“I need a better description”: An Investigation Into User Expectations For Differential Privacy

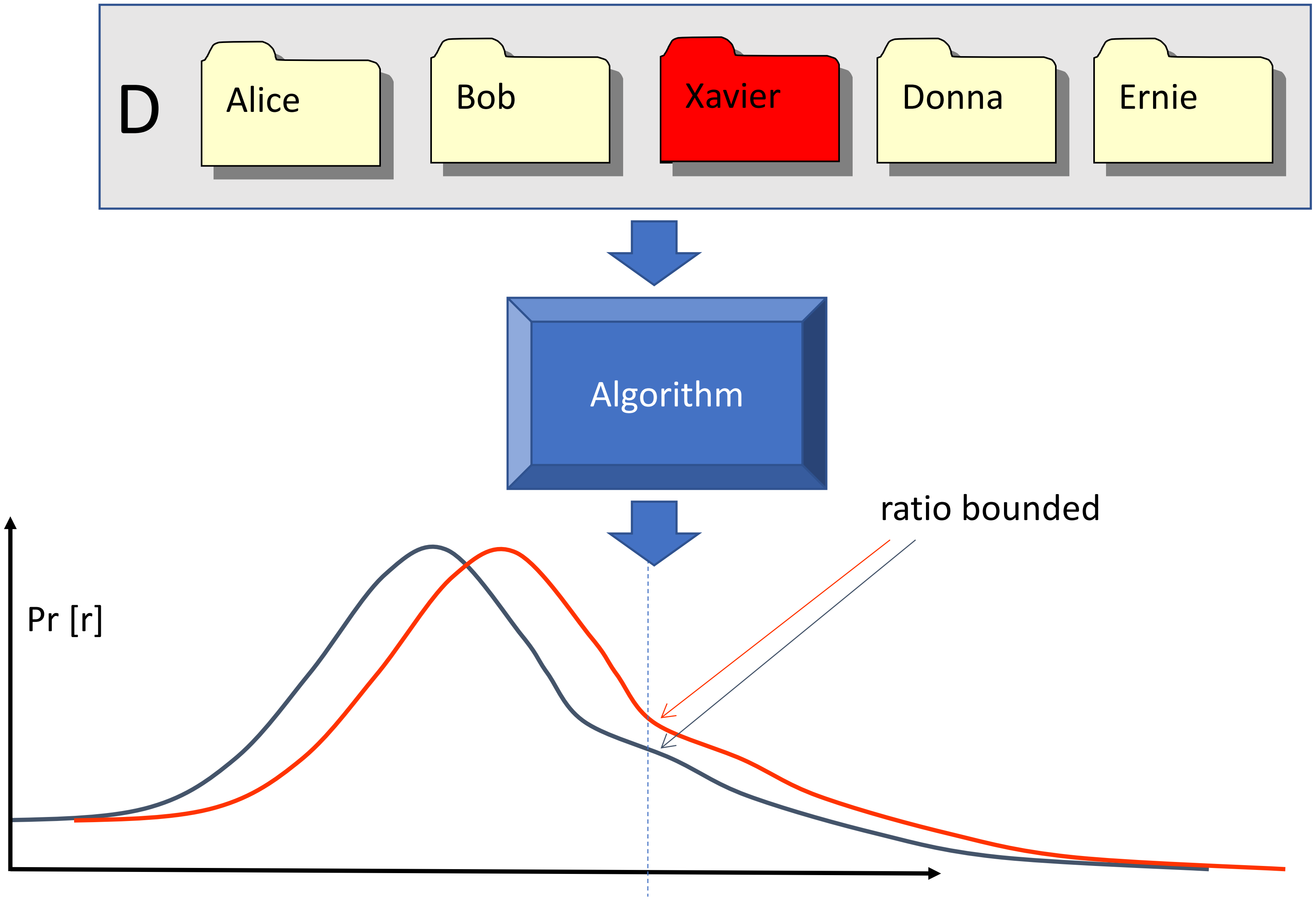
Gabriel Kaptchuk (Boston University)

joint work with

Rachel Cummings (Columbia University)

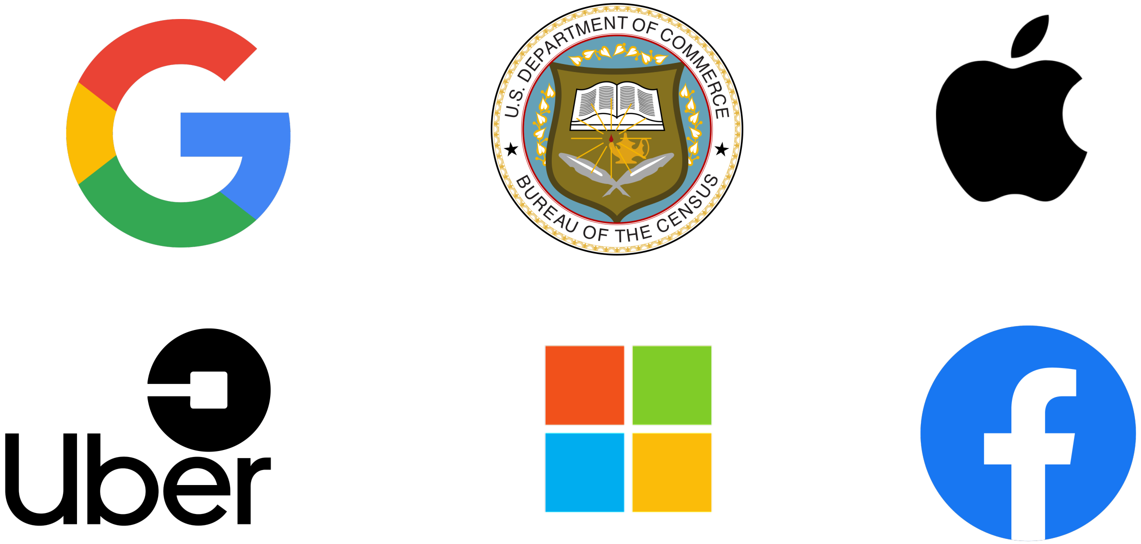
Elissa M Redmiles (Max Plank Institute for Software Systems)

Differential privacy [DMNS ‘06]



Differential privacy [DMNS ‘06]

- Differential privacy is deployed in practice by major tech companies and government organizations:



- How should these organizations explain differential privacy to their users?

You should sign up for my product. It does cool things!

