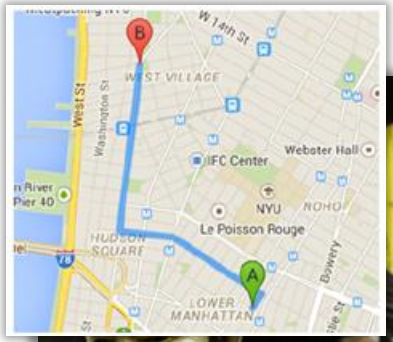
Privacy beyond confidentiality, data science beyond spying:

From movement data and data privacy towards a wider fundamental rights discourse

Bettina Berendt
https://people.cs.kuleuven.be/~bettina.berendt

Annual Privacy Forum 2019

Rome, 13 June 2019



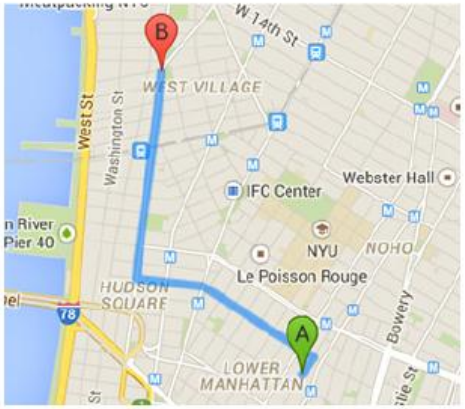
Plan

- 2 case studies related to data (protection)
  - Movement data, privacy, and other rights
- Why is this interesting for us as APFers?
  - (In)visibility, privacy, and other rights
- What can we do?

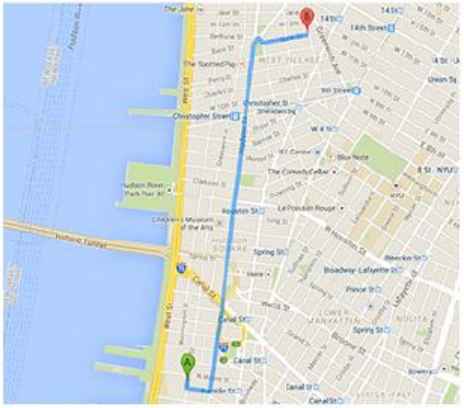
Two case studies related to data (protection): Movement data, privacy, and other rights

Case study 1: New York City taxi rides

- Dataset released in 2014, in response to Freedom of Information request
- Fields:
  - Pseudonymised taxi ID and driver ID,
  - startpoint,
  - endpoint,
  - times,
  - fare, tip, …



JANUARY 6, 2013 • 3:29 PM - 3:38 PM
78 CROSBY ST. TO 580 HUDSON ST.
$7.50 FARE • $2 TIP • ©SPLASH



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

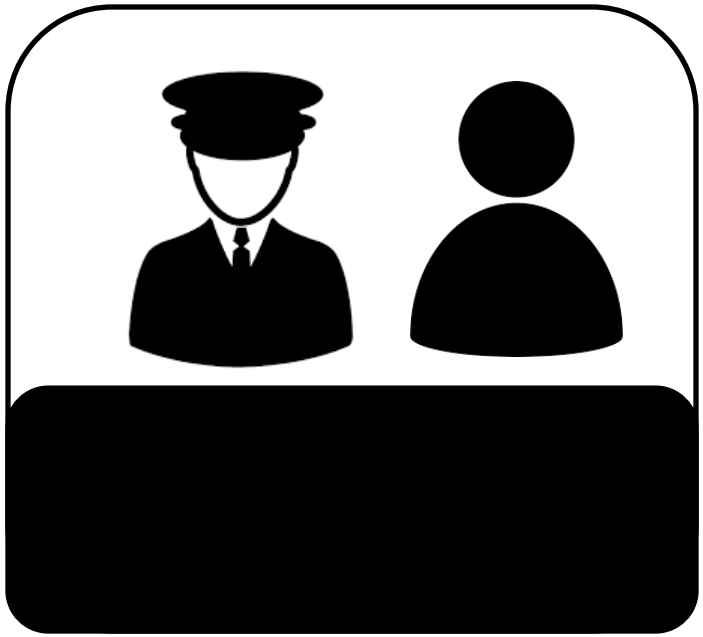
Useful data!

