Module 3: User Privacy and HCI

Privacy for Data Analysis and ML

CS848 Fall 2024

UNIVERSITY OF WATERLOO | DSG Data Systems Group

Logistics

- Project
  - Project ideas has been posted on Learn (this Tue noon)
  - Start brainstorm your project
  - Choose project due is Sep 24
  - Project proposal due is Oct 3

- Paper reading and presentation
  - Site: https://uauw-fall2024privacy.hotcrp.com/
  - Bidding completed (Sep 18)
  - Assignment by this weekend [hotcrp, course website]
  - Start paper review/presentation/discussion in “Legal Privacy” next Thur:
    - *L2:* M. Nouwens, I. Liccardi, M. Veale, D. Karger, and L. Kagal, “*Dark Patterns after the GDPR: Scraping Consent Pop-ups and Demonstrating their Influence: CHI 2020*

Recap

- Module 1: Empirical Privacy
  - Design an algorithmic privacy attack

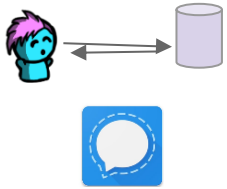
- Module 2: Semantic Privacy
  - Differential privacy (DP)
  - DP primitives
  - DP composition
  - In-class exercise

Consider:

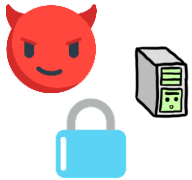
- What is <technical topic of choice>?
- How would you explain it to someone?
- Who do you need to explain it to?
- What do you need to explain to ensure that **it is used correctly**?
- What would you say to give the general intuition of it to <insert curious family member's name here>

DP

Usability



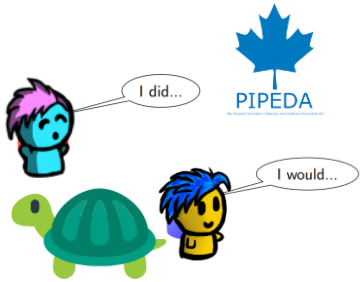
Functionality



Deployability and Verifiability

“Accessibility”

“Efficiency”



Trust and Perceptions

You may already be familiar with a “usability” based design principle

Module 3: User Privacy and HCI

Bailey Kacsmar

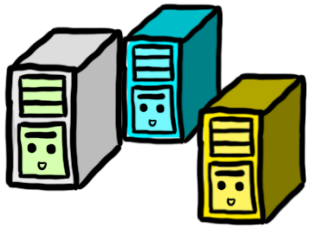
- Why (and how) do we “need” to consider usability? [30 mins]
  - Example: Why Johnny Can’t Encrypt: Usability and PGP
- Usability based analysis [25 mins]
  - Mini-Crash Course on some human research methodologies for CS Students
- Using analysis towards cryptography [45 mins]
  - Example: HCI and PSI
- In-class exercises

Why (and How) do we need to consider Usability?

Example: Why Johnny Can’t Encrypt: Usability and PGP

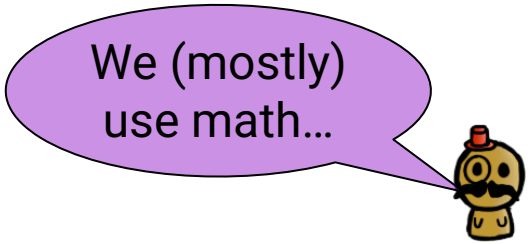
Base Cryptography - Writing “secret” messages

