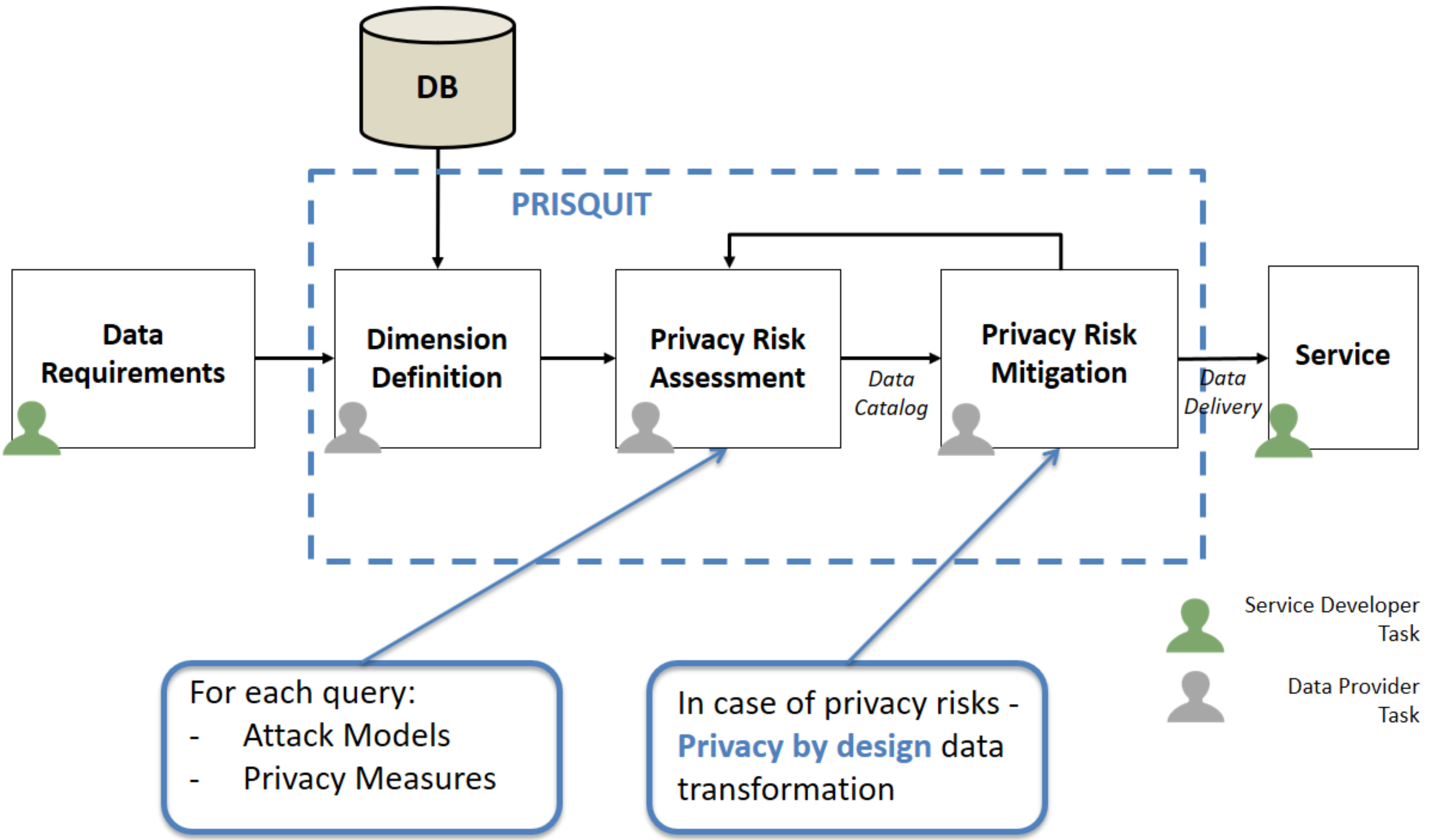


RAC$_U$ and RAC$_D$ varying the **frequency** and fixing #location and grid





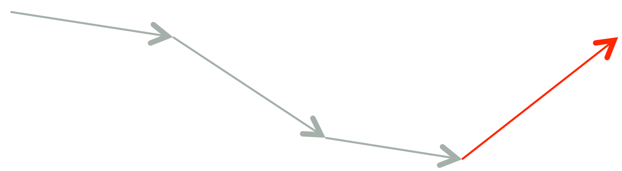
Privacy-by-Design & Risk Assessment in Big Data Analytics



Privacy-by-Design requirements

Data Dimensions

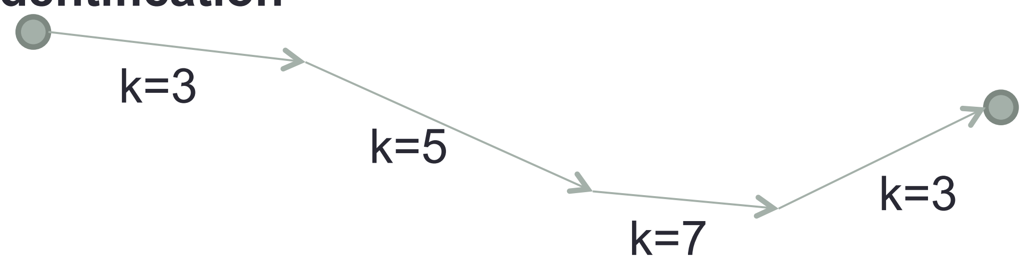
Background Knowledge Dimensions



For each combination we simulate an attack and empirically quantify the privacy risk

Probability of re-identification

Risk of re-identification



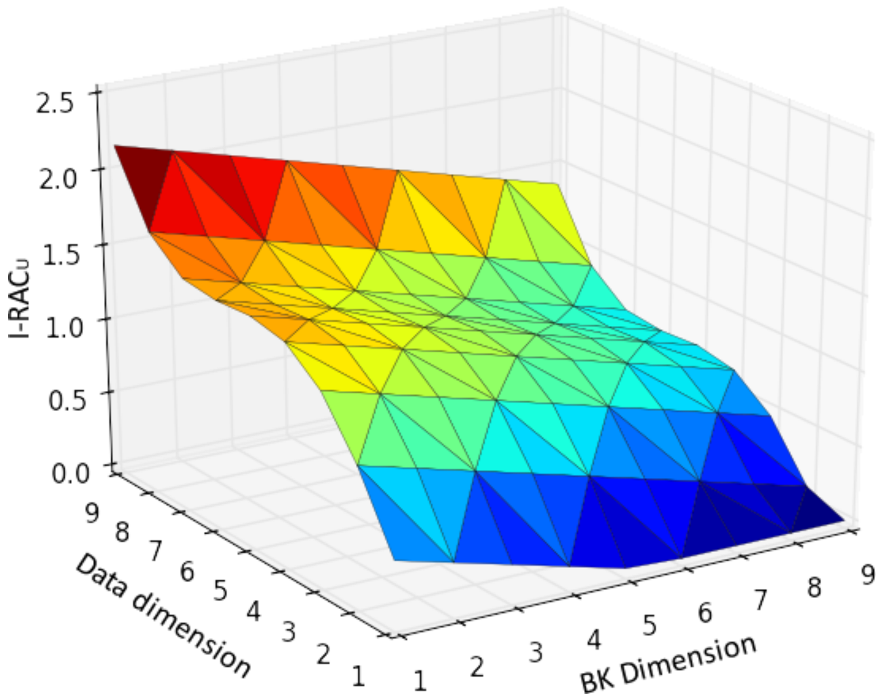
Data Catalog

For each:

- **Data Format**, i.e., the data needed for the service
- **Risk Assessment Setting**, i.e., the set of pre-processing and Background Knowledge

The Data Catalog provides:

- **Quantification of Privacy Risk**, i.e., the evaluation of the real risk of re-identification
- **Quantification of Data Quality**, i.e., the quality level we can achieve with private data, compared with the data quality of original data.



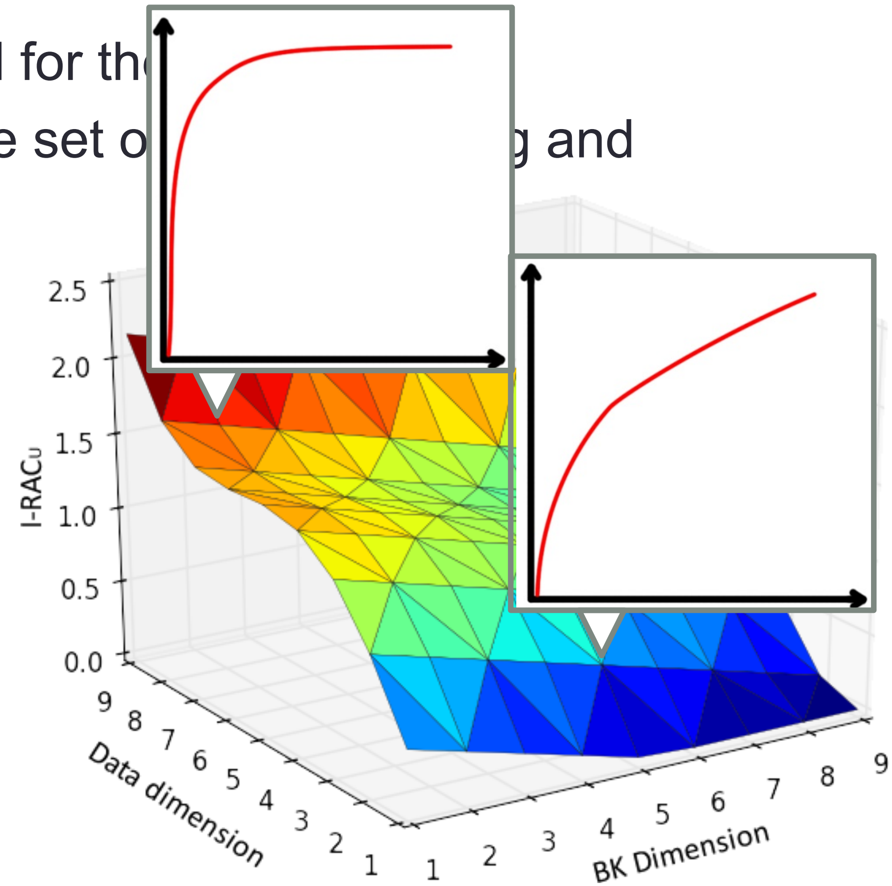
Data Catalog

For each:

- **Data Format**, i.e., the data needed for th
- **Risk Assessment Setting**, i.e., the set o g and Background Knowledge

The Data Catalog provides:

- **Quantification of Privacy Risk**, i.e., the evaluation of the real risk of re-identification
- **Quantification of Data Quality**, i.e., the quality level we can achieve with private data, compared with the data quality of original data.



Risk Assessment in Mobile phone socio-meters Analysis

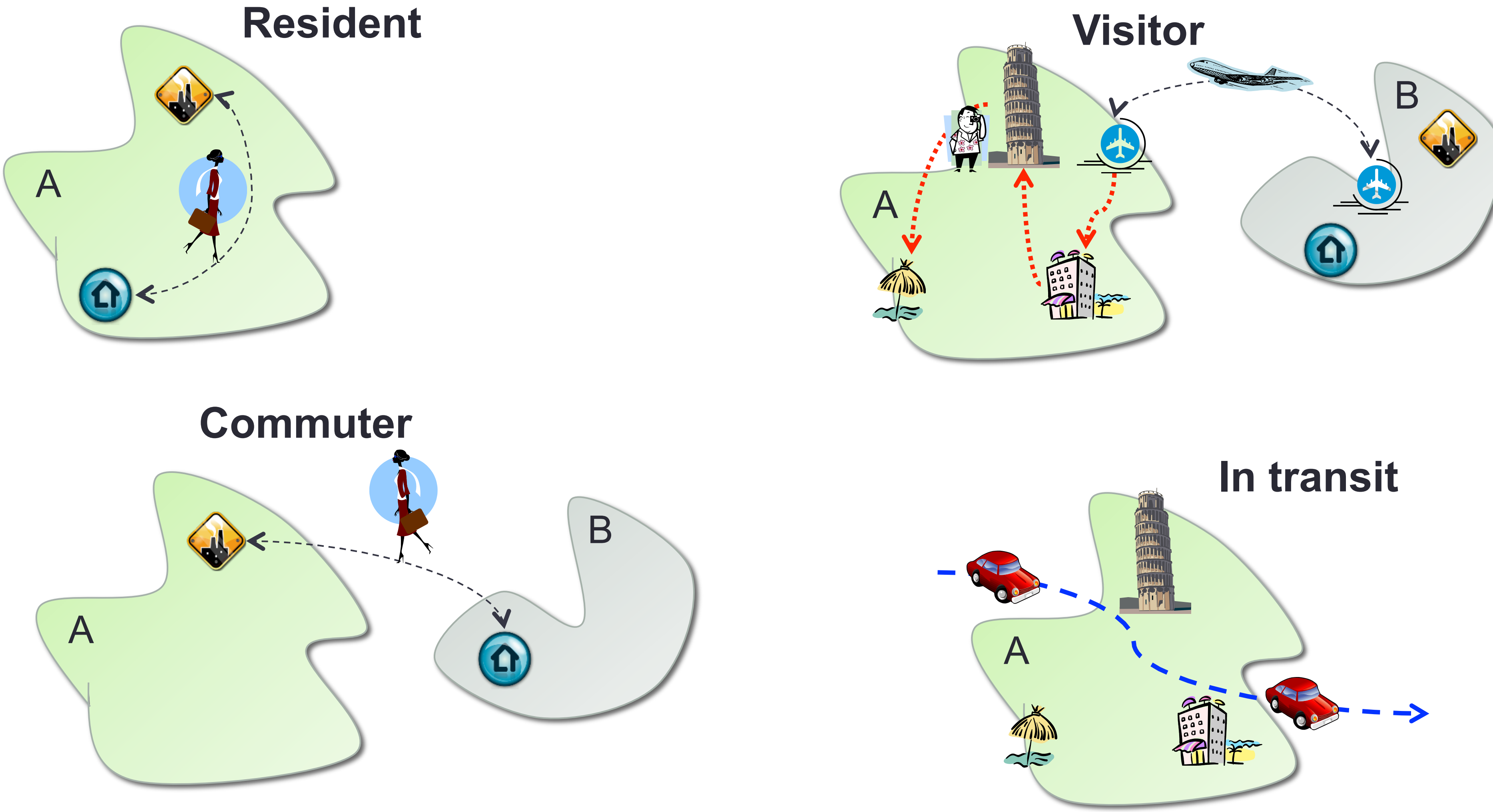
A. Monreale, S. Rinzivillo, F. Pratesi, F. Giannotti, D. Pedreschi

Knowledge Discovery and Delivery Lab
(ISTI-CNR & Univ. Pisa)
www-kdd.isti.cnr.it

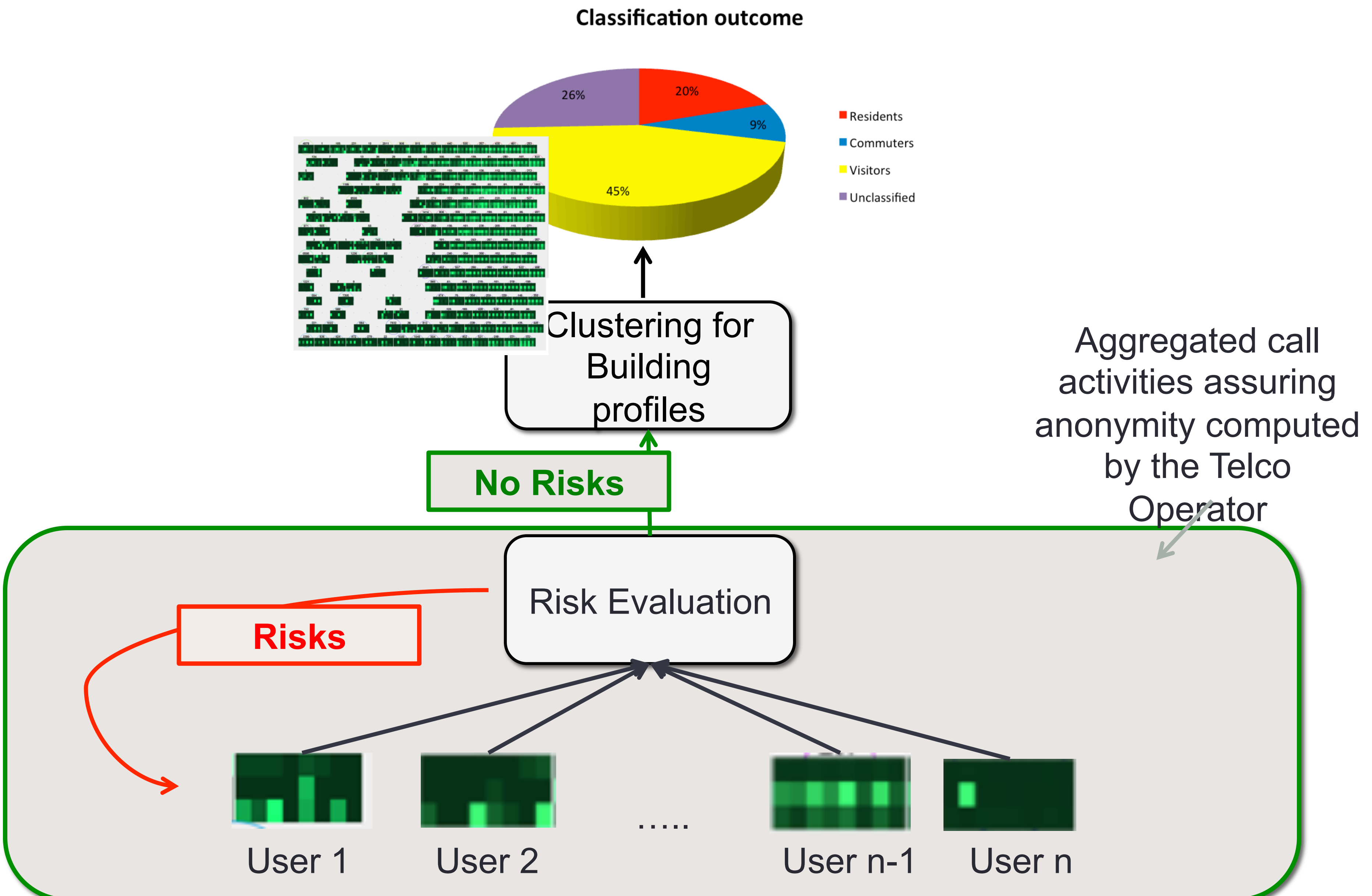
Objective

- To partition of the users tracked by GSM phone calls into the following main categories:
  - **Residents**
  - **Commuters**
  - **Visitors/Tourists**
  - **People in transit**

User Categories



Privacy-Aware socio-meter



Attack risk based on Call Activities

Analyst working on GSM data of 2M users with access to their call profiles

From: 02/11/15

To: 22/11/15

Apriori knowledge:
3 weeks of her boy-friend's call activity

Inference:
his activities in Pisa during the remaining week

From: 23/11/15

To: 29/11/15

Assumption: the attacker knows the user is one of the profiles

Example of the attack

Attacker knows *exactly* the call made by U in the first 3 weeks

	week 1		week 2		week 3		week 4	
morning	1						?	?
afternoon		2			1	1	?	?
evening	1		3	1	2		?	?

