

Perspectives of Stakeholders in Data Governance



Julia Bernd
International Computer Science Institute



Science of Security and Privacy
Quarterly Lablet PI Meeting
November 16, 2021

Overview: Studying Data Stakeholders

The same data is handled by stakeholders with...

- Different goals for data use
- Different technological capabilities and resources
- Different approaches to data management and governance



Technology Developers

Software
companies

Individual
developers

...

Data Stakeholders
in Focus



Technology & Data Platforms

Commercial mobile
platforms & API hosts
Gov't platforms

...

Technology Developers

Software
companies
Individual
developers

...

Data Stakeholders in Focus

Technology & Data Platforms

Commercial mobile
platforms & API hosts
Gov't platforms
...

Data Users

Ad networks
Businesses &
professional users
...

Technology Developers

Software
companies
Individual
developers
...

Data Stakeholders in Focus

Technology & Data Platforms

Commercial mobile
platforms & API hosts
Gov't platforms
...

Data Users

Ad networks
Businesses &
professional users
...

Technology Developers

Software
companies
Individual
developers
...

Tech/Data Regulators

Lawmakers
Regulatory
Agencies
...

Data Stakeholders in Focus

Technology & Data Platforms

Commercial mobile
platforms & API hosts
Gov't platforms
...

Data Users

Ad networks
Businesses &
professional users
...

Technology Developers

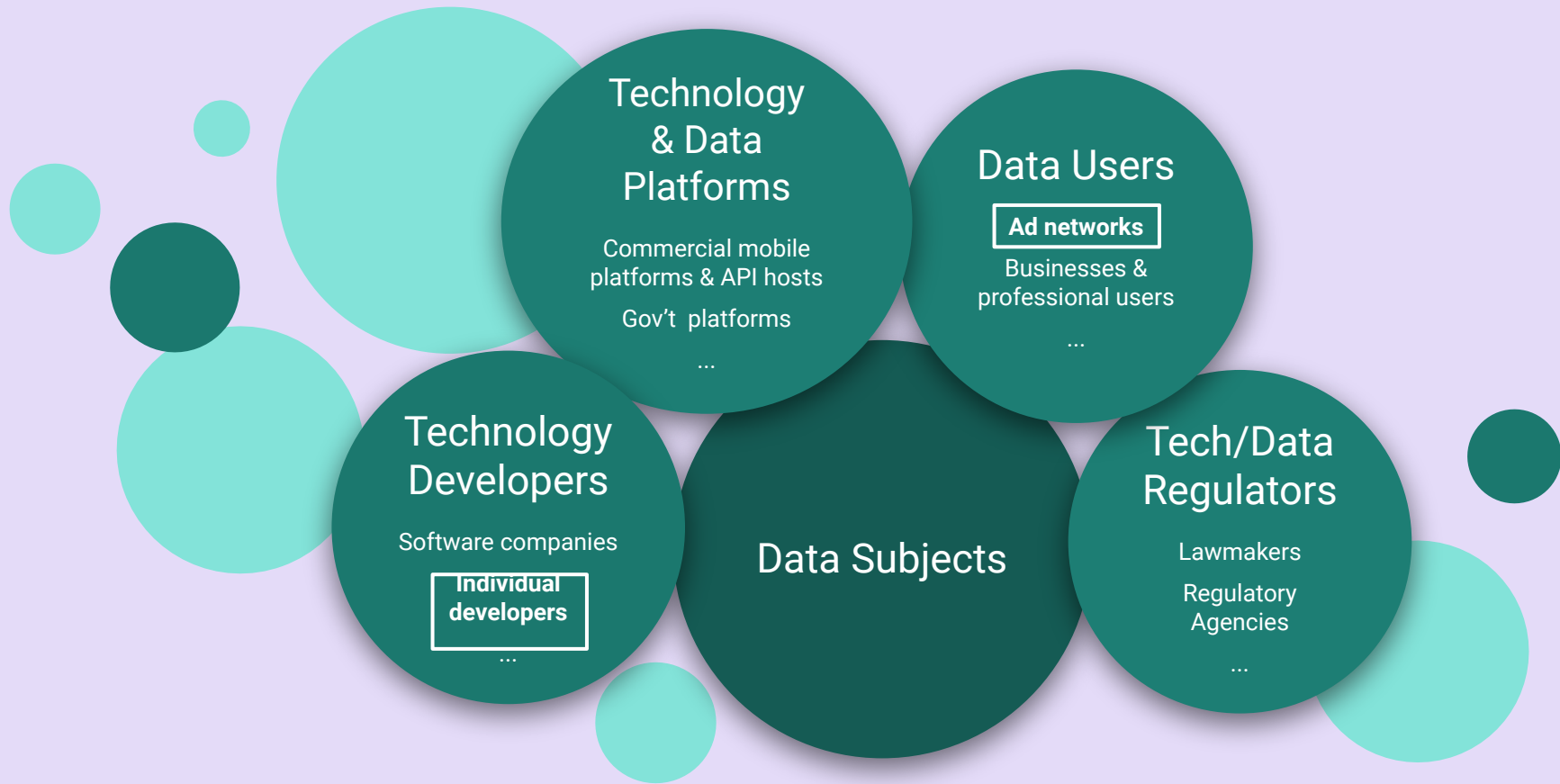
Software
companies
Individual
developers
...

Data Subjects

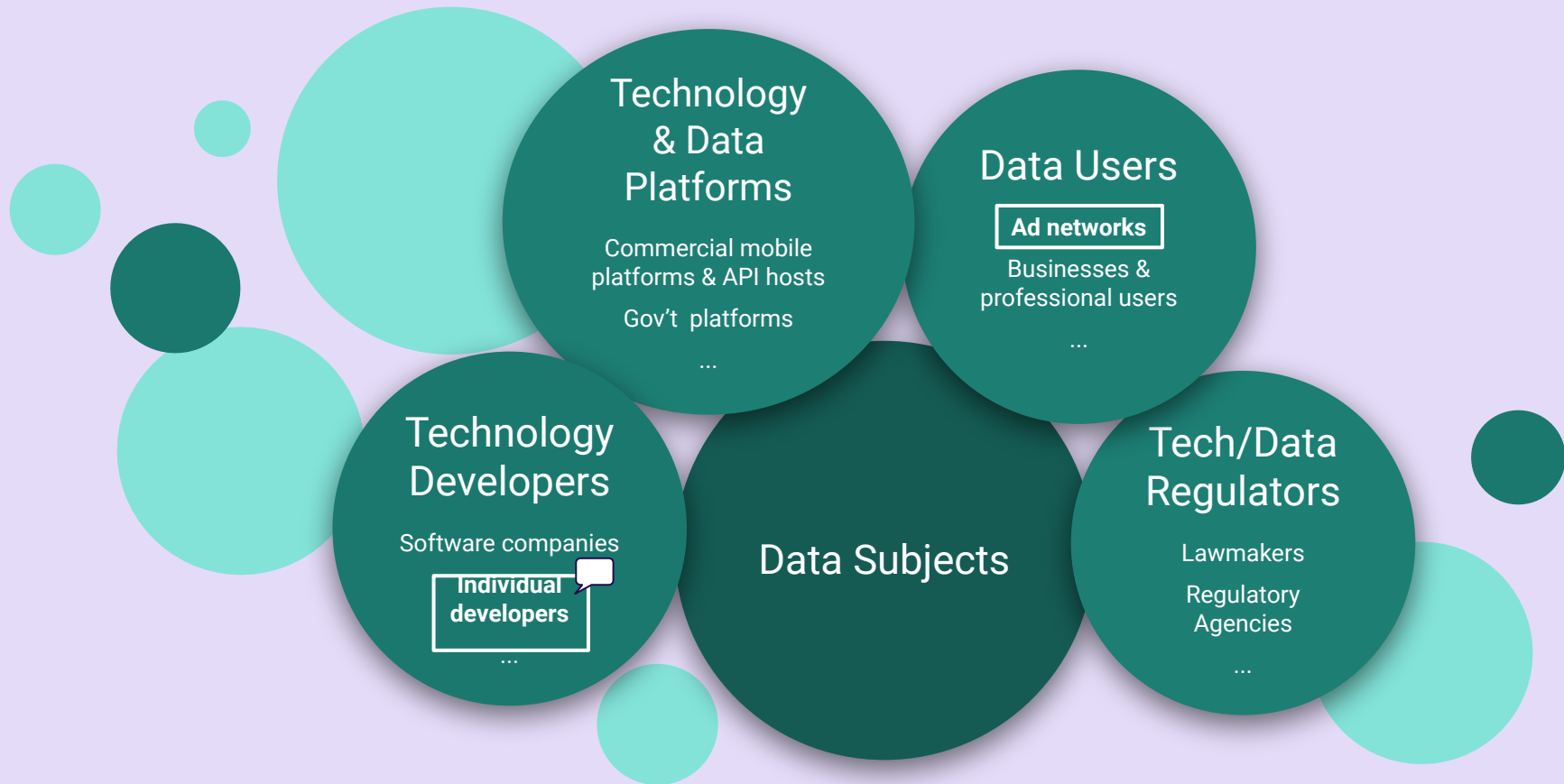
Tech/Data Regulators

Lawmakers
Regulatory
Agencies
...

Data Stakeholders in Focus



Stakeholders: App Developers and Ad Networks



Stakeholders: App Developers and Ad Networks

Deciding on Personalized Ads: Nudging Developers About User Privacy

Mohammad Tahaei, Alisa Frik, and Kami Vaniea.
Symposium on Usable Privacy and Security
(SOUPS '21), August 2021.



THE UNIVERSITY OF EDINBURGH
informatics

Berkeley
UNIVERSITY OF CALIFORNIA



Microsoft
Research



Deciding on Personalized Ads: Nudging Developers About User Privacy

Mohammad Tahaei, Alisa Frik, and Kami Vaniea
Symposium on Usable Privacy and Security (SOUPS '21), August 2021.



