Engaging Content
Engaging People

Investigating Conditional Data Value Under GDPR

Harshvardhan J. Pandit [TCD], Plamen Petkov [DCU],
Declan O'Sullivan [TCD], Dave Lewis [TCD]
ADAPT Centre - [TCD] Trinity College Dublin - [DCU] Dublin City University
https://openscience.adaptcentre.ie/ → GDPR work
pandith@tcd.ie | @coolharsh55
email
twitter

The ADAPT Centre is funded under the SFI Research Centres Programme (Grant 13/RC/2106) and is co-funded under the European Regional Development Fund.

Data Value

Layman's Guide to Data Value

- consider data as
  - an asset
  - an entity
- track of data
  - cycles
  - flows
  - chains
  - networks

Availability

core & implicit assumption that data will always be available

BUT

does not apply to open data!

What if the data that is available

now

have value today

may not be available

later

loss of value tomorrow

??? ??? ???

When can this happen?

Assumptions / Conditions / Specificity

1. consider dealing with personal data
2. personal data is governed by data protection laws
3. laws such as the GDPR affect data availability

→ personal data becomes unusable from the legal basis p.o.v.

Is this a new problem?

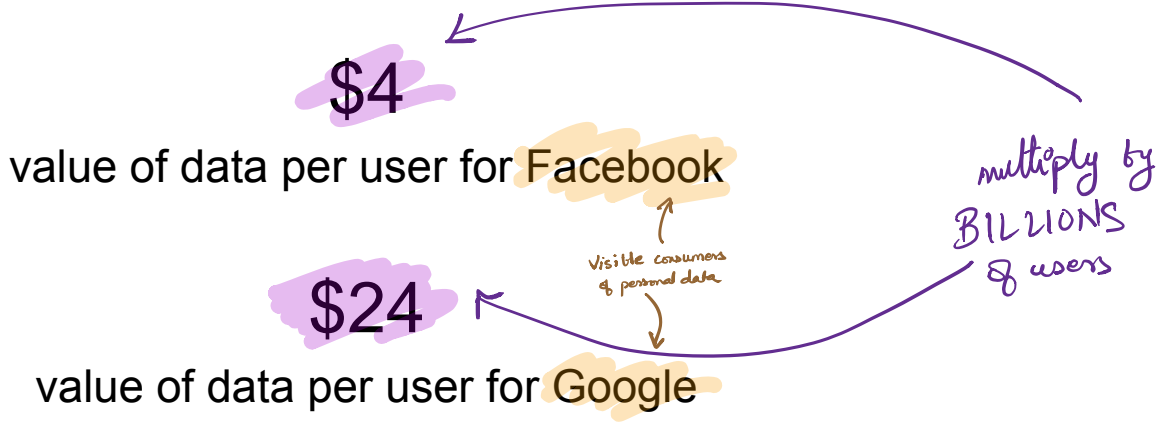
NOT A NEW PROBLEM

PRE-GDPR

JPMorganChase

study conducted in 2012

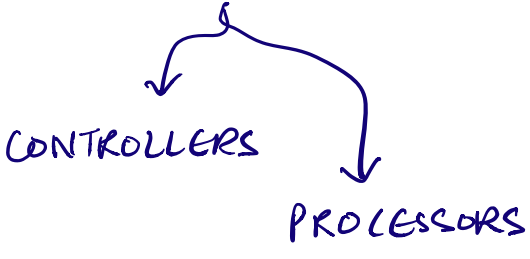
+ 6 years, the value has increased significantly



"Investigating Conditional Data Value Under GDPR"
http://openscience.adaptcentre.ie/ Harshvardhan J. Pandit pandith@tcd.ie @coolharsh55
presented at Governing Value: The Practice of Exploiting Data Value Workshop Co-located with SEMANTiCS 2018

More about GDPR

- Given Consent → prior to data collection
- Right to Withdraw Consent → at any time!
- Right to be forgotten → delete retrospective data
- (Pseudo-)Anonymisation → enforced legally
- Obligations → on data consumers



Incorporating Conditionality

1 **Availability**

> is the data available?

2 **Conditionality**

> when is the data available?

> what events trigger these conditions?

> how to incorporate them into data value frameworks?

methods
approaches
systems
pipelines

Availability as a Metric?

based on

[12] Lim, C., Kim, K.H., Kim, M.J., Heo, J.Y., Kim, K.J., Maglio, P.P., 2018. **From data to value: A nine-factor framework for data-based value creation in information-intensive services.** International Journal of Information Management 39, 121–135.

