gesis Leibniz Institute for the Social Sciences

A Vignette Experiment on Determinants of Health Data Sharing Behavior

Willingness to Donate Sensor Data, Medical Records, and Biomarkers

Henning Silber, Frederic Gerdon, Ruben Bach, Christoph Kern, Florian Keusch, & Frauke Kreuter

Member of Leibniz Association

Importance of Health Data Sharing

- Empirical data is crucial for evidence-based political decision making
- Dependency on people's willingness to share their data
- Technological developments made large-scale collection and analyses of health data easier
- Occurrence of new challenges regarding data privacy and ethics
- Decisions whether to share data are highly context dependent

Theory of Contextual Integrity (CI)

- *Contextual informational norms* define which data flows are appropriate
- Situational parameters that define a data flow include:
  - Data type
  - Involved actors
    - Data subject
    - Data sender
    - Data recipient
  - Transmission principles

Nissenbaum (2018)

Privacy Calculus

- Individuals consider potential benefits and risks in privacy-related decision making (Culnan & Armstrong 1999)

- Privacy is an economic good that can be traded for benefits, such as other goods or services (Smith et al. 2011, Kehr et al. 2015)

Motivation

Finding of a vignette experiment conducted by Gerdon et al. (2021) published in *Harvard Data Science Review*

- Comparison of data sharing regarding three different data types (health data, energy consumption, and geographical location)

- Surprisingly, sharing data with a public institution was less accepted than with a private organization, possibly due to fears of use in another context or fears of government surveillance

- This finding has worrisome consequences if it generalizes

Data Type

- Different data types are often associated with different levels of data sensitivity and different levels of data sharing effort (e.g., Silber et al. 2021)

- Sharing data passively via a smartphone app is little effort when the app is installed, while sharing biomarkers (e.g., blood donation) is likely to be perceived as more effortful

Data Type: Hypothesis

- Our vignette experiment will include:
  - ▶ Sensor data
  - ▶ Medial records
  - ▶ Biomarkers

- H1: People are more likely to share their sensor data and medical records than their biomarkers

Data Recipient

- Gerdon et al. (2021) found higher acceptability of sharing health data with a private organization than with a public institution

- Research shows different trust levels across public institutions (Krause et al. 2019), which we plan to investigate more in detail with respect to data sharing to put the previous finding in perspective

Data Recipient: Hypotheses

- Our vignette experiment will include:
  - Public health agency
  - University research center
  - Private company

- H2.1: The acceptability levels of data sharing are higher for university research centers than for the private companies and public health agencies
- H2.2: A higher level of individual trust in a respective institution results in higher willingness to share data

Purpose

- There can be a specific benefit for the data subject (e.g., personal health recommendation)

- It may appear straightforward that people are more likely to share data if they perceive a personal benefit than a public benefit

- However, data use for public interest is accepted in many contexts as well, and even oftentimes deemed as a prerequisite (Waind 2020)

Purpose: Hypotheses

- Our vignette experiment will include:
  - Personal benefit
  - Public benefit

- H3.1: People are more willing to share their data if they expect a personal benefit

- H3.2: People with higher scores on altruism are more likely to be willing to share their health data for a public purpose

- H3.3: People who perceive public duties as important are more likely to be willing to share their health data for a public purpose

- H3.4: People who trust in science are more likely to share data when it is for public purpose

Interaction Hypotheses

- H4.1: For private recipients, we expect that people are less likely to share data that are associated with specific health contexts (medical records and biomarkers) than sensor data

- H4.2: For public recipients, we expect that people are more likely to agree to share their biomarkers and medical records than their sensor data

- H4.3: People might be concerned if their data is used out of context. Thus, links between public recipient and public benefit and private recipient and personal benefit are most acceptable

Additional Explorative Measures

- Personal health condition

- Medical history

- Smartphone and smartwatch usage

- Technical affinity

- Social trust

- Political ideology

Research Design

- German online panel
- 750 respondents (based on a power analysis)
- Expected field time: May 2022
- 18 unique vignettes (3x3x2)
  - 3 data types
  - 3 data recipients
  - 2 purposes
- Topic: cancer research
- Every respondent receives 3 vignettes (one vignette per data type)

Vignettes: Sensors

Sensors installed on smartphones, smartwatches, or other wearable devices collect data that can be used to assess the health conditions of people. With the consent of a person, these data are transmitted to a German public health agency [private company; university research center]. This public health agency [private company; university research center] uses these data for a research program to fight cancer. [This public health agency [private company, university research center] uses these data to provide people with personal recommendations on their health behavior with respect to protection against cancer.] The public health agency [private company; university research center] guarantees that the data are safe, anonymous, and protected from misuse.