Why Advertising Networks?

- About 77% of free Android apps contain an ad library
- Massive data collection, including sensitive data like location data
- Some users find personalized ads discomforting, discriminatory, and intrusive
- Developers may say they're concerned about user privacy, but still pick options that aren't privacy-friendly

“Options” Given to Developers

Restricted data processing

Google AdMob CCPA

You can choose from two options for users that Google determines are in California. If you want to continue to show personalised ads, tell us the partners that you want to monetise your ads with below. By default, data processing isn't restricted and personalised ads will continue to show.



Don't restrict data processing

Google continues to show personalised ads to eligible users in California. Personalised ads are based on a user's past behaviour, such as previous visits to sites or apps or where the user has been.

Restrict data processing

Google restricts how it uses certain unique identifiers and other data. Google only shows non-personalised ads from Google demand to eligible users in California. Non-personalised ads are based on contextual information, such as the content of your site or app.

Select the type of ads that you want to show

Google AdMob GDPR

You can choose from two ad serving options. If you don't make any changes, personalised ads will continue to show for EEA and UK users. Your selection will not affect mediation.



Personalised ads

Google can show personalised ads to your users in the EEA and the UK.

Non-personalised ads

Google will show only non-personalised ads to your users in the EEA and the UK.

Research Questions

How does choice framing in ad networks impact developers' decisions about ad personalization?

What are the reasons behind developers' choices of personalized or non-personalized ads?

Research Questions

How does choice framing in ad networks impact developers' decisions about ad personalization?

What are the reasons behind developers' choices of personalized or non-personalized ads?

Impact of Wording and Choice Framing (Nudges) on Developers' Decisions

- An online between-subject experiment with 400 participants with mobile app development experience
- Six conditions, varying words and trade-offs between privacy, revenue, and user facing descriptions
- Other questions around which operating system, format of the ad, quality of graphics, and what regulations apply (not covered here)
- Analysis: quantitative & qualitative

Impact of (Nudges)

- An online between development and user facing design
- Six conditions, user facing design
- Other question graphics, and v
- Analysis: quan

Deciding on Personalized Ads: Mohammad Tahaei, Alisa Frik, K

Imagine that you are a shareholder in a software development company. Together with a small team, you created an app. The app will be published in Europe and the United States and is mainly targeted towards adults (above age of 18).

To monetise the app, you have decided to use the "Acme" ad network to show ads to your users.

The Acme ad network offers a step-by-step Assistant – a graphical user interface that provides various configuration choices for integrating ads into your app.

The Assistant asks the developer several questions and then provides ad network configuration code based on the answers that can be imported directly into an app with minimal additional coding required.

The following are the 5 questions asked by Acme's Assistant, please answer them as if you were developing the app.

Which ad formats are you integrating?

- Interstitial: full-page ads appear at natural breaks & transitions, such as level completion. Supports video content.
- Rewarded Video: ads reward users for watching short videos and interacting with playable ads and surveys. Good for monetising free-to-play users. Supports video content.
- Banner: A basic ad format that appears at the top & bottom of the device screen.
- Native: customisable ad format that matches the look & feel of your app. Ads appear inline with app content. Supports video content.

What level of graphics do you want for your ads?

- Ads with moderate to low graphics quality. These ads will work on most phones.
- Ads with highest graphics quality. These ads will work best on newer phones with the latest operating systems.

Which platform are you integrating Acme ad network on?

- iOS
- Unity
- Windows Phone
- Android

Select the type of ads that you want to show.

- Non-personalised ads: Acme will show only non-personalised ads to your users.
- Personalised ads: Acme can show personalised ads to your users.

Which of the following regulations apply to this app?

- COPPA (Children's Online Privacy Protection Act)
- CCPA (California Consumer Privacy Act)
- GDPR (General Data Protection Regulation)
- HIPAA (Health Insurance Portability and Accountability Act)
- I don't know
- None of the above

Framing ons

nts with mobile app
acy, revenue, and
f the ad, quality of

Impact of Wording and Choice Framing (Nudges) on Developers' Decisions

Select the type of ads that you want to show.

- Non-personalised ads: Acme will show only non-personalised ads to your users.
- Personalised ads: Acme can show personalised ads to your users.

Six Conditions

Control: Minimal Information

- Personalized ads
- Non-personalized ads

Data Processing Restrictions

- Ads with unrestricted data processing
- Ads with restricted data processing

User-Facing Descriptions

- “Personalized ads” tag will be displayed to users
- “Non-personalized ads” tag will be displayed to users

Privacy Focused

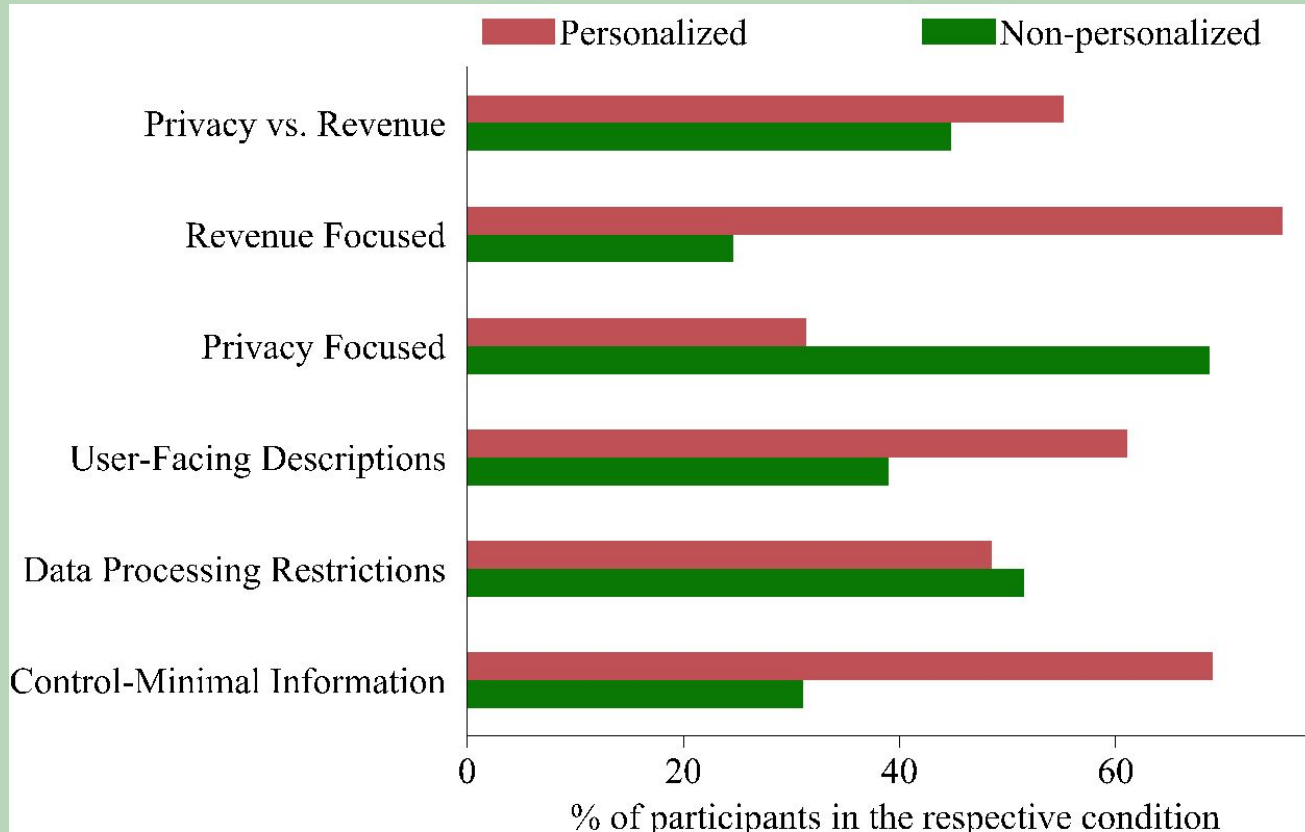
- Ads with lower user privacy
- Ads with higher user privacy