E.g. “optimization of the revenue of NYC Taxi Service using Markov Decision Processes"

State authorities

Civil society

Researchers



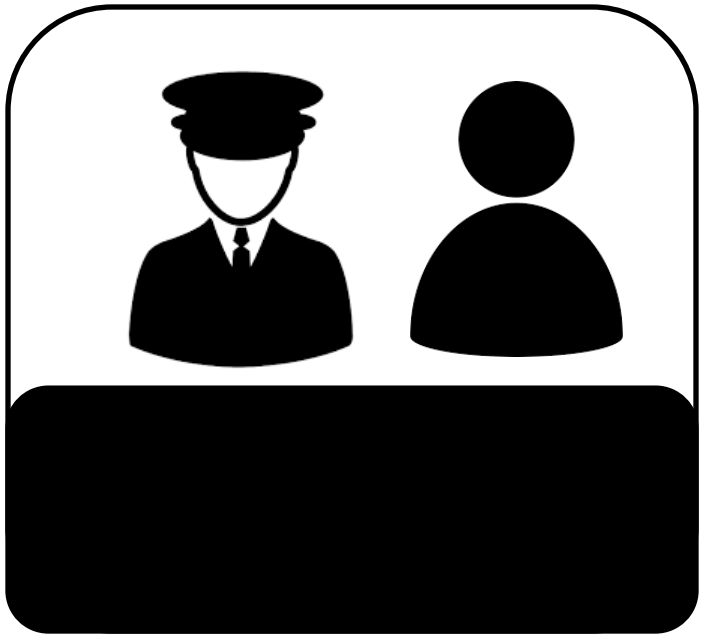
De-de-identification!