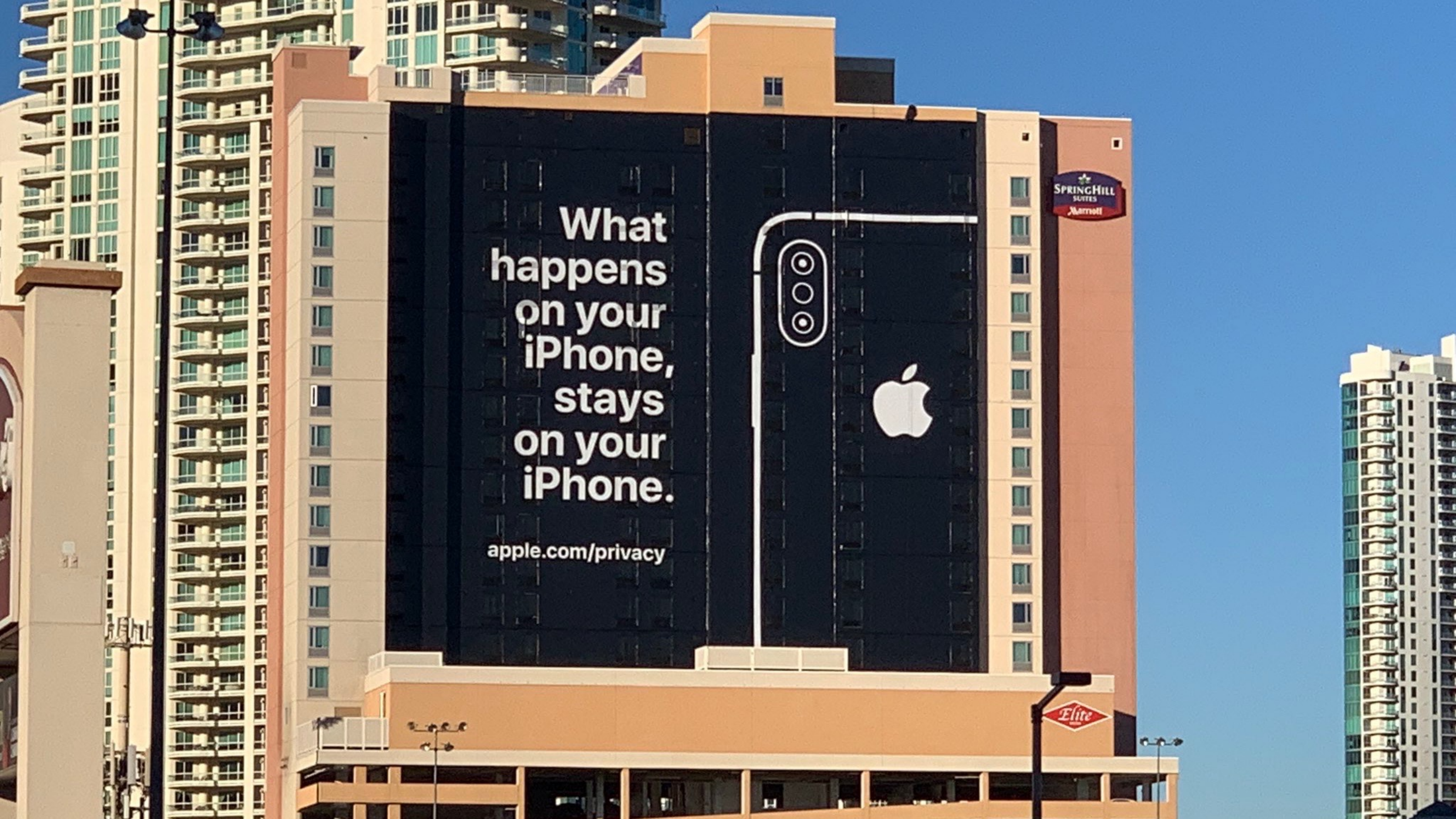
I want to, but I'm concerned about privacy!

We use differential privacy!

What's that??

good company

Good Company is the logo of a Ski company – No endorsement implied



What happens on your iPhone, stays on your iPhone.
apple.com/privacy
SpringHill Suites
Marriott
Elite

Research questions

Does DP “work” for users?

How do users understand DP when they encounter it “in-the-wild?”

Research questions

(RQ1) Do users care about the type of protections provided by DP?

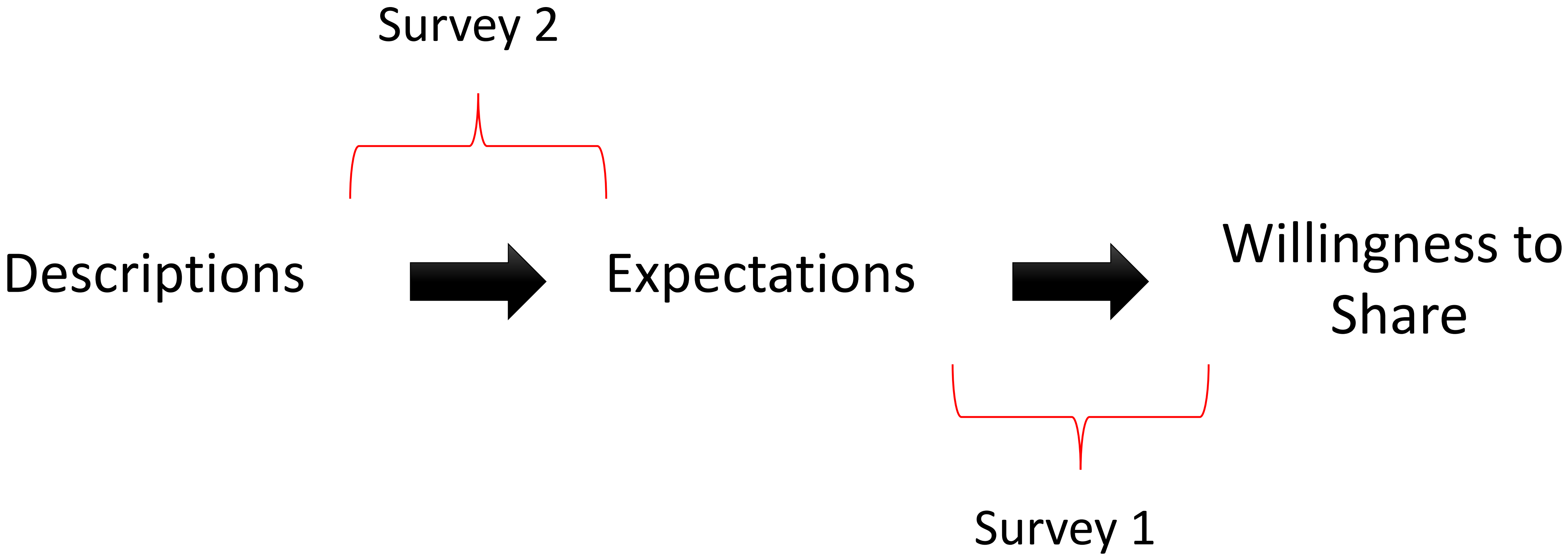
(RQ2) Are users more willing to share their data with reduce information disclosure risks?

(RQ3) How does the way DP is described impact users’ expectations of protection against information disclosures?

(RQ4) How does the way DP is described impact users’ willingness to share their data?

Research plan / Outline

- Vignette-based surveys to elicit preferences/perceptions (n=2,424)
- To address RQ1 and RQ2,
  - Present information-sharing scenario, and query privacy concerns
  - Set privacy expectations for those concerns, and query willingness to share data
  - Measures how users' privacy concerns align with the protections provided by DP
- To address RQ3 and RQ4,
  - Collect descriptions of differential privacy
  - Present information-sharing scenario and a DP description
  - Query privacy expectations and willingness to share data
  - Measures how accurately and effectively DP descriptions set user expectations