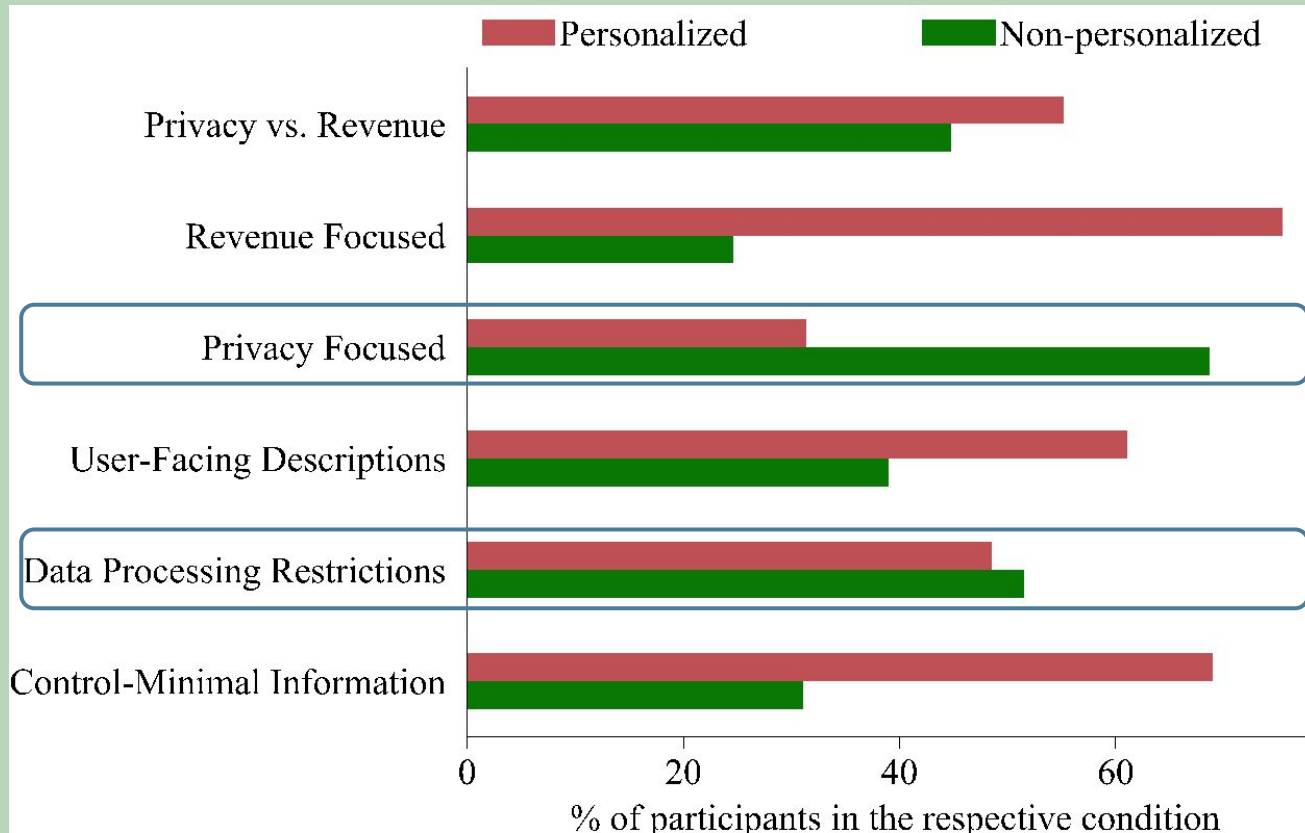
Revenue Focused

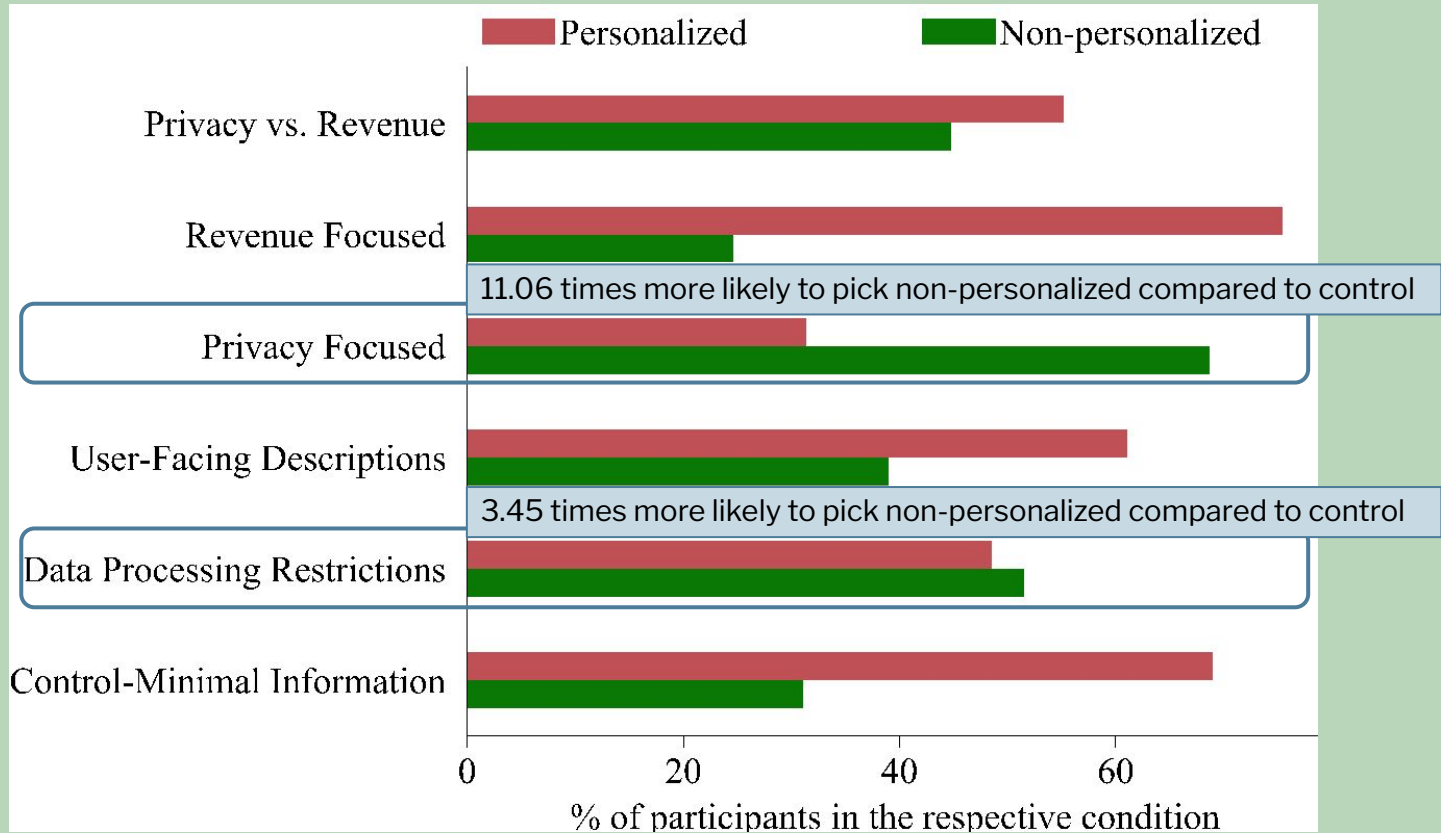
- Ads with higher revenue
- Ads with lower revenue

Privacy vs. Revenue

- Ads with higher revenue
- Ads with higher user privacy







Reasons for Picking Personalized vs. Non-Personalized Ads

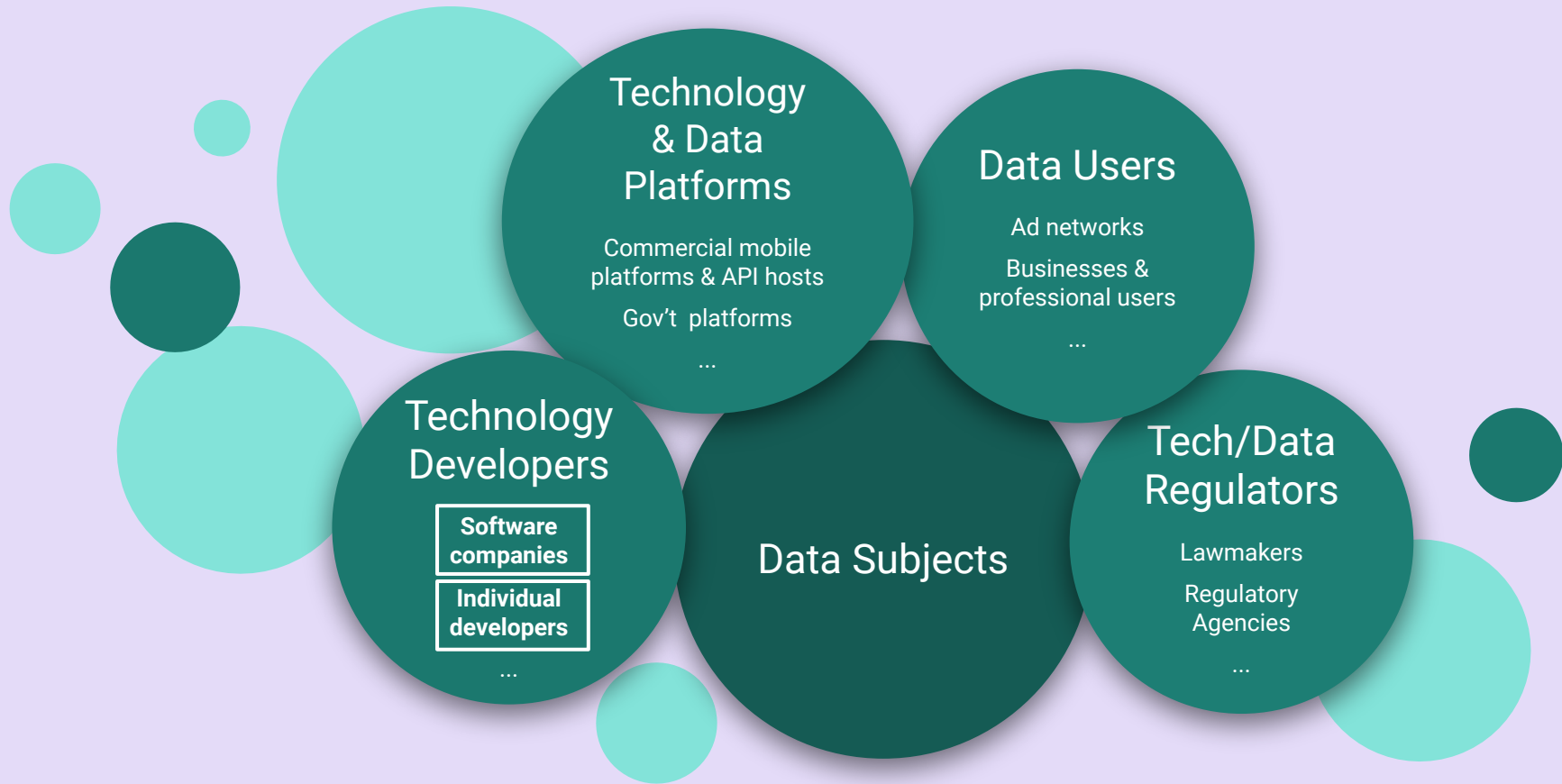
- Impact on revenue (42%)
- User privacy (40%)
 - *Users don't care vs. protecting users' sensitive data, gaining their trust, complying with privacy regulations, gaining competitive advantage*
- Relevance to users (39%)
 - *Relevant ads are more useful vs. relevant ads are more distracting*
- User experience (15%)
 - *Personalized ads are less annoying, more enjoyable, and of higher quality vs. non-personalized ads are less invasive and creepy*
- Other reasons (up to 2%)