Initial attempt using data-based value framework

expanded into

(i) data collection
(ii) information creation
(iii) value creation
(iv) distribution through the provider network

from [12]

(1) data source
(2) data collection
(3) data (artefact)
(4) data analysis
(5) information on the data source
(6) information delivery
(7) customer (information user)
(8) value in information use
(9) provider network

⇒ **(10) data availability.**

PROPOSED ADDITIONAL METRIC

Possible Application

activities that depend or are enriched by this data

1. Identify processes that use conditional data

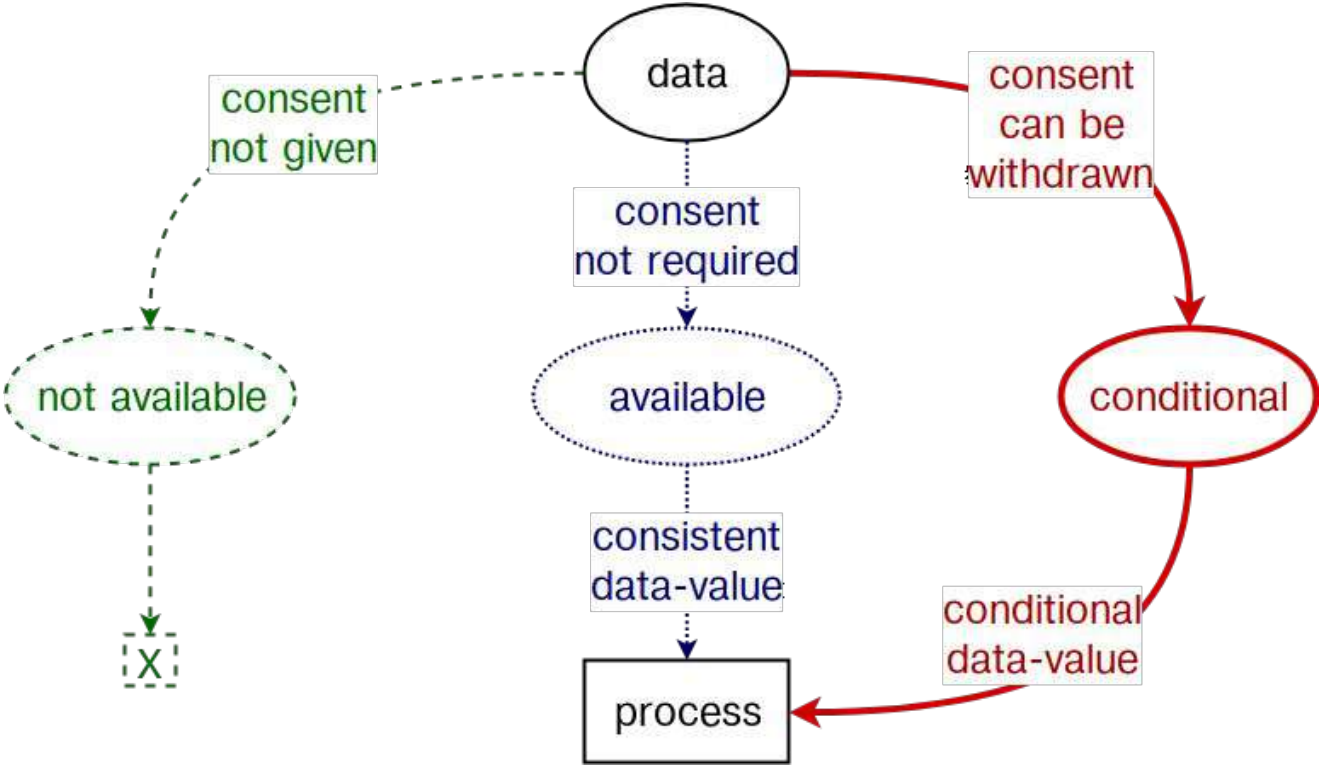
events that trigger

2. Quantify (if possible) the conditionality and availability

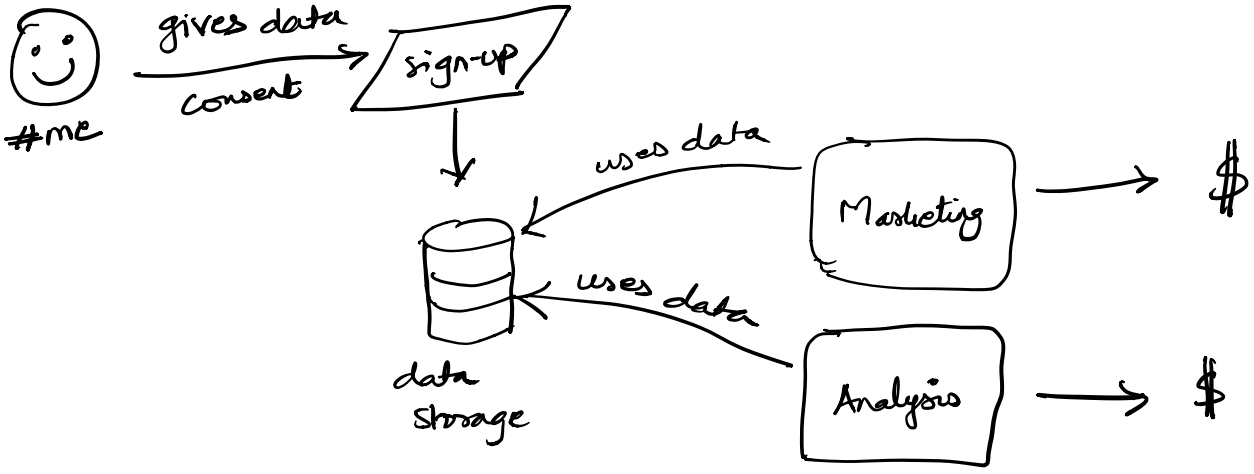
3. Find alternate means to maximise value