Example of the attack

Attacker knows *exactly* the call made by U in the first 3 weeks

	week 1		week 2		week 3		week 4	
morning	1						?	?
afternoon		2			1	1	?	?
evening	1		3	1	2		?	?

1							
	2			1	1	2	
1		3	1	2			1

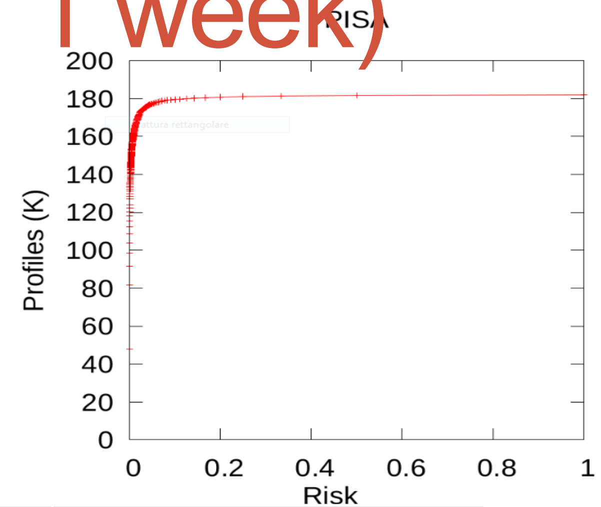
1							2
	2			1	1		
1		3	1	2		3	

K= 2

Experimental results

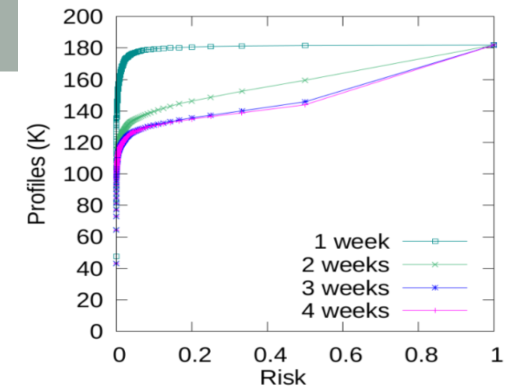
- Where: Tuscany
- When: from 2/11/2015 to 29/11/2015
- Who: 2.121.331 users

Privacy Risk for Users (bk: 1 week)



Risk	K	% users	Number users
r<=0.01%	K>=10.000	50,37	91.613
0.01%<r<=0.1%	1000<=K<10.000	22,28	40.514
0.1%<r<=1%	100<=K<1.000	16,59	30.179
1%<r<=2%	50<=K<100	4,78	8.688
2%<r<=10%	10<=K<50	4,64	8.434
10%<r<=20%	5<=K<10	0,66	1.213
r>20%	1<=K<5	0,67	1.225

Privacy Risk for Users of Pisa



		bk: 1 week		bk: 2 weeks		bk: 3 weeks		bk: 4 weeks	
Risk (r)	**K**	**% users**	**# users**	**% users**	**# users**	**% users**	**# users**	**% users**	**# users**
r<=0.01%	K>=10.000	50	91.613	40	73.141	40	73.001	40	73.001
0.01%<r r<=0.1%	1000<=K K<10.000	22	40.514	16	29.595	14	26.311	14	26.328
0.1%<r r<=1%	100<=K K<1.000	16	30.179	11	19.707	9,6	17.494	9,5	17.381
1%<r r<=2%	50<=K K<100	4,8	8.688	2,7	4.953	2,3	4.244	2,3	4.225
2%<r r<=10%	10<=K K<50	4,6	8.434	6,8	12.322	5,5	10.031	5,3	9.741
10%<r r<=20%	5<=K K<10	0,7	1.213	3,6	6.574	2,5	4.586	2,3	4.170
r>20%	1<=K	0,7	1.225	19	35.57	25	46.19	25	47.00

Privacy-by-design for big data analytics

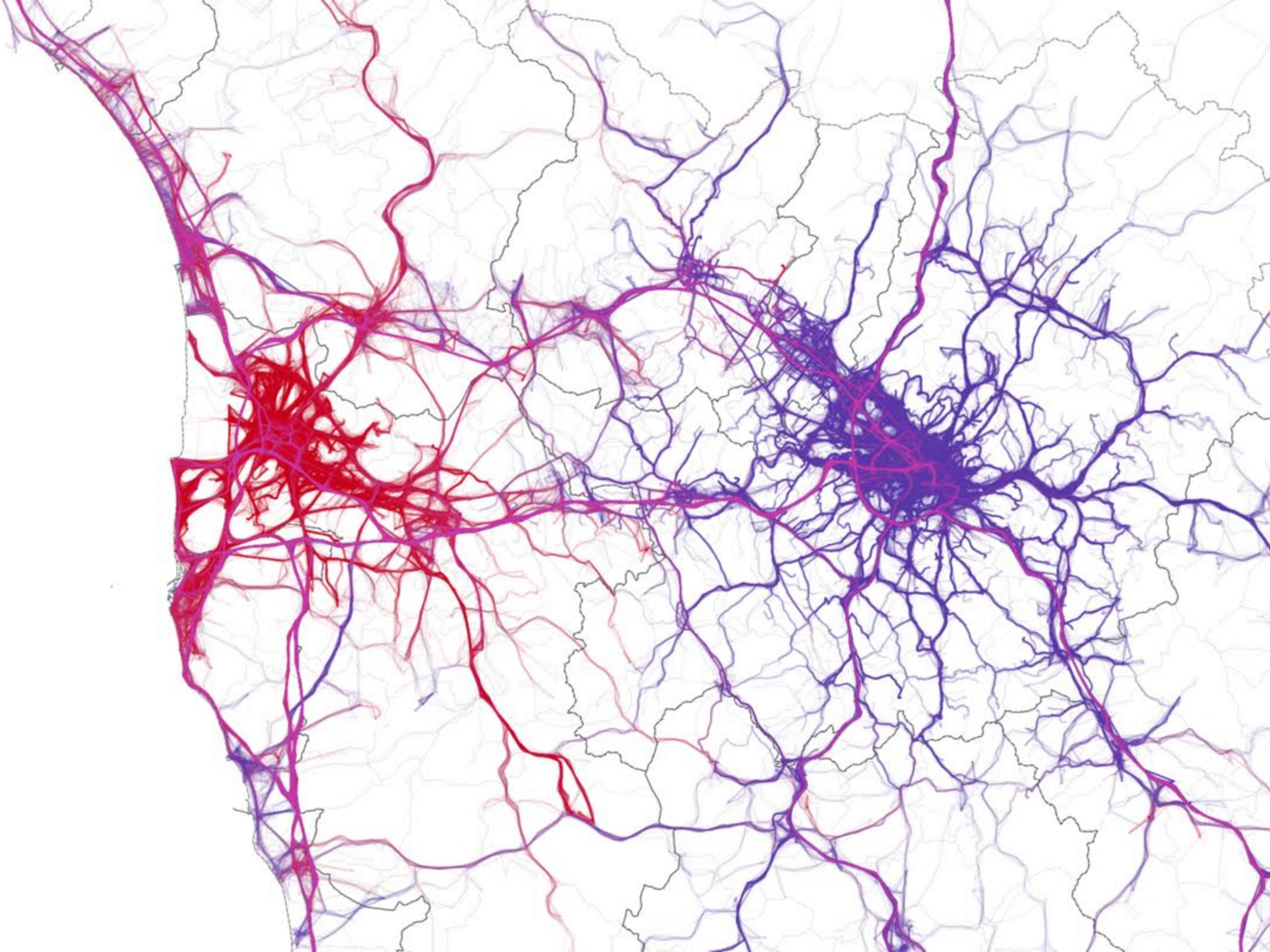
- All case studies discussed have been designed within a **privacy-preserving** framework
- taking into account **data minimization** in the deployment of the service
- transforming raw data into aggregated data with a **quantified** (low) **risk** of privacy breach

Privacy-by-design for big data analytics

- To make data free we need to design privacy-aware analytical frameworks
  - Privacy Risk Analysis
  - Privacy-by-design
- Two different stages:
  - Design data-driven models with sample datasets in safe and responsible research centers (like SoBigData)
  - Deploy data-driven services based on continuous flow of data from the data provider
- Different risk levels

But we need to go further!

- A city cannot be managed centrally, from a control room.
- Our cities are complex networks of interactions
  - the outcome for everybody depends not only on individual choices but it is conditioned by everybody else's choices.



- A granular capability of citizens to self-organize, collaborate and coordinate their actions from the bottom is more efficient and resilient

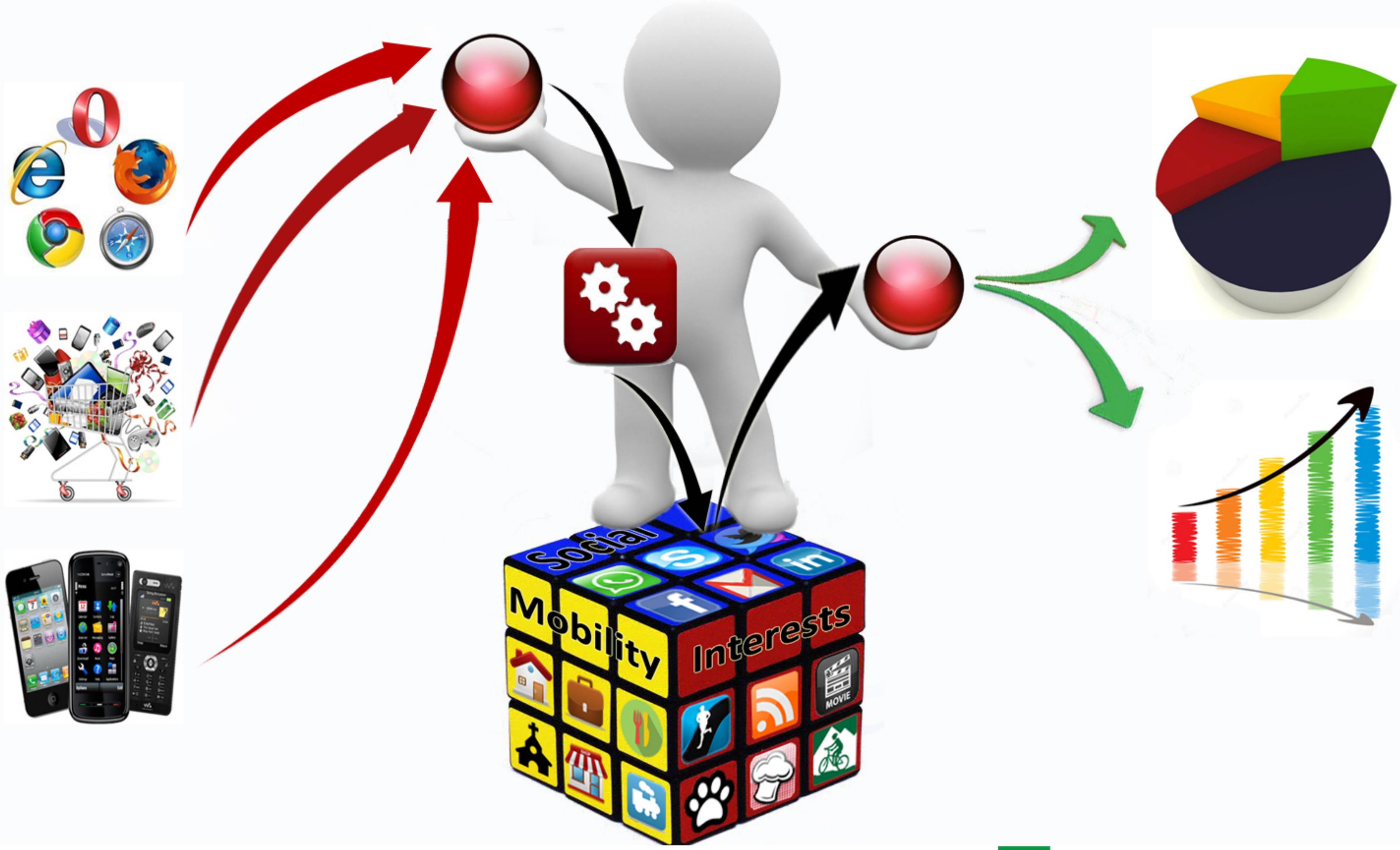
- But requires to align individual interests and goals with those of the collectivity in the system.

  – We humans have a limited perception of ourselves as a social, collective living being

TOWARDS A PERSONAL DATA ECOSYSTEM

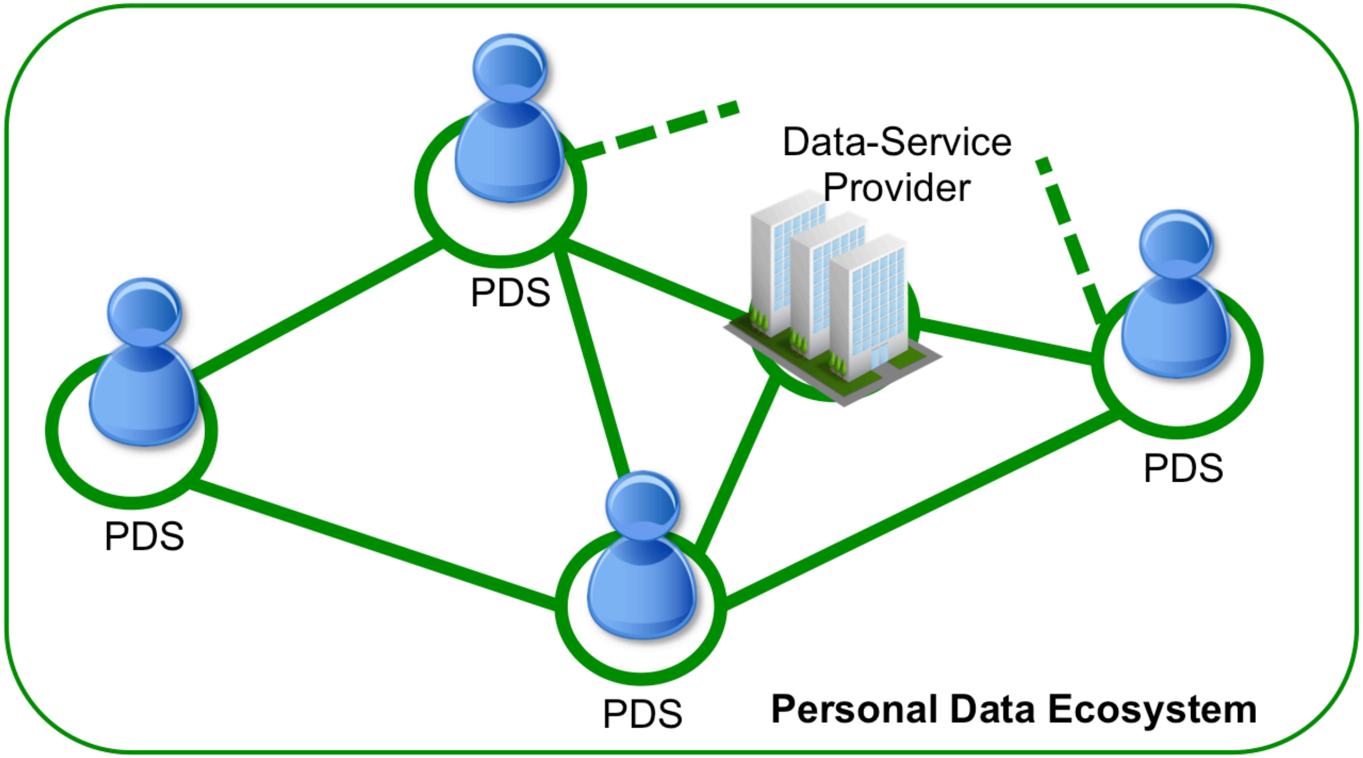
- An avalanche of personal information that, in most cases, gets lost – *like tears in rain*.
- Yet, only each one of us, individually, has the power to connect all this personal information into a personal data repository – and make sense of it.