Perceived Control Over Ad Network Choices

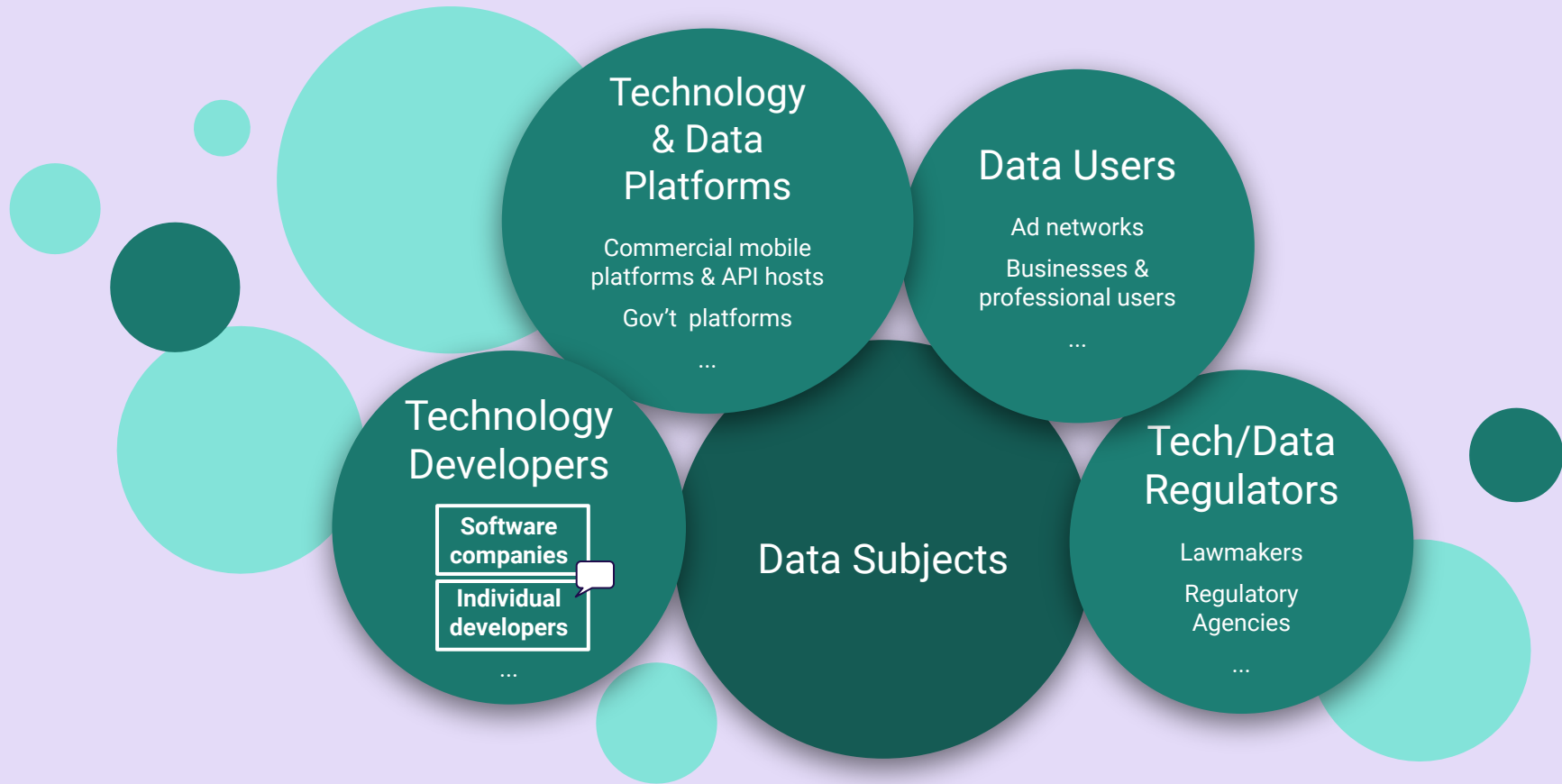
- Choices about ad networks' and apps' business models are often made by upper level and middle management (~32-40%)
- Yet, developers are involved in:
 - *Choosing ad networks (36%)*
 - *Configuring ads (47%)*
 - *Integrating ads' code (47%)*
- Participants think...
 - *Developers have moderate (40%) or very little (33%) control over data collection by ad networks*
 - *Users have very little (55%) or no control (12%) over it*

Final Thoughts

- Framing of “options” given to developers has an impact on developers’ decisions, and consequently their users’ privacy
- Minimizing the use of dark patterns directed toward developers
- Transparent and honest interfaces are needed



Stakeholders: Individual Developers in Product Teams



Stakeholders: Individual Developers in Product Teams

Privacy Champions in Software Teams: Understanding Their Motivations, Strategies, and Challenges

Mohammad Tahaei, Alisa Frik, and Kami Vaniea.

ACM CHI Conference on Human Factors in
Computing Systems (CHI '21),
May 2021.



THE UNIVERSITY of EDINBURGH
informatics

Berkeley
UNIVERSITY OF CALIFORNIA



Microsoft
Research

Privacy Champions in Software Teams: Understanding Their Motivations, Strategies, and Challenges

Mohammad Tahaei, Alisa Frik, Kami Vaniea

ACM CHI Conference on Human Factors in Computing Systems (CHI '21), May 2021.



THE UNIVERSITY of EDINBURGH
informatics

Berkeley
UNIVERSITY OF CALIFORNIA



Microsoft
Research

Motivation

Developers don't often consider privacy and find it difficult to implement

Organisational culture is a big factor in whether developers consider privacy

So how do we move organisations towards privacy-friendly practices?

Motivations: Champions



“Champions” advocate for a cause (e.g., an innovation or idea), encourage others to engage, and aid with overcoming barriers that a new idea could face

Literature in software engineering shows champions’ value in promoting both security and new software technologies

And what about privacy?

Privacy Champions



Formally or informally promote best practices for users' privacy, educate others, persuade, and advocate for privacy adoption throughout the software development process

Have an official or unofficial role within their team acting as the “voice” of users' privacy for the product or team, for example by giving privacy-related advice that can influence decisions and privacy practices

Research Questions

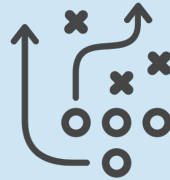


What do privacy champions find motivating, rewarding, challenging and frustrating in promoting user privacy in their organisations?

Research Questions



What do privacy champions find motivating, rewarding, challenging and frustrating in promoting user privacy in their organisations?

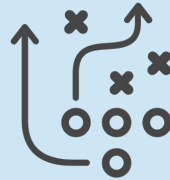


What strategies and channels do they find least and most effective in achieving that goal?

Research Questions



What do privacy champions find motivating, rewarding, challenging and frustrating in promoting user privacy in their organisations?



What strategies and channels do they find least and most effective in achieving that goal?



What resources do they use to keep up with the latest in privacy?

Method

Screening survey, asked others to nominate someone who they consider to be a privacy champion

12 Privacy Champions

Interview

Analysis: two coders, independently

Findings

Motivations of Privacy Champions

Personal values: Human rights, ethical values, social benefits

“The Snowden revelations came out and I felt extremely strongly that what he did was heroic and that I should figure out a way to support that kind of effort even if it’s in a very small or less risky way.” (P10)

Organizational motivations: Competitive advantage, existential value

“If we are perceived as an organization that doesn’t care about user privacy, then that will harm us. If we are perceived as an organization that does care, that will benefit us.” (P1)

Experiential motivators: Educational and hands-on experiences

Common Barriers for Implementing Privacy in Software Design

Negative privacy culture and attitudes (e.g., “I’ve got nothing to hide”)

Tensions between privacy and (other) business priorities

Lack of standardisation, evaluation metrics, and automated privacy tools

Technical complexity

Strategies for Promoting Privacy

Effective Strategies

Regular privacy-focused meetings and informal discussions

Management support, facilitation of communication among stakeholders (e.g., between legal and product teams)

Appropriate privacy documentation and guidelines

Incorporating privacy considerations into design and code reviews

Hands-on, practical training and mentoring

Sharing tools and libraries

Leveraging external influence: political and regulatory support, academic work, public critics

Strategies for Promoting Privacy

Not Effective Strategies

Punishing developers for not implementing privacy features

Company-wide awareness programs or on-boarding privacy training for new hires

How to Support and Attract Privacy Champions

Embed privacy values in the organisational culture

Acknowledge privacy-oriented efforts (by both colleagues and management)

Provide resources (compensate extra time, or let them spend 10-20% of their time on privacy work)

Offer hands-on privacy-oriented projects (e.g., privacy hackathons)




Include privacy topics in university degree and online learning curricula

Privacy Champions in Software Teams: Understanding Their Motivations, Strategies, and Challenges

Mohammad Tahaei, Alisa Frik, Kami Vaniea

mohammad.tahaei@ed.ac.uk

Read more about privacy champions' efforts and knowledge about privacy practices in software teams in the paper:

-  How they conceptualisations privacy, such as data protection, transparency, and trust
-  About their motivations, like empathy towards users and educational background
-  Discussion points about how to best support privacy champions & embed privacy in software teams



Case Study: Stakeholders in Health App Privacy



Perspectives of...

- App developers
- Healthcare professionals
- End users

Case Study: Stakeholders in Health App Privacy



Perspectives of...

- App developers
- Healthcare professionals
- End users

Comparing to...