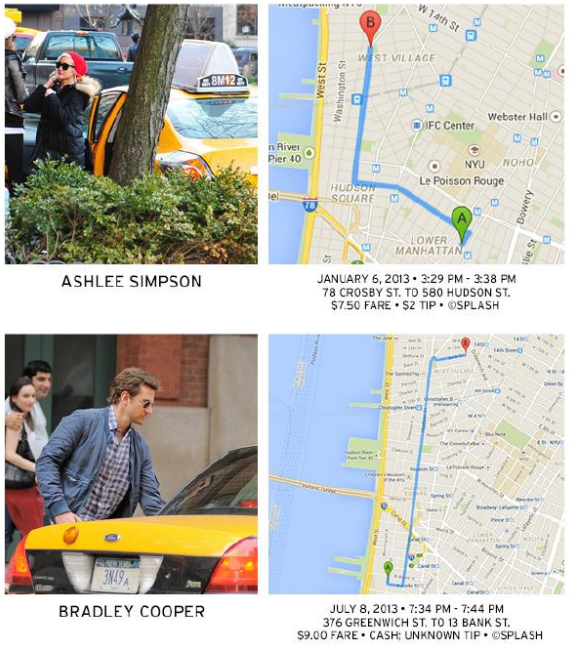
State authorities

Paparazzi

Civil society

Researchers





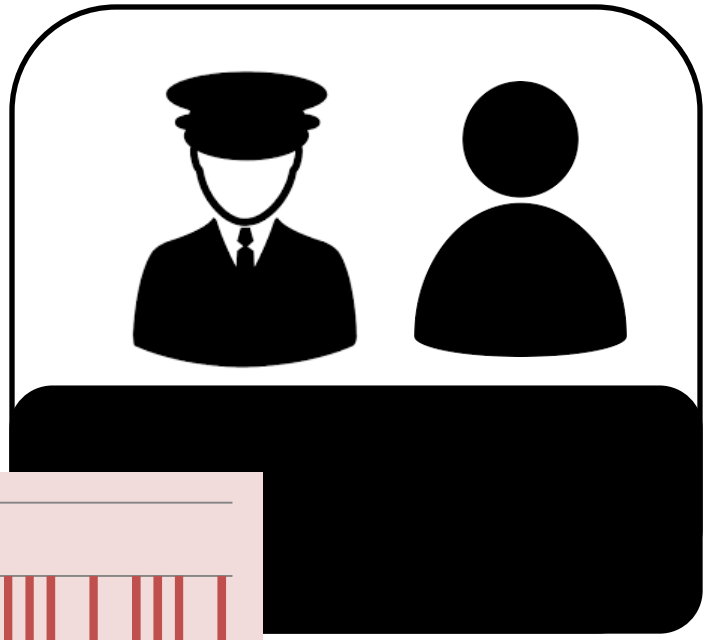
De-de-identification! Profiling!