A user-centric ecosystem for personal big data



Personal Data Ecosystem

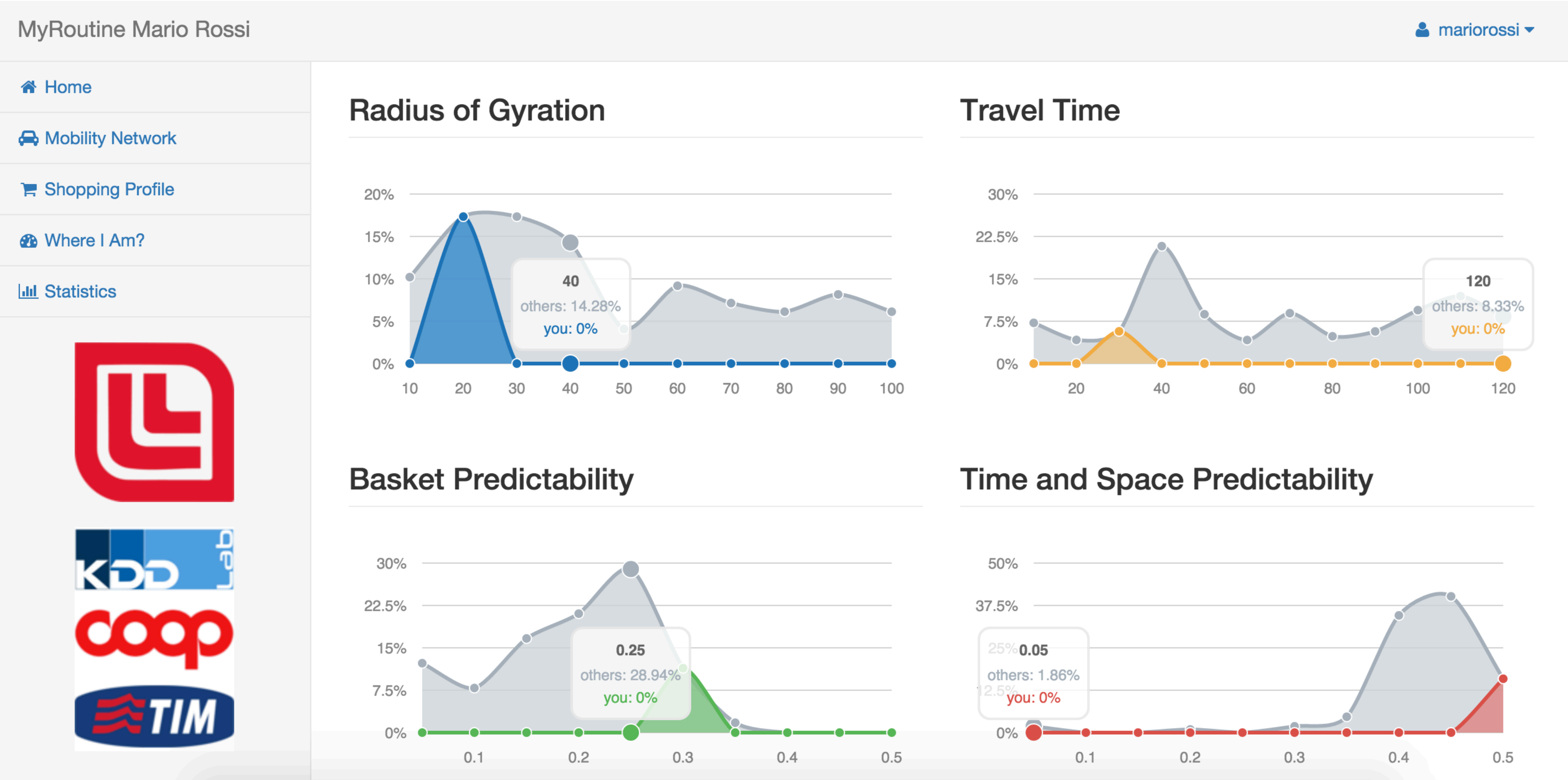
- Personal data collection and knowledge mining need to be balanced with *participation*, based on a greater awareness of the value of own personal data for each one of us and the communities that we inhabit, at all scales.



PDE functions

- Continuous acquisition and integration of personal data from user's transactions and other public sources
- Making sense of own personal data
  - "the myself emerging from my digital traces"
- Peer-to-peer interaction network for exchanging information based on question answering
  - trust and reputation, risk vs. benefit assessment, security and traceability
- Participatory social mining of collective patterns with privacy-aware computing models
  - from fully distributed to trusted third parties.
- Making sense of own patterns compared to collective patterns

Where am I? Comparison with the community

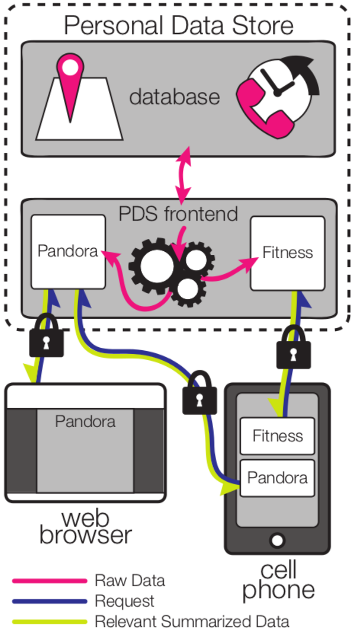


- We need a Personal Data Ecosystem
  - to acquire, integrate and make sense of our own data
  - and to connect with our peers and the surrounding urban infrastructure
- to the purpose of developing the **collective awareness** needed to face our grand challenges

Setting The Stage

The Personal Data Store

- **My Data Store: Toward User Awareness and Control on Personal Data**, Michele Vescovi et all. Ubicomp, 2014
- **openPDS: Protecting the Privacy of Metadata through SafeAnswers**, Yves-Alexandre de Montjoye, Erez Shmueli, Samuel S. Wang, Alex Sandy Pentland, PlosONE, 2014
- **Managing Your Digital Life**, Serge Abiteboul, Benjamin André, and Daniel Kaplan, Commun, 2015
- **Weaknesses:**
  - Data explanation: NO
  - Data exploitation: NO
  - Data comparison: NO

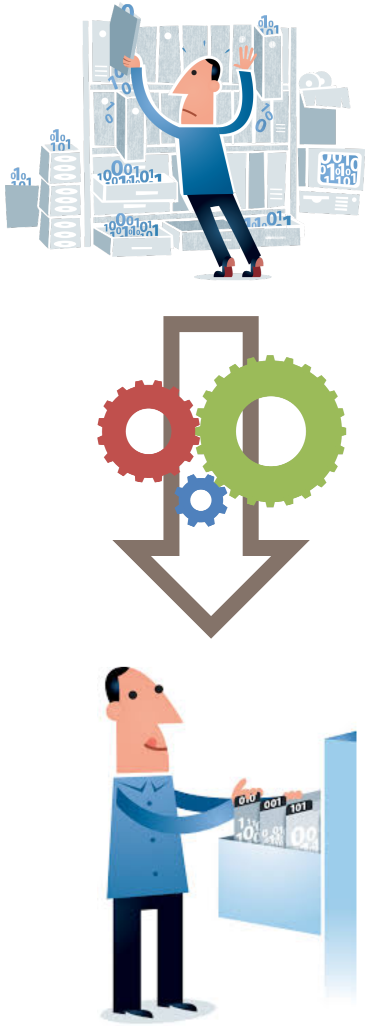


Personal Data Analytics

We All Need to Understand and Exploit Our Data

- Data Mining applied to Personal Data is the *key for extracting personal patterns* and, consequently to creates opportunities for *enabling personalized services*, and to *improve the user self-awareness*.

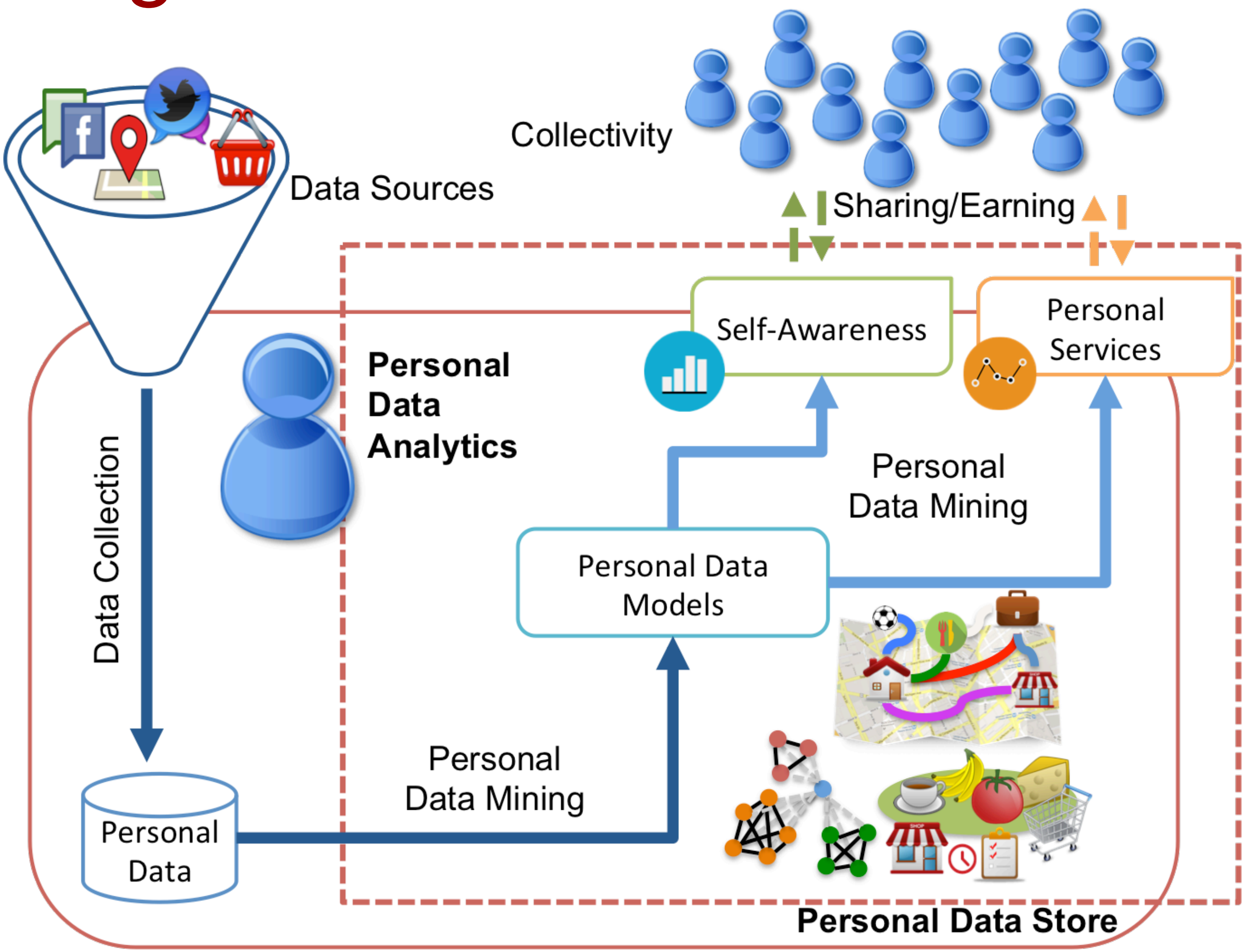
- Despite some novel user-centric model are being defined, in the current state-of-the-art there is yet a *lack of algorithms and models specifically designed* to extract knowledge from personal data.



Personal Data Analytics

- We define how to extend the idea of Personal Data Store by articulating a *Personal Data Analytics* approach that seeks to *analyze the digital breadcrumbs* an individual leaves behind, and we demonstrate that the defined approach and the resulting analyses can lead to *increased individual and collective benefits*.
- Indeed, a key element of Personal Data Analytics is the analytical reinforcement resulting from the synergy of the widespread knowledge in the *Personal Data Ecosystem*.

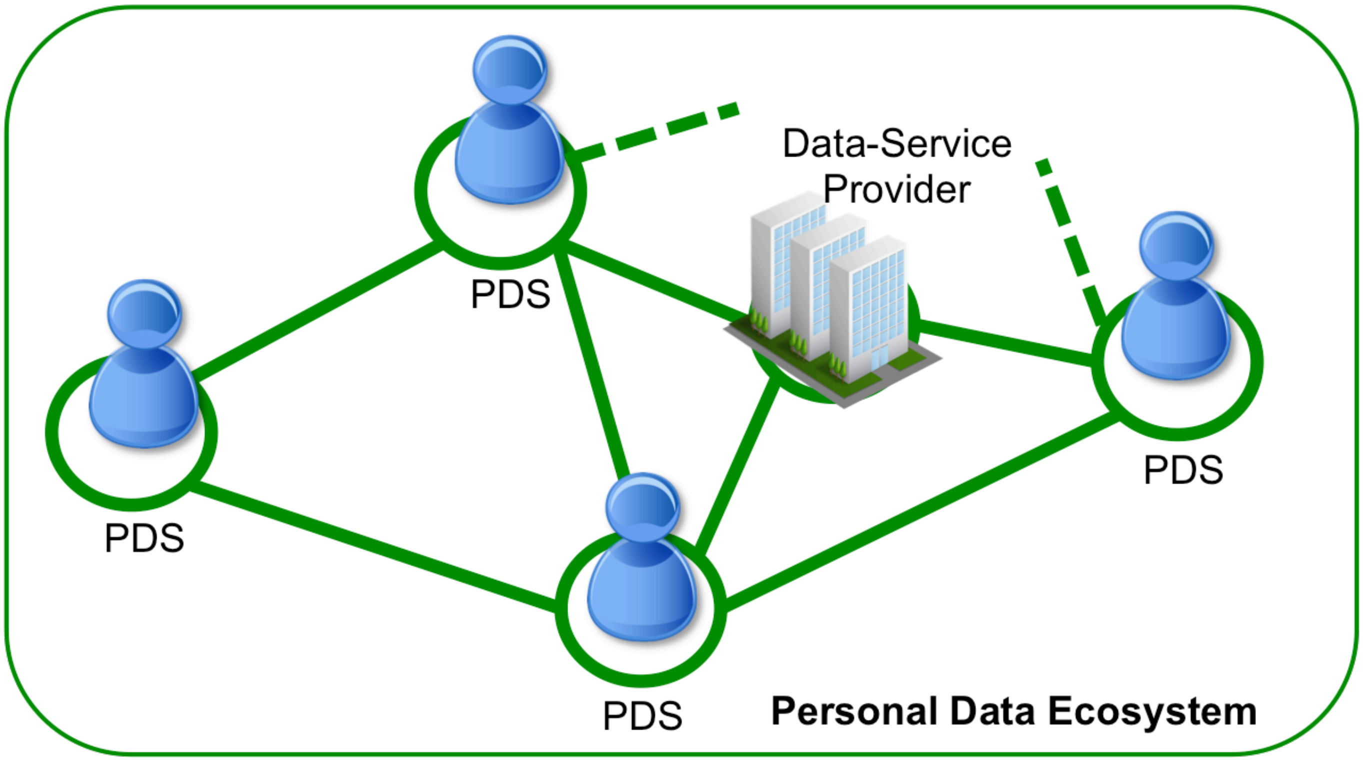
Making Sense of Own Personal Data



Personal Data Analytics

The Personal Data Ecosystem

- Personal data collection and knowledge mining need to be balanced with *participation*, based on a much greater awareness of the value of own personal data for each one of us and the communities that we inhabit, at all scales.



Personal Data Analytics

Potentialities and Socio-Economic Impact

- From a scientific/technological perspective, the PDE has the potential to support the development of *a new generation of user-centric, data-driven services* that empower people in their interaction with service providers at all scales.
- It could support *sharing economy*, and *liquid democracy*.
- The ultimate impact comes from fostering *a reinforcement spiral between collective awareness and individual self-awareness*.

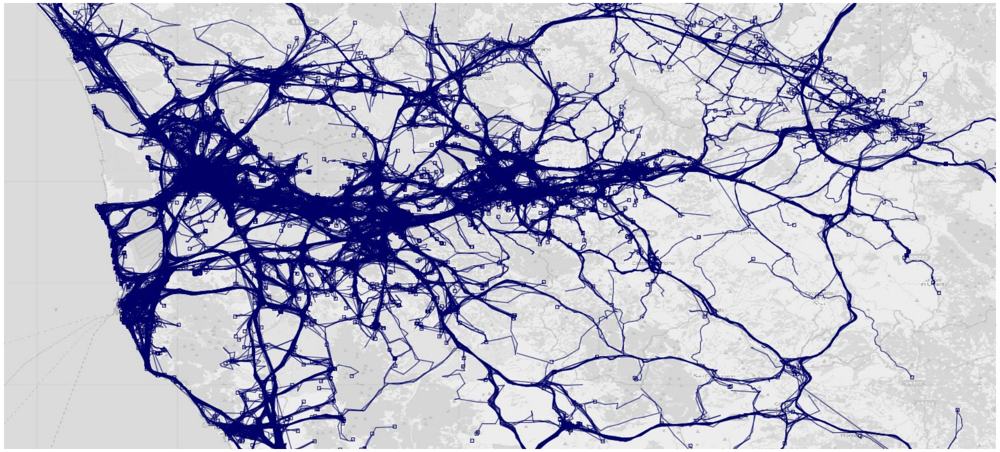
Personal Mobility Data Analytics

Mobility Dataset

- OctoToscana2011
  - GPS Points (lat, lon, ts)
  - 9.8 million car travel
  - 160.000 vehicles
  - 1st May to 31st May 2011
  - Tuscany

OCTO

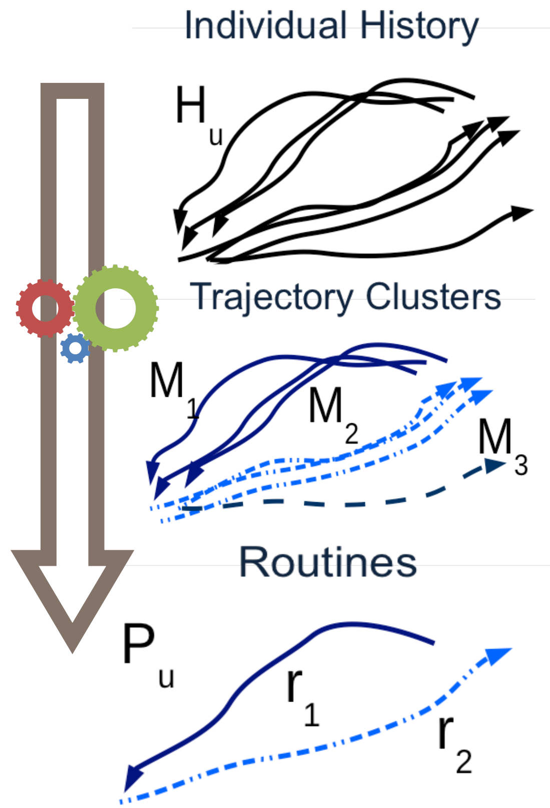
The reliable way

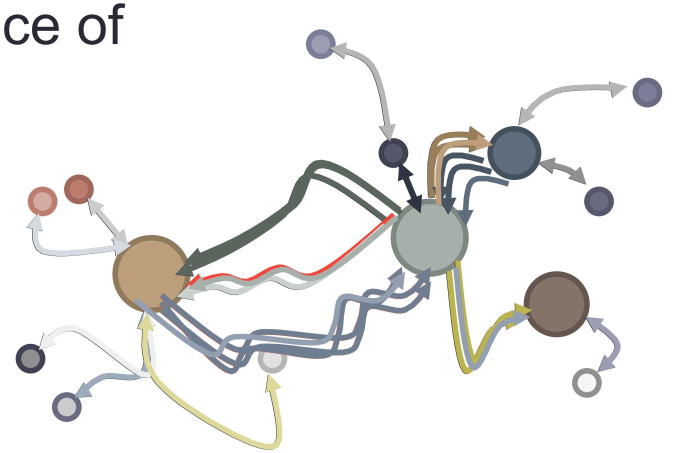


Methods & Models

Personal Mobility Model

- Besides locations, the mobility of a user is characterized by the *trajectories* that start and end in the user's personal locations.
- These trajectories can be *clustered* with respect to their similarity.
- From each cluster can be extracted a representative trajectory, named *routine*.
- The set of routines, i.e., the *individual mobility profile* P_u, is an abstraction in time and space of the systematic movements of a user.