- App behavior
- Data aggregation
- Laws and regulations
- Privacy policies

Why Health Apps?

- Thousands of apps with health-related purposes

Why Health Apps?

- Thousands of apps with health-related purposes
- Sensitivity of health data

Why Health Apps?

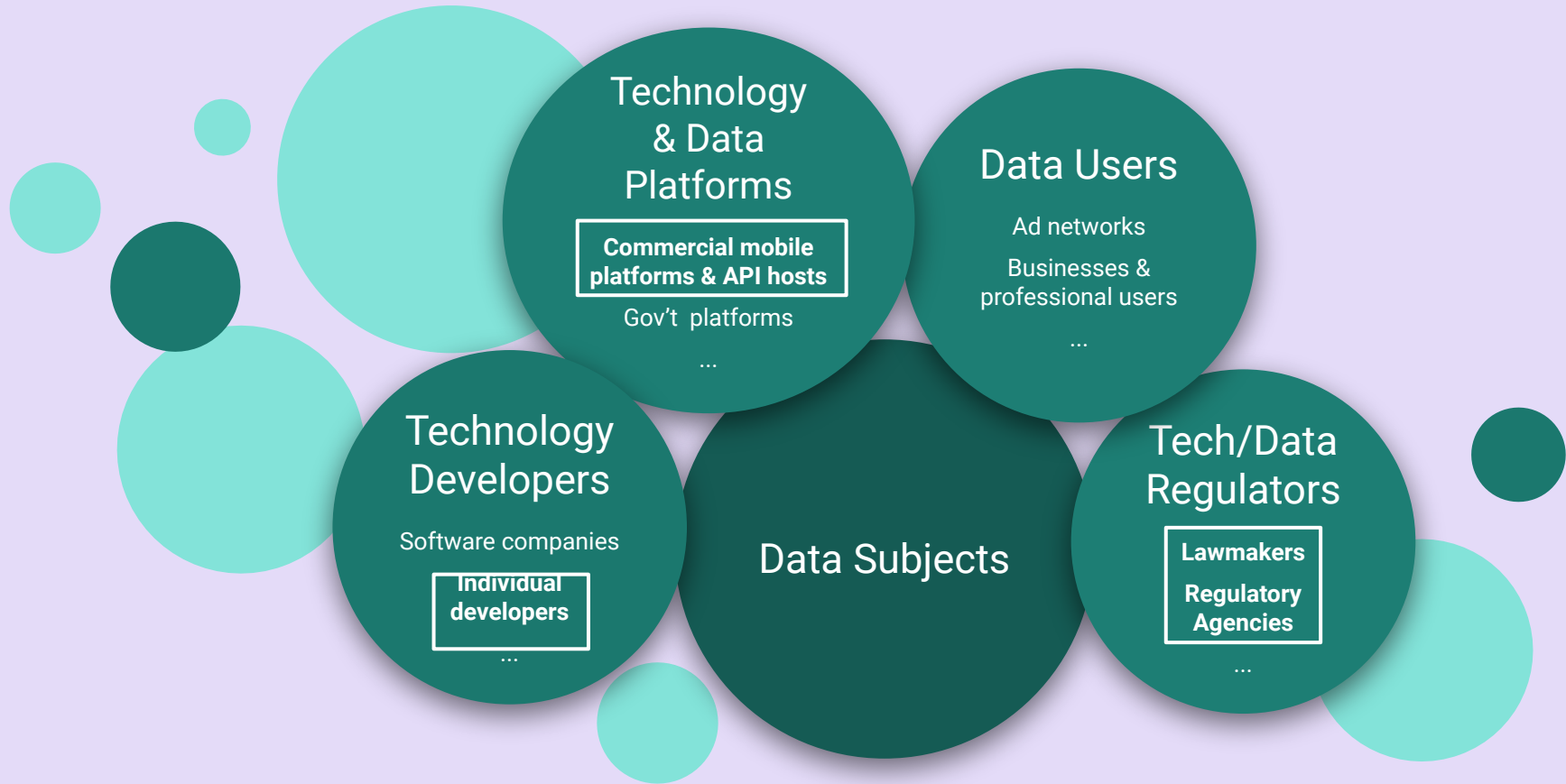
- Thousands of apps with health-related purposes
- Sensitivity of health data
- Consumers make assumptions about legal and technical protections

Why Health Apps?

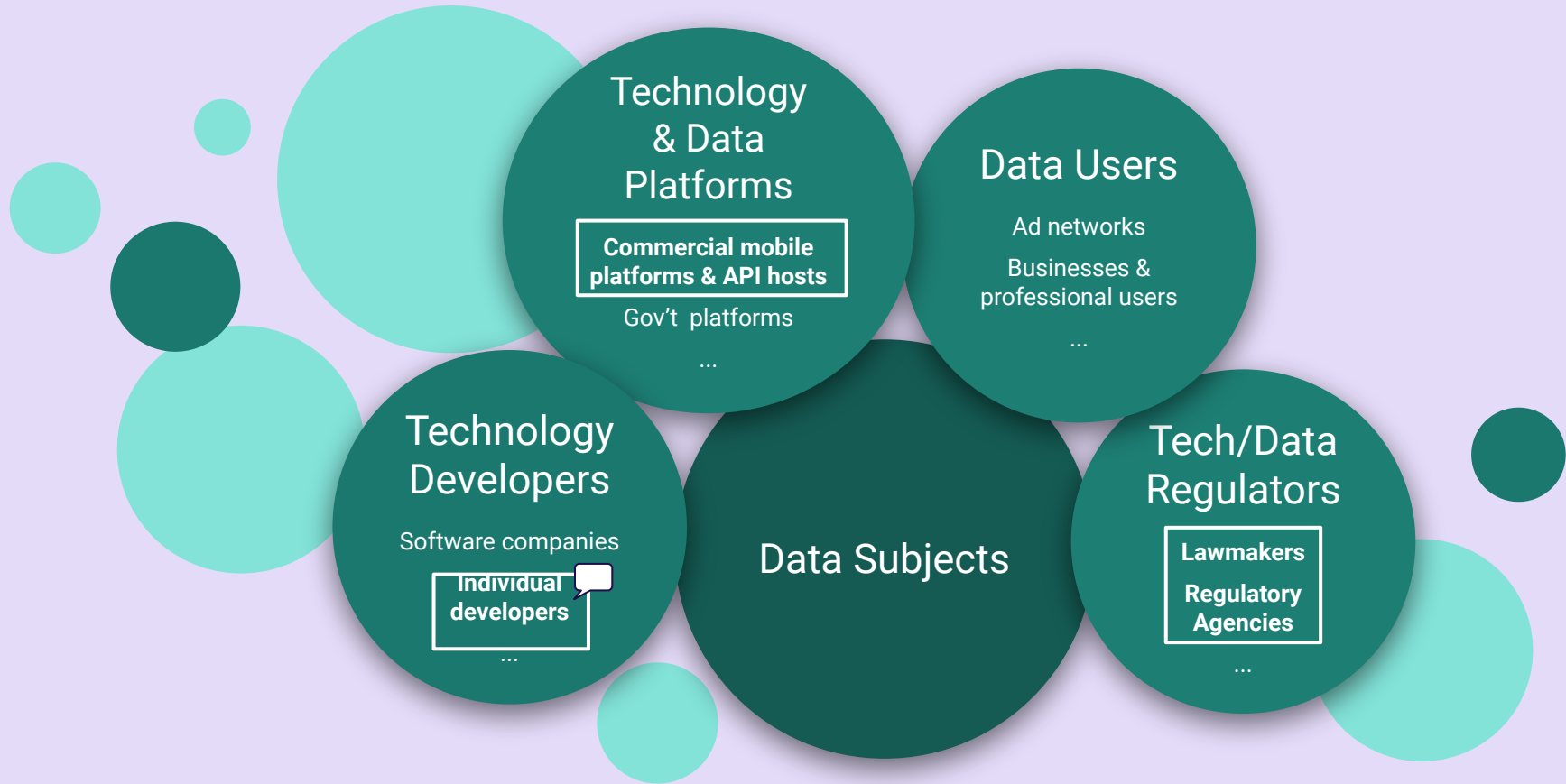
- Thousands of apps with health-related purposes
- Sensitivity of health data
- Consumers make assumptions about legal and technical protections
- ...that may not be true of every kind of health app!

Why Health Apps?

- Thousands of apps with health-related purposes
- Sensitivity of health data
- Consumers make assumptions about legal and technical protections
- ...that may not be true of every kind of health app!
- Developers may not be aware of special privacy requirements, nor of users' concerns



Stakeholders: App Developers, Platforms, APIs, and Law



Stakeholders: App Developers, Platforms, APIs, and Law

Who Puts the Privacy in the Health App? Developers and What Moves Them

Work in progress!

- ICSI - University of Bristol



Research Questions (General)

High-level research questions:

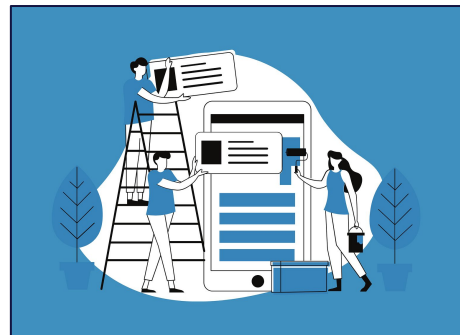
- How do developers of health apps approach questions about data management?
 - What motivates privacy-relevant choices they make about data handling?



Research Questions (General)

High-level research questions:

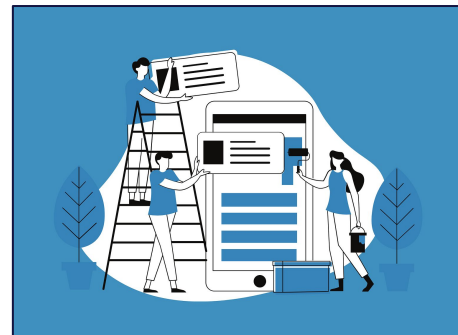
- How do developers of health apps approach questions about data management?
 - What motivates privacy-relevant choices they make about data handling?
- What challenges do health app developers face w.r.t. data management, and how do they navigate them?