Cryptography for Security and Privacy



Someone wants to complete a task

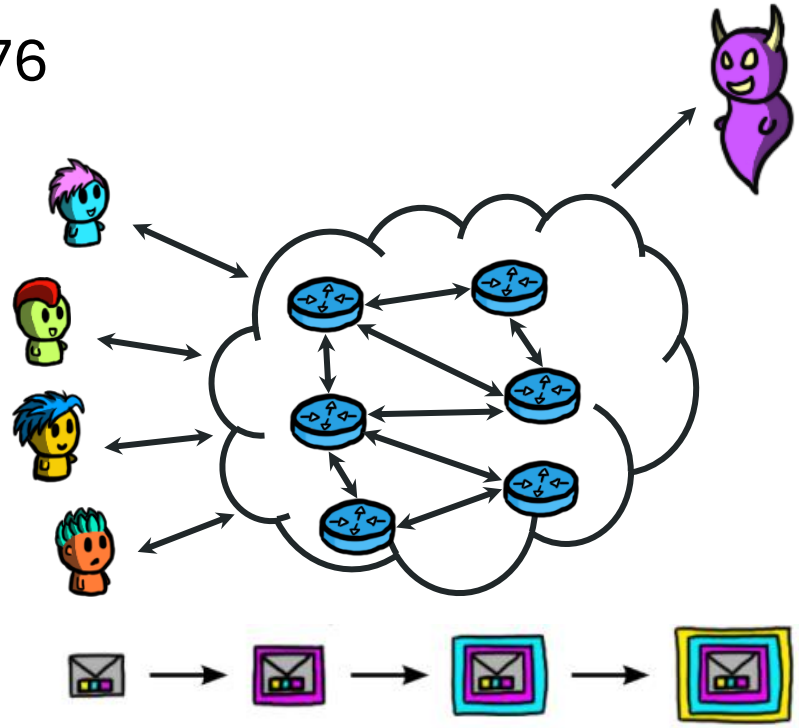
But there are privacy implications and risk from that task



Researchers develop technical solutions

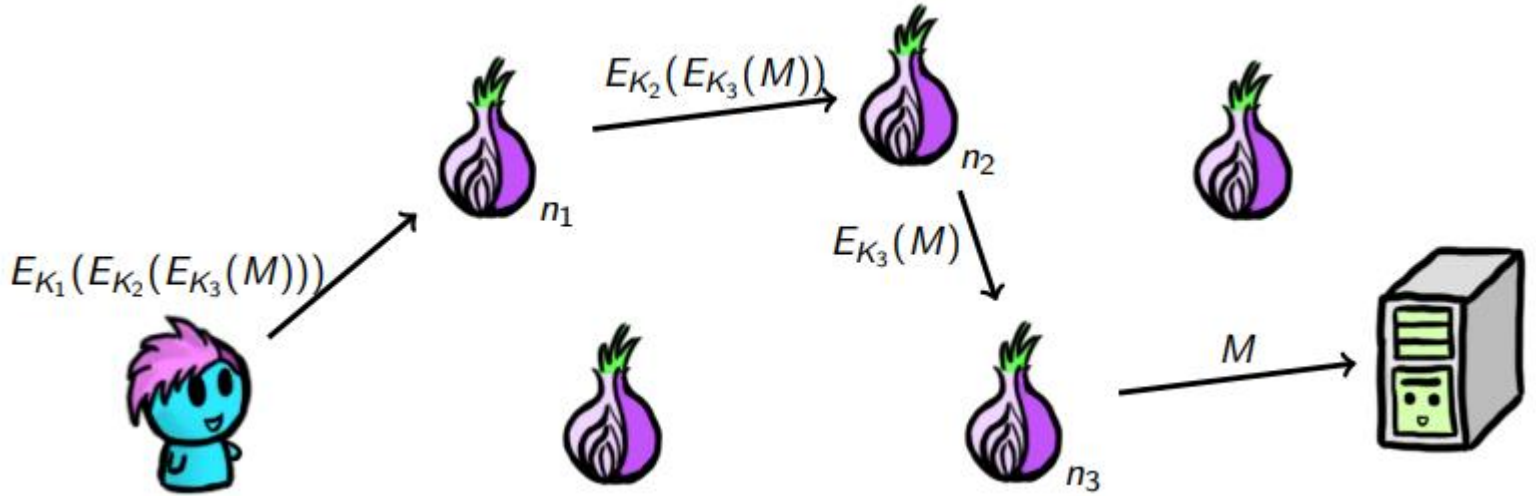
Cryptography for Communications?