Survey 2

Descriptions

Expectations

Willingness to Share

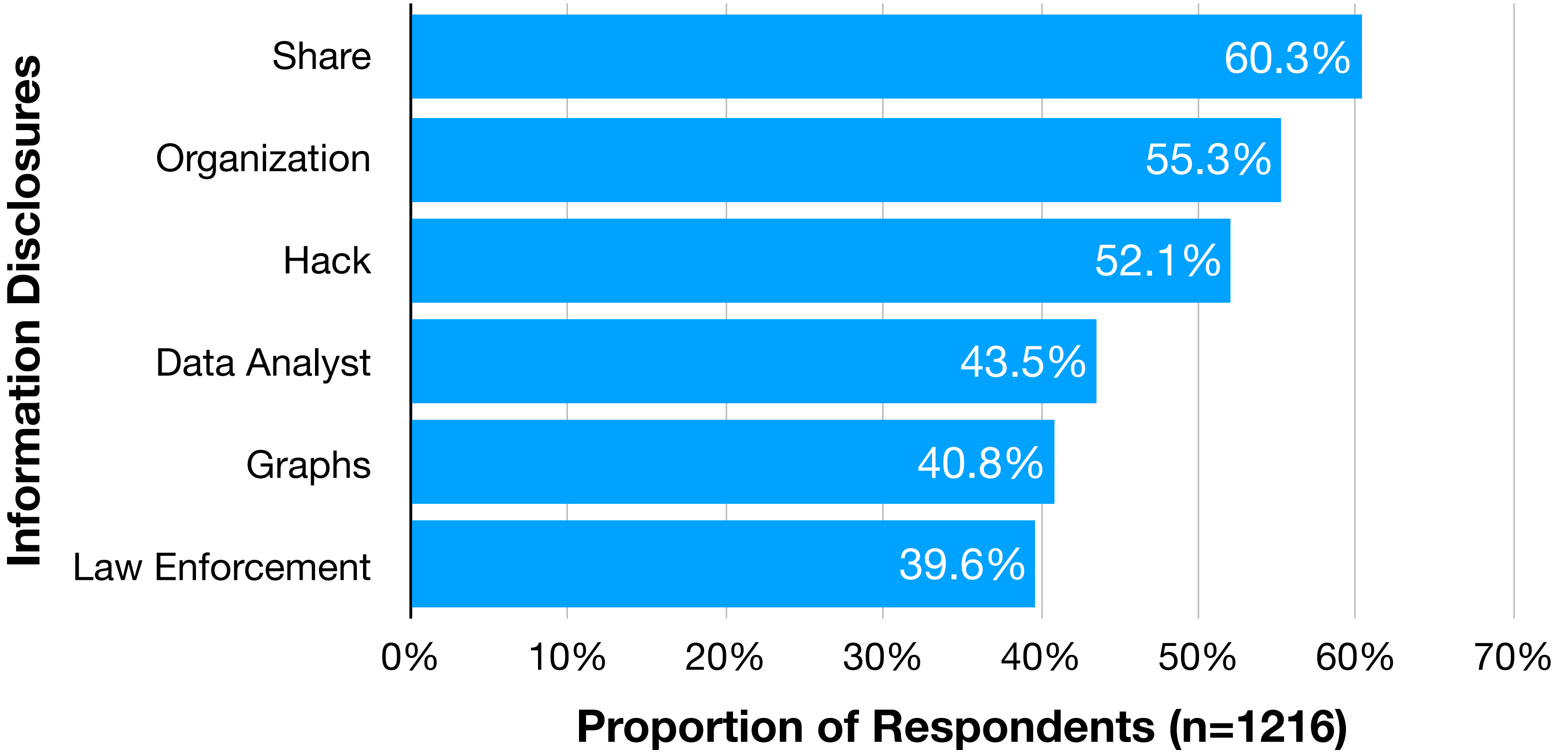
Survey 1

For information of survey methods, please consult the paper

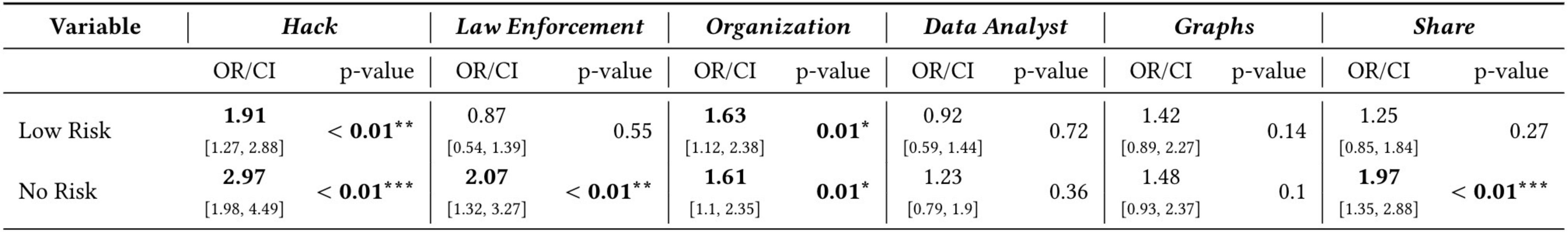
Information disclosure possibilities

Name	Expectation
Hack	A criminal or foreign government that hacks the organization could learn my data
Law Enforcement	Law enforcement could access my data with a court order
Organization	My friend will learn my data/ Organization will store data
Data Analyst	A data analyst working for the organization could learn my data
Graphs	Graphs or informational charts created using the collected information could reveal my data
Sharing	Data that the organization shares with other organizations could reveal my data

RQ1: What information disclosures concern users?

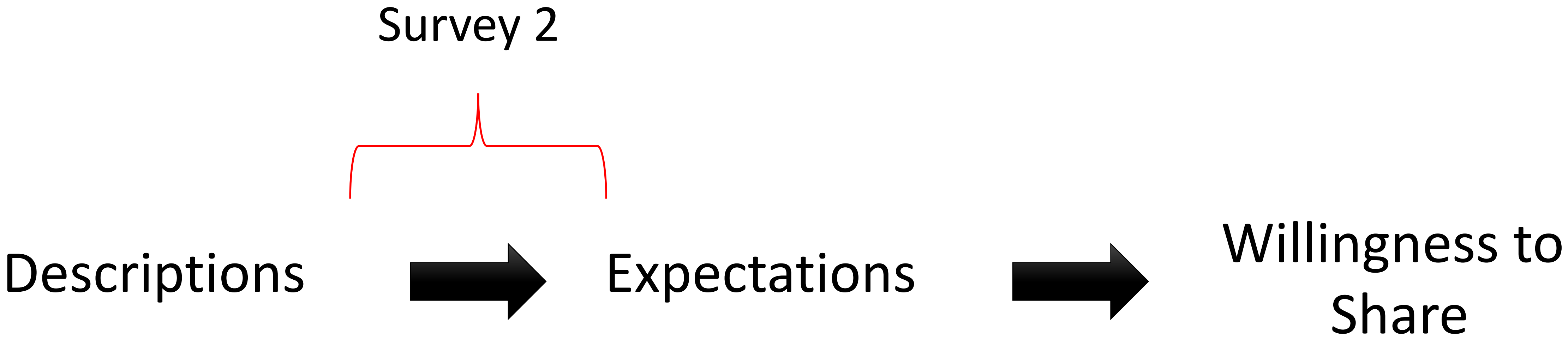


RQ2: How does the probability of information disclosures affect data sharing?



Variable	*Hack*		*Law Enforcement*		*Organization*		*Data Analyst*		*Graphs*		*Share*	
	OR/CI	p-value	OR/CI	p-value	OR/CI	p-value	OR/CI	p-value	OR/CI	p-value	OR/CI	p-value
Low Risk	**1.91** [1.27, 2.88]	< **0.01****	0.87 [0.54, 1.39]	0.55	**1.63** [1.12, 2.38]	**0.01***	0.92 [0.59, 1.44]	0.72	1.42 [0.89, 2.27]	0.14	1.25 [0.85, 1.84]	0.27
No Risk	**2.97** [1.98, 4.49]	< **0.01*****	**2.07** [1.32, 3.27]	< **0.01****	**1.61** [1.1, 2.35]	**0.01***	1.23 [0.79, 1.9]	0.36	1.48 [0.93, 2.37]	0.1	**1.97** [1.35, 2.88]	< **0.01*****

- Logistic regression model for respondents who cared about disclosure
- Odds Ratio measures change in odds of sharing data, compared against High Risk condition (OR>1 $\Longrightarrow$ increased sharing), CI: 95%
- Read as: “decreased risk of <disclosure> increased chances of sharing by X”



Survey 2

Descriptions

Expectations

Willingness to Share

“differential privacy,” the new **gold standard in data privacy** protection.

Differential privacy works by algorithmically **scrambling individual user data** so that it cannot be traced back

‘‘differential privacy,’’ which alters the numbers but **does not change core findings** to protect the identities of individual respondents.

We use differential privacy!

When a differential privacy algorithm is applied to a data set, those **links get blurred**, and bits of data can no longer be traced to their source.

In short, differential privacy **allows general statistical analysis** without revealing information about a particular individual in the data

In ideal implementations, this **risk remains close to zero**, guaranteeing… virtually no adverse effect on them from an informational standpoint.

Differential Privacy Descriptions Gathered

37
<u>Industry</u>

U.S. DEPARTMENT OF COMMERCE
BUREAU OF THE CENSUS
integrate.ai
Uber
WorkFusion
IMMUTA
Georgian Partners
IMMUTA
Bluecore
thunder
Experience Cloud
ad exchanger

30
<u>Press</u>

theguardian
The Boston Globe
engadget
msn
The Washington Post
BBC
FOX NEWS
Reveal
from The Center for Investigative Reporting
TechCrunch

10
<u>Academic</u>

Resulting Descriptions:

1. Unsubstantial

2. DP Techniques

3. DP enables analysis

4. DP is widely trusted

5. DP reduces user risk

6. Technical description of DP

Resulting Descriptions:

1. Unsubstantial

2. DP Techniques

3. DP enables analysis

Differential privacy…

is the gold standard in data privacy and protection and is widely recognized as the strongest guarantee of privacy available.

4. DP is widely trusted

5. DP reduces user risk

6. Technical description of DP

Resulting Descriptions:

1. Unsubstantial

2. DP Techniques

3. DP enables analysis

Differential privacy…

injects statistical noise into collected data in a way that protects privacy without significantly changing conclusions.