this is already being done, but not formalised from the data value POV

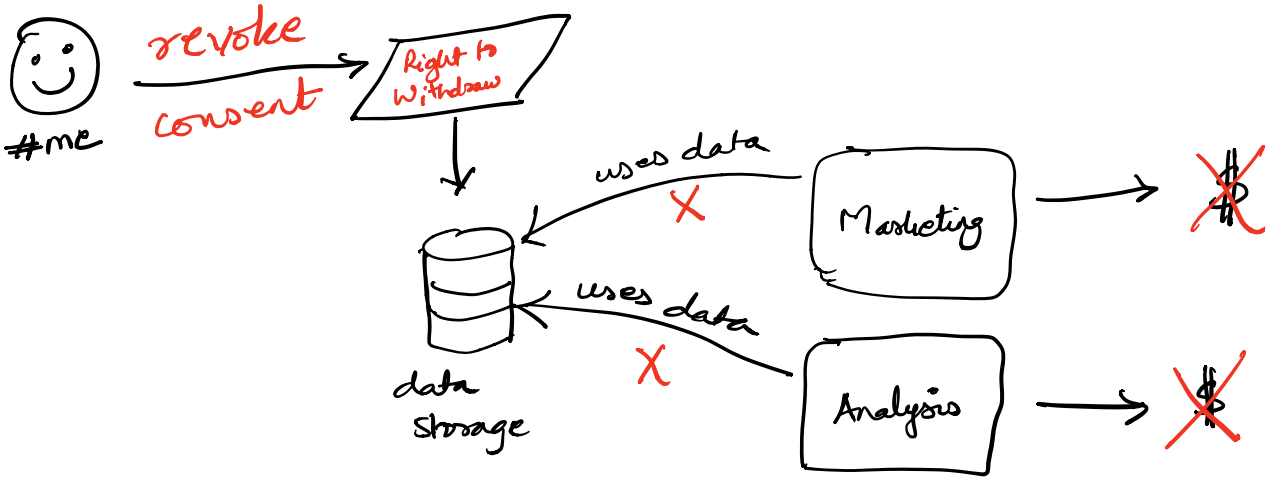
GDPR-Specific Application

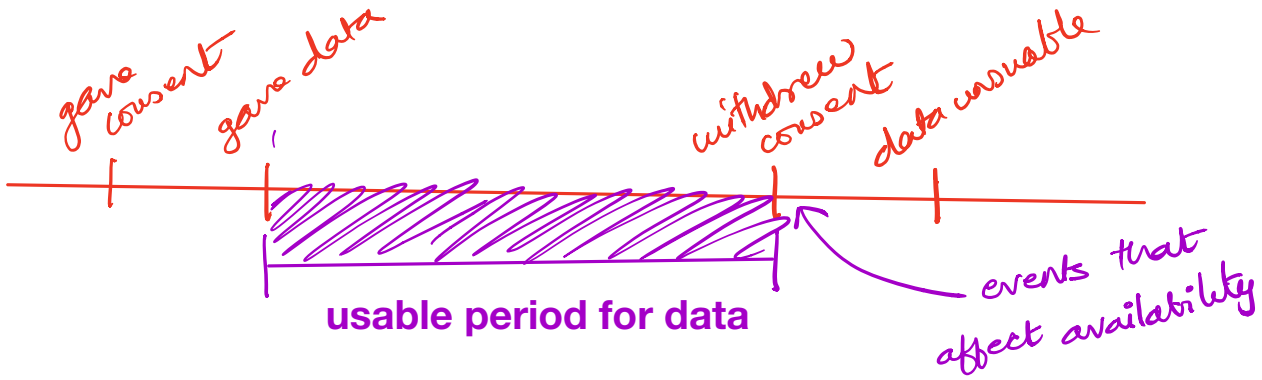


GDPR-Specific Application



GDPR-Specific Application





Data is the new oil….? Let's take that example!

data is the world's most valuable resource

"Investigating Conditional Data Value Under GDPR"
http://openscience.adaptcentre.ie/ Harshvardhan J. Pandit pandith@tcd.ie @coolharsh55

presented at Governing Value: The Practice of Exploiting Data Value Workshop Co-located with SEMANTiCS 2018

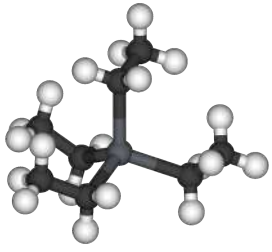
Fallback analogy to oil

All petrol/gasoline stations only provide ↓

WHY?

and what does this have to do with DATA VALUE?

