

Privacy laws and pervasive sensing / big data: forever incommensurate?

Trusted Personal Data Management Panel, TdW Conference
Max Mühlhäuser

Technische Universität Darmstadt





EU Charter of fundamental rights:

Art. 8: Protection of personal data

Everyone has the right to the protection of personal data concerning him or her.

Such data must be processed fairly **for specified purposes** and on the basis of the **consent of the person concerned** or some other legitimate basis laid down by law

Everyone has the right of access to data which has been collected concerning him or her, and the right to have it rectified. Compliance with these rules shall be subject to control by an independent authority.



anonymization



data thrift



Big Data Incommensurate w/ Data Thrift



TECHNISCHE
UNIVERSITÄT
DARMSTADT

Two of the 'big data' principles:

collect more data¹ && don't throw any data away²

Why? Just two examples (given by Prof. Mark Whitehorn, Univ. Dundee²):

“Even information which might seem insignificant, can be incredibly useful with the correct use. Google decided to keep all the information from **users' spelling mistakes**... They looked into what was typed and what the user was trying to say. ... they have established that they can still direct users to where they want to go... They have effectively created the most **powerful spellchecker in the world** just by using **data that others would have thrown away.**”

“When you put your PIN into a cash machine, you **put it in at a very precise speed** because you know it. If someone steals your card, they're highly unlikely to enter it at a similar speed so banks can use this as a further method of authentication.”

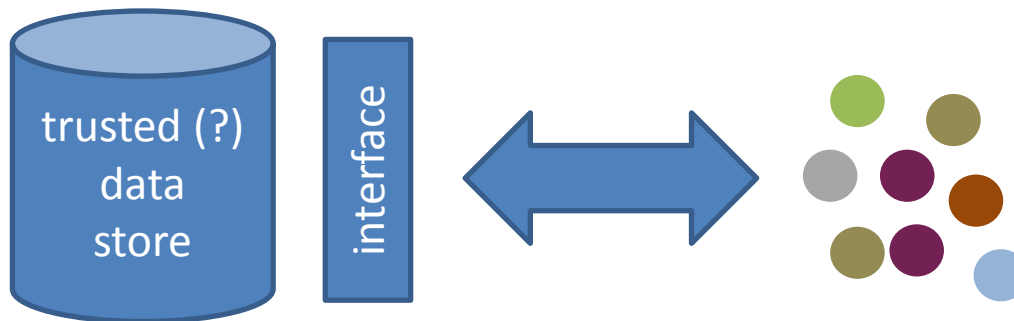
1: S. Nelson: 10 Tips for Better Big Data Analysis, cf. <http://www.dummies.com/how-to/content/10-tips-for-better-big-data-analysis.html>

2: C. Bailey, Blog about Big Data Summit 2012, cf. <http://christianbailey.net/uncategorized/dont-throw-any-data-away/>



The Data Anonymization Myth

- Privacy Laws are about PII (Personally Identifiable Information)
- Prof. Anind Narayanan (Princeton U)¹:
 - Priv. laws: “scrub ‘PII’ in a way that *prevents the possibility of re-identification*”
 - Anind: “... *essentially impossible* ... in a foolproof way w/o losing the utility of the data”
- *Consequence:*
 - *privacy is not an issue of data but of data processing!!*
 - *needed: **trusted data store** + well-controlled **interface** + case-based **rules***

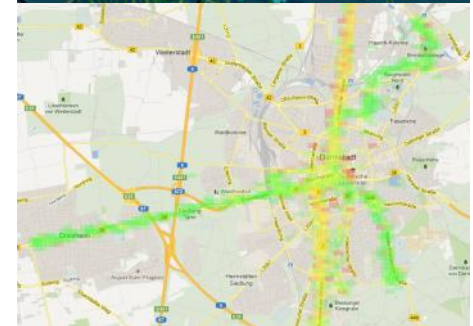
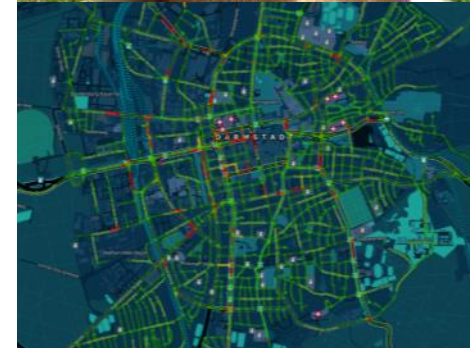