4. DP is widely trusted

5. DP reduces user risk

6. Technical description of DP

Resulting Descriptions:

1. Unsubstantial

2. DP Techniques

3. DP enables analysis

Differential privacy…

allows analysts to learn useful information from large amounts of data without compromising an individual's privacy.

4. DP is widely trusted

5. DP reduces user risk

6. Technical description of DP

Resulting Descriptions:

1. Unsubstantial

2. DP Techniques

3. DP enables analysis

Differential privacy…

is a novel mathematical technique to preserve privacy which is used by companies like Apple and Uber.

4. DP is widely trusted

5. DP reduces user risk

6. Technical description of DP

Resulting Descriptions:

1. Unsubstantial

2. DP Techniques

3. DP enables analysis

Differential privacy…

protects a user’s identity and the specifics of their data, meaning individuals incur almost no risk by joining the dataset.

4. DP is widely trusted

5. DP reduces user risk

6. Technical description of DP

Resulting Descriptions:

1. Unsubstantial

2. DP Techniques

3. DP enables analysis

Differential privacy…

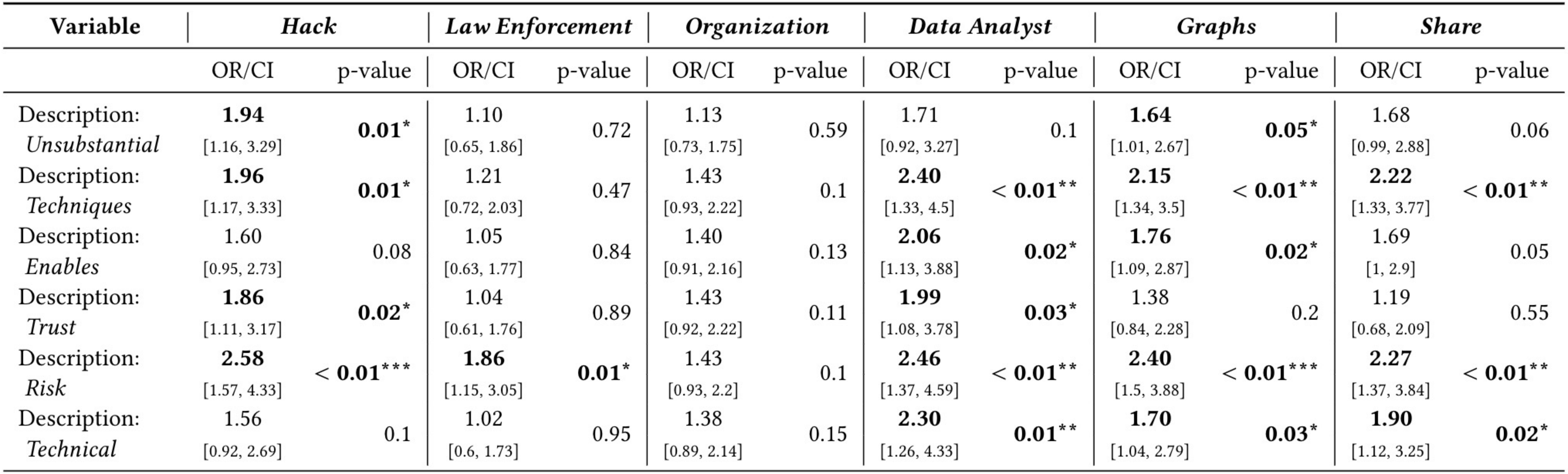
ensures that the removal or addition of a single database item does not (substantially) affect the outcome of any analysis. It follows that no risk is incurred by joining the database, providing a mathematically rigorous means of coping with the fact that distributional information may be disclosive. [Dwork08]

4. DP is widely trusted

5. DP reduces user risk

6. Technical description of DP

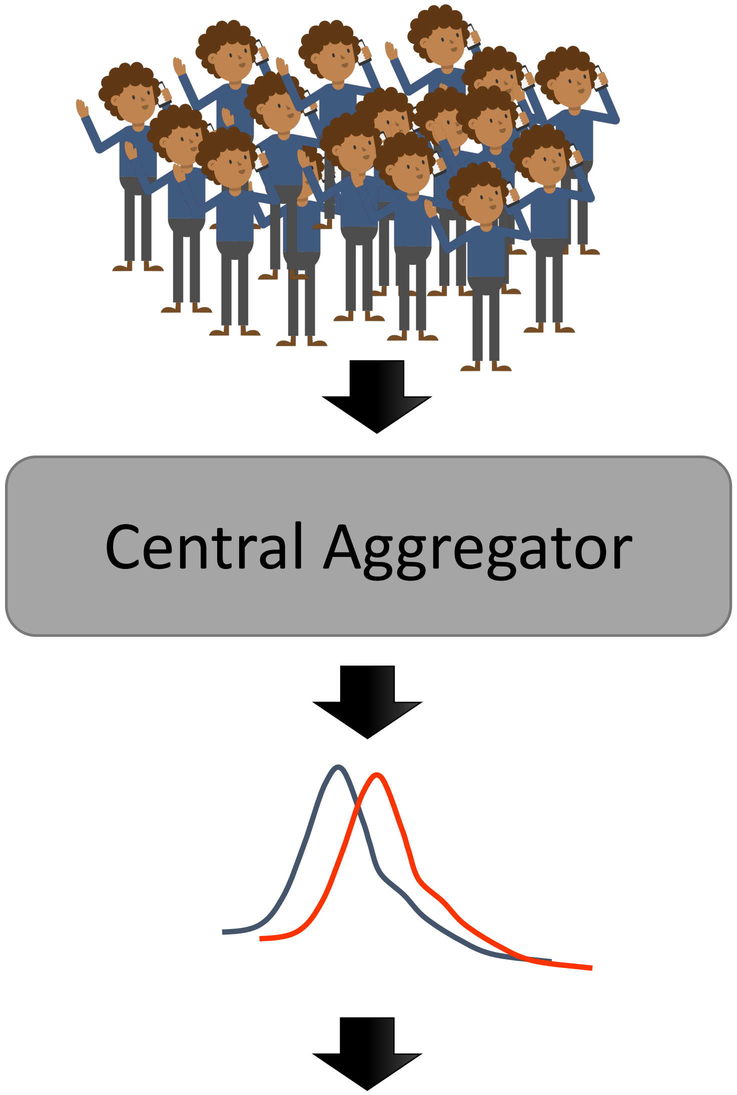
RQ3: How do differential privacy descriptions affect privacy expectations?



Variable	*Hack*		*Law Enforcement*		*Organization*		*Data Analyst*		*Graphs*		*Share*	
	OR/CI	p-value	OR/CI	p-value	OR/CI	p-value	OR/CI	p-value	OR/CI	p-value	OR/CI	p-value
Description: *Unsubstantial*	**1.94** [1.16, 3.29]	**0.01***	1.10 [0.65, 1.86]	0.72	1.13 [0.73, 1.75]	0.59	1.71 [0.92, 3.27]	0.1	**1.64** [1.01, 2.67]	**0.05***	1.68 [0.99, 2.88]	0.06
Description: *Techniques*	**1.96** [1.17, 3.33]	**0.01***	1.21 [0.72, 2.03]	0.47	1.43 [0.93, 2.22]	0.1	**2.40** [1.33, 4.5]	< **0.01****	**2.15** [1.34, 3.5]	< **0.01****	**2.22** [1.33, 3.77]	< **0.01****
Description: *Enables*	1.60 [0.95, 2.73]	0.08	1.05 [0.63, 1.77]	0.84	1.40 [0.91, 2.16]	0.13	**2.06** [1.13, 3.88]	**0.02***	**1.76** [1.09, 2.87]	**0.02***	1.69 [1, 2.9]	0.05
Description: *Trust*	**1.86** [1.11, 3.17]	**0.02***	1.04 [0.61, 1.76]	0.89	1.43 [0.92, 2.22]	0.11	**1.99** [1.08, 3.78]	**0.03***	1.38 [0.84, 2.28]	0.2	1.19 [0.68, 2.09]	0.55
Description: *Risk*	**2.58** [1.57, 4.33]	< **0.01*****	**1.86** [1.15, 3.05]	**0.01***	1.43 [0.93, 2.2]	0.1	**2.46** [1.37, 4.59]	< **0.01****	**2.40** [1.5, 3.88]	< **0.01*****	**2.27** [1.37, 3.84]	< **0.01****
Description: *Technical*	1.56 [0.92, 2.69]	0.1	1.02 [0.6, 1.73]	0.95	1.38 [0.89, 2.14]	0.15	**2.30** [1.26, 4.33]	**0.01****	**1.70** [1.04, 2.79]	**0.03***	**1.90** [1.12, 3.25]	**0.02***

