cerre Centre on Regulation in Europe

The role of data for digital markets contestability

Case studies and data access remedies

cerre.eu

Improving network and digital industries regulation

01

VALUE OF DATA AND HARMS

THREE CASE STUDIES HIGHLIGHT IMPORTANCE OF USER DATA

ONLINE SEARCH

Detailed user data especially relevant for search quality regarding new search terms.

E-COMMERCE

Detailed user data especially relevant for recommendations of *new products* and/or recommendations for *new users*

MEDIA PLATFORMS

Detailed user data especially relevant for personalisation and targeted advertising

ECONOMIC VALUE CREATION FROM DATA

Number of users increase 'breadth' of user profiles

- Data often created as by-product of usage (clicks, searches)
- Positive but diminishing returns from broader data sets

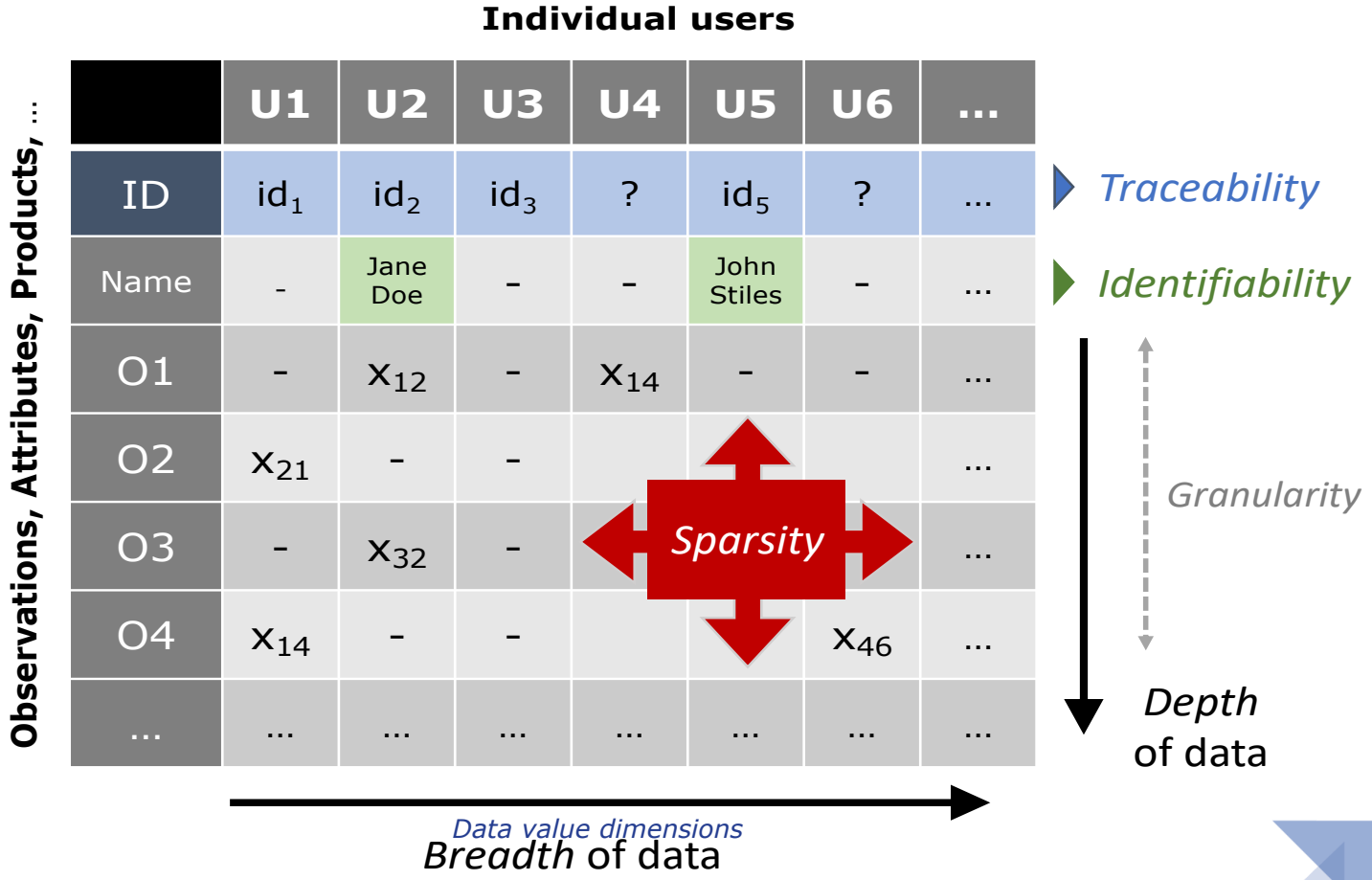
More data on individual users increases 'depth' of user profiles

- More data on user behaviour gradually improves quality of algorithmic learning
- Reinforces benefits of broader data sets