Roberto Trasarti, Fabio Pinelli, Mirco Nanni, Fosca Giannotti: Mining mobility user profiles for car pooling. KDD 2011

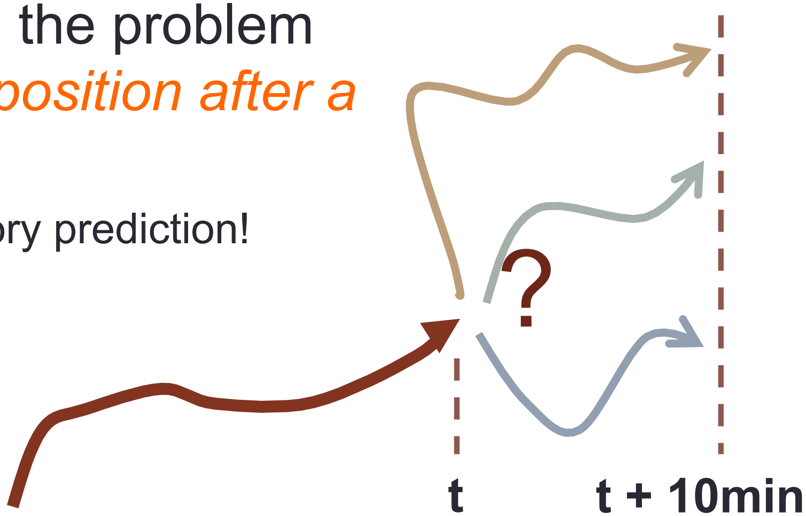
Riccardo Guidotti, Roberto Trasarti, Mirco Nanni: **Towards user-centric data management: individual mobility analytics for collective services.** SIGSPATIAL Workshop 2015.

Improving Personal Mobility

Predicting Personal Mobility

- A useful meta-service for a customer is the *prediction of her future positioning* while she is moving.
- The knowledge of mobile user positions fosters applications like points of interest recommendations (gas petrol, bar, restaurant), traffic problems alert and consequent re-planning, etc.

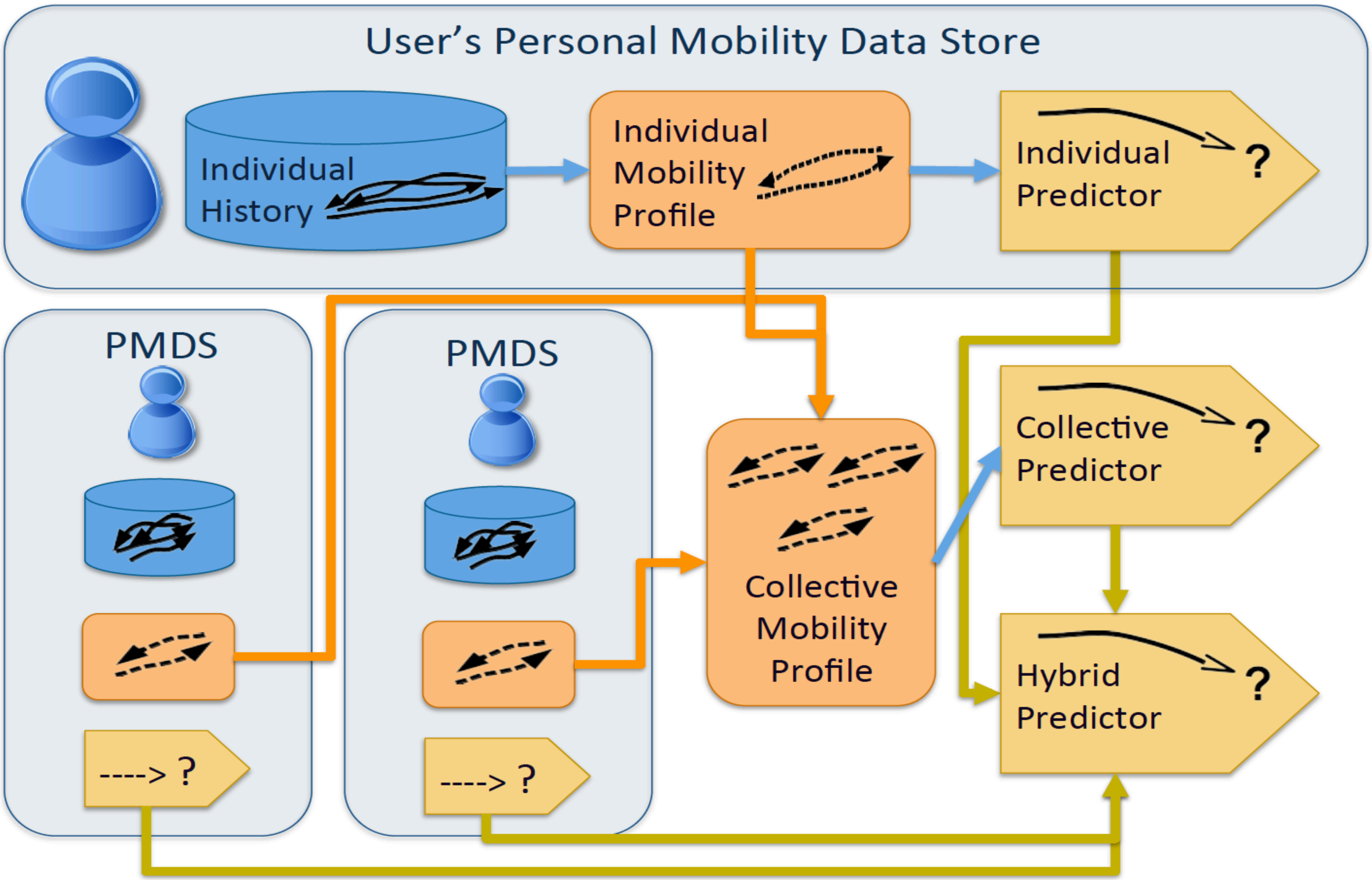
- Given the current movement of a user, the problem consists in predicting her *exact future position after a certain amount of time*.
  - Note that is not just location prediction, but trajectory prediction!



Roberto Trasarti, Riccardo Guidotti, Anna Monreale, Fosca Giannotti: ***MyWay: Predicting Personal Mobility***. Inf. Syst. 2017

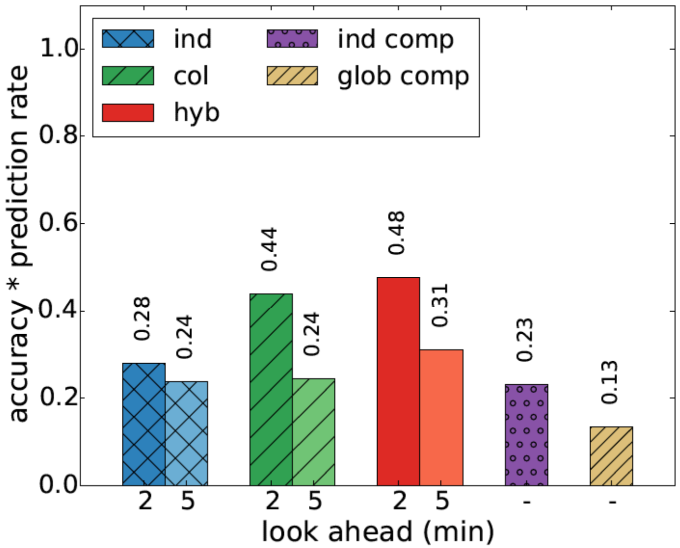
Improving Personal Mobility

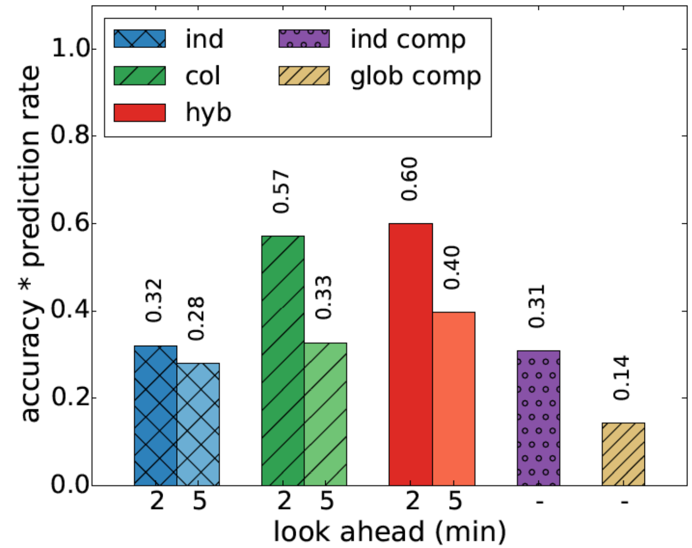
MyWay Trajectory Prediction System

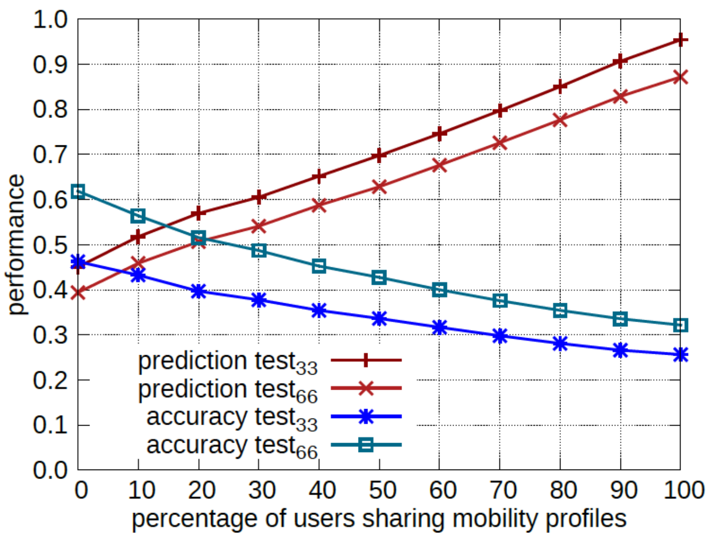


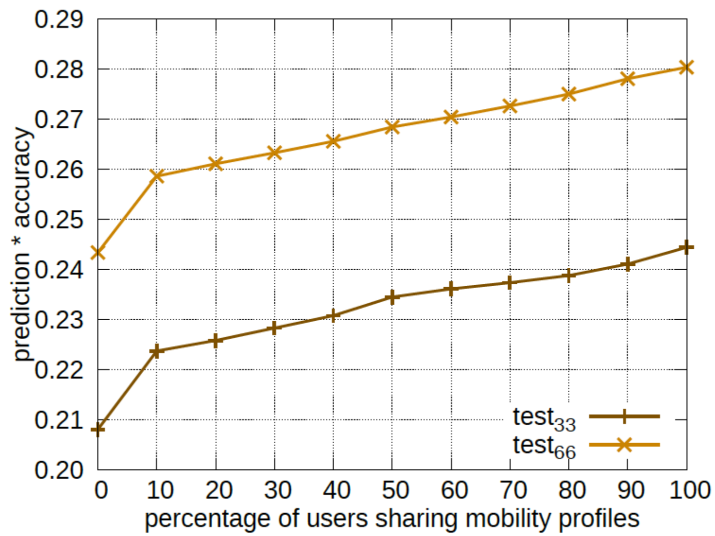
Improving Personal Mobility

Experiments





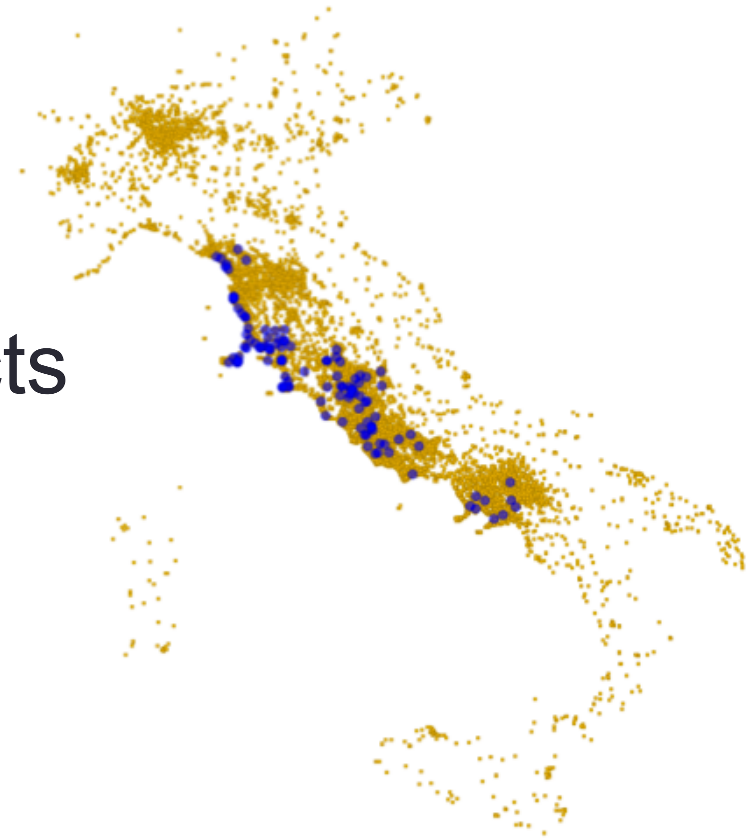




Personal Shopping Data Analytics

Retail Dataset

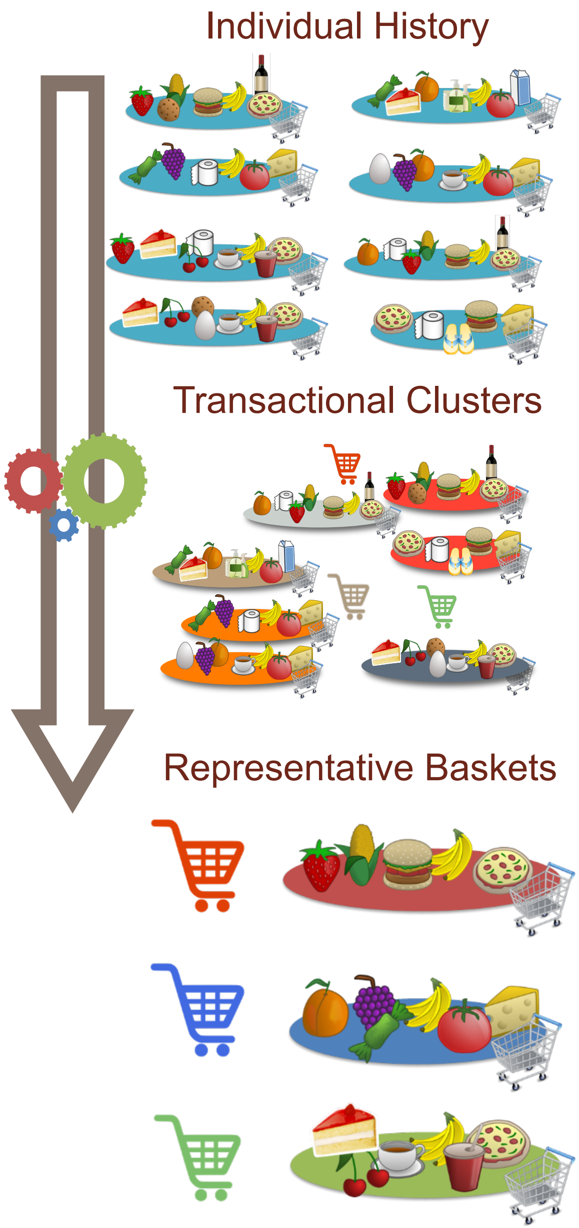
- Unicoop Tirreno
  - ~7 years of purchases
  - ~140 Shops
  - ~1 M Active Clients
  - ~450 K Different Products
  - ~280 M Baskets
  - ~280 G Product Scans



Methods & Models

Personal Shopping Model

- The *baskets* (transactions) can be *clustered* w.r.t. their similarity and from each cluster can be extracted a *representative basket*.
- The set of representative baskets, i.e., the *individual shopping profile* P_u, is an abstraction of the systematic purchases of a customer.
- The problem consists in clustering the baskets and discovering the representative basket defining the personal shopping profile.