- Read as: “Describing DP with <description> increases privacy expectations for <disclosure> by multiplicative factor of X”

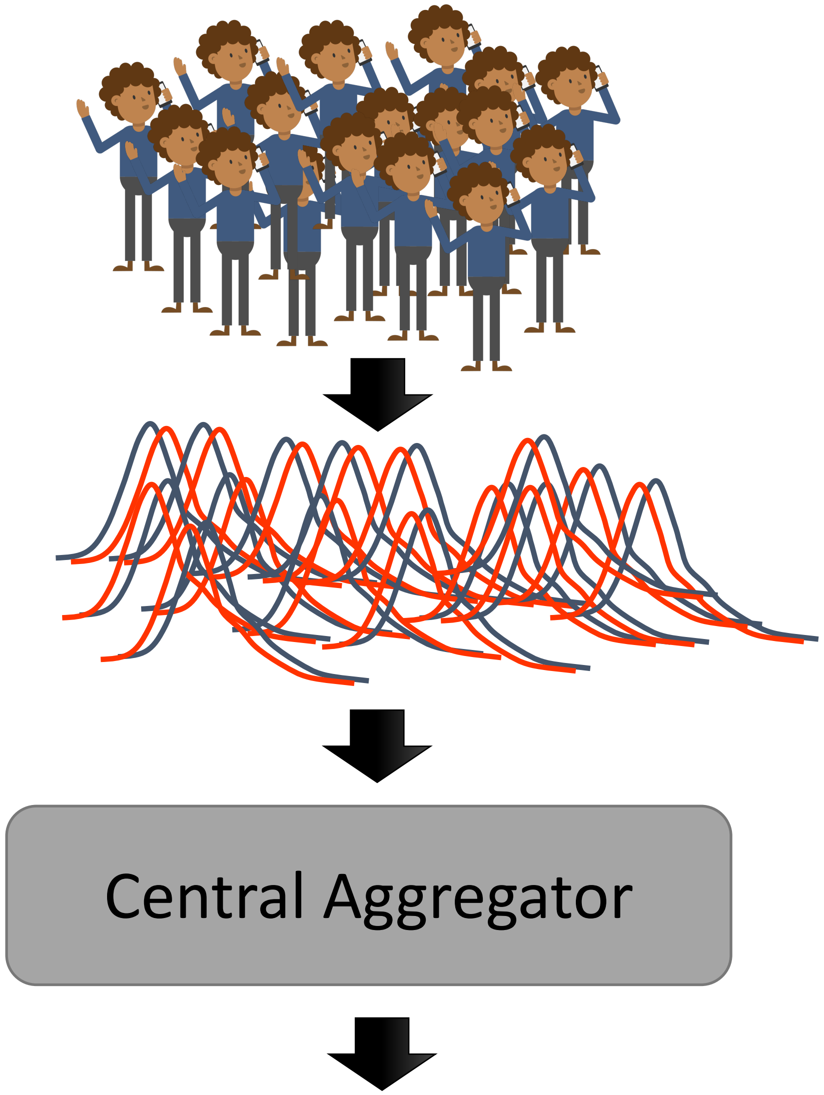
Central Model



Central Aggregator

Private Output

Local Model



Central Aggregator

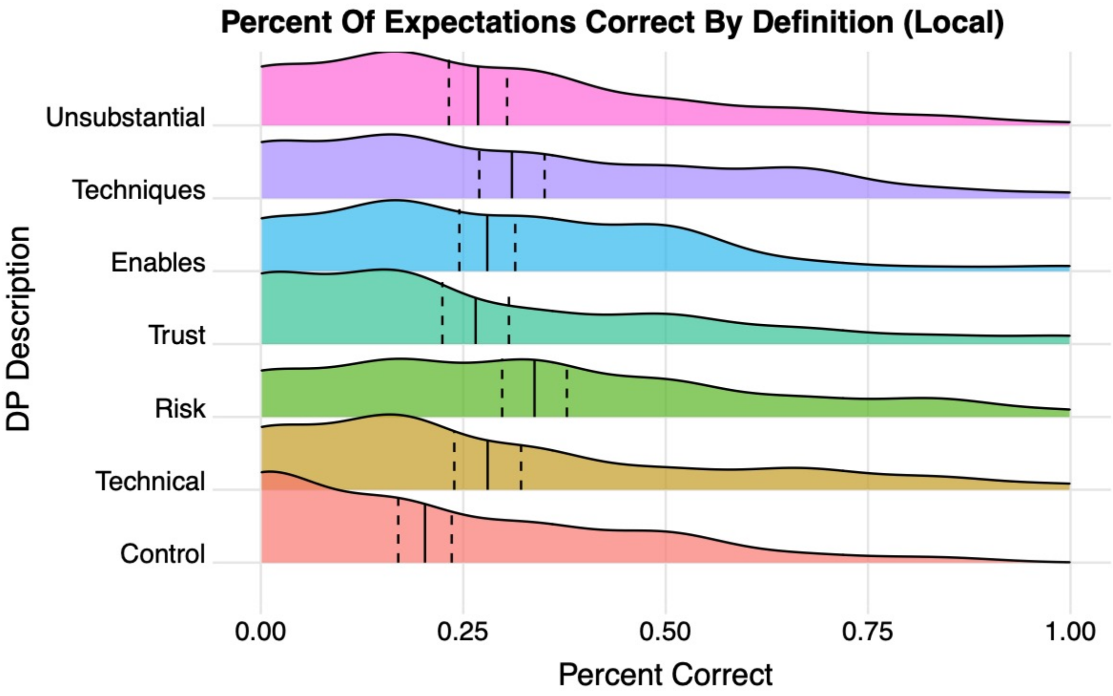
Private Output

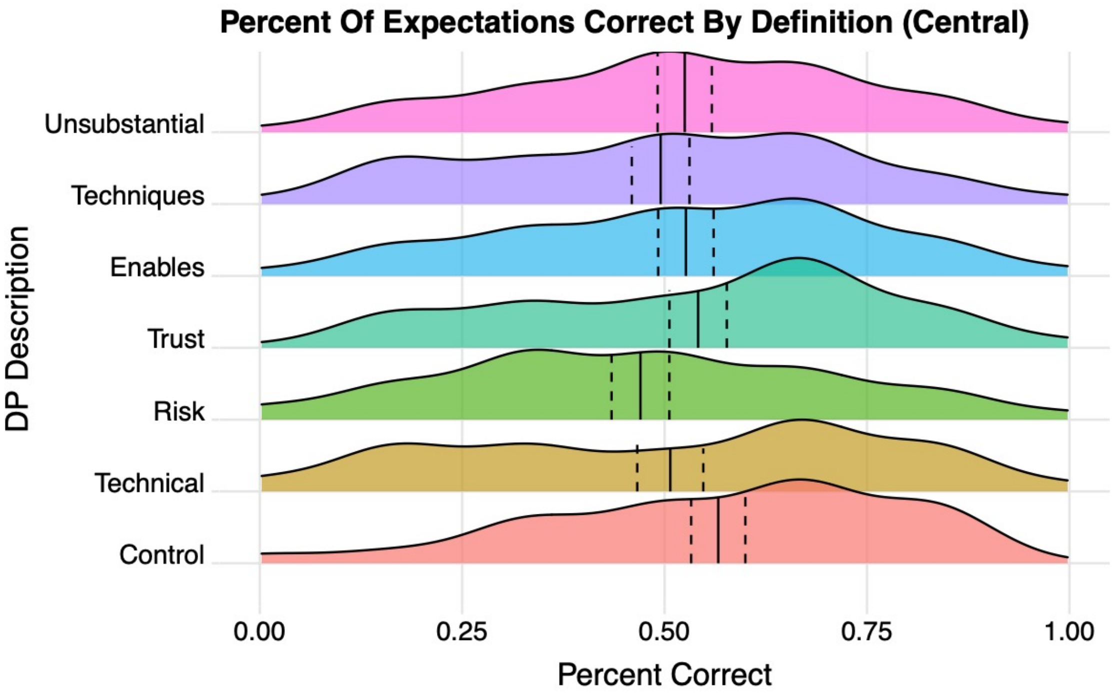
Ground truth* vulnerabilities – Local vs central DP