- Diffie-Hellman Key Exchange, 1976
- RSA Encryption, 1977
- Shamir secret sharing, 1979
- PGP, Pretty good privacy, 1991
- …



Application Example: Sending Messages with Tor

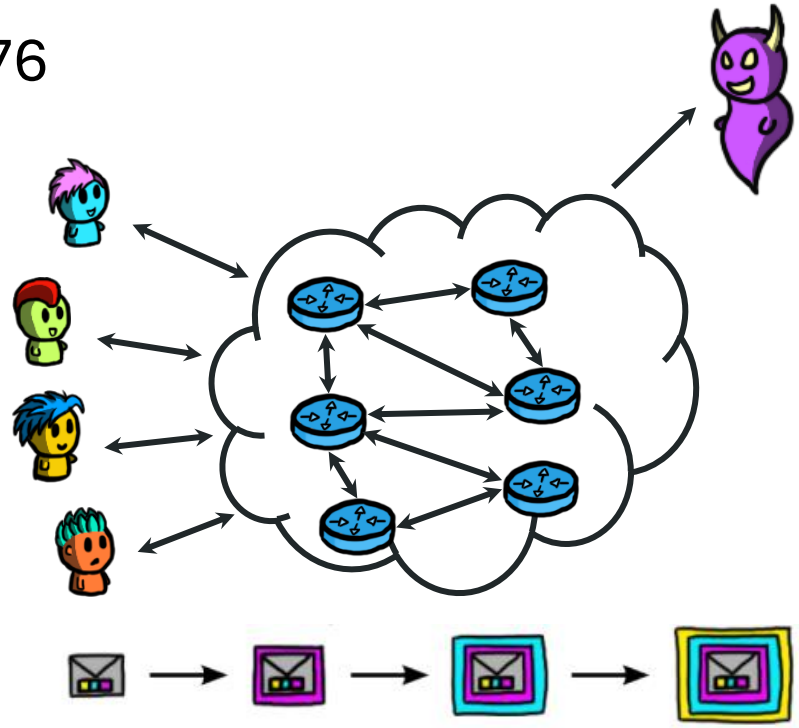
Alice (after many steps of PKC) encrypts her message “like an onion”; each node peels a layer off and forwards it to the next step



If connecting to a web server, M is encrypted (e.g., TLS)

Cryptography for **Everyday**

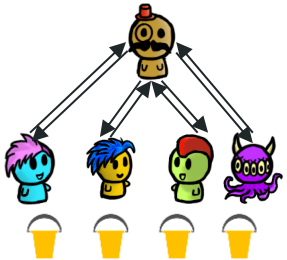
- Diffie-Hellman Key Exchange, 1976
- RSA Encryption, 1977
- Shamir secret sharing, 1979
- PGP, Pretty good privacy, 1991
- …