Economic value also influenced by timeliness, accuracy and granularity of data

Complementary inputs needed to derive economic value from data

- computing resources, skilled labour, algorithms


Individual users
Observations, Attributes, Products, ...
U1
U2
U3
U4
U5
U6
...
ID
Name
O1
O2
O3
O4
...
id_1
id_2
id_3
?
id_5
?
...
-
Jane Doe
-
-
John Stiles
-
...
x_{12}
x_{14}
x_{21}
x_{32}
x_{14}
x_{46}
Sparsity
Traceability
Identifiability
Granularity
Depth of data
Data value dimensions
Breadth of data

FEEDBACK LOOPS FROM DATA
DATA-DRIVEN NETWORK EFFECTS

Data-driven **network effects** facilitate market tipping

Data advantage can spill-over to related markets and **induce tipping**

Ancillary data services even allow for **data collection in other markets** without need for competition in these markets

Data-driven network effects create **entry barriers** that lead to **reduced innovation** and ‘kill zones‘ for venture capital

Inherent **efficiencies in data-driven network effects** which can benefit consumers

POLICY OBJECTIVES

- Short-run **contestability of incumbents** through data sharing is unrealistic
- Rather long-run objective of **niche-entry-and-growth**
- Especially **deep and broad behavioural user data** bottleneck resource
- **Data sharing** allows to maintain efficiency of incumbent, while increasing efficiency and ability to innovate of entrants
- Goal: **reduce concentration** of collection of user data

02

POSSIBLE DATA ACCESS REMEDIES

THREE TYPES OF ACCESS REMEDIES

1

LIMITING
THE USE
OF DATA

2

FACILITATING
BULK SHARING
OF AGGREGATED
USER DATA

3

FACILITATING
PORTABILITY
OF INDIVIDUAL
USER DATA

1. LIMITING THE USE OF DATA

- Data silos / data walls
- Shorter data retention periods
- Prohibit buying into defaults
- Line of business restrictions
  - Vertical separation
  - Ancillary data services (authentication services, payment services, web tracking services)
- Privacy Enhancing Technologies

1. LIMITING THE USE OF DATA
CONCLUSIONS

- Limiting the efficiency of the incumbent, rather than enhancing the efficiency of entrants
- Also sceptical about effectiveness/usefulness of
  - Data siloing
  - Shorter data retention periods
  - Prohibit buying into defaults
  - Vertical separation
- However, worth exploring further:
  - Line of business restrictions for ancillary data services
  - Privacy Enhancing Technologies (case-by-case)

2. BULK SHARING OF AGGREGATED USER DATA
PRINCIPLES

- Only **raw user data** (observed and volunteered) may have to be shared
- Only data that was created as a **by-product of consumers' usage** (not volunteered data at the core of the service)
- **Secure** and sufficiently **anonymised**
- **Real-time and continuous** sharing through APIs

2. BULK SHARING OF AGGREGATED USER DATA
TRADE-OFFS

PROTECTING LEGITIMATE BUSINESS INCENTIVES vs. **PROMOTING COMPETITION**

Data as a by-product vs. data as a main product
Existence of viable commercial data offers

USERS' PRIVACY vs. **USEFULNESS OF DATA SET OF ALGORITHMIC LEARNING**

Anonymisation
Data trusts and data sandboxing
Unlawfulness of de-anonymisation

2. BULK SHARING OF AGGREGATED USER DATA
EXAMPLE: SEARCH LOGS

- Publicly shared data with stronger anonymisation
- Individually shared data
  - More detailed data set compiled case-by-case
  - Requires vetting of data recipient
  - Subject to liability and other safeguards

2. BULK SHARING OF AGGREGATED USER DATA
EXAMPLE: SEARCH LOGS

DATA ON THE QUERY	DATA ON THE SEARCH RESULTS PAGE (SERP)	DATA ON THE USER
Keywords (e.g., raw search string, synthetic search string)	Clicked URLs (first clicked result, last clicked result, all clicked results)	Unique identifier
Timestamp (e.g., week, day, hour, seconds)	Zero-Click search (yes/no)	Device metadata (e.g., mobile/ desktop, browser metadata)
Connected queries in the same session	Results ranking (top 3, top 5, top 10)	Location data (IP-address, GPS)
	Layout of the SERP (sponsored results, one-boxes)	Other available user attributes (e.g., age and gender from account data)

3. PORTABILITY OF INDIVIDUAL USER DATA
PRINCIPLES

- Only raw user data (observed and volunteered)
  - Same scope as Art. 20 GDPR
- Requires consumer consent on fine-granular level
  - Disallow extortion of consent through commercial incentives or disincentives
- Real-time, continuous portability using APIs
- Open and secure standards with high reliability and performance

03

CONCLUSIONS

CONCLUSIONS

Especially **observed behavioural user data** important input for algorithmic learning

Data-driven network effects create **significant entry barriers**

Policy focus should be on **niche-entry-and-growth**

CONCLUSIONS

Data sharing preferred over
limiting data use and collection

2 modes of mandated data sharing recommended

'broad' user data in bulk | 'deep' user data (individual consent)

Combination sharing (deep & broad) user data feasible

Careful **case-by-case analysis** required
before mandating bulk data sharing

Centre on Regulation in Europe

Avenue Louise, 475 (box 10)
1050 Brussels, Belgium

+32 2 230 83 60

info@cerre.eu

cerre.eu

@CERRE_ThinkTank

cerre.eu

Improving network and digital industries regulation