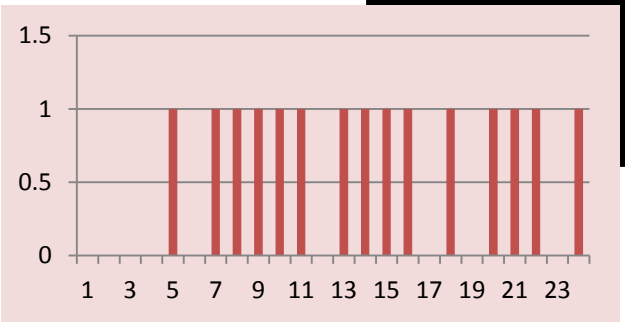
State authorities

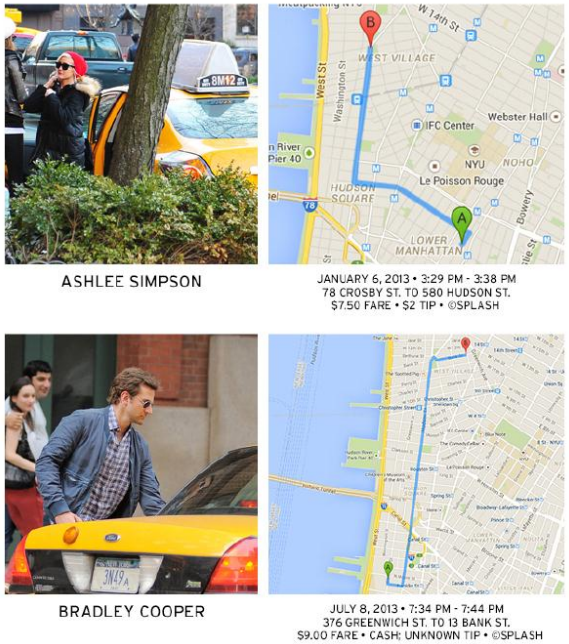
Paparazzi

Civil society

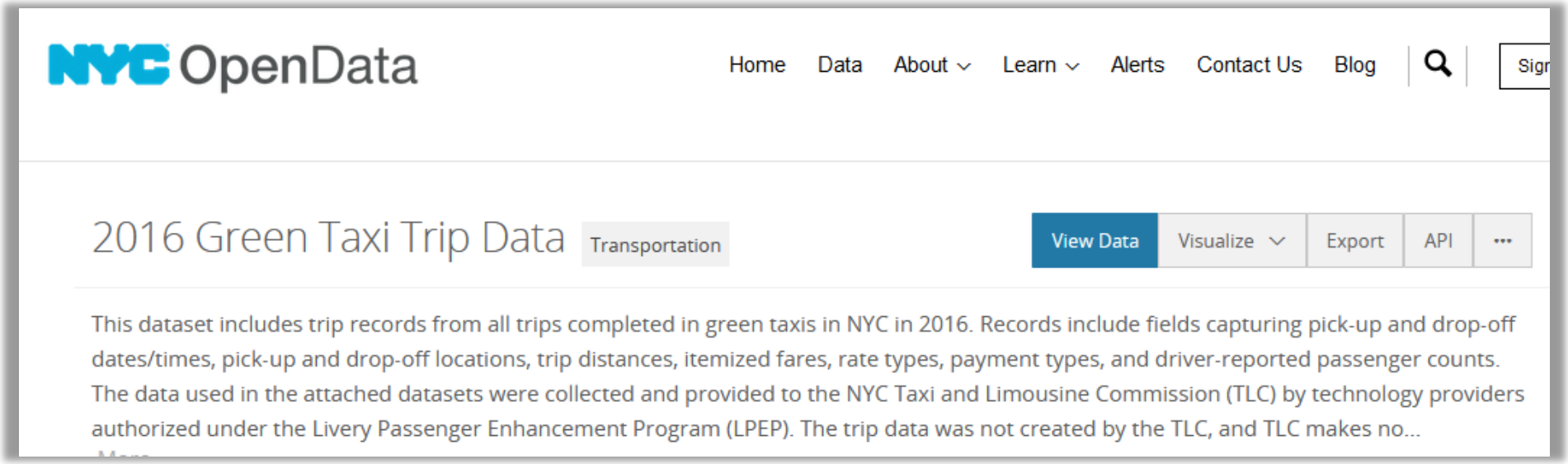
Researchers







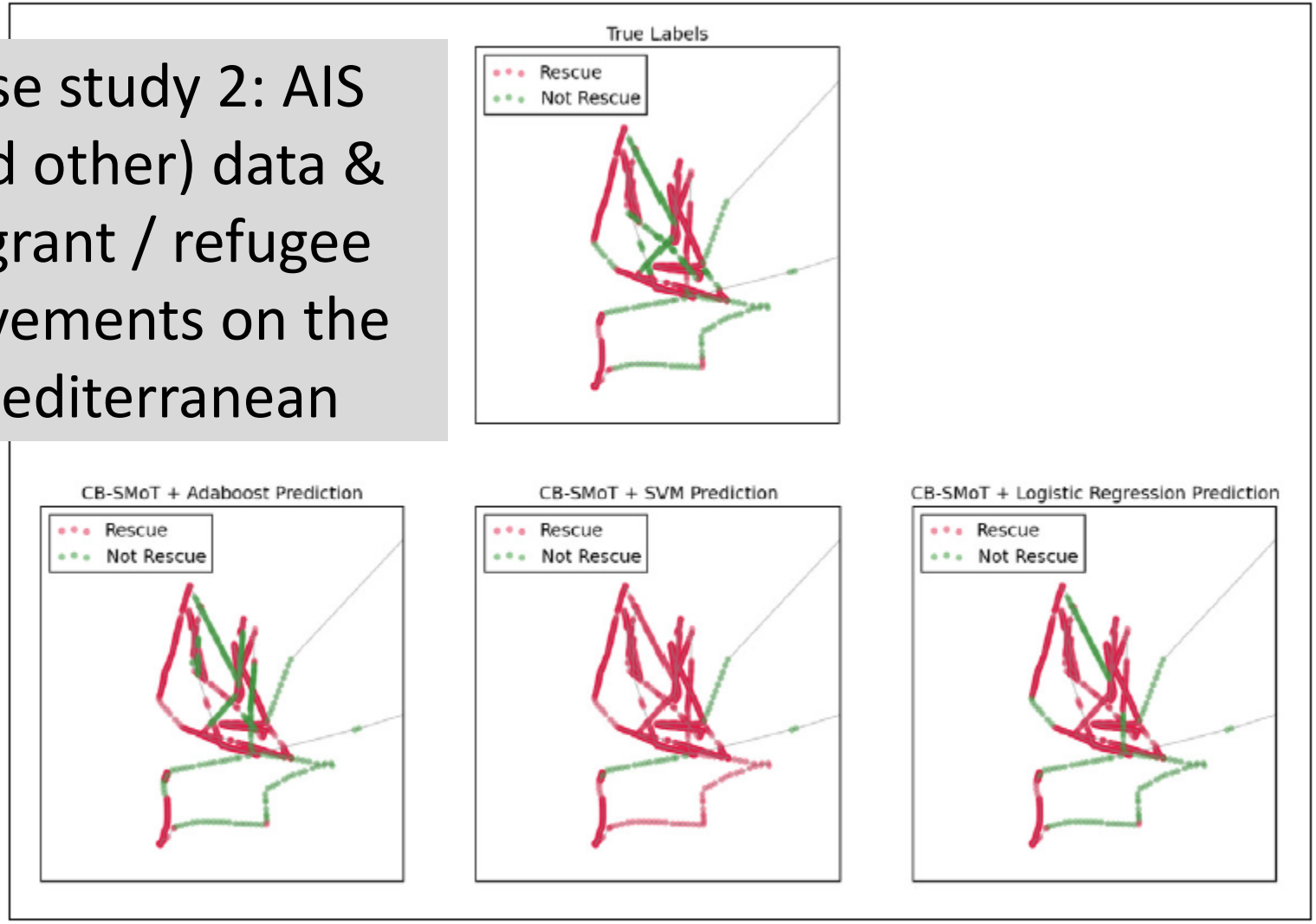
Solution approach? – Data suppression



- “removal of taxi identifiers […] adversely impact[s] certain types of analysis on the data“
- may / should European researchers use the 2013 dataset?

Figure 10. Classification of a sample trajectory

Case study 2: AIS (and other) data & migrant / refugee movements on the Mediterranean



Source: UN Global Pulse, 2017.

(Hoffmann et al., 2017)

Note: The top cell describes a "ground truth" trajectory from a single ship, where rescue and non-rescue data points have been manually tagged through a visual inspection of ship location, speed and behaviour. A clustering algorithm, CB-SMoT, was used to automatically group similar points within the trajectory based on how far apart they were in distance and time (Palma