Disclosure	Local	Central
Hack	False	True
Law Enforcement	False	True
Organization	False	True
Data Analyst	False	True
Graphs	False	False
Sharing	False	True

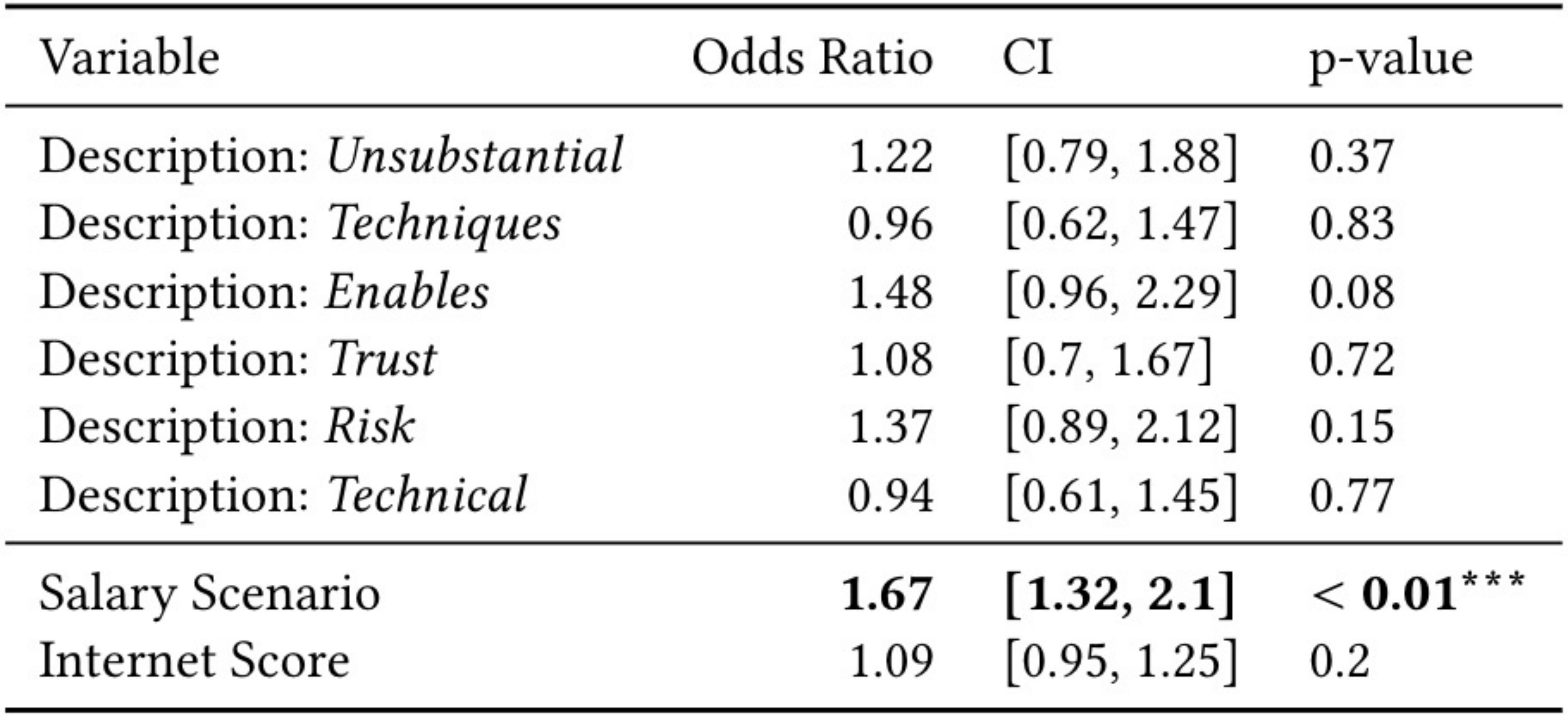
*For "typical" implementation. Actual ground truth will depend on implementation details, including privacy parameters

Correctness of expectations – Local vs central DP



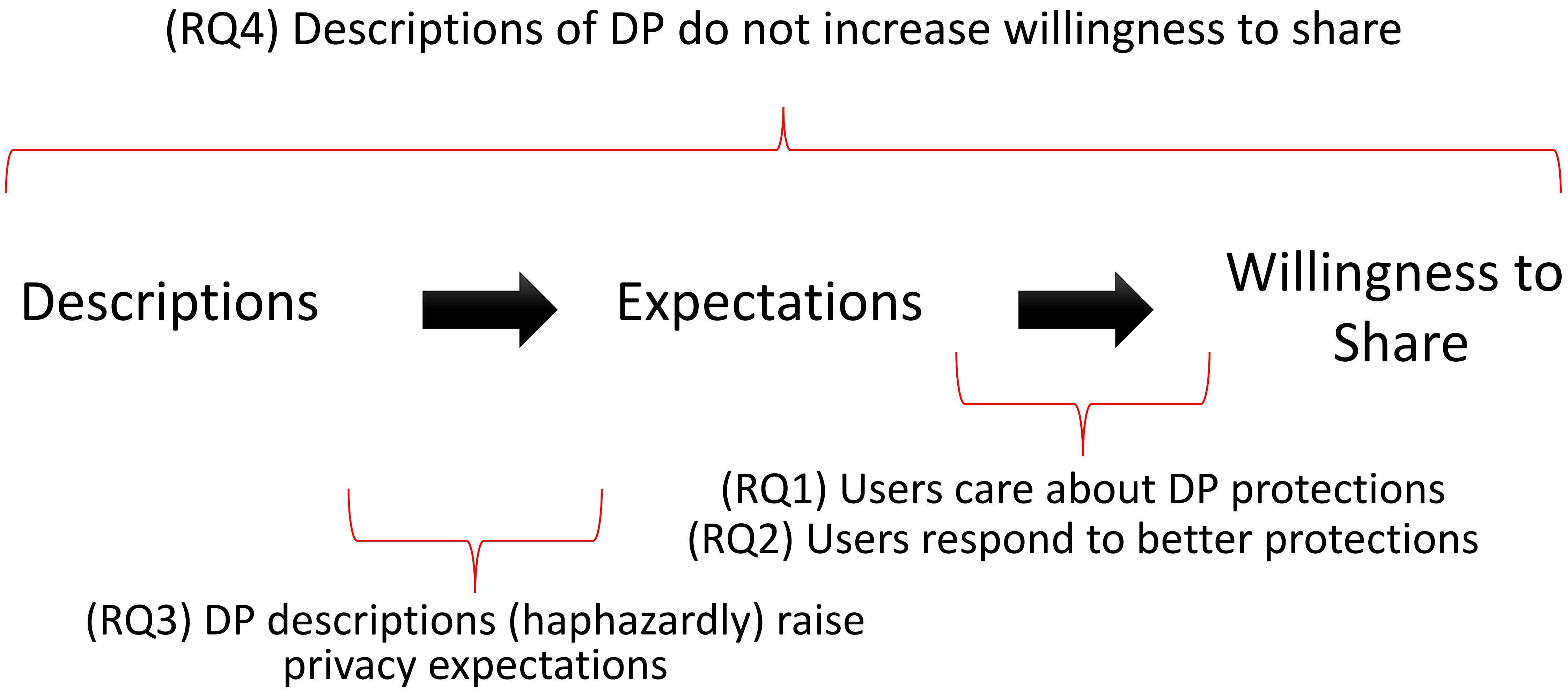


RQ4: How do descriptions of DP affect sharing?