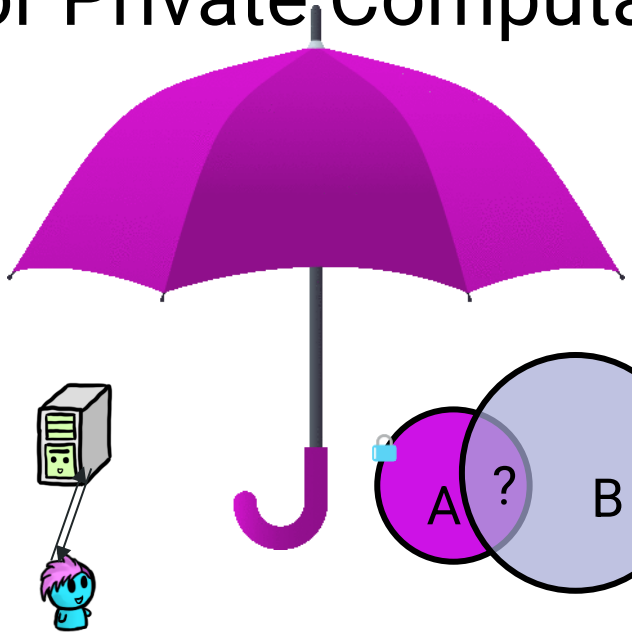
Cryptography for Private Computations

Balancing Privacy and Utility

Cryptography for Private Computations

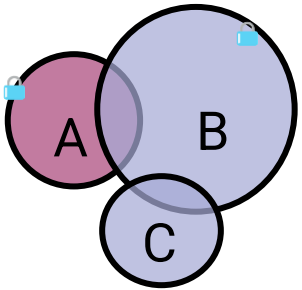


Private Machine Learning



Private Query Processing

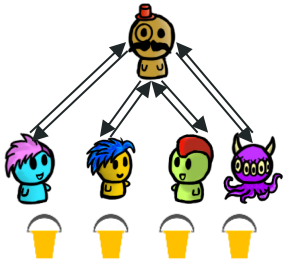
Private Set Intersection



Multiparty Computations

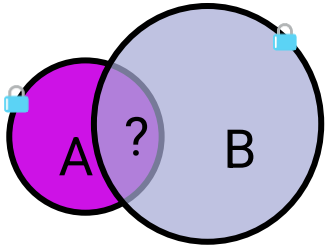
Private Computations Class

Define, **what** is being protected, from **whom**,
and under what **conditions** this protection will hold.

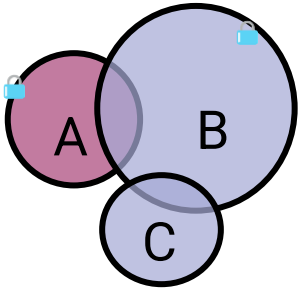


Private Machine Learning

Private Query Processing



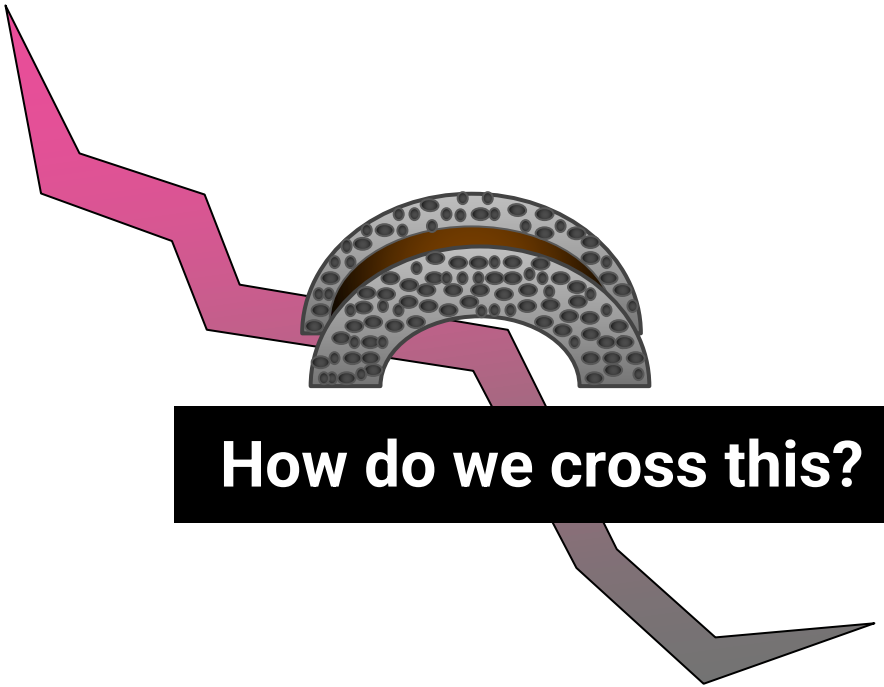
Private Set Intersection



Multiparty Computations

A Tale as Old as Time…

Academic
Cryptography



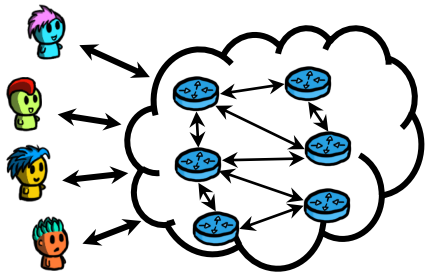
Correctly Deployed
Cryptography

Utility, the Usability Scapegoat

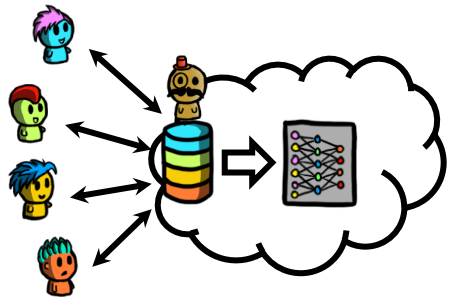
Definition: the benefit that users (and the provider) get from using the system.