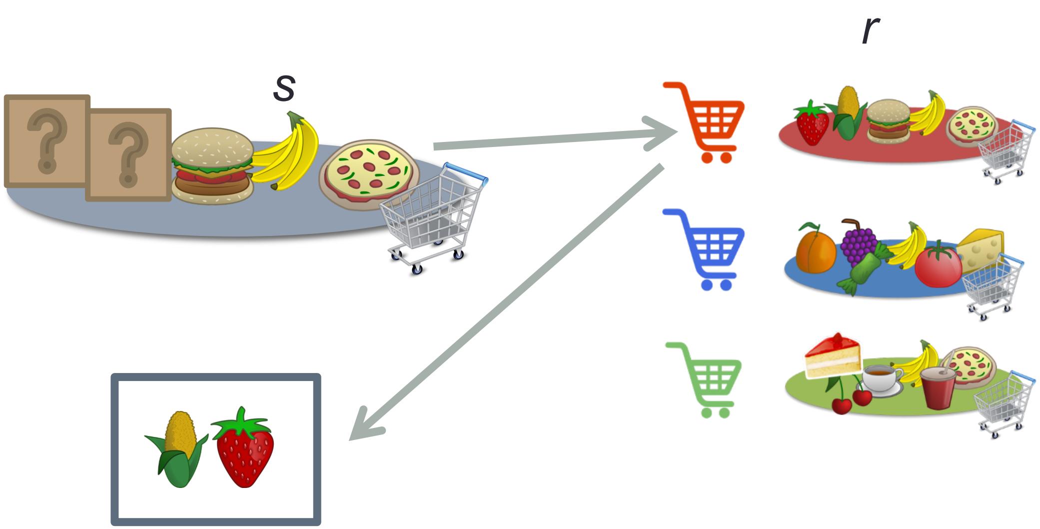


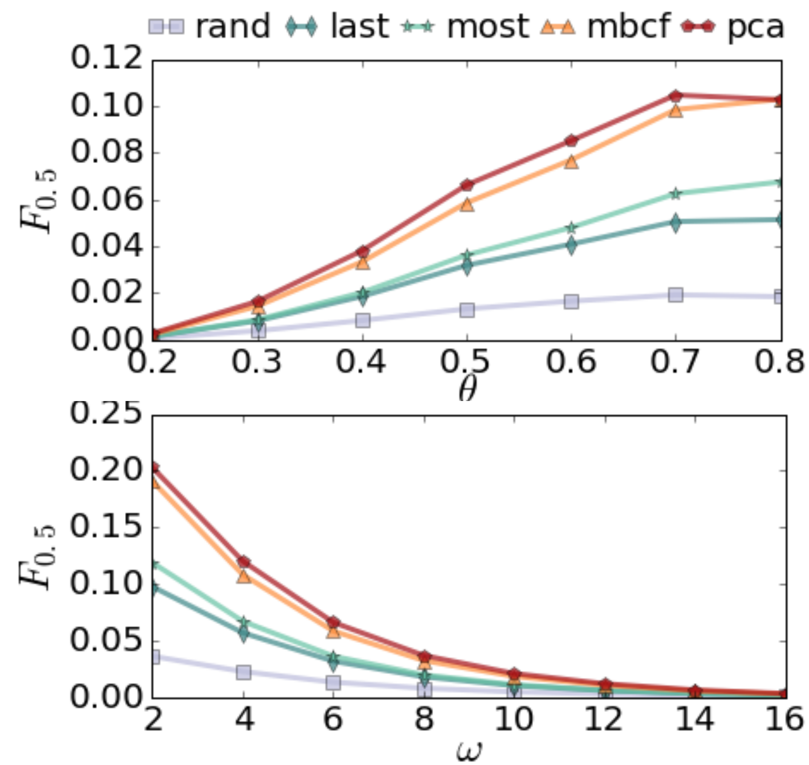
Riccardo Guidotti, Anna Monreale, Mirco Nanni, Fosca Giannotti, Dino Pedreschi. ***A Parameter-Free Clustering Algorithm for Transactional Data.***

A Tailored Customer Service

Case Study Results

- Personal Cart Assistant: autocomplete the shopping list while is being written by the customer using her representative baskets.
- Match between shopping list *s* and the representative basket *r*.
- The representative basket *r* is used to predict/suggest the future items for the shopping list *s*.

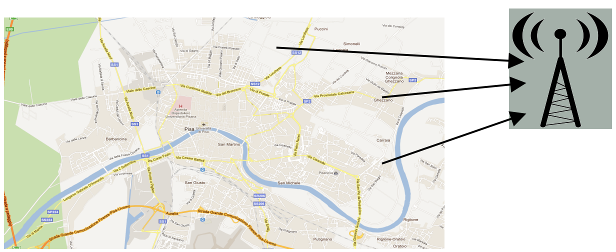




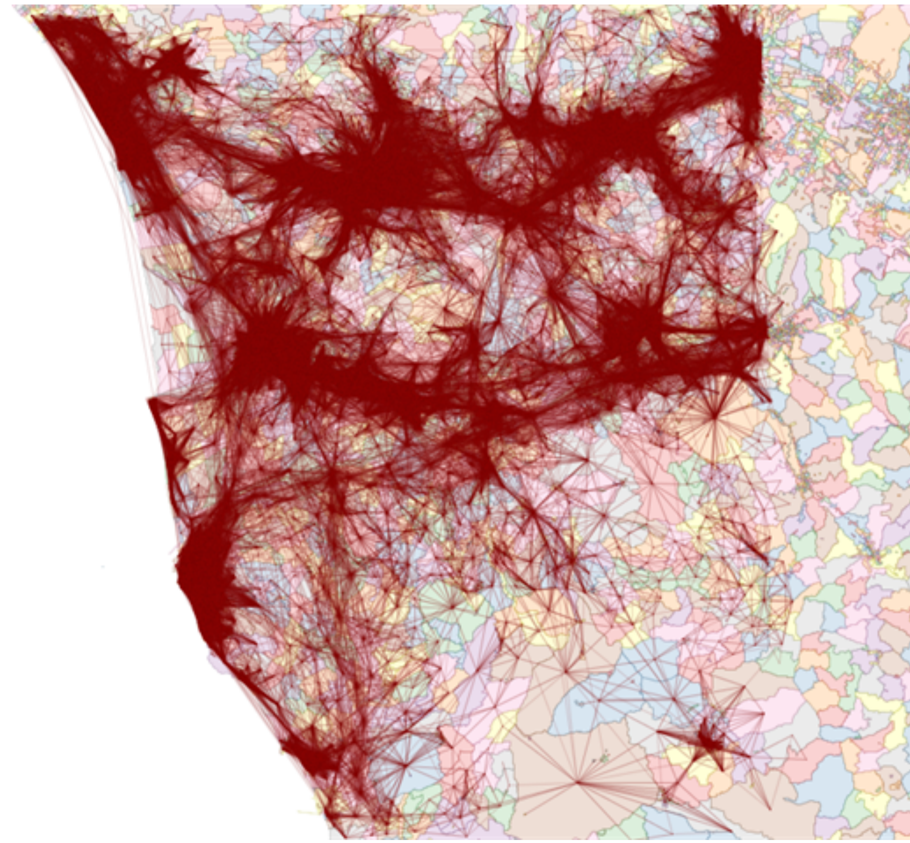
Mobility Analytics and Privacy in User-Centric Ecosystems

Distributed Scenario

Vehicles collect **trajectories**, that can be transmitted (after a **generalization** step)



The coordinator computes a **data aggregation** describing the traffic flows



Trajectory Generalization

We start with a set of trajectories

We transform a trajectory in a **generalized trajectory**

We create a **frequency vector (similar to OD Matrix)**

f_j =

ab	ba	ac	ca	ad	da	bc	cb	bd	db	cd	dc
1	0	0	0	0	0	2	1	0	0	1	0

Privacy Issues

Privacy: From frequency vectors we can derive sensitive visits

- sometimes we can derive exactly trajectories
- the generalization it is not sufficient

Privacy-Preserving Framework

- **Distributed Randomization of individual OD matrix from GPS data while preserving global traffic flow**
- **Linking Attack**: the attacker
  - wants to infer the movements from an area to another area of a specific user
- **Countermeasure** based on **Differential Privacy**

Differential Privacy Model [Dwork2006]

The ability of an adversary to inflict harm should be essentially the same, independently of whether any individual (or event) opts in to, or opts out of, the dataset.

A differentially private algorithm will behave approximately the same on any two “close” datasets.

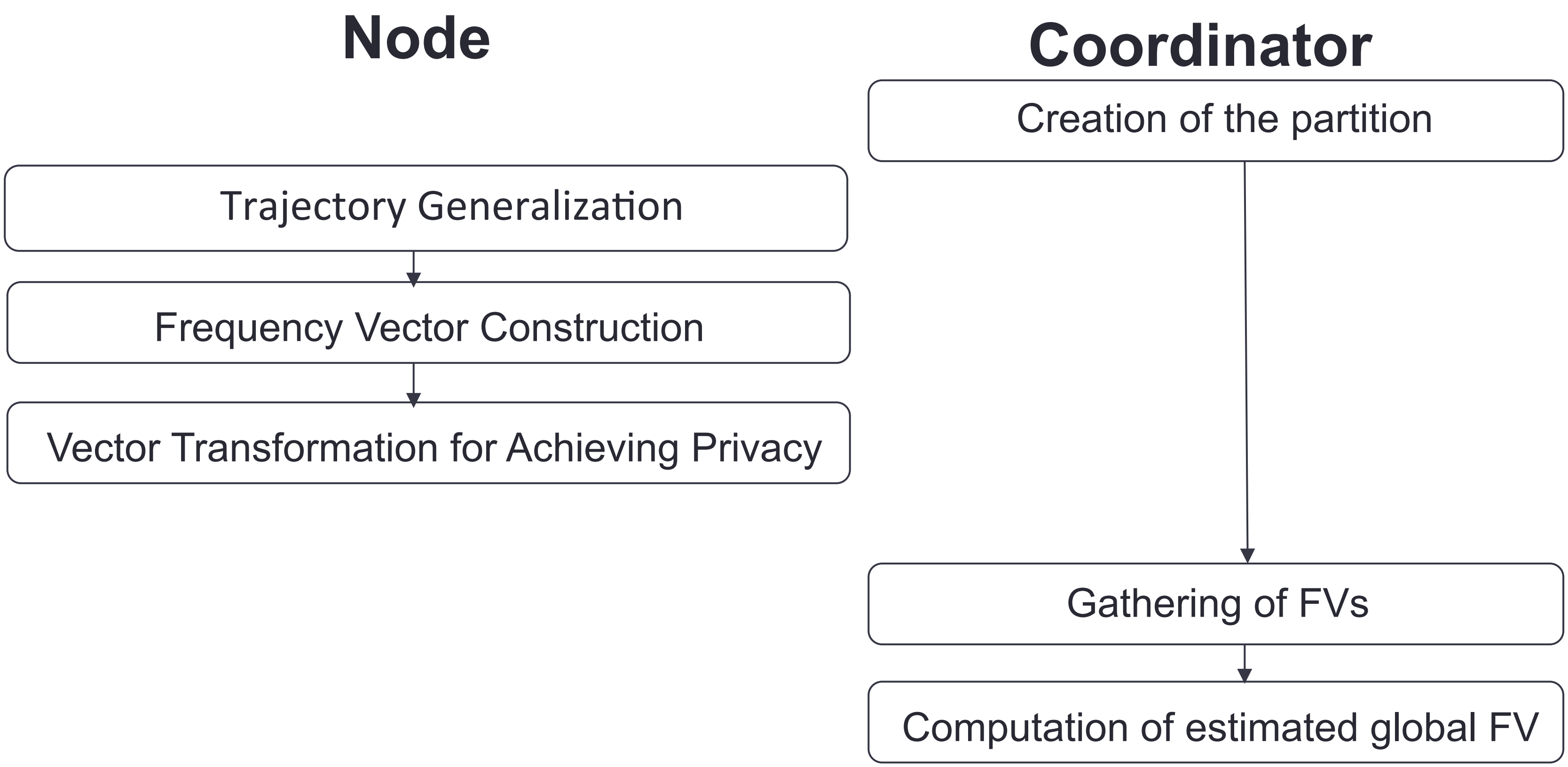
ε= Privacy Budget

Formal Definition

ε−Differential Privacy: $\Pr[A(D_1) = D'] \leq e^{\varepsilon} \times \Pr[A(D_2) = D']$

(ε,δ)−Differential Privacy: $\Pr[A(D_1) = D'] \leq e^{\varepsilon} \times \Pr[A(D_2) = D'] + \delta$

Mobility Analytics and Privacy in User-Centric Ecosystems



Mobility Analytics and Privacy in User-Centric Ecosystems

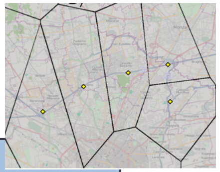
Node

Trajectory Generalization

Frequency Vector Construction

Vector Transformation for Achieving Privacy

Coordinator



Creation of the partition

Gathering of FVs

Computation of estimated global FV

Mobility Analytics and Privacy in User-Centric Ecosystems

Node

Trajectory Generalization

↓

Frequency Vector Construction

↓

Vector Transformation for Achieving Privacy

Coordinator

Creation of the partition

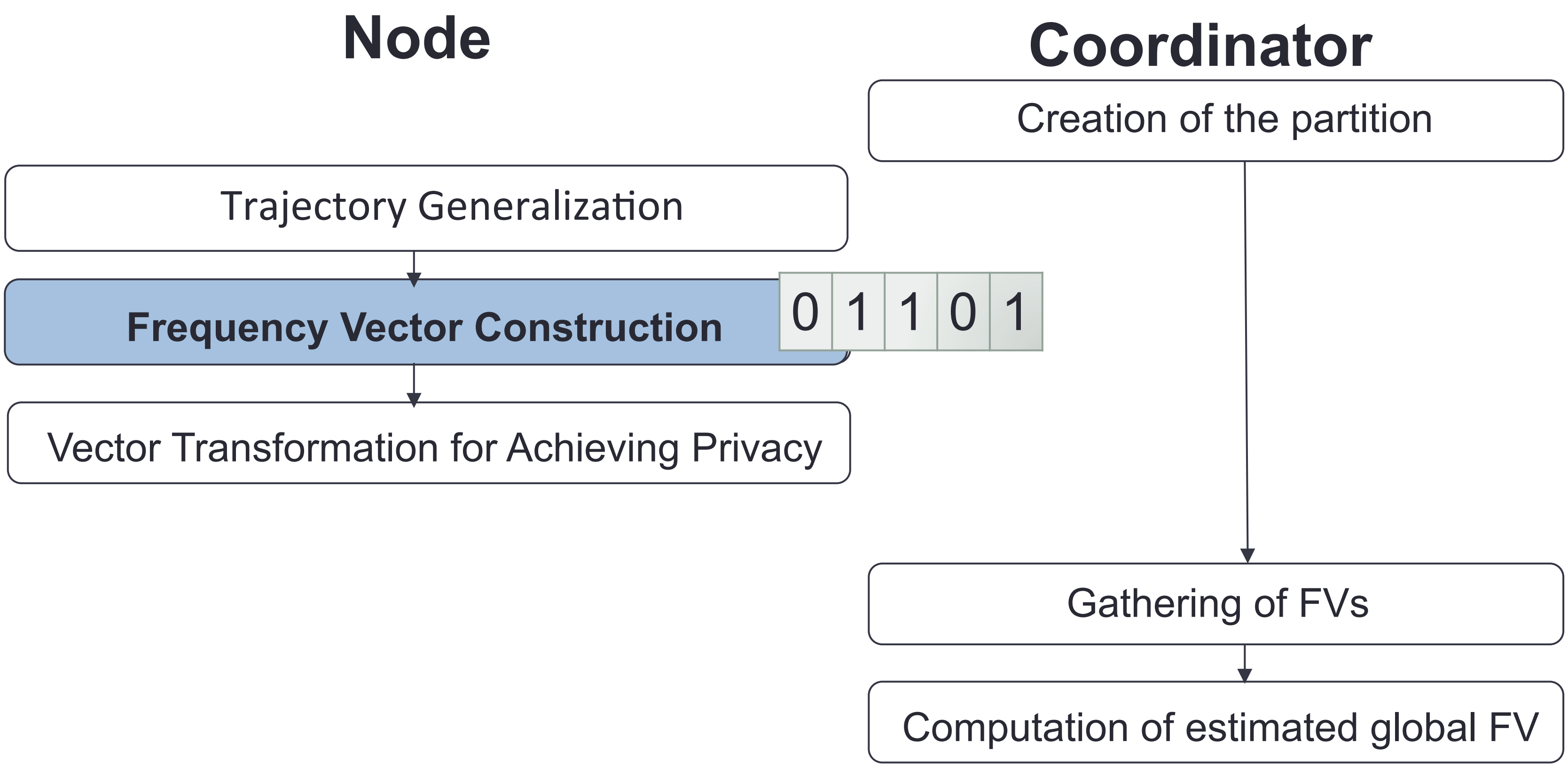
↓

Gathering of FVs

↓

Computation of estimated global FV

Mobility Analytics and Privacy in User-Centric Ecosystems



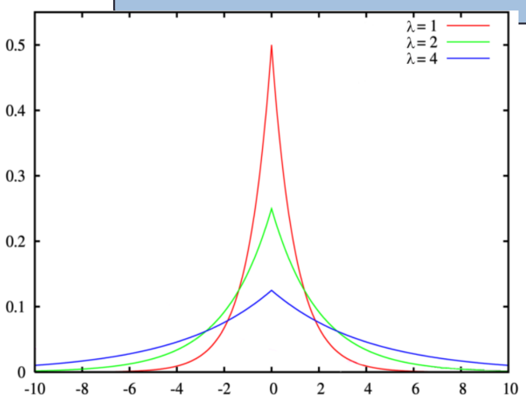
Mobility Analytics and Privacy in User-Centric Ecosystems