Variable	Odds Ratio	CI	p-value
Description: *Unsubstantial*	1.22	[0.79, 1.88]	0.37
Description: *Techniques*	0.96	[0.62, 1.47]	0.83
Description: *Enables*	1.48	[0.96, 2.29]	0.08
Description: *Trust*	1.08	[0.7, 1.67]	0.72
Description: *Risk*	1.37	[0.89, 2.12]	0.15
Description: *Technical*	0.94	[0.61, 1.45]	0.77
Salary Scenario	**1.67**	**[1.32, 2.1]**	$< \mathbf{0.01}^{***}$
Internet Score	1.09	[0.95, 1.25]	0.2

- DP descriptions have no significant effect
- Information sharing scenario does effect sharing intention



(RQ4) Descriptions of DP do not increase willingness to share

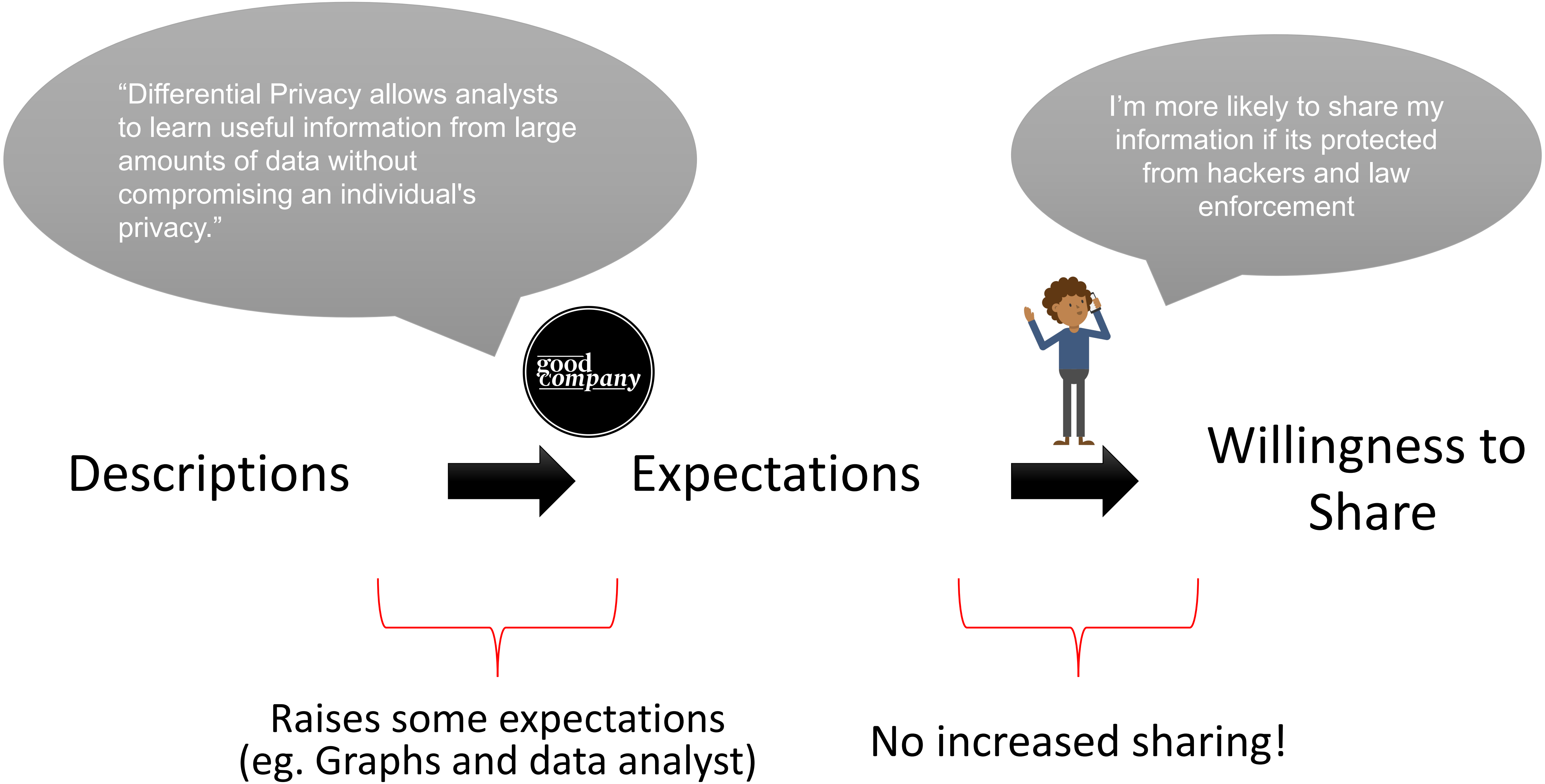
Descriptions

Expectations

Willingness to Share

(RQ1) Users care about DP protections
(RQ2) Users respond to better protections

(RQ3) DP descriptions (haphazardly) raise privacy expectations



“Differential Privacy allows analysts to learn useful information from large amounts of data without compromising an individual's privacy.”

I’m more likely to share my information if its protected from hackers and law enforcement