Communications system:

- For users: being able to communicate



Data Science:

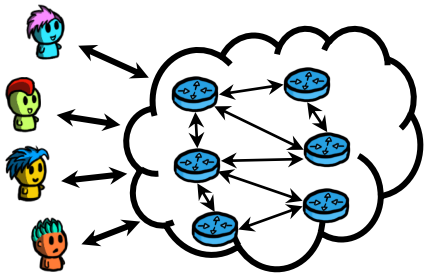


- For participants: maybe they get compensation?
- For data owner: it can sell access to model/analysis for revenue
- Analysts: they pay to get benefits from the model's outputs
- General public: maybe the model outputs are good for society?

Quantifying Utility the Scapegoat

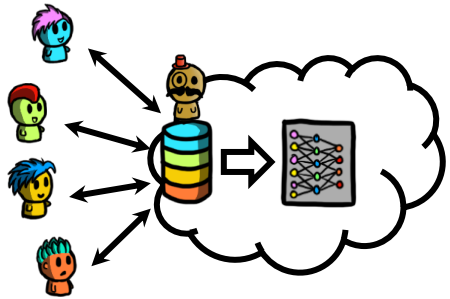
Q: How do we *quantify* utility?

Communications system:



- Low packets dropped
- High bandwidth/throughput
- Low latency/delay…

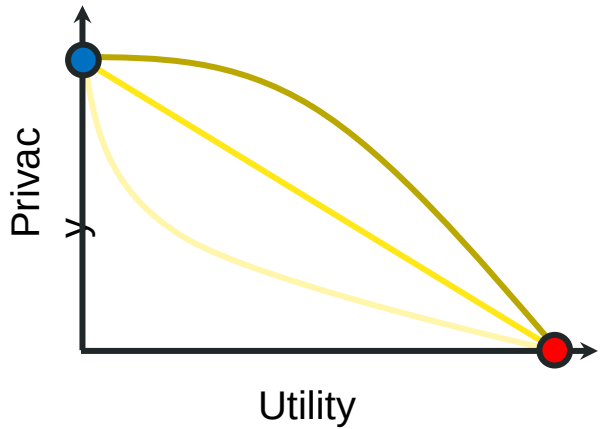
Machine learning:



- Useful model (high test accuracy)
- Unbiased model (low disparity among subpopulations)
- Low computational requirements to build the model
- Fast training algorithm…

The Privacy-Utility trade-off

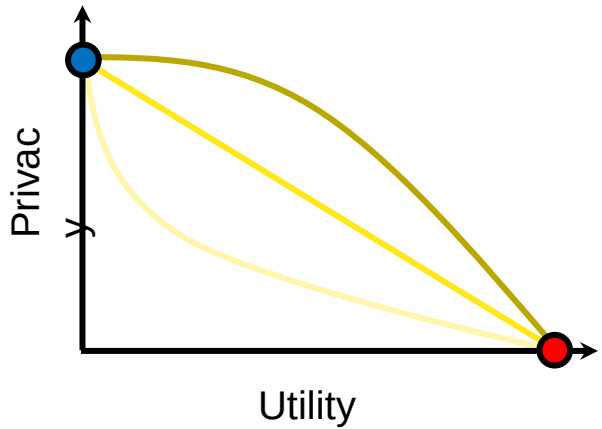
- Given any metric for privacy and for utility, they are usually at odds:



- **Q:** How do you design a system that provides maximum utility?
- **Q:** How do you design a system that provides maximum privacy?
- Designing a system that provides a good privacy-utility trade-off is hard!