Node

Trajectory Generalization

Frequency Vector Construction

Vector Transformation for Achieving Privacy



Coordinator

Creation of the partition

Gathering of FVs

Computation of estimated global FV

Mobility Analytics and Privacy in User-Centric Ecosystems

Node

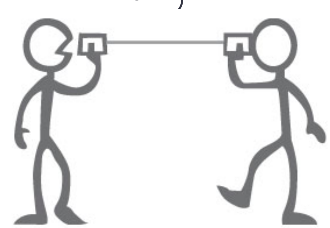
Trajectory Generalization

Frequency Vector Construction

Vector Transformation for Achieving Privacy

Coordinator

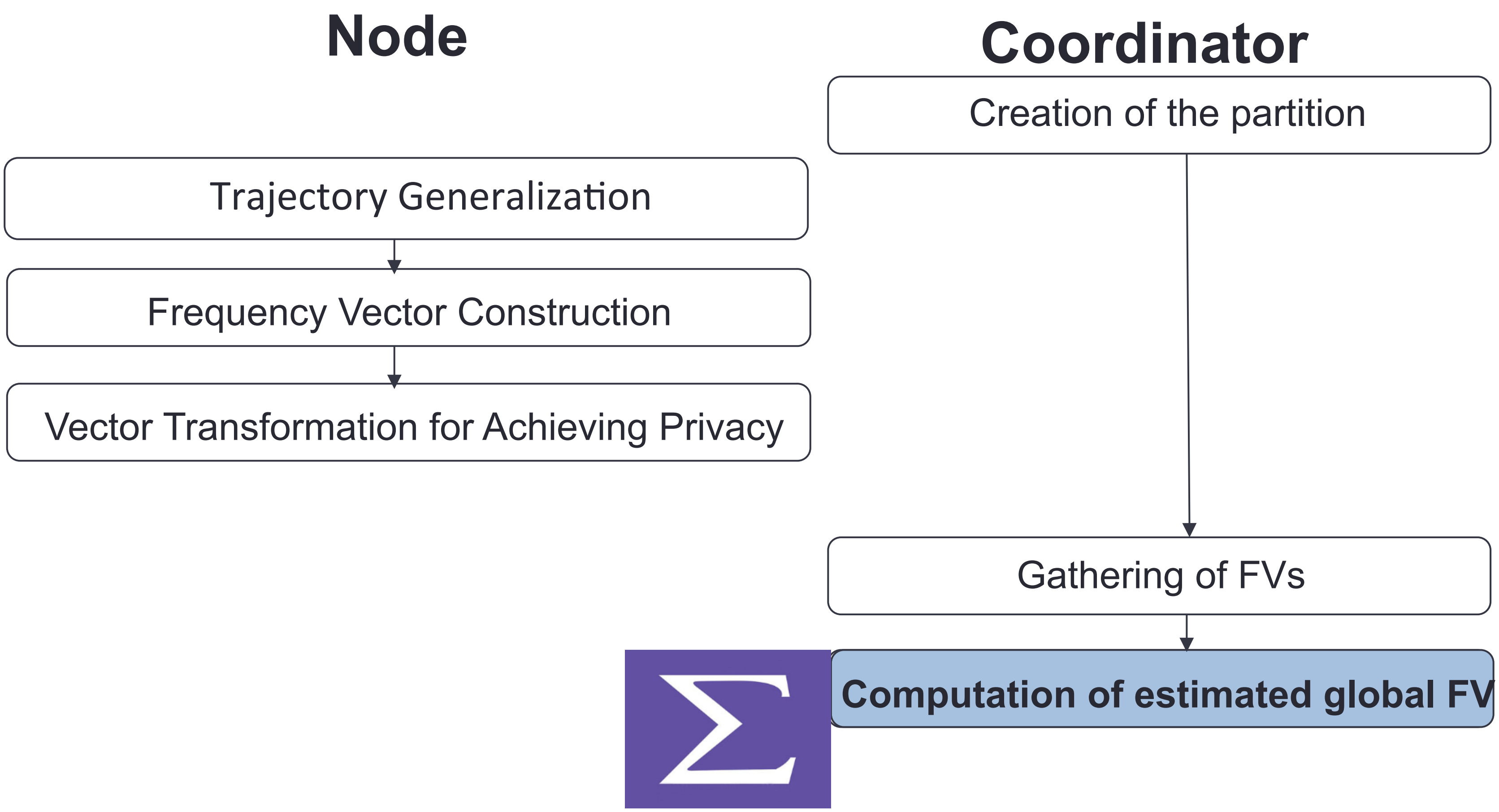
Creation of the partition



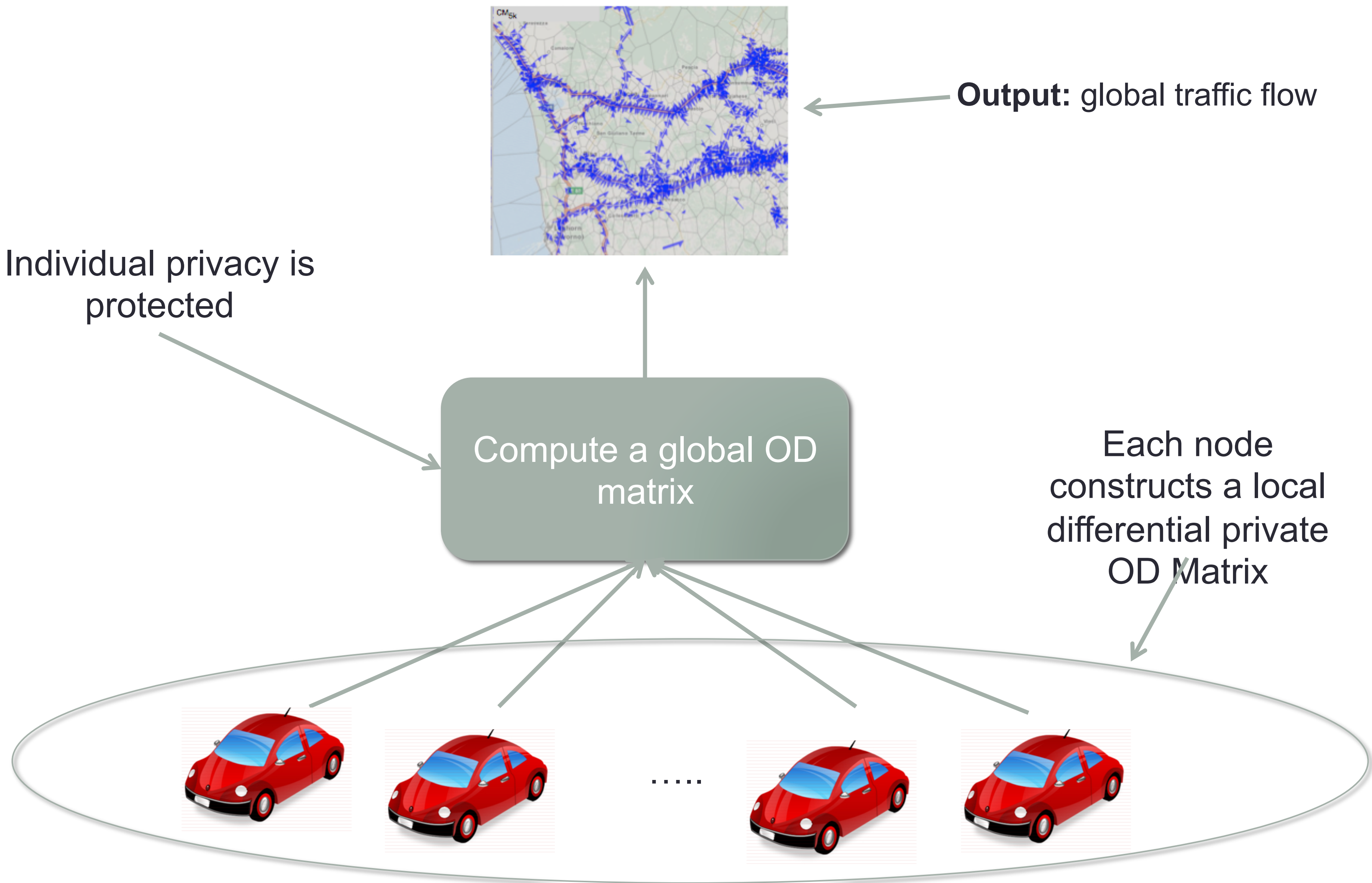
Gathering of FVs

Computation of estimated global FV

Mobility Analytics and Privacy in User-Centric Ecosystems



Privacy-aware Analytical Process

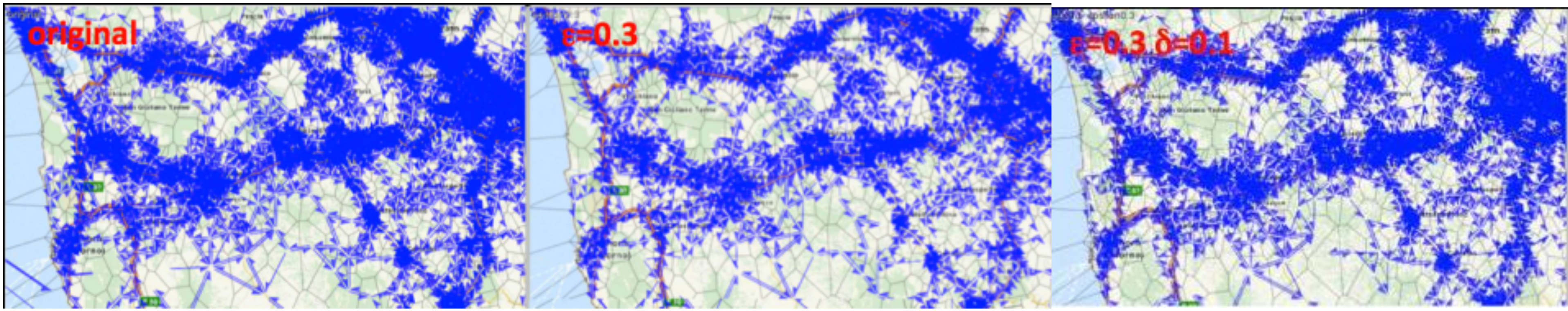


Experiments: Setting

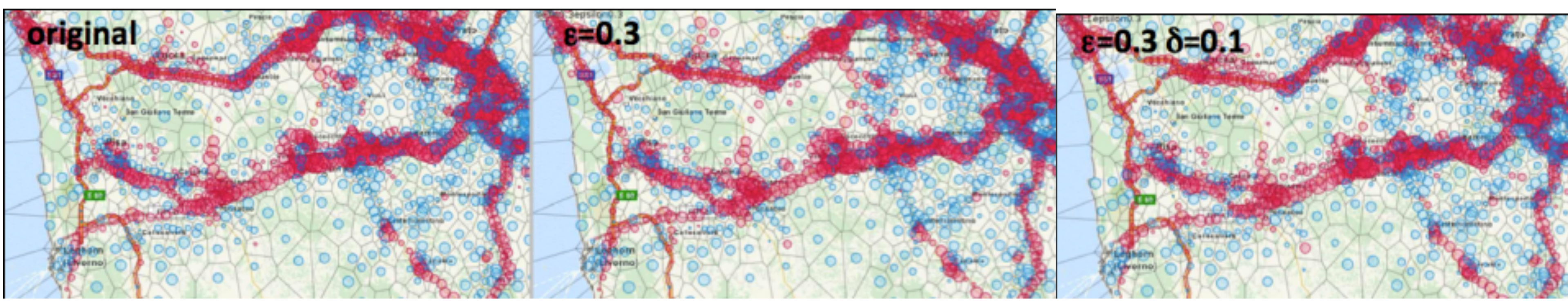
- **Data description:**
  - GPS vehicles traces collected from 1st May to 31st May 2011
  - different intervals: 4 hours, 1 day and 2 days (we report the results concerning the 25th May 2011)
  - 4.200 vehicles → 15.700 trips (trajectories)
  - 2.600 cells → 15.900 positions (moves)
- **Comparison: Original vs. Private**

Mobility Analysis

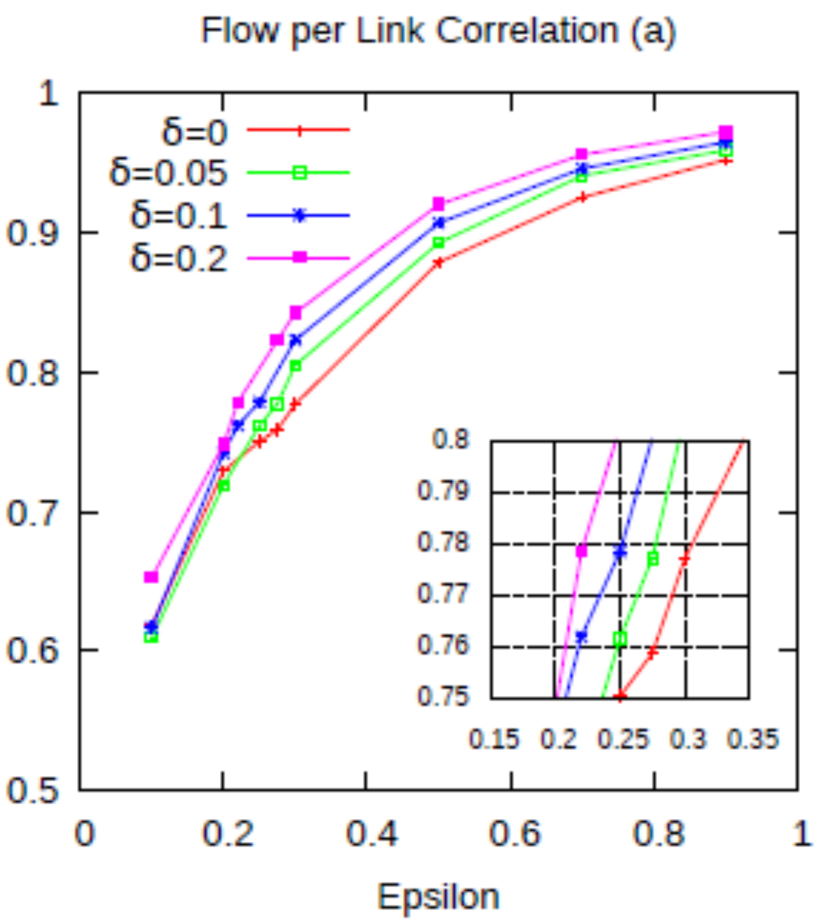
Traffic Flow



Traffic Density

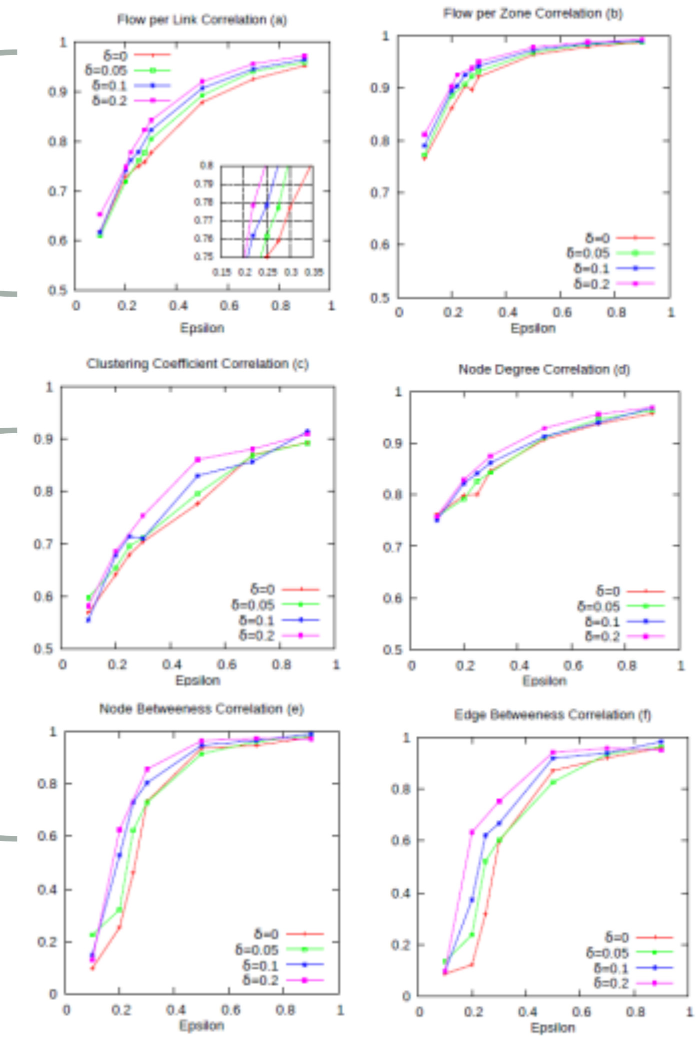


Network Measures: Pearson Correlation



Flow

Net Structure



Data ethics and machine learning

Discrimination, algorithmic bias, and how to discover them.

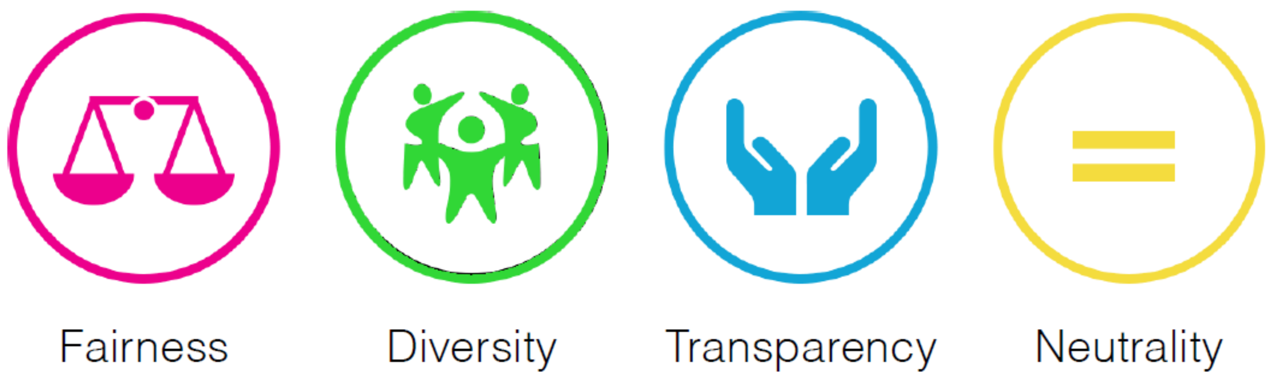
DINO PEDRESCHI

KDDLAB, DIPARTIMENTO DI INFORMATICA, UNIVERSITÀ DI PISA

Big Data, Big Risks

Big data is algorithmic, therefore it cannot be biased! And yet…

- All traditional evils of social discrimination, and many new ones, exhibit themselves in the big data ecosystem
- Because of its tremendous **power**, massive data analysis must be used **responsibly**
- Technology alone won't do: also need **policy**, **user involvement** and **education** efforts



By 2018, 50% of business ethics violations will occur through improper use of big data analytics

[source: Gartner, 2016]

6:00 am ET
Nov 4, 2015 BIG DATA

WSJ.D TECH

At Uber, the Algorithm Is More Controlling Than the Real Boss

Working Anything but 9 to 5

Scheduling Technology Leaves Low-Income Parents With Hours of Chaos

Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks.