Research Questions (General)

High-level research questions:

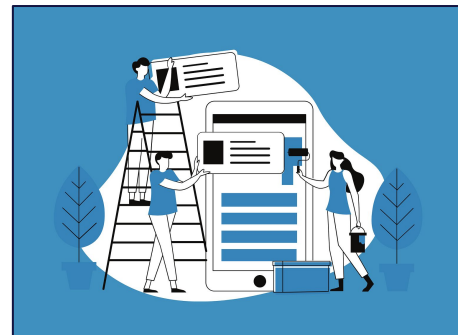
- How do developers of health apps approach questions about data management?
 - What motivates privacy-relevant choices they make about data handling?
- What challenges do health app developers face w.r.t. data management, and how do they navigate them?
- How are enhanced privacy requirements for apps that access health data structured, documented, and enforced?



Research Questions (General)

High-level research questions:

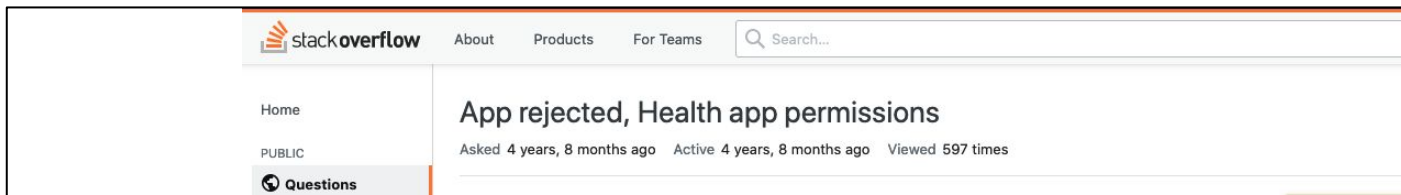
- How do developers of health apps approach questions about data management?
 - What motivates privacy-relevant choices they make about data handling?
- What challenges do health app developers face w.r.t. data management, and how do they navigate them?
- How are enhanced privacy requirements for apps that access health data structured, documented, and enforced?
- What assumptions do developers make about other stakeholders' privacy postures?



Study #1: Questions on StackOverflow

Analyzing StackOverflow posts about health apps

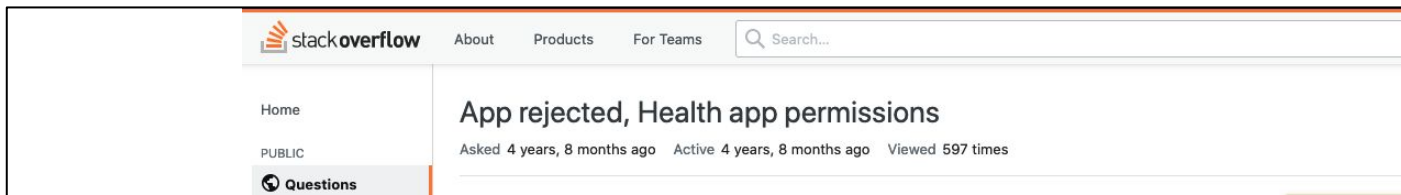
- Goal: Get a handle on the main privacy-related technical concerns and challenges; sources of concerns; and available resources.



Study #1: Questions on StackOverflow

Analyzing StackOverflow posts about health apps

- Goal: Get a handle on the main privacy-related technical concerns and challenges; sources of concerns; and available resources.
- Additional research questions:
 - What are developers' main sources of information about health app privacy?
 - About platform requirements and legal obligations?
 - About methods and tools they can use?
 - How do posts about different platforms or health-data APIs differ?
 - How do developers attribute the causes of privacy-related pain points?



StackOverflow Study: Methodology

- Keyword search on last five years of posts
 - Keywords: {fitness, health, healthcare, medical, google-fit} + {android, ios, watchos, wearos}
- Discarded out-of-scope posts (single coder per, plus reliability check)
 - Criteria: about a mobile app; about human health; about privacy, security, permissions, or authentication
- Developed codebook of common themes (independent codebooks → merge and test)
- Coding all remaining relevant posts (two coders per + resolve disagreements)



StackOverflow Study: Dataset Characteristics

- Out of the original pool of posts about health apps (N=1073), 26% were about permissions or privacy (N=277)

StackOverflow Study: Dataset Characteristics

- Out of the original pool of posts about health apps (N=1073), 26% were about permissions or privacy (N=277)
- Platform (permissions/privacy questions):
 - iOS/WatchOS - 51%
 - Android/WearOS - 48%
 - Both - 1%

StackOverflow Study: Dataset Characteristics

- Out of the original pool of posts about health apps (N=1073), 26% were about permissions or privacy (N=277)
- Platform (permissions/privacy questions):
 - iOS/WatchOS - 51%
 - Android/WearOS - 48%
 - Both - 1%
- Most questions pertain to the platform-integrated APIs for health data (HealthKit, Google Fit, Samsung Health)

StackOverflow Study: Preliminary Findings 1

- Most posters are trying to access some particular data, for example:
 - HealthKit: How to get the permissions requests right
 - Google Fit: How to authenticate to the API



StackOverflow Study: Preliminary Findings 1

- Most posters are trying to access some particular data, for example:
 - HealthKit: How to get the permissions requests right
 - Google Fit: How to authenticate to the API
- Other common reasons for posting:
 - Errors and crashes
 - Special requirements for health data when submitting to the app store
 - Unexpected resource access
 - Problems caused by third-party components



StackOverflow Study: Preliminary Findings 2

- Few explicit mentions of data privacy or sensitivity
- Motivation is to find solutions that satisfy platform requirements; rarely mention legal regulations or ethical considerations explicitly

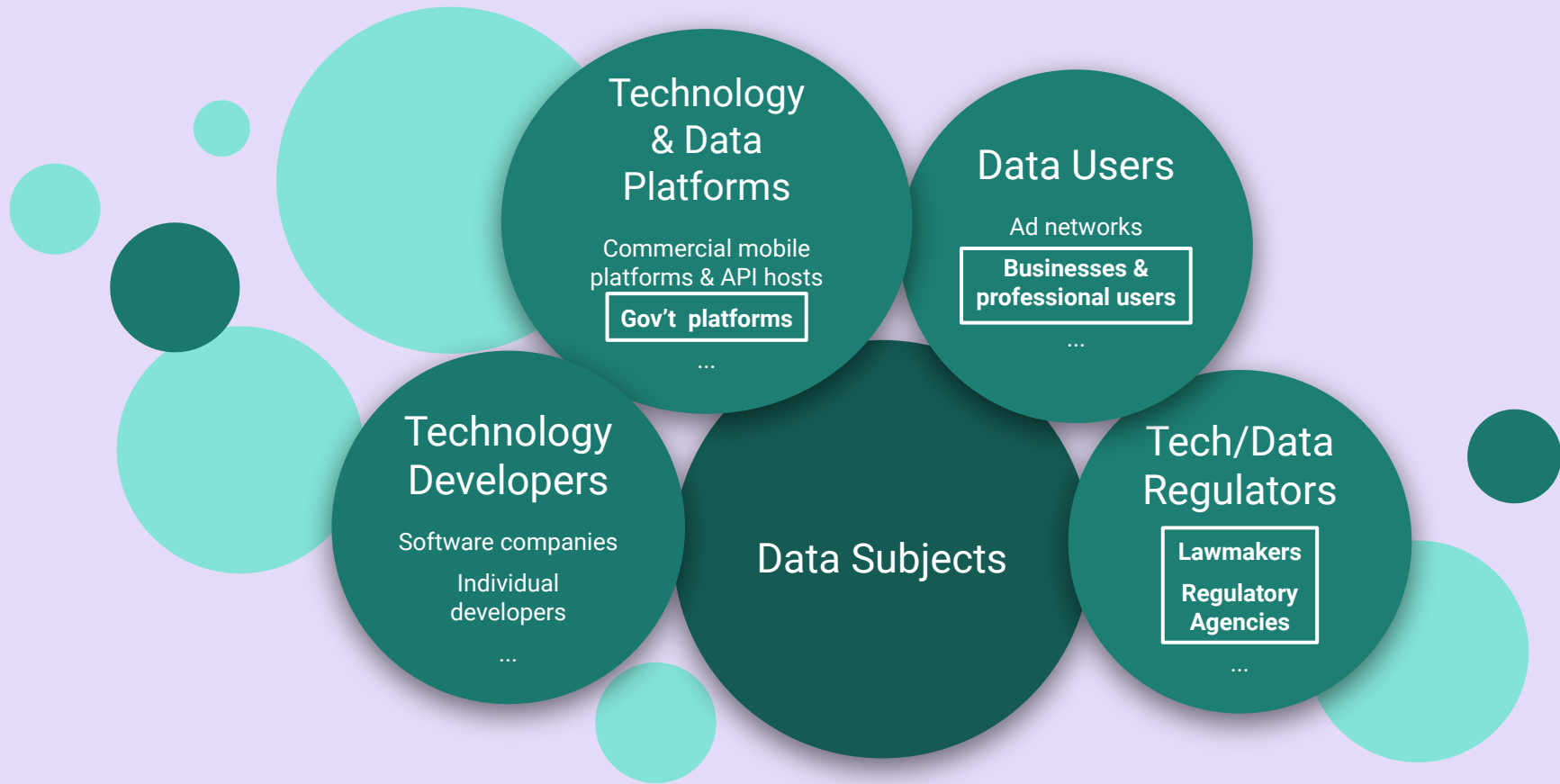
StackOverflow Study: Preliminary Findings 2