"How likely or unlikely would you agree to share your health data for this purpose?" (1 very unlikely to (7) very likely)

Vignettes: Medical History

Health records obtained from doctors' offices can be used to assess the health conditions of people. With the consent of a person, these data are transmitted to a German public health agency [private company; university research center]. This public health agency [private company; university research center] uses these data for a research program to fight cancer. [This public health agency [private company; university research center] uses these data to provide people with personal recommendations on their health behavior with respect to protection against cancer.] The public health agency [private company; university research center] guarantees that the data are safe, anonymous, and protected from misuse.

Vignettes: Biomarkers

Blood samples that are collected for biobanks can be used to assess the health conditions of people. With the consent of a person, these data are transferred to a German public health agency [private company; university research center]. This public health agency [private company; university research center] uses these data for a research program to fight cancer. [This public health agency [private company, university research center] uses these data to provide people with personal recommendations on their health behavior with respect to protection against cancer.] The public health agency [private company; university research center] guarantees that the data are safe, anonymous, and protected from misuse.

Data Analyses

- Multi-level model to account for the hierarchical data structure (vignettes nested in respondents)

  ▸ Continuous dependent variable and random-effect model

  ▸ Hypotheses (main and interaction effects) will be tested two-sided

  ▸ Two robustness checks

    - Fixed-effect model

    - Ordinal dependent variable

Next steps

- Pre-registered report currently under peer-review
  - ▶ Theory and hypotheses
  - ▶ Study design
  - ▶ Analyses code in R
- Study implementation
- Adding results and discussion to the manuscript

Spring meeting, DGS methods section, March 25-26, 2022

Henning Silber, Frederic Gerdon,
Ruben Bach, Christoph Kern,
Florian Keusch, & Frauke Kreuter

A Vignette Experiment on Determinants of Health Data Sharing Behavior: Willingness to Donate Sensor Data, Medical Records, and Biomarkers

Contact: henning.silber@gesis.org

References

Culnan, M. J., & Armstrong, P. K. (1999). Information privacy concerns, procedural fairness, and impersonal trust: An empirical investigation. *Organization science*, 10(1), 104-115.

Gerdon, F., Nissenbaum, H., Bach, R. L., Kreuter, F., & Zins, S. (2021). Individual acceptance of using health data for private and public benefit: Changes during the COVID-19 pandemic. Special Issue COVID-19: Unprecedented Challenges and Chances. H*arvard Data Science Review*, advanced online access.

Kehr, F., Kowatsch, T., Wentzel, D., & Fleisch, E. (2015). Blissfully ignorant: the effects of general privacy concerns, general institutional trust, and affect in the privacy calculus. *Information Systems Journal*, 25(6), 607-635.

Krause, N. M., Brossard, D., Scheufele, D. A., Xenos, M. A., & Franke, K. (2019). Trends—Americans' trust in science and scientists. *Public Opinion Quarterly*, 83(4), 817-836.

Nissenbaum, H. (2018). Respecting context to protect privacy: Why meaning matters. *Science and Engineering Ethics*, 24(3), 831–852.

Silber, H., Breuer, J., Beuthner, C., Gummer, T., Keusch, F., Siegers, P., Stier, S., and Weiß, B. (2021), "Linking surveys and digital trace data: Insights from two studies on determinants of data sharing behavior", *SocArXiv*, preprint.

Smith, H. J., Dinev, T., & Xu, H. (2011). Information privacy research: an interdisciplinary review. *MIS quarterly*, 35(4), 989-1015.

Waind, E. (2020). Trust, security and public interest: striking the balance. A narrative review of previous literature on public attitudes towards the sharing, linking and use of administrative data for research. *International Journal of Population Data Science*, 5(3), 1368.