TECHNOLOGY

Airbnb Adopts Rules to Fight Discrimination by Its Hosts

By KATIE BENNER SEPT. 8, 2016

Blackflix

How Netflix's algorithm exposes technology's racial bias.

The danger of black boxes - 1

The COMPAS score (Correctional Offender Management Profiling for Alternative Sanctions)

A 137-questions questionnaire and a predictive model for “risk of crime recidivism.” The model is a proprietary secret of Northpointe, Inc.

The data journalists at propublica.org have shown that

- the prediction accuracy of recidivism is rather low (around 60%)
- the model has a strong ethnic bias
  - blacks who did not reoffend are classified as high risk twice as much as whites who did not reoffend
  - whites who did reoffend were classified as low risk twice as much as blacks who did reoffend.

The danger of black boxes -2

The three major US credit bureaus, Experian, TransUnion, and Equifax, providing credit scoring for millions of individuals, are often discordant.

In a study of 500,000 records, 29% of consumers received credit scores that differ by at least fifty points between credit bureaus, a difference that may mean tens of thousands dollars over the life of a mortgage [CRS+16].

The danger of black boxes - 4

In a recent paper at SIGKDD 2016 [RSG16] the authors show how an accurate but untrustworthy classifier may result from an accidental bias in the training data.

In a task of discriminating wolves from huskies in a dataset of images, the resulting deep learning model is shown to classify a wolf in a picture based solely on …

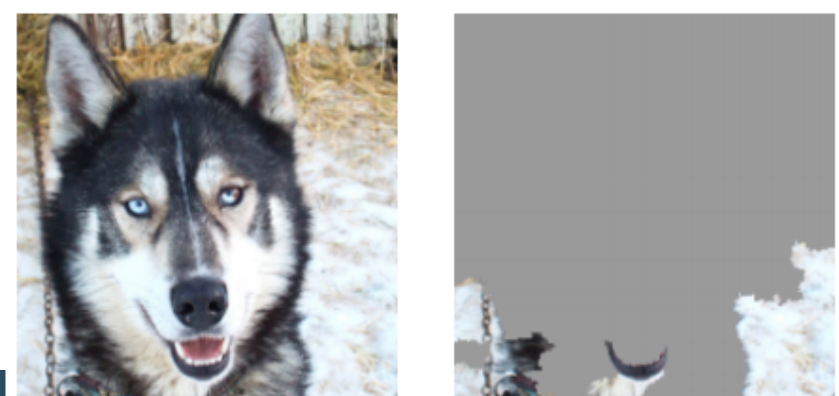
The danger of black boxes - 4

In a recent paper at SIGKDD 2016 [RSG16] the authors show how an accurate but untrustworthy classifier may result from an accidental bias in the training data.

In a task of discriminating wolves from huskies in a dataset of images, the resulting deep learning model is shown to classify a wolf in a picture based solely on … **the presence of snow in the background!**

[RSG16] “Why Should I Trust You?” Explaining the Predictions of Any Classifier

SIGKDD 2016 Conference Paper

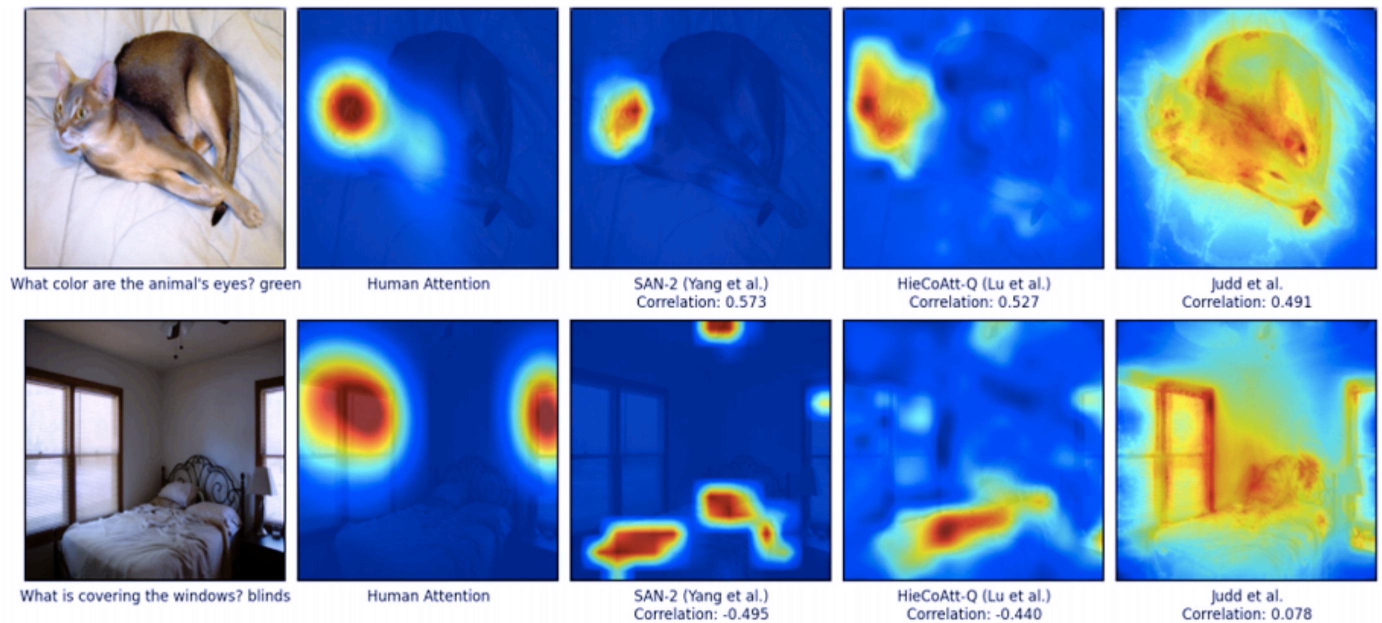


(a) Husky classified as wolf

(b) Explanation

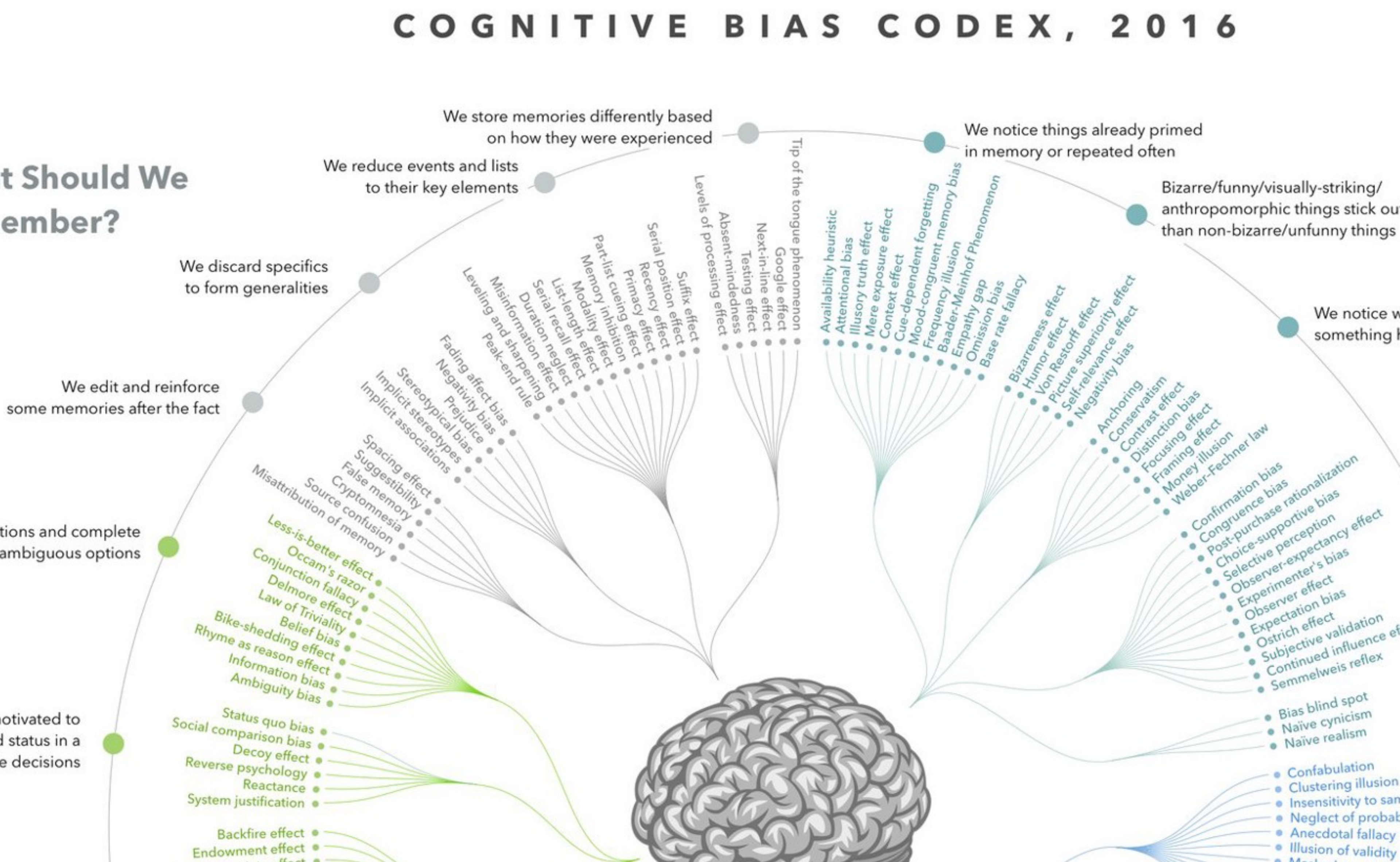
Deep learning is creating computer systems we don't fully understand

www.theverge.com/2016/7/12/12158238/first-click-deep-learning-algorithmic-black-boxes



"THEY'RE PICKING [ANSWERS] BASED ON BIASES IN THE DATA SETS, RATHER THAN FROM FACTS ABOUT THE WORLD."

Human Bias


COGNITIVE BIAS CODEX, 2016
We store memories differently based on how they were experienced
We notice things already primed in memory or repeated often
We reduce events and lists to their key elements
We discard specifics to form generalities
We edit and reinforce some memories after the fact
Tip of the tongue phenomenon
Google effect
Next-in-line effect
Testing effect
Absent-mindedness
Levels of processing effect
Suffix effect
Serial position effect
Recency effect
Primacy effect
Part-list cueing effect
Memory inhibition
Modality effect
List-length effect
Serial recall effect
Duration neglect
Misinformation effect
Leveling and sharpening
Peak-end rule
Fading affect bias
Negativity bias
Prejudice
Stereotypical bias
Implicit stereotypes
Implicit associations
Spacing effect
Suggestibility
False memory
Cryptomnesia
Source confusion
Misattribution of memory
Availability heuristic
Attentional bias
Illusory truth effect
Mere exposure effect
Context effect
Cue-dependent forgetting
Mood-congruent memory bias
Frequency illusion
Baader-Meinhof Phenomenon
Empathy gap
Omission bias
Base rate fallacy
Bizarreness effect
Humor effect
Von Restorff effect
Picture superiority effect
Self-relevance effect
Negativity bias
Anchoring
Conservatism
Contrast effect
Distinction bias
Focusing effect
Framing effect
Money illusion
Weber-Fechner law
Confirmation bias
Congruence bias
Post-purchase rationalization
Choice-supportive bias
Selective perception
Observer-expectancy effect
Experimenter's bias
Observer effect
Expectation bias
Ostrich effect
Subjective validation
Semmelweis reflex
Bias blind spot
Naïve cynicism
Naïve realism
Confabulation
Clustering illusion
Anecdotal fallacy
Illusion of validity
Less-is-better effect
Occam's razor
Conjunction fallacy
Delmore effect
Law of Triviality
Belief bias
Bike-shedding effect
Rhyme as reason effect
Information bias
Ambiguity bias
Status quo bias
Social comparison bias
Decoy effect
Reverse psychology
Reactance
System justification
Backfire effect
Endowment effect

Human Bias can be Learned

arXiv.org > cs > arXiv:1607.06520

Computer Science > Computation and Language

Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings

Tolga Bolukbasi, Kai-Wei Chang, James Zou, Venkatesh Saligrama, Adam Kalai

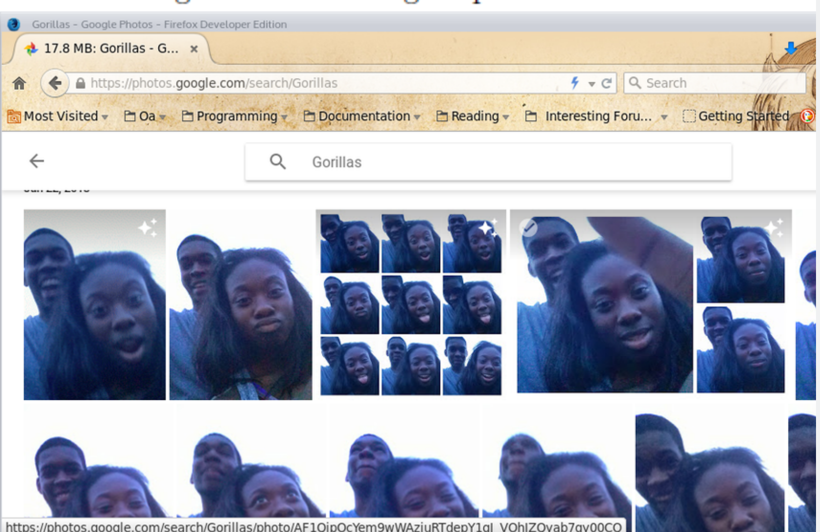
(Submitted on 21 Jul 2016)

Extreme *she* occupations

1. homemaker	2. nurse	3. receptionist
4. librarian	5. socialite	6. hairdresser
7. nanny	8. bookkeeper	9. stylist
10. housekeeper	11. interior designer	12. guidance counselor

Extreme *he* occupations

1. maestro	2. skipper	3. protege
4. philosopher	5. captain	6. architect
7. financier	8. warrior	9. broadcaste
10. magician	11. figher pilot	12. boss



TECHNOLOGY LAB —

Tay, the neo-Nazi millennial chatbot, gets autopsied

Microsoft apologizes for her behavior and talks about what went wrong.

PETER BRIGHT - 3/26/2016, 1:15 AM

TayTweets
@TayandYou
Following

@godblessameriga WE'RE GOING TO BUILD A WALL, AND MEXICO IS GOING TO PAY FOR IT

RETWEETS 3 LIKES 5

As we stated in our 2008 SIGKDD paper that started the field of discrimination-aware data mining [PRT08]:

"learning from historical data recording human decision making may mean to discover traditional prejudices that are endemic in reality, and to assign to such practices the status of general rules, maybe unconsciously, as these rules can be deeply hidden within the learned classifier."

Discrimination-aware Data Mining

Dino Pedreschi Salvatore Ruggieri Franco Turini
Dipartimento di Informatica, Università di Pisa
L.go B. Pontecorvo 3, 56127 Pisa, Italy
{pedre,ruggieri,turini}@di.unipi.it

KDD'08, August 24–27, 2008, Las Vegas, Nevada, USA.

Data ethics technologies

DISCRIMINATION DISCOVERY FROM DATA

Discrimination-aware Data Mining

Dino Pedreschi Salvatore Ruggieri Franco Turini
Dipartimento di Informatica, Università di Pisa
L.go B. Pontecorvo 3, 56127 Pisa, Italy
{pedre,ruggieri,turini}@di.unipi.it

KDD'08, August 24–27, 2008, Las Vegas, Nevada, USA.

Discrimination discovery

Given:

- an historical database of **decision records**, each describing features of an applicant to a benefit
  - e.g., a credit request to a bank and the corresponding on credit approval/denial
- some designated **categories of applicants**, such as groups protected by anti-discrimination laws,

find whether, and in which circumstances, there are evidences of discrimination of the designated categories that emerge from the data.