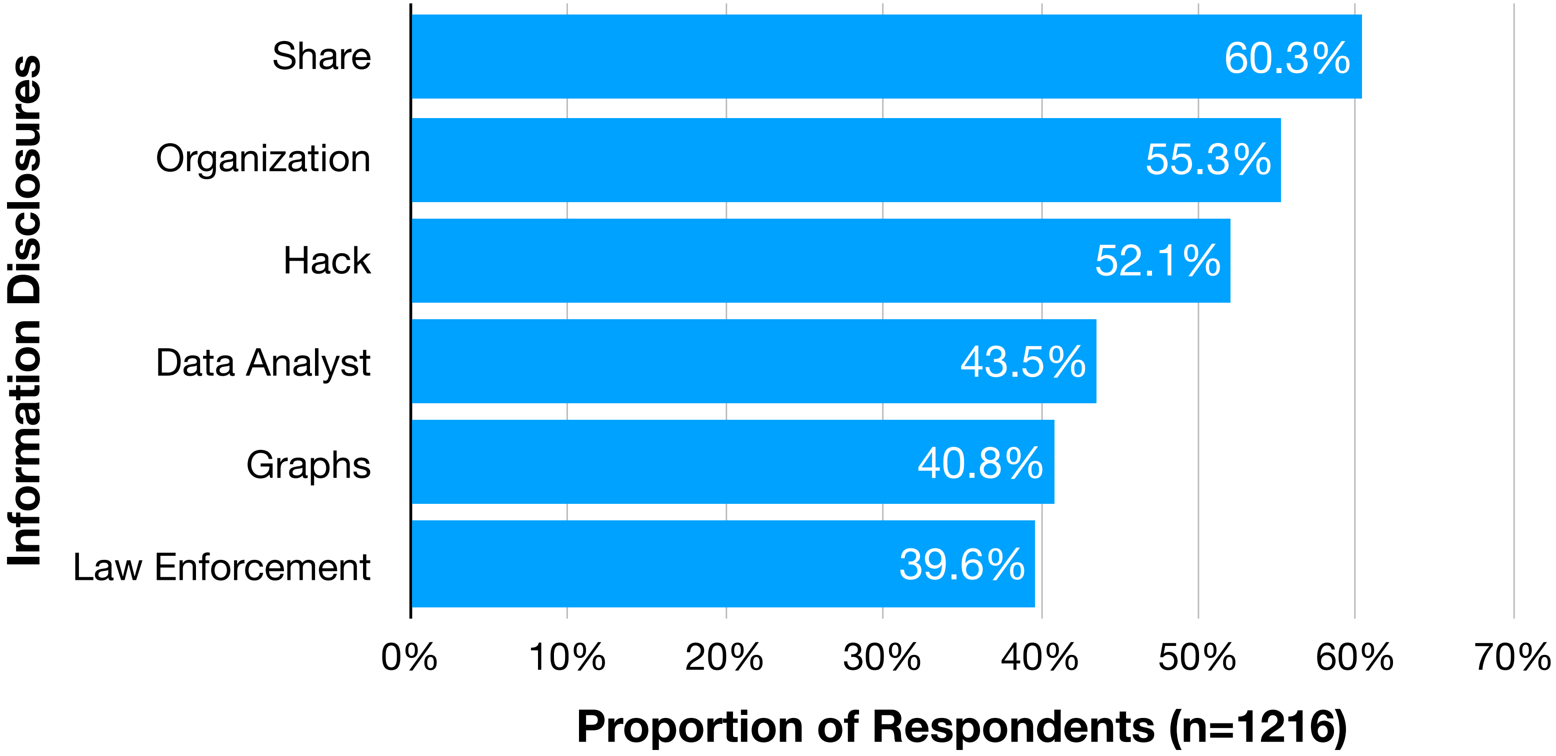
good company

Descriptions → Expectations → Willingness to Share

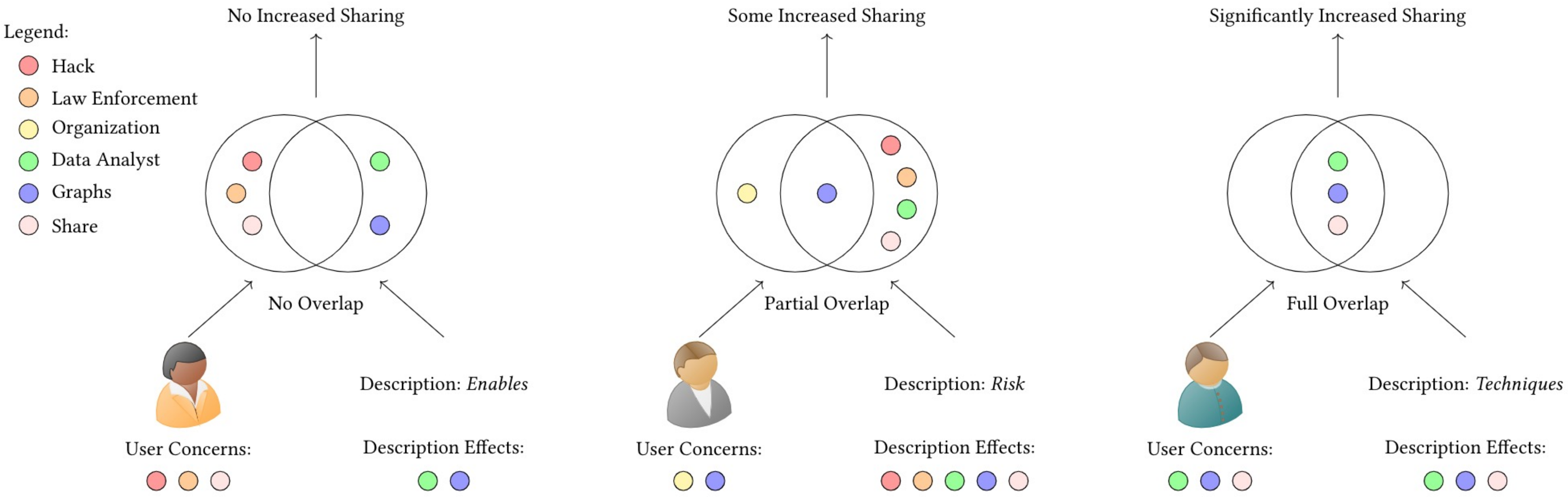
Raises some expectations (eg. Graphs and data analyst)

No increased sharing!

RQ1: What information disclosures concern users?



Misalignment of User Concern and Description Effects



Legend:

- 🔴 Hack
- 🟠 Law Enforcement
- 🟡 Organization
- 🟢 Data Analyst
- 🔵 Graphs
- ⚪ Share

No Increased Sharing

No Overlap

Description: *Enables*

User Concerns: 🔴 🟠 ⚪

Description Effects: 🟢 🔵

Some Increased Sharing

Partial Overlap

Description: *Risk*

User Concerns: 🟡 🔵

Description Effects: 🔴 🟠 🟢 🔵 ⚪

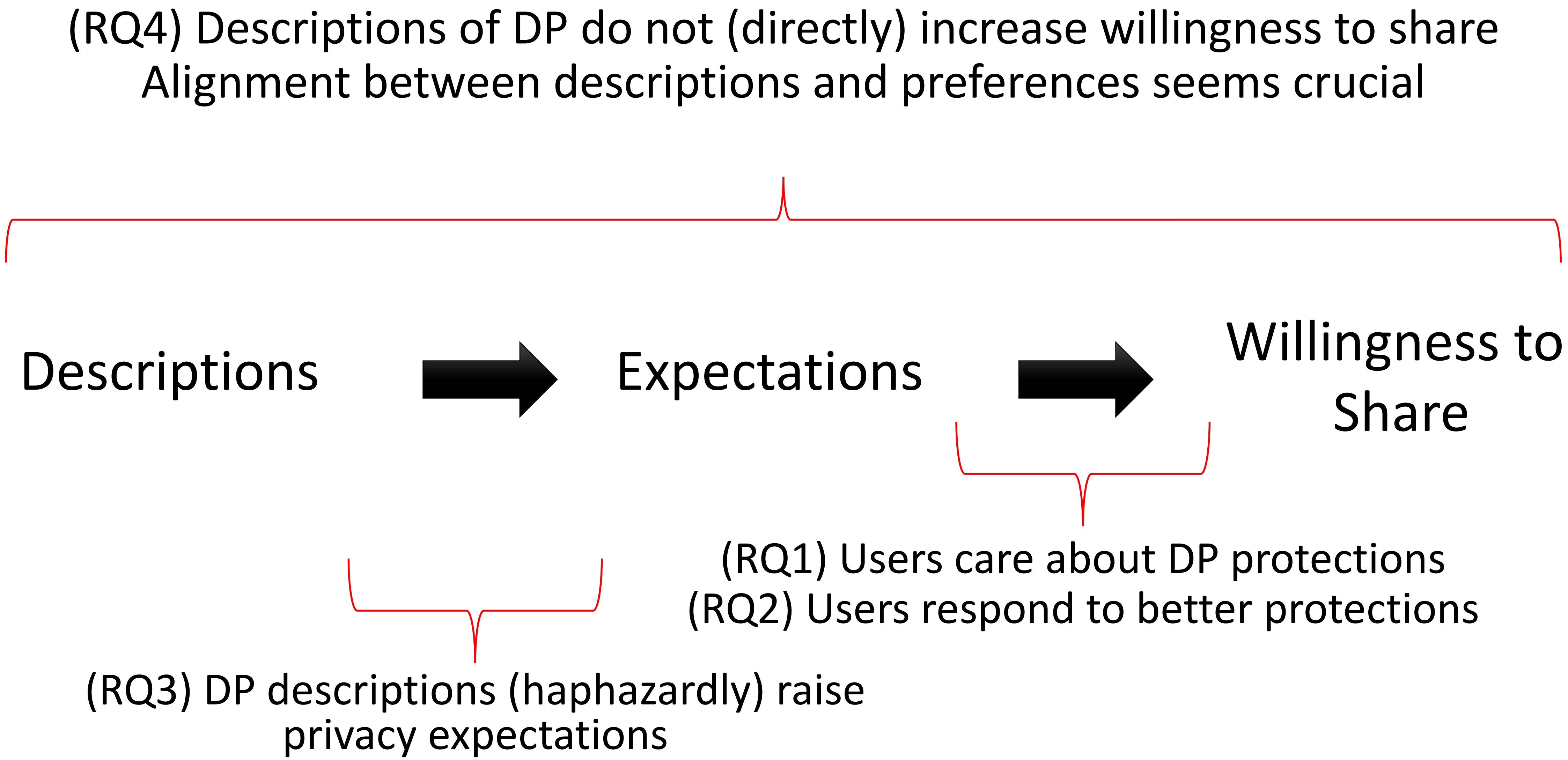
Significantly Increased Sharing

Full Overlap

Description: *Techniques*

User Concerns: 🟢 🔵 ⚪

Description Effects: 🟢 🔵 ⚪



(RQ4) Descriptions of DP do not (directly) increase willingness to share
Alignment between descriptions and preferences seems crucial

Descriptions

Expectations

Willingness to Share

(RQ1) Users care about DP protections
(RQ2) Users respond to better protections

(RQ3) DP descriptions (haphazardly) raise privacy expectations

Thanks!

Gabriel Kaptchuk, Boston University

Big thanks to my collaborators Rachel Cummings and Elissa Redmiles