et al., 2008). Three standard classification methods (AdaBoost, support vector machines and logistic regression) were then applied to these clusters; the results are shown in the second row of the figure.

Useful data!

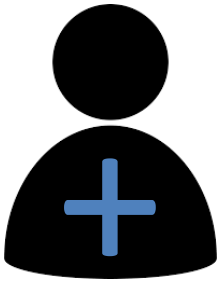

Roma
Napoli
Bari
Shqipëria
Θεσσαλονίκη
Tekirdağ
İstanbul
Bursa
Balıkesir
Kütahya
İzmir
Afyonkarahi
Konya
Aydın
Denizli
Antalya
València
Palma
Palermo
Alger
Constantine
Tiaret
Djelfa
Batna
تونس
سوسة
طرابلس
بنغازي
El Bayadh
البيض
Ghardaia
غرداية
Ouargla
ورقلة
نالوت
الجبل الغربي
سرت
البطنان
مطروح
الواحات
الجفرة
وادي الشاطئ
Illizi
إليزي
الإسكندرية
بورسعيد
القاهرة
السويس
المنيا

Useful data! But …?

Why is this interesting for us?
(In)visibility, privacy, and other rights

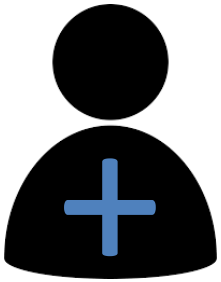
De Hert & Gutwirth 2006: interests in visibility (+) resp. invisibility (–)

Power, unchecked

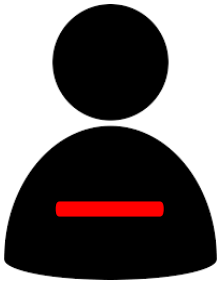


De Hert & Gutwirth 2006: interests in visibility (+) resp. invisibility (–)