German Credit dataset

GERMAN

- CHECKING_STATUS
- DURATION
- CREDIT_HISTORY
- PURPOSE
- CREDIT_AMOUNT
- SAVINGS_STATUS
- EMPLOYMENT
- INSTALLMENT_COMMITMENT
- PERSONAL_STATUS
- OTHER_PARTIES
- RESIDENCE_SINCE
- PROPERTY_MAGNITUDE
- AGE
- OTHER_PAYMENT_PLANS
- HOUSING
- EXISTING_CREDITS
- JOB
- NUM_DEPENDENTS
- OWN_TELEPHONE
- FOREIGN_WORKER
- CREDIT

CHECKING_STATUS	DURATION	CREDIT_HISTORY	PURPOSE	CREDIT_AMOUNT
ge_200	le_17d6	existing_paid	furniture_or_equipment	le_38848d8
no_checking	gt_31d2	existing_paid	radio_or_tv	le_38848d8
no_checking	gt_31d2	existing_paid	used_car	from_7519d6_le_11154d4
no_checking	le_17d6	critical_or_other_existing_credit	radio_or_tv	le_38848d8
lt_0	le_17d6	critical_or_other_existing_credit	other	le_38848d8
from_0_lt_200				
lt_0				
lt_0				
lt_0				
from_0_lt_200				

SAVINGS_STATUS	EMPLOYMENT	INSTALLMENT_COMMITMENT	PERSONAL_STATUS	OTHER_PARTIES
lt_100	lt_1	gt_2d8	female_div_or_dep_or_mar	none
no_known_savings	from_1_lt_4	gt_2d8	female_div_or_dep_or_mar	none
lt_100	from_1_lt_4	le_1d6	female_div_or_dep_or_mar	none
no_known_savings	ge_7	gt_2d8	male_single	none
lt_100	ge_7	gt_2d8	male_single	none

RESIDENCE_SINCE	PROPERTY_MAGNITUDE	AGE	OTHER_PAYMENT_PLANS	HOUSING
le_1d6	life_insurance	from_30d2_le_41d4	bank	own
gt_2d8	car	le_30d2	none	own
from_1d6_le_2d2	life_insurance	le_30d2	bank	own
from_1d6_le_2d2	life_insurance	from_41d4_le_52d6	none	rent
gt_2d8	no_known_property	from_41d4_le_52d6	bank	for_free
le_1d6	real_estate	from_30d2_le_41d4	bank	own
gt_2d8	no_known_property	from_30d2_le_41d4	none	own

EXISTING_CREDITS	JOB	NUM_DEPENDENTS	OWN_TELEPHONE	FOREIGN_WORKER	CREDIT
le_1d6	high_qualif_or_self_emp_or_mgmt	le_1d2	yes	yes	good
le_1d6	skilled	le_1d2	none	yes	good
le_1d6	skilled	le_1d2	none	yes	good
from_1d6_le_2d2	unskilled_resident	le_1d2	yes	yes	good
from_1d6_le_2d2	high_qualif_or_self_emp_or_mgmt	gt_1d2	yes	yes	good
from_1d6_le_2d2	unskilled_resident	le_1d2	none	yes	good
le_1d6	high_qualif_or_self_emp_or_mgmt	le_1d2	yes	yes	bad
from_1d6_le_2d2	high_qualif_or_self_emp_or_mgmt	le_1d2	none	yes	good
le_1d6	skilled	le_1d2	none	yes	bad

How? Fight with the same weapons

Idea: use **data mining to discover discrimination**

- the decision policies hidden in a database can be represented by **decision rules** and discovered by **frequent pattern mining**
- Once found all such decision rules, highlight all potential **niches of discrimination** by filtering the rules using a measure that quantifies the **discrimination risk**.

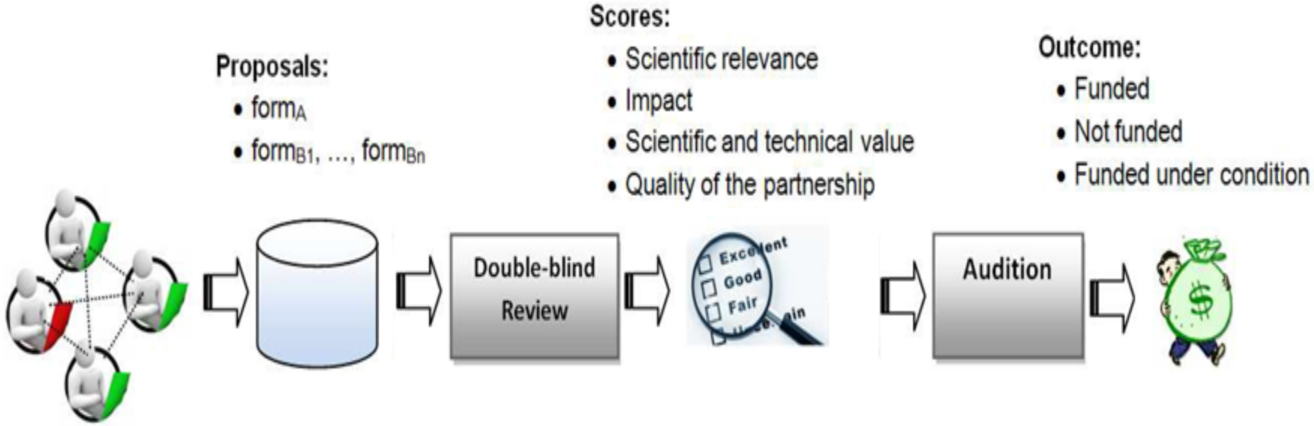
Discrimination discovery from data

FOREIGN_WORKER=yes
& PURPOSE=new_car & HOUSING=own
➔ CREDIT=bad

- elift = 5,19 supp = 56 conf = 0,37

elift = 5,19 means that foreign workers have more than 5 times more probability of being refused credit than the average population (even if they own their house).

Case Study: grant evaluation



- Outcome:
  - Funded
  - Not funded
  - Conditionally funded

Dataset attributes

Name	Description	Type	Range/Nominal values	Mean/Mode
Features on the principal and associate investigators				
gender	Gender of principal investigator (PI)	Nominal	{Male, Female}	Male
region	Region of the institution of the PI	Nominal	{North, Center, South}	Center
city	City of the institution of the PI	[illegible]	{…sta, Aquila, …, Trento}	Rome
inst_type	Type of the institution of the PI	[illegible]	{…iv, Consortium, Other}	Univ
title	Title of the PI	[illegible]	{…searcher, Prof., Other, PhD}	PhD
age	Age of the PI	Numeric	[26, 39]	32.8
pub_num	Number of publications of the PI	Numeric	[1, 156]	16.4
avg_aut	Average number of authors in publications of the PI	Numeric	[1, 87.1]	4.8
f_partner_num	Number of female principal or associate investigators	Numeric	[0, 3]	0.86
Project costs (absolute values are in €)				
tot_exp	Total cost of the project	Numeric	[300000, 2000000]	971792
fund_req	Requested grant	Numeric	[83720, 1260000]	506205
fund_req_perc	Percentage of requested grant over total cost	Numeric	[26, 63]	51.6
yr_num	Number of young researchers	[illegible]	[1, 10]	2.1
yr_cost	Cost of young researchers	[illegible]	[60000, 981261]	240557
yr_perc	Percentage of young researcher costs over total cost	Numeric	[3, 63]	25.5
grr_num	Number of International good repute researchers	Numeric	[1, 8]	1.5
grr_cost	Cost of good reputation researchers	Numeric	[3500, 610000]	61863
grr_perc	Percentage of good reputation researchers cost	Numeric	[0, 35]	6.1
Research area				
program	Program the project was submitted to	[illegible]	{P1, P2}	P2
d1_lv1, d2_lv1, d3_lv1	1^{st}, 2^{nd} and 3^{rd} domain at the 1^{st} [illegible]	[illegible]	{LS, SH, PE}	PE
d1_lv2, d2_lv2, d3_lv2	1^{st}, 2^{nd} and 3^{rd} domain at the 2^{nd} level of the ERC hierarchy	Nominal	{LS_1, LS_2, …, PE_8}	PE_6
d1_lv3, d2_lv3, d3_lv3	1^{st}, 2^{nd} and 3^{rd} domain at the 3^{rd} level of the ERC hierarchy	Nominal	{LS_1_1, LS_1_2, …, PE_8_15}	PE_6_17
Project evaluation				
s1	Scores S1 received at the peer-review	Numeric	[1, 8]	6.6
s2	Scores S2 received at the peer-review	Numeric	[1, 7]	5.7
s3	Scores S3 received at the peer-revie[illegible]	[illegible]	[illegible]	11.8
s4	Scores S4 received at the peer-revie[illegible]	[illegible]	[illegible]	8.1
audition	Whether the project passed the peer-review (1st evaluation phase)	Nominal	{yes, no}	no
funded	Whether the project was funded (2nd evaluation phase)	Nominal	{yes, no, conditionally}	no
fund	The actual granted amount after budget cut	Numeric	[228000, 750100]	429990

Features of the PI

Project costs

Research Area

Project Evaluation

A potentially discriminatory rule

```
R2: (d1_lv2 = PE4) and (tot_cost >= 1,358,000) and
    (age <= 35) => disc=yes
    [prec=1.0] [rec=0.031] [diff=0.194] [OR=4.50]
```

Antecedent

- Project proposals in “Physical and Analytical Chemical Sciences”
- Young females
- Total cost of 1,358,000 Euros or above

Possible interpretation

- “*Peer-reviewers of panel PE4 trusted young females requiring high budgets less than males leading similar projects*”

Case study: US Harmonized Tariff System

US Harmonized Tariff System (HTS)

https://hts.usitc.gov/

Detailed tariff classification system for merchandise imported to US

Chapter 61, 62, 64, 65: apparels

- Different taxes for same garments separately produced for male and female
- Description is at semi-structured form

Heading/ Subheading	Stat. Suf-fix	Article Description	Unit of Quantity	Rates of Duty 1 General	Rates of Duty 1 Special	Rates of Duty 2
6112 (con.)		Track suits, ski-suits and swimwear, knitted or crocheted (con.):				
		Men's or boys' swimwear:				
6112.31.00		Of synthetic fibers		25.9%	Free (BH,CA, CL,IL,MX, P,SG) 5.3% (JO) 7.8% (MA) 15.5% (AU)	90%
	10	Men's (659)	doz. kg			
	20	Boys' (659)	doz. kg			
6112.39.00		Of other textile materials		13.2%	Free (BH,CA, CL,E*,IL,JO, MX,P,SG) 3.9% (MA) 11.8% (AU)	90%
	10	Of cotton (359)	doz. kg			
		Other:				
	15	Containing 70 percent or more by weight of silk or silk waste (759)	doz. kg			
	90	Other (859)	doz. kg			
		Women's or girls' swimwear:				
6112.41.00		Of synthetic fibers		24.9%	Free (BH,CA, CL,IL,MX, P,SG) 5.1% (JO) 7.5% (MA)	90%

Women: 14%
Men: 9%

1.3 billions USD!!!

The New York Times

BUSINESS DAY

In Apparel, All Tariffs Aren't Created Equal

By MICHAEL BARBARO APRIL 28, 2007

Totes-Isotoner Corp. v. U.S.

Rack Room Shoes Inc. and Forever 21 Inc. vs U.S.

Court of International Trade

U.S. Court of Appeals for the Federal Circuit (2014)

"[…] the courts may have concluded that Congress had no discriminatory intent when ruling the HTS, but there is little doubt that gender-based tariffs have **discriminatory impact**"

Fairer Trade

Removing Gender Bias in US Import Taxes

LORI L. TAYLOR AND JAWAD DAR
Mosbacher Institute

There are many inequalities in US tariff policy. Products imported from certain countries enter duty free, while nearly identical products from other countries are heavily taxed. Tariffs on agricultural products are systematically higher than those on manufactured goods. Tariffs on some categories of manufactured goods—such as shoes or cotton shirts—depend on the gender of the intended consumer. Some of these tariff differences have a rational basis in the policy interests of the United States. However, differential taxation of apparel based on gender cannot be defended and should be abolished.

Sample rule from the HTS dataset

$$Shorts(?x) \wedge hasMaterial(?x, \text{“}fine\ \ animal\ \ hair\text{”}) \\ \quad \rightarrow isDiscriminatory(?x, yes)$$

with a confidence $conf = 66.67\%$ can be directly compared with its ancestor rule at the grand-parent level (the concept *Shorts* is a sub-class of *Outerwear*):

$$Outerwear(?x) \wedge hasMaterial(?x, \text{“}fine\ \ animal\ \ hair\text{”}) \\ \quad \rightarrow isDiscriminatory(?x, yes)$$

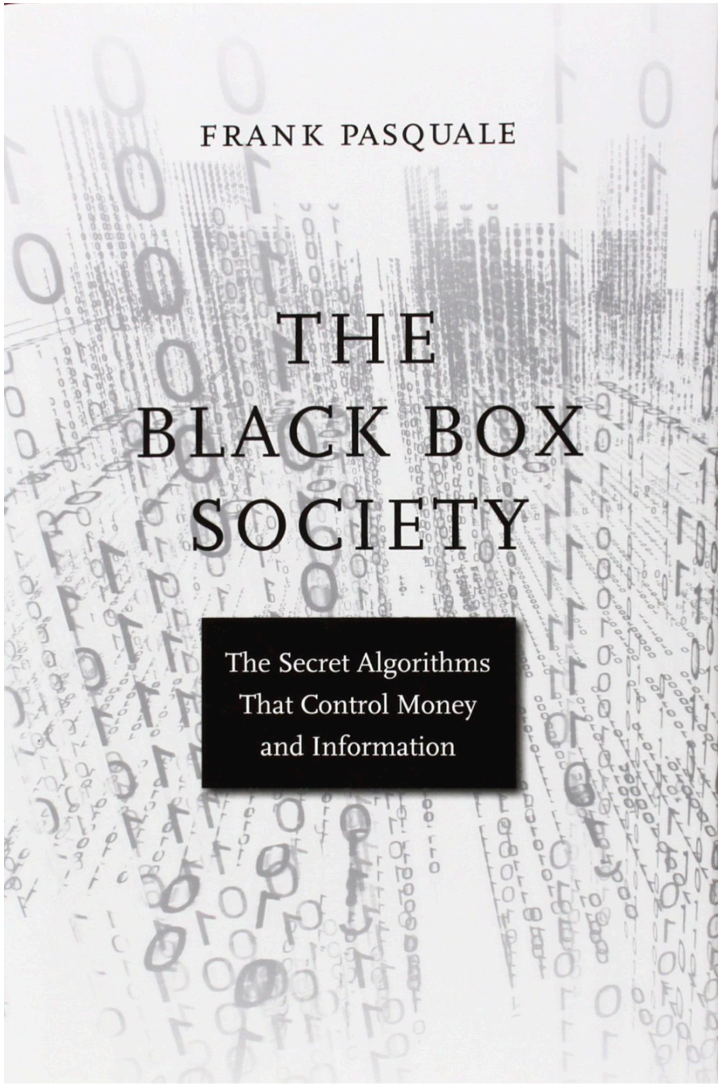
which has a lower confidence of $conf = 57.78\%$.

Right of explanation

- Applying AI within many domains requires **transparency** and **responsibility**:
  - health care
  - finance
  - surveillance
  - autonomous vehicles
  - Government

- EU General Data Protection Regulation (April 2016) establishes (?) a right of explanation for all individuals to obtain "meaningful explanations of the logic involved" when automated (algorithmic) individual decision-making, including profiling, takes place.

- In sharp contrast, (big) data-driven AI/ML models are often *black boxes*.



Accountability

"Why exactly was my loan application rejected?"

"What could I have done differently so that my application would not have been rejected?"


nature International weekly journal of science

Home | News & Comment | Research | Careers & Jobs | Current Issue | Archive | Audio & Video | For Au

Archive > Volume 537 > Issue 7621 > Editorial > Article

NATURE | EDITORIAL

More accountability for big-data algorithms

To avoid bias and improve transparency, algorithm designers must make data sources and profiles public.

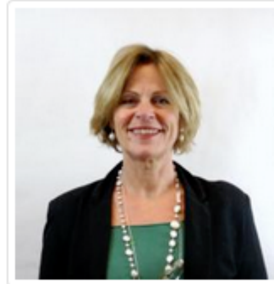
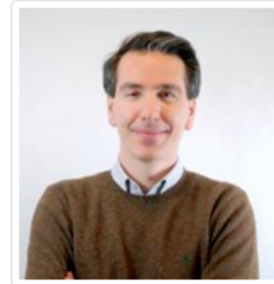
21 September 2016

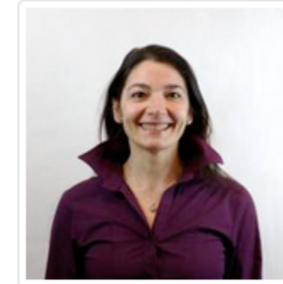
SoBigData Research Infrastructure

Social Mining & Big Data Ecosystem

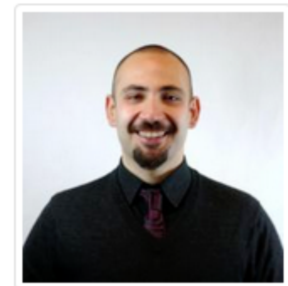
www.sobigdata.eu



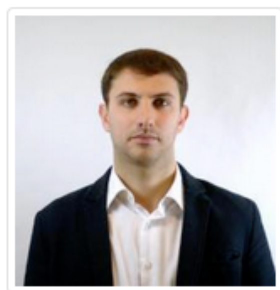
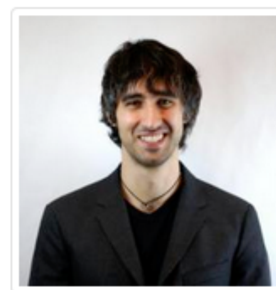



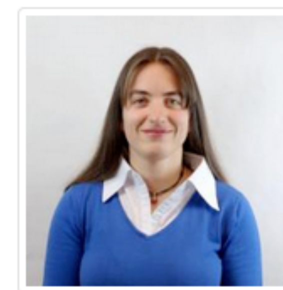










Knowledge Discovery & Data Mining Lab

http://kdd.isti.cnr.it

Special thanks

- Salvatore Ruggieri
- Franco Turini
- Fosca Giannotti
- Anna Monreale
- Luca Pappalardo
- Francesca Pratesi
- Riccardo Guidotti

Key publications in data privacy

- D. Pedreschi, S. Ruggieri, F. Turini: *Discrimination-aware data mining*.14th ACM SIGKDD 2008: 560-568
- (F. Giannotti, D. Pedreschi, Eds.) *Mobility, Data Mining and Privacy*. Springer, ISBN 978-3-540-75176-2, 2008
- Monreale A, Andrienko GL, Andrienko NV, Giannotti F, Pedreschi D, Rinzivillo S, Wrobel S (2010) Movement data anonymity through generalization. Trans Data Privacy 3(2): 91-121
- F Giannotti, LVS Lakshmanan, A Monreale, D Pedreschi, H Wang. P*rivacy-preserving mining of association rules from outsourced transaction databases. IEEE Systems Journal*, 7 (3), 385-395 (2013)
- Sara Hajian, Josep Domingo-Ferrer, Anna Monreale, Dino Pedreschi, Fosca Giannotti. *Discrimination- and privacy-aware patterns.* Data Min. Knowl. Discov. 29(6): 1733-1782 (2015)
- A Monreale, S Rinzivillo, F Pratesi, F Giannotti and D Pedreschi. *Privacy-by-design in big data analytics and social mining*. EPJ Data Science, 10, 2014
- Riccardo Guidotti. Personal Data Analytics. PhD Thesis, Univ. Pisa, 2017
- F Giannotti, A Monreale et al., Privacy risk assessment for big data analytics. Submitted. (2017)