- Few explicit mentions of data privacy or sensitivity
- Motivation is to find solutions that satisfy platform requirements; rarely mention legal regulations or ethical considerations explicitly
- Few posters question the necessity for enhanced permissions structures
- Frequent complaints about lack of clarity and documentation about how to satisfy requirements

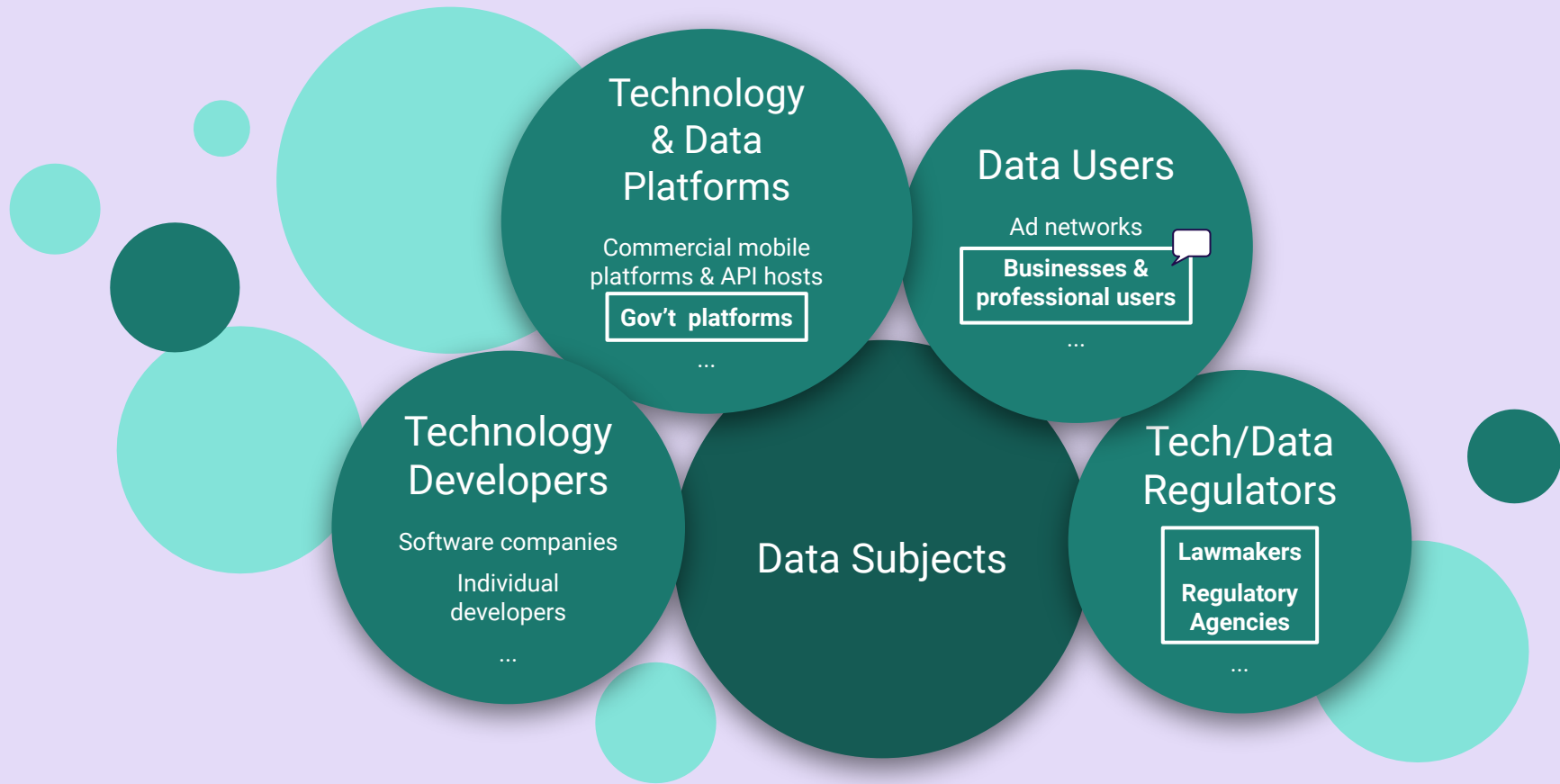


Possible Research Directions

- Surveys of health app developers
 - Challenges, motivations, beliefs
- Walkthroughs of submitting a health app
- Deep dive on available resources and documentation
 - Quality? Currency? Focus?
- Cross-industry studies
 - What issues and factors are unique to health? What's more general?



Stakeholders: Professional Users, Platforms, and Law



Stakeholders: Professional Users, Platforms, and Law

Do You Trust Dr. App? Healthcare Professionals' Views on Health App Privacy

Work in progress!

- ICSI - Aalto University



Research Questions

- What are healthcare professionals' expectations about health apps' data collection, use, and sharing practices? Their concerns?
 - How do expectations and concerns vary across different types of apps?

Research Questions

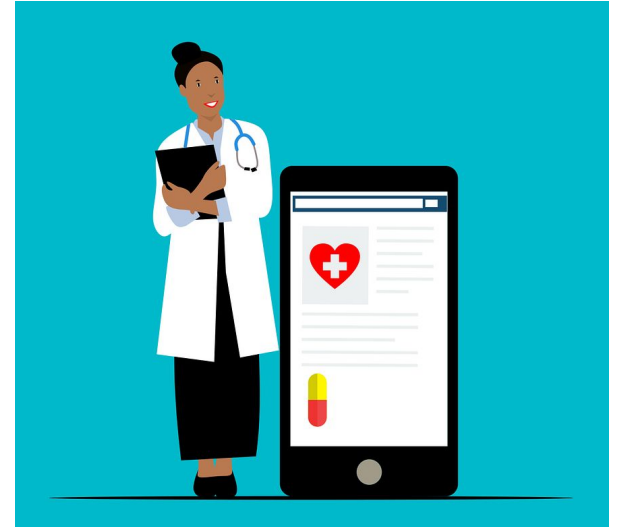
- What are healthcare professionals' expectations about health apps' data collection, use, and sharing practices? Their concerns?
 - How do expectations and concerns vary across different types of apps?
- What data privacy questions do healthcare professionals consider when deciding whether to recommend an app?
 - Where do they seek answers to those questions?

Research Questions

- What are healthcare professionals' expectations about health apps' data collection, use, and sharing practices? Their concerns?
 - How do expectations and concerns vary across different types of apps?
- What data privacy questions do healthcare professionals consider when deciding whether to recommend an app?
 - Where do they seek answers to those questions?
- How do privacy expectations, concerns, questions, and resources vary across healthcare providers in different countries?
 - Are those differences related to different structures and requirements for handling patients' medical data?