Power, unchecked



Society under the rule of privacy and data protection law



Some stakeholders' interests in visibility (+) resp. invisibility (–): Privacy and other reasons

Civil society

+

State authorities

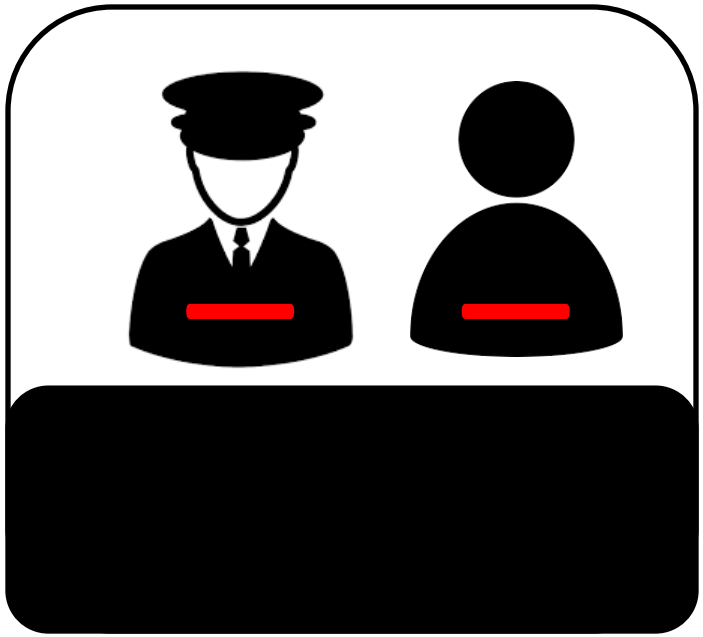
+

Paparazzi

+

Researchers

+ –



Interests in visibility (+) resp. invisibility (–): Privacy, security and other reasons

Vessel owners

* **Coastguard** +
* **NGO** +
* **commercial** ?