1: Privacy and Security: Myths and Fallacies of “Personally Identifiable Information” .CACM 53 (6), 2010, pp. 23-25



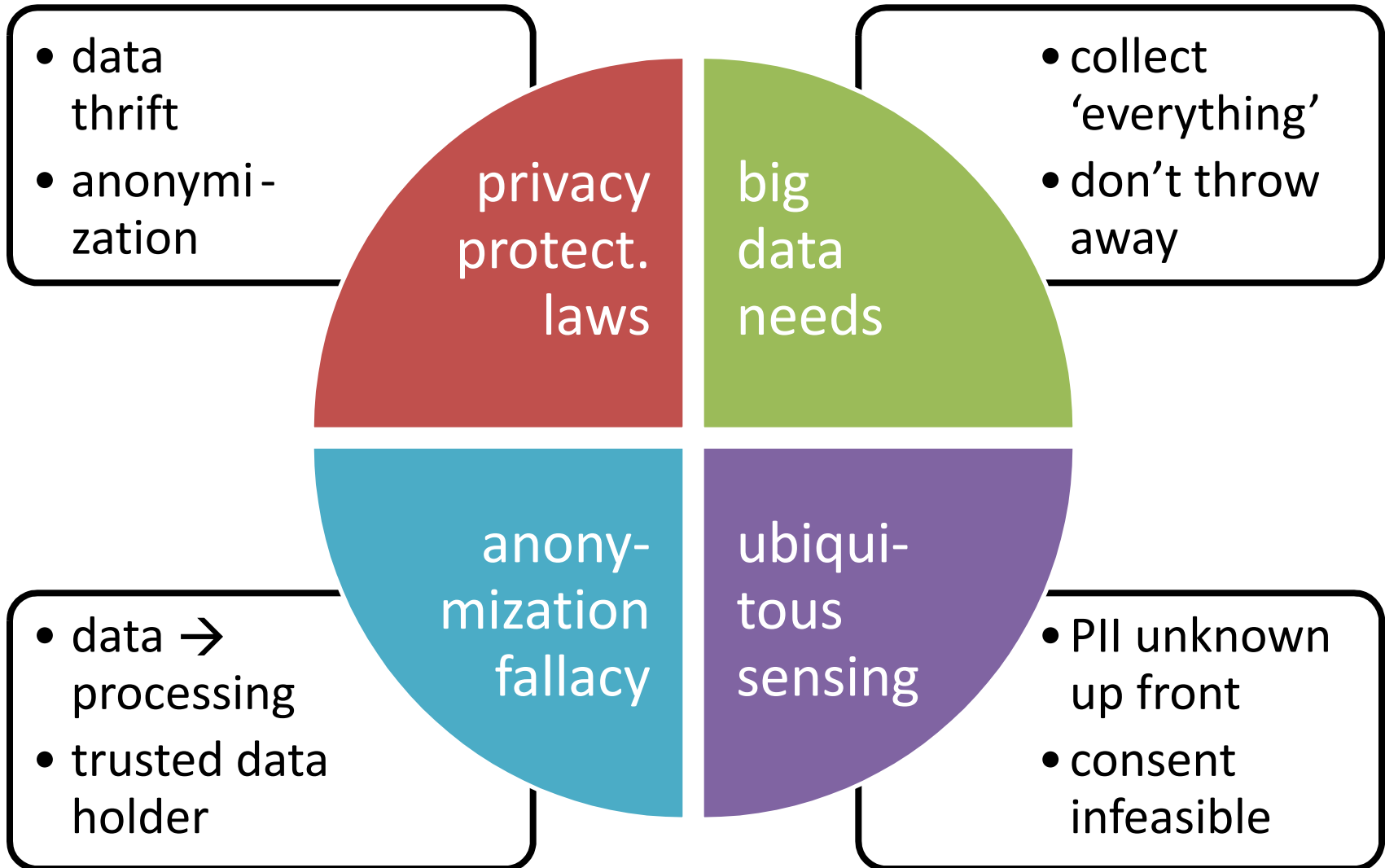
Sensor Data: The End of Privacy?

- many, many, many devices ...
- ... collecting data that is *considered* “non-PII”
- ... but that *can be linked* (today, tomorrow)
- ... often *concerns by-passers* (consent illusory!)
- ... and that is *stored @ zillions of systems*





Dilemma summary = discussion base





THANK YOU!