The Sign of
Data Externalities

Data Externalities Tutorial - FAccT 2021
Charles Cui, Andreas Haupt, Mihaela Curmei, Yixin Wang, Rediet Abebe

FAccT 2021

Session goals

- Discuss environments where data externalities arise.

- Enable cross-disciplinary conversation about potential solutions.

Quick Round of Intro

1. Name
2. Pronouns
3. Positions (grad student, undergraduate, professor, industry, etc)
4. Research Interest

Network Externalities vs. Data Externalities

Network externalities: the effects on the value of a product or service caused by the change in the number of users.
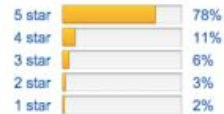
Examples of Network Externalities

1. Road Network   (*negative externalities*)
2. Telecommunication Network   (*negative externalities*)
3. Online Markets   (*positive externalities*)
4. Social Media Platforms   (*positive externalities*)

Data Externalities

Data externalities: the phenomenon that data of some consumers reveal information about others.
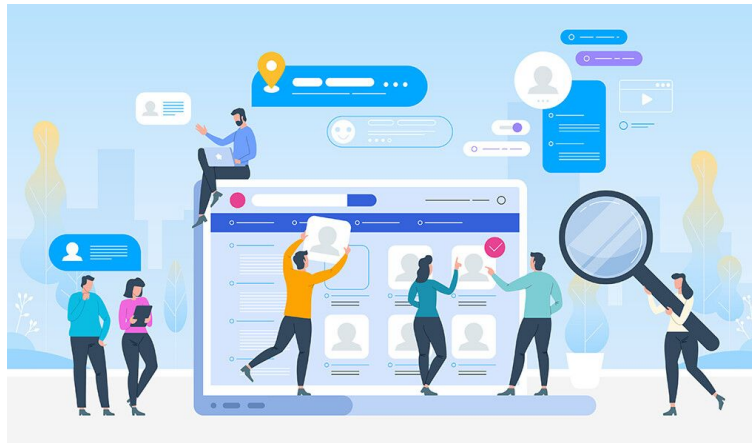
Data Externalities

Data externalities: the phenomenon that data of some consumers reveal information about others.

Q: What are some situations/environments where data externalities are present?

*please use the link in the zoom chat to enter your answers.

Environments with Data Externalities

- Social Media

Environments with Data Externalities

- Social Media
- Online Marketplaces

Environments with Data Externalities

- Social Media
- Online Marketplace
- Almost everywhere involving data sharing and where a central data processor can learn about one consumer from others' data.

Case Studies

Use the second section of the etherpad discussion space

- \+   for positive data externality
- \-   for negative data externality

16	Are data externalities positive or negative in the following environments?
17	
18	*Case 1:* + - + + - +
19	
20	*Case 2:*
21	
22	*Case 3:*
23	
24	*Case 4:*
25	

Examples of Data Externalities

Online Market

In an online marketplace like Amazon, are data externalities present? If so, are they positive or negative?

(Use etherpad link in zoom chat)

Frequently bought together

Total price: **$219.61**

[Add all three to Cart]

[Add all three to List]

ⓘ These items are shipped from and sold by different sellers. Show details

☑ **This item:** NEO Chair Office Chair Computer Headrest Desk Chair- Head Rest Business Ergonomic High Chair Cushion… $89.98

☑ OFM ESS Collection Chair Mat with Lip for Carpet, 36" x 48", Clear $29.64

☑ JOISCOPE Home Office Computer Desk, Study Writing Desk with Wooden Storage Shelf,2-Tier Industrial… $99.99

Examples of Data Externalities

Online Market

In an online market like Amazon, Amazon can learn about a single user from other users who share similar characteristics, which makes individual data less valuable. In this case, data externalities for the users are negative.
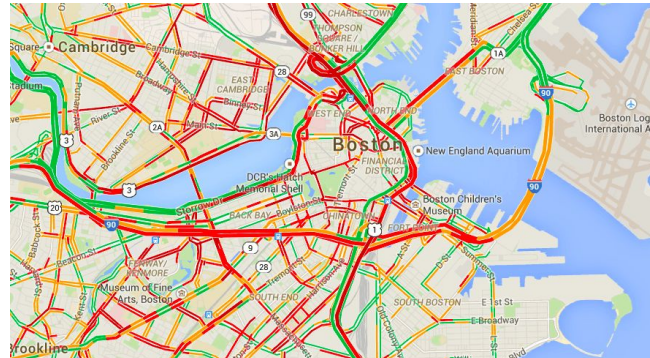
Examples of Data Externalities

Navigation Services

In a navigation service app like Google Maps, are data externalities present? If so, are they positive or negative?

(Use etherpad link in zoom chat)

Examples of Data Externalities

Navigation Services

If a driver reports the traffic situation, only in combination with other users' reports, this can be seen as a validated accident. Data from others makes data more useful, not more replaceable, and others' sharing of data strengthens the data contributor's bargaining position. In this situation, positive data externalities arise.

Examples of Data Externalities

Social Media Networks

In a social media platform like Facebook, are data externalities present? If so, are they positive or negative?

(Use etherpad link in zoom chat)

Examples of Data Externalities

Social Media Networks

When someone posts on Facebook about the purchase of a product, they convey information that Facebook can use to predict the probability that people in similar social groups will buy that product.

Consumers might obtain some benefit from targeted advertising, but Facebook reaps the most profit from it.

Examples of Data Externalities

Crowd-sourced Review Platform

In a crowd-sourced review platform like Yelp, are data externalities present? If so, are they positive or negative?

(Use etherpad link in zoom chat)

Examples of Data Externalities

Crowd-sourced Review Platform

In a platform like Yelp, one review can become valuable with other reviews validating it. Therefore, data externalities are positive in this environment.

Discussion

Q: How can we *identify* environments where data externalities are positive?

Discussion

Q: How can we *identify* environments where data externalities are positive?

- **A**: Complementary data vs. replaceable data.
- **A**: The amount of data already available to the data processor.

Discussion

Q: How can we *design* systems where data externalities are weakly positive?

Thank you!

Charles Cui:
charlescui@u.northwestern.edu
https://yccui.github.io/