State authorities

–

Pirates

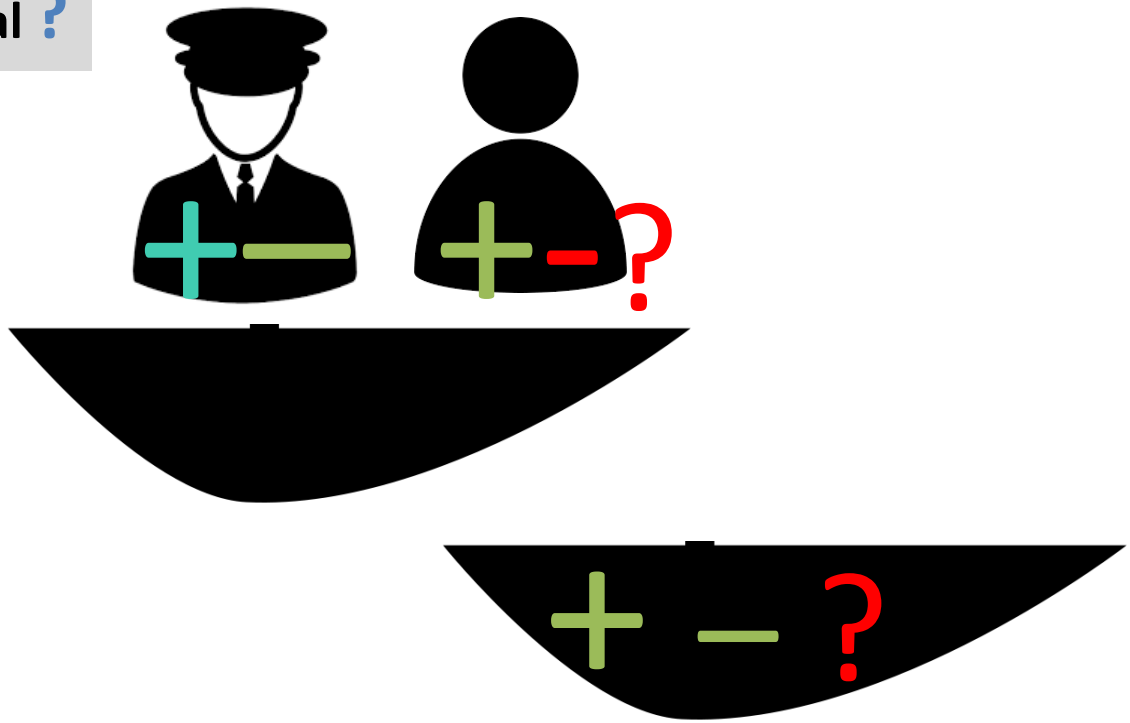
+

Researchers

?

Civil society

+



What happens when actors proliferate?

Vessel owners

*** Coastguard**

*** NGO**

*** commercial**

State authorities

Pirates

Researchers

Civil society

From *invisibility ~ privacy* to *visibility ~ the right to have rights* – on the interdependency of data protection and other areas of law

From *invisibility ~ privacy* to *visibility ~ the right to have rights* – on the interdependency of data protection and other areas of law

Various places, throughout the 2010s

From *invisibility ~ privacy* to *visibility ~ the right to have rights* – on the interdependency of data protection and other areas of law

Various places, throughout the 2010s

MIT, 2017+

From *invisibility ~ privacy* to *visibility ~ the right to have rights* – on the interdependency of data protection and other areas of law

Various places, throughout the 2010s

Rome / Italian parliament, 28 Nov 2018

MIT, 2017+

From *invisibility ~ privacy* to *visibility ~ the right to have rights* – on the interdependency of data protection and other areas of law

Various places, throughout the 2010s

Rome / Italian parliament, 28 Nov 2018

MIT, 2017+

Berlin, 2019: Inviolability also of the refugee home

What can we do?
Data, tools and narratives

Research and narratives

- Every data modelling,
- every privacy / data protection modelling,
- every tool (including PETs)

tells a story.

Solution approach? – What to represent in holistic trajectory data?

- The **context**?
  - E.g. AIS, EUROSUR
- Factors that **influenced** the data
  - E.g. who gave instructions (Italians or Libyans)
- **Impacts** of the data / represented world
  - E.g. costs and incentive effects