Fallback analogy to oil



tetraethyl-lead

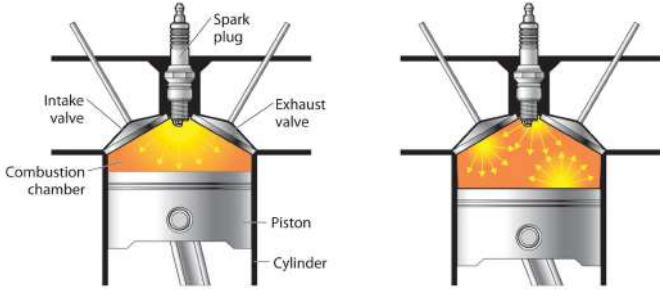
↓

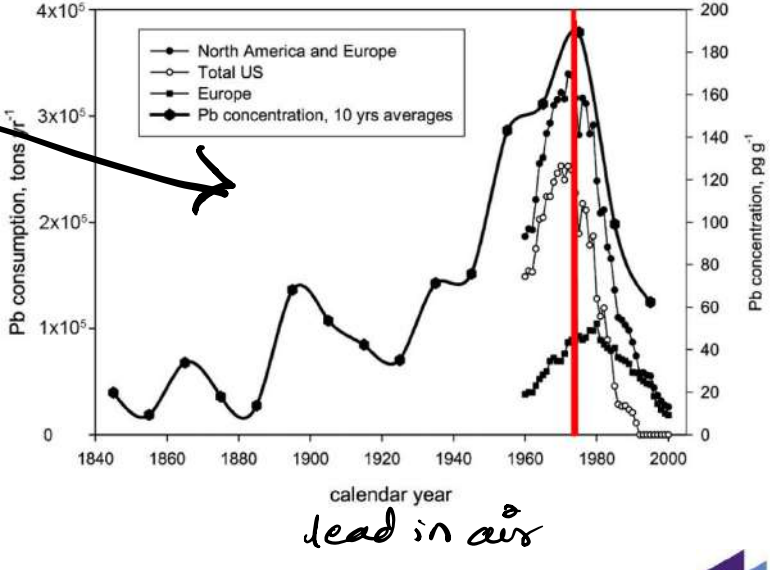
ENVIRONMENTAL HAZARD

↓

BANNED IN ~1970's

REDUCED → KNOCKING





lead in air

Lessons to learn

(a) value of petrol

(b) innovation

Alternate Value of Personal Data

will have some loss of value

Anonymisation
and
Pseudo-Anonymisation

example mechanisms that highlight alternate approaches to extract value

Personalisation / Recommendation / Analytics

Discussion!

How to → ① identify events that affect data value
→ ② formalise them

something other than personal data ⇒ YES!

CONDITIONALITY
- temporal
- geographical
- random
- cost

Can it be applied to
- data that becomes 'stale'?
- data that has a temporal lifespan?