The Privacy-Utility trade-off

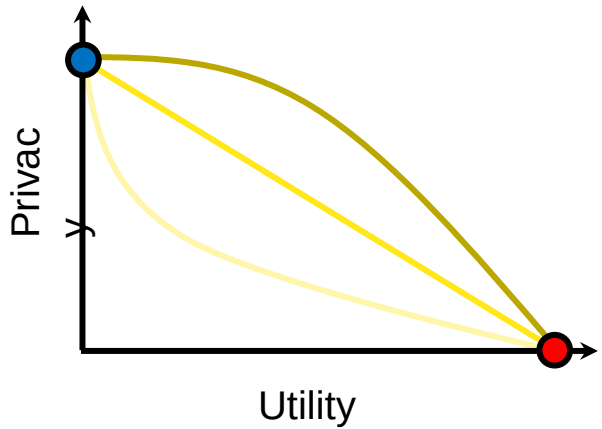
- Given any metric for privacy and for utility, they are usually at odds:



- How do you design a system that provides maximum utility?
  - You design it without privacy in mind
- How do you design a system that provides maximum privacy?
  - ..?
- Designing a system that provides a good privacy-utility trade-off is hard!

The Privacy-Utility trade-off

- Given any metric for privacy and for utility, they are usually at odds:



- How do you design a system that provides maximum utility?
  - You design it without privacy in mind
- How do you design a system that provides maximum privacy?
  - You don’t design it
- Designing a system that provides a good privacy-utility trade-off is hard!

The Entanglement, Beyond Utility Alone

Cryptography for privacy or even security is entangled **with humans**

Beyond Data the Abstraction

Google and Mastercard Cut a Secret Ad Deal to Track Retail Sales

Google found the perfect way to link online ads to store purchases: credit card data

By Mark Bergen and Jennifer Surane
August 30, 2018, 3:43 PM EDT *Updated on August 31, 2018, 12:40 PM EDT*

washingtonpost.com

Now for sale: Data on your mental health

Drew Harwell

These retailers share customer data with Facebook's owner. Customers may not have been told | CBC News

Thomas Daigle · CBC News · Posted: Feb 07, 2023 4:00 AM EST | Last

Home Depot didn't get customer consent before sharing data with Facebook's owner, privacy watchdog finds | CBC News

Catharine Tunney · CBC News · Posted: Jan 26, 2023 9:53 AM
Updated: January 27

Double-double tracking: How Tim Hortons knows where you sleep, work and vacation

James McLeod ✉ June 15, 2020 In : Canada Privacy 0 1,169 11 min read

Beyond Data the Abstraction

Google and Mastercard … Deal to Tra… …ental

Google found the … card data

By Mark Bergen and Jennife…
August 30, 2018, 3:43 PM ED…

Home Depot … consent befo… Facebook's o… finds | CBC Ne…

Catharine Tunney · CBC Ne…
Updated: January 27

ADOBE / CREATORS / TECH

Adobe's new terms of service aren't the problem – it's the trust

/ The reaction from Adobe's customers to a small update highlights the growing lack of faith surrounding big tech companies and their AI tools.

Creatives are fearful of how Adobe's adoption of generative AI will impact their privacy and rights over their work. Illustration by Haein Jeong / The Verge

By Jess Weatherbed, a news writer focused on creative industries, computing, and internet culture. Jess started her career at TechRadar, covering news and hardware reviews.

Jun 7, 2024, 1:37 PM MDT

11 Comments (11 New)

If you buy something from a Verge link, Vox Media may earn a commission. See our ethics statement.

…and Vacation

…ne 15, 2020 In : Canada Privacy 0 1,169 11 min read

Utility?

Communication?

Accessibility?

Usability?

Computation?

Hardware?

Intuition?

What does usability mean for cryptography???

This Security Trope…

People are the **weakest link** in the chain

Reject this Security Trope

People are the **weakest link** in the chain

– but it is **not that simple**, nor is that fair

Why Johnny Can’t Encrypt - 1999

Set the stage:

- We have crypto...
- We have crypto tools...

Whitten and Tygar. "Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0." USENIX Security Symposium