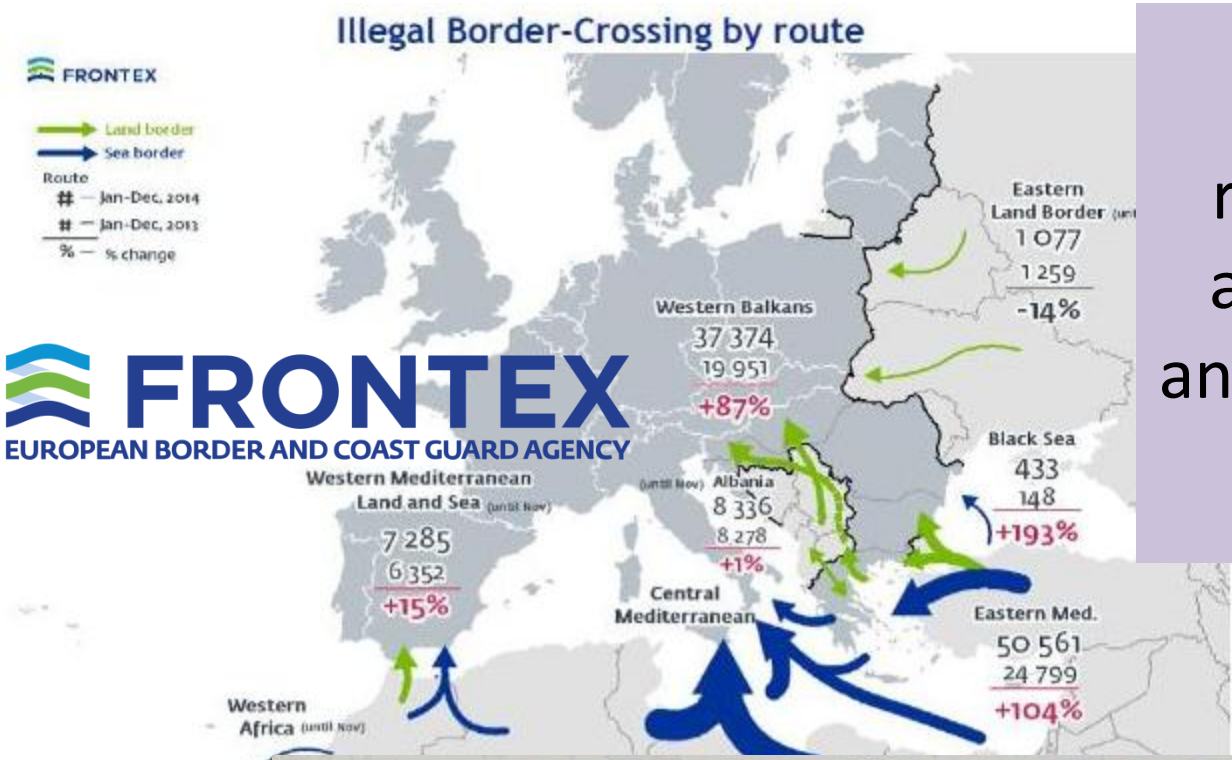
- Who has **access** to the data (and does or does not react to alerts, and if so, how)?
  - E.g. Libyan coastguard

Spanish trawler causes chaos during Libya Coast Guard rescue operation

Sunday, November 25, 2018 - 21:27

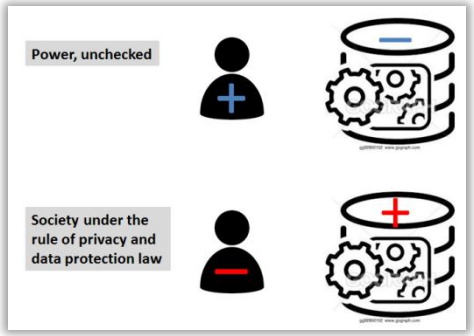
A Libyan Coast Guard patrol rescued on Friday 27 illegal immigrants off Khmos city while they were en route to the Italian shores.

What *ARE* the represented events anyway? Narratives and counter-narratives with Big Data

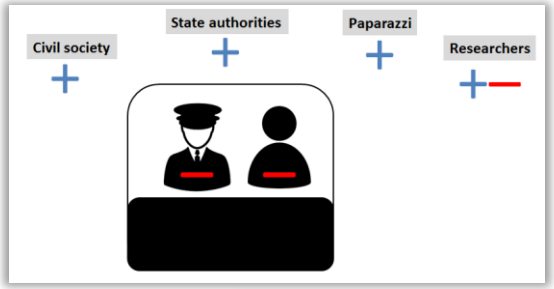


Future work: Towards more diverse research narratives around data, data protection, and fundamental rights and freedoms

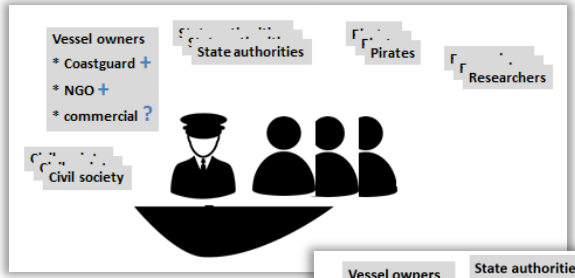
- Data protection against power imbalances



- Data for (e.g.) revenue, and the right to be let alone



- Group privacy, group data protection



- Data for Social Good, and the freedom from unreasonable constraints on the construction of one's own identity