



Leonie Reichert (HU Berlin), Björn Scheuermann (TU Darmstadt)

# An Analysis of Requirements and Privacy Threats in Mobile Data Donations

# Intro





# Check out some apps built for research.

[View more apps in App Store](#) ↗



## Apple Research

Delivering studies at scale.  
Apple



## VascTrac

A research study of peripheral artery disease.  
Stanford University



## MyHeart Counts

Improves our understanding of heart health.  
Stanford University



## MyGeneRank

Exploring how genetic factors influence disease.  
The Scripps Research Institute



## EpiWatch

A research study for people with epilepsy.  
Johns Hopkins Digital



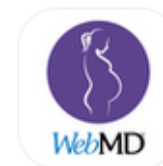
## Corrie Health

An app designed to help patients recover from a heart attack.  
Johns Hopkins Digital



## MS Mosaic

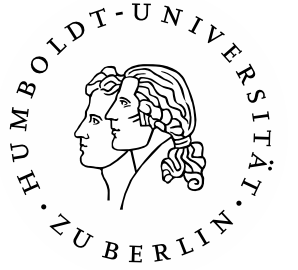
Seeks to understand Multiple Sclerosis from daily experiences.



## WebMD Pregnancy

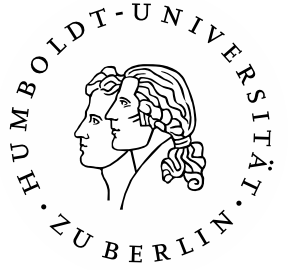
Track a baby's growth and development, week by week.  
WebMD

Apple Researchkit Apps



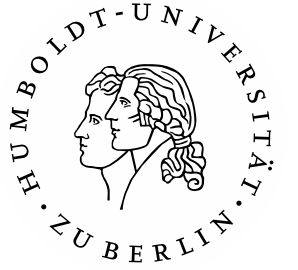
# Focus of this Talk

- Analysis of risks when **crowdsourcing** of health data
  - Focus on research context
  - Is privacy ensured? What type of data is at risk?
- Outline
  - What data do apps collect?
  - Model data collection process
  - Analyse risks



# 1

## Analysis of Existing Apps



## Provided Functionalities

- Pubmed Search:
  - Clinical trials for mobile health
  - Using smartphones and apps
- Top 100 entries retrieved
- 74 papers accessible and relevant
- Provided functionalities were sorted into categories
  - Multiple categories possible



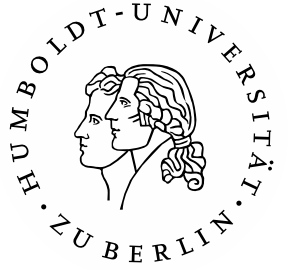
## Provided Functionalities

- Informing and educating 45
- Reminders and notifications 37

- **Communication** with professionals 19
- **Communication** between donors 6

- Self-tracking 24
- Questionnaires 23
- App interaction 24
- Feedback 29

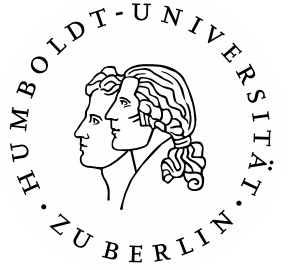
- External sensors 10
- Wearables 9
- Internal sensors 2
- Digital health/tracking 4



# 2

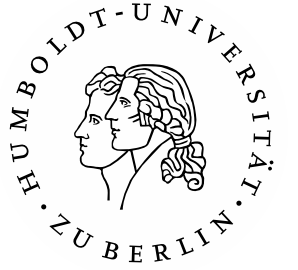
## Privacy Threat Modeling





## Linddun Framework

- Linkability (L)
- Identifiability (I)
- Non-repudiation (N)
- Detectability (D)
- Disclosure of information (D)
- Unawareness (U)
- Non-compliance (N)



## Modeling Steps

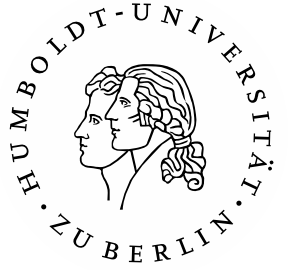
- Create data flow diagram
  - Actors
  - Data stores
  - Processes
  - Data flows
- Identify risks with threat catalog

Researchers

Data Donors

App Store

Professionals



# Modelling App Distribution

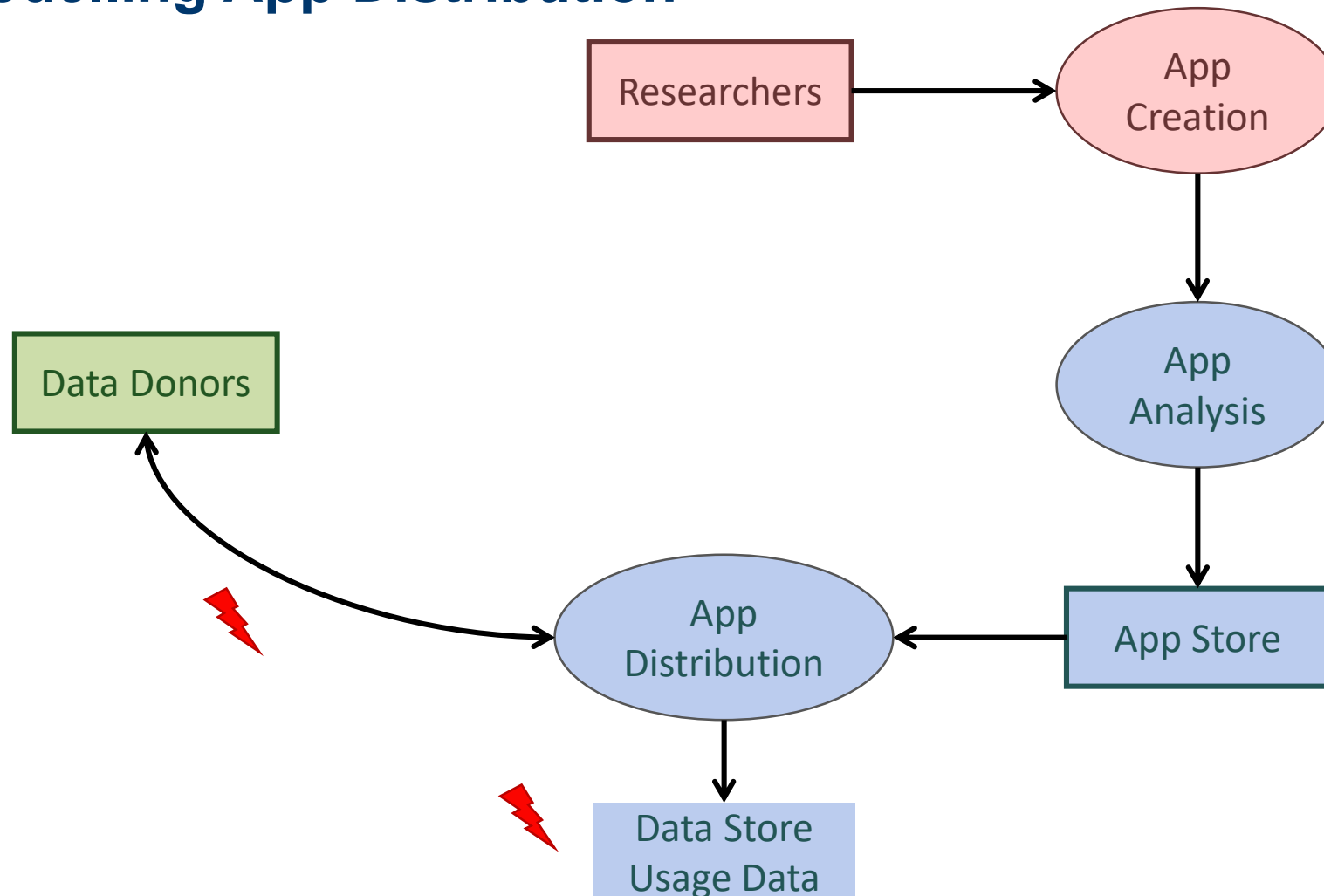
Data Donors

Researchers

App Store

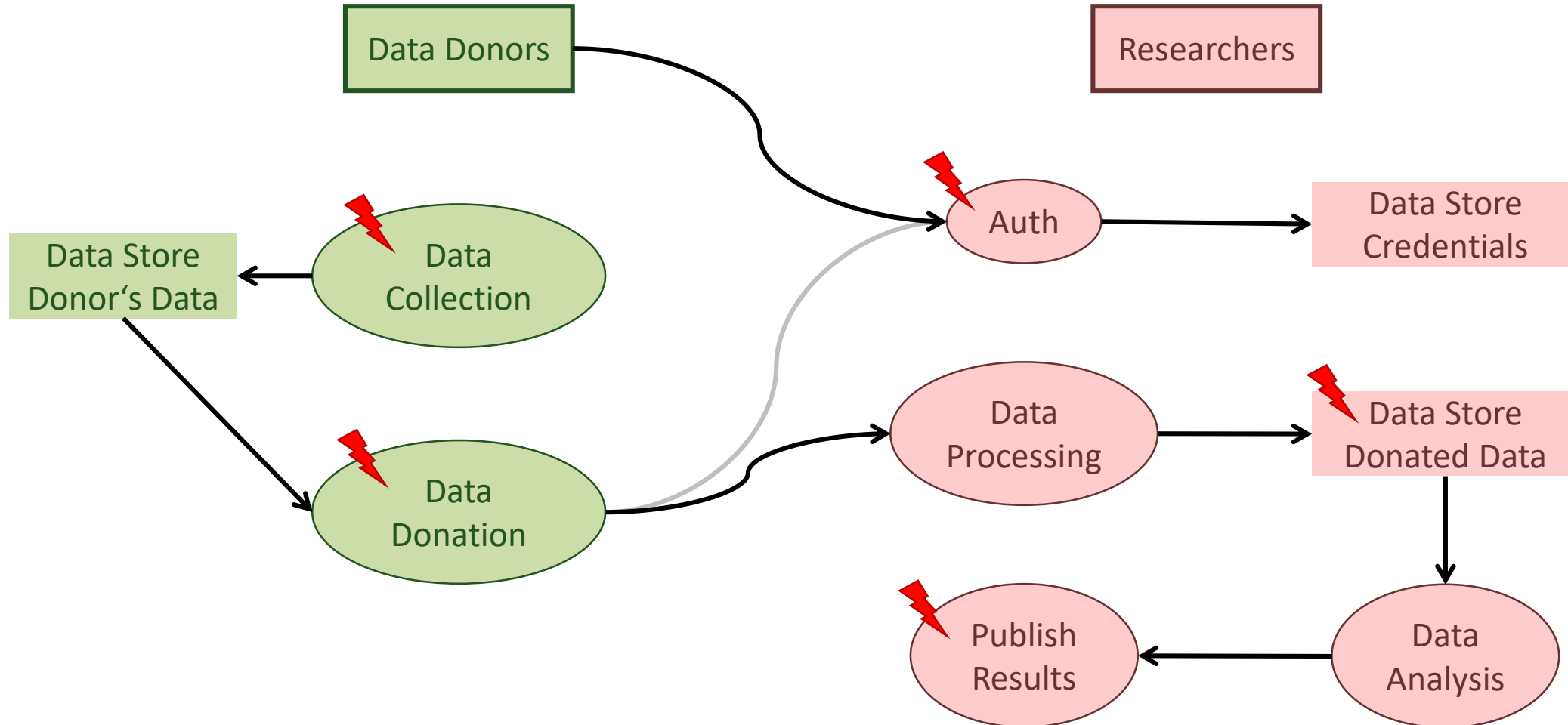


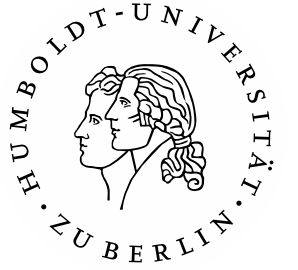
# Modelling App Distribution





# Modelling Data Collection





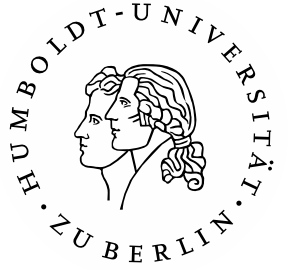
## More Notable Attack Vectors

- Leak of communication data
  - Even if encrypted and self-hosted
  - Metadata, communication patterns, message types...
- Threats to Functionality:
  - Fake researchers
  - Manipulation of studies
- ...



## Conclusion

- Many apps exist for crowdsourcing medical data
- Used **infrastructures leak various types of data**
- Solutions to some threats in the paper
- Suitable privacy preserving platforms for **wide-spread** use needed



# Questions?

Contact: [reicleon@hu-berlin.de](mailto:reicleon@hu-berlin.de)





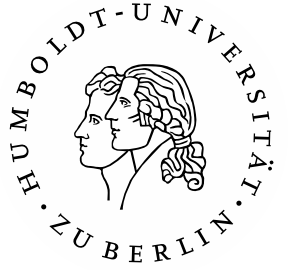
## References

- Dariusz Sankowski. Image. <https://pixabay.com/de/photos/iphone-6s-plus-ich-schaue-apfel-1032784>. Accessed: 2023-06-27.
- Apple Inc. What you can do with ResearchKit. <https://www.researchandcare.org/researchkit>. Accessed: 2023-06-27.
- National Library of Medicine. Pubmed. <https://pubmed.ncbi.nlm.nih.gov>, Accessed: 2023-06-27.
- DistriNet Research Group. LINDDUN. <https://linddun.org>, Accessed: 2023-06-27.



## References

- The New York Times. Cambridge Analytica and Facebook: The scandal and the fallout so far. <https://www.nytimes.com/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html>. Accessed:2023-06-30.
- DENG, Mina, et al. A privacy threat analysis framework: supporting the elicitation and fulfillment of privacy requirements. *Requirements Engineering*, 2011, 16. Jg., Nr. 1, S. 3-32.



## Modelling Communication

- Communication types
  - Donors to professionals
  - Between donors: Directly or in groups
- Communication infrastructure
  - Self-hosted or external third party
  - Data is encrypted but meta-data leaks



# Communication Donor to Donor

