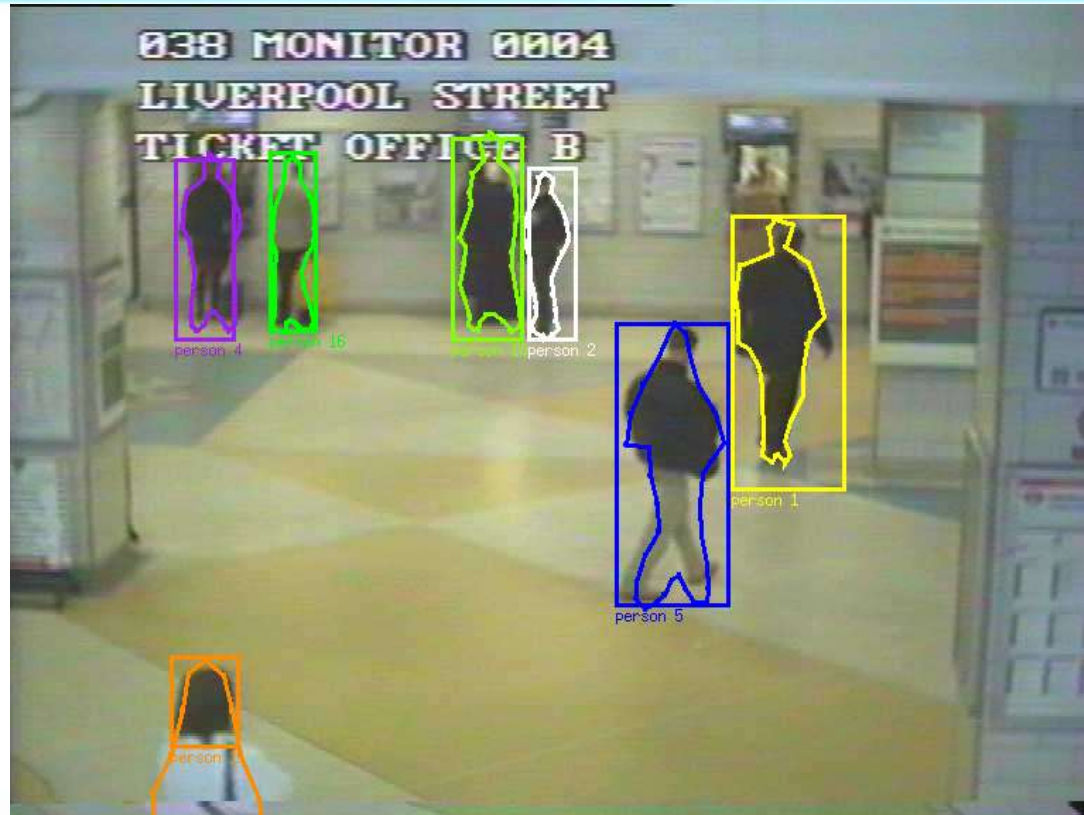


# User Privacy in Mobility Data

**Chiara Renso, Beatrice Rapisarda**

ISTI - National Research Council of Italy (CNR)

# We are tracked everywhere and everytime



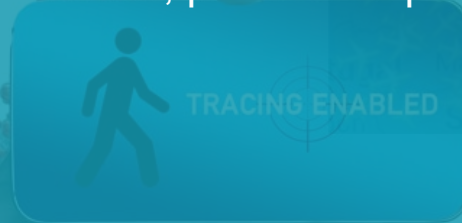






# Mobility data analysis: benefits for the society

- + **Climate Change:** analyse transport data to reduce emissions
- + **Health:** monitoring and containing the outbreak.
- + **Socio-economic** and behavioural research
- + **Support tourism** monitoring and recommendation
- + **Urban Planning:** security, traffic control, public transportation management
- + ..... many others!!



# Mobility Data and Privacy

According to **GDPR**, location data are personal data (Article 4).

Location data become sensitive data when, from the analysis of movement, we can infer **sensitive information** like religious preferences or health status.



From location data, even when de-identified, it is sometimes possible to infer the identity of the person: **our tracks identify us.**

It has been shown that 4 spatio-temporal locations are sufficient to identify the 95% of mobile phone individuals individuals. Two points are sufficient to identify 50% of users(\*).

(\*) *Unique in the Crowd: The privacy bounds of human mobility*  
Yves-Alexandre de Montjoye, César A. Hidalgo, Michel Verleysen, Vincent D. Blondel. *Nature*, Volume 3, Article Number: 1376 (2013)





# Re-identification case

**FROM** de-identified dataset of NY City Taxi **TO** identification of travellers using **only** public data

De-identified NY City Taxi rides released (pick-up, drop-off, time, fare amount, tips)

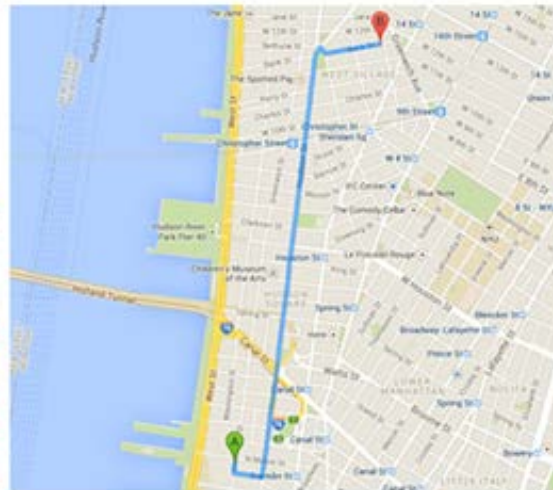
Someone mapped the information using public images of celebrities getting in or out of taxis.

They used this information to isolate journey details for Bradley Cooper and Jessica Alba.

→ We know the tips given and where they stopped



BRADLEY COOPER



JULY 8, 2013 • 7:34 PM - 7:44 PM  
376 GREENWICH ST. TO 13 BANK ST.  
\$9.00 FARE • CASH; UNKNOWN TIP • ©SPLASH

People can do the same for a relative or a colleague with **a small amount of auxiliary knowledge**, allowing anyone to figure out the locations to which a given individual has travelled.

**The data de-identification sometimes is not enough!**



# We live in a world with highly interlinked data

## MASTER

Multiple Aspects Trajectory Management and Analysis  
H2020-MSCA-RISE-17 GA 777695

We study methods for analysing trajectory data linked to several additional aspects (e.g. social media, weather, traffic conditions, news, etc).

We consider a **privacy - by - design** approach where new anonymization methods need to be developed for highly linked data

We are assisted by an **Independent Ethics Advisor - Prof. Bettina Berendt** - to support us on ethics issues related to the analysis of these data



This project has received funding from the European Commission's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement N. 777695. The content of this presentation reflects only the author's view and the Agency is not responsible for any use that may be made of the information it contains.



# Privacy - by - design



## New mobility data anonymization methods that preserve the privacy

Under some conditions, from the anonymized dataset **it is not possible** to re-identify the individual.

- K-anonymity/l-diversity/t-closeness
- Obfuscation
- Differential Privacy
- ... many others ....

# The privacy debate

Should we renounce to our privacy for the public good (e.g. COVID)?



Put in other words, which is a good trade off between privacy and data utility?

In many cases we can **preserve the privacy of the individuals** while still benefiting the society, thanks to new privacy preserving technologies, algorithms and analysis methods



# The role of the new technologies and the advancements of research



new  
technologies  
increase the  
privacy  
preservation



new  
technologies  
increase the re-  
identification  
risk

# Collect as little data as possible

COVID -19: tracing apps use the Bluetooth Exposure Notification → no trajectories, no location, no personal data, encrypted IDs, only information about a possible exposure to positive individuals.



**This is the minimum information we need for the objective of COVID contact tracing.**

Sometimes we just need to apply a kind of *Occam razor* concept to data collection and analysis!

# Creating Trust circle



We need to create trust  
in technology: positive  
circle!



# Ethics: in the law and beyond the law (\*)



**The laws that protect data (such as the GDPR) embody many ethical principles**

→ giving users access to their data, deleting data on request, respecting consent and purpose limitation

**.... but just because something is legal, doesn't mean it's ethical!**

→ you may get consent from data subjects for a processing of data with a deeply unethical goal.

**We need to look at the bigger picture of the goals and methods of mobility data analysis.**

**-- “Ethics by design?” --**

(\*) Thanks to the contribution of our IEA Prof. Bettina Berendt



DATA SCIENCE FOR MOBILITY

# Check out MASTER next events!

## Data Science for Mobility summer school

To be held in Santorini in April 2021 (.... depending on the COVID restrictions...)

with an interesting panel on Emerging issues on mobility data science  
(including Privacy and Ethics)

<http://master-school.isti.cnr.it>



This project has received funding from the European Commission's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement N. 777695. The content of this presentation reflects only the author's view and the Agency is not responsible for any use that may be made of the information it contains.



# MASTER

Multiple Aspects Trajectory Management and Analysis  
H2020-MSCA-RISE-17 GA 777695 - 2018-2022



Master Project H2020



[twitter.com/masterh2020](https://twitter.com/masterh2020)



<http://www.master-project-h2020.eu/>



This project has received funding from the European Commission's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement N. 777695. The content of this presentation reflects only the author's view and the Agency is not responsible for any use that may be made of the information it contains.





# Thank you!

For further information please contact:

[chiara.renso@isti.cnr.it](mailto:chiara.renso@isti.cnr.it)

[beatrice.rapisarda@iit.cnr.it](mailto:beatrice.rapisarda@iit.cnr.it)

**ESOF2020**  
EUROSCIENCE OPEN FORUM  
**TRIESTE**