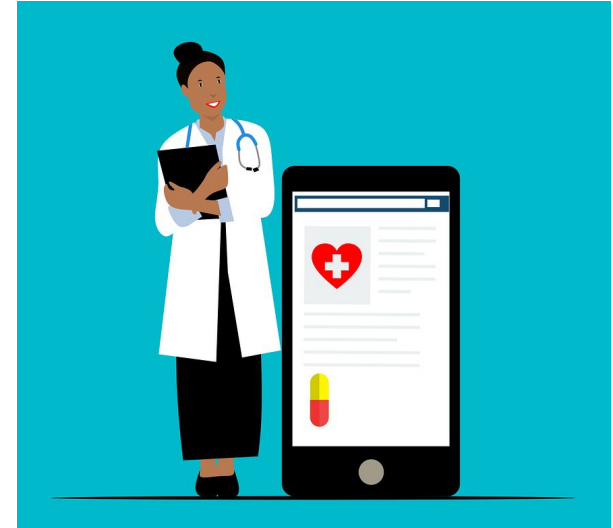
Study Design

- Interviews with ~40 healthcare professionals



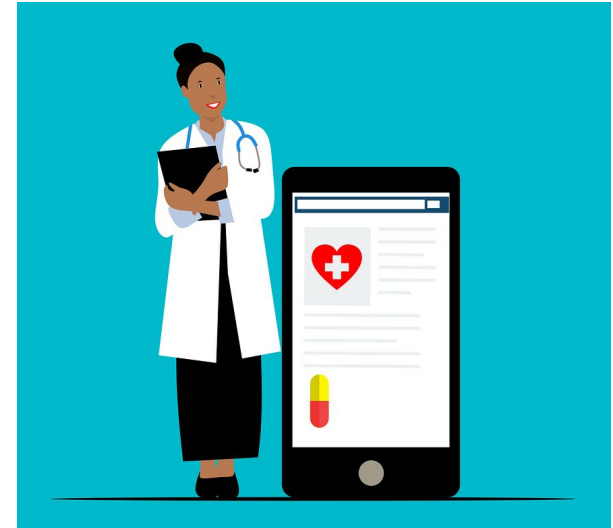
Study Design

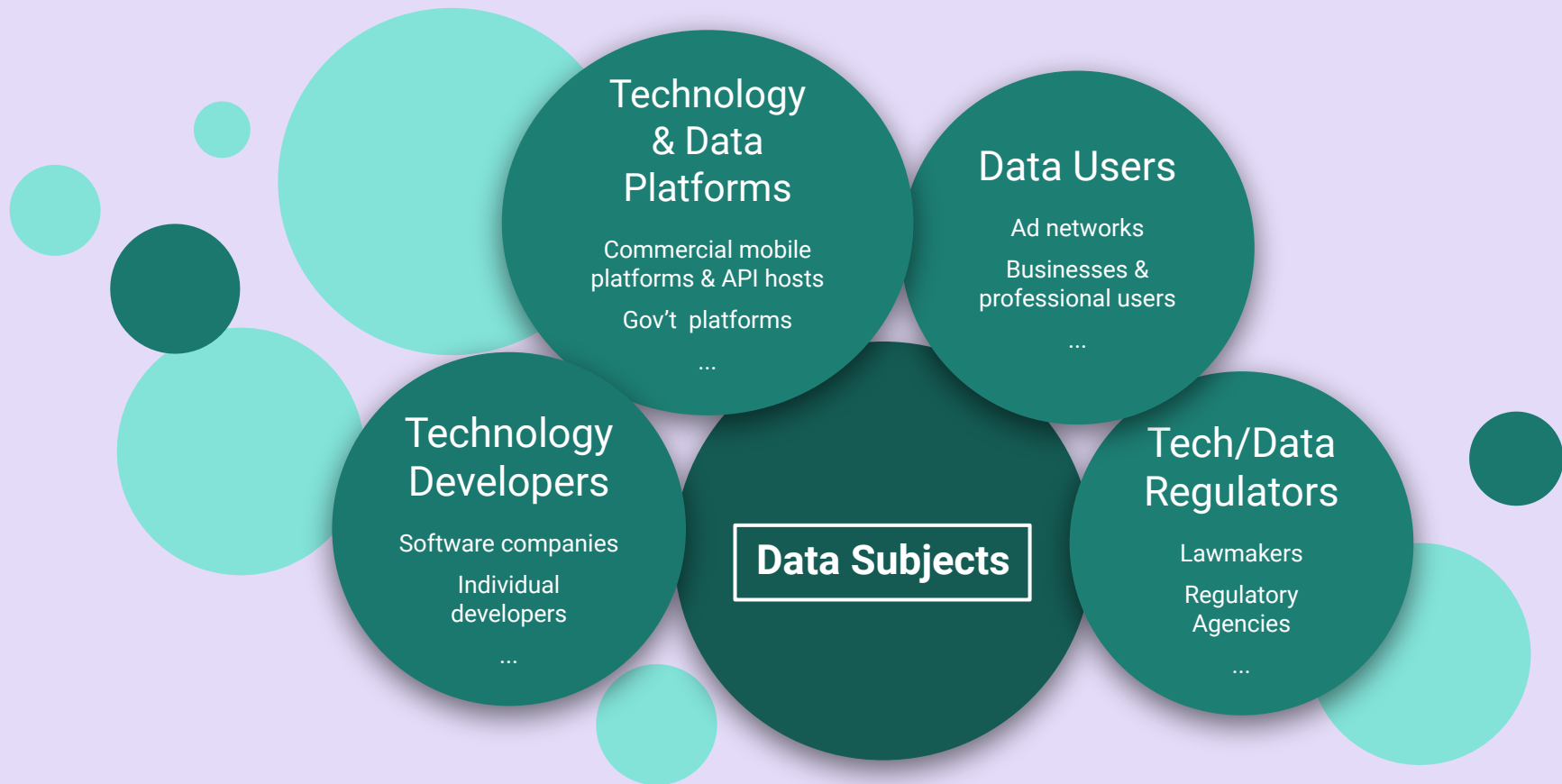
- Interviews with ~40 healthcare professionals
- Recruit participants from 5 countries
 - U.S., Finland, Singapore, Sri Lanka, Sweden
 - Variance in legal and ethical protections, centralization and control of medical data, and availability of government-sponsored health platforms



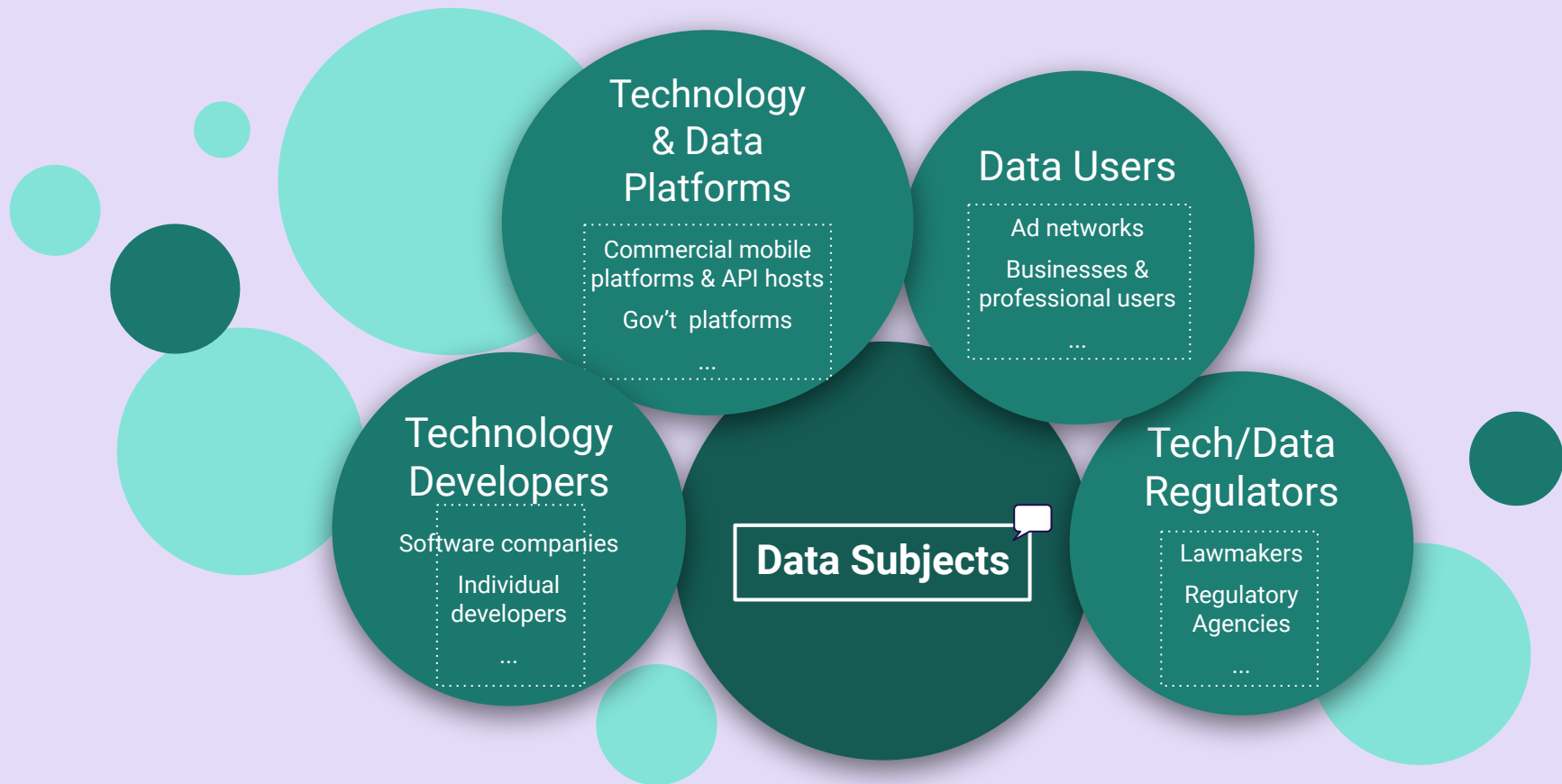
Study Design

- Interviews with ~40 healthcare professionals
- Recruit participants from 5 countries
 - U.S., Finland, Singapore, Sri Lanka, Sweden
 - Variance in legal and ethical protections, centralization and control of medical data, and availability of government-sponsored health platforms
- Currently awaiting IRB/ethics approval





Stakeholders: Individual End Users



Stakeholders: Individual End Users

Users' Reasoning About Health App Privacy... And Whether Reason Matches Reality

Work in progress!

- ICSI - St. Mary's University - UC Berkeley



Research Questions

- What are users' expectations about health apps' data handling practices and privacy protections?

Research Questions

- What are users' expectations about health apps' data handling practices and privacy protections?
- What are users' expectations about legal protections w.r.t. health app data?

Research Questions

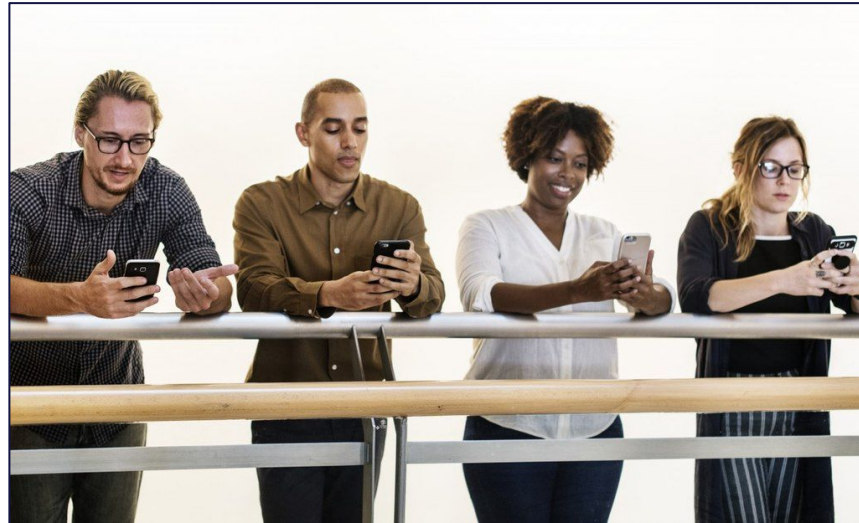
- What are users' expectations about health apps' data handling practices and privacy protections?
- What are users' expectations about legal protections w.r.t. health app data?
- What factors shape users' privacy expectations regarding health apps?

Research Questions

- What are users' expectations about health apps' data handling practices and privacy protections?
- What are users' expectations about legal protections w.r.t. health app data?
- What factors shape users' privacy expectations regarding health apps?
- How do users' expectations differ for different types of apps? Different types of data?

Study Design

- Large survey study
 - App store mock-ups for different types of apps
 - Vignette questions about expectations



Thank you!

jbernd@icsi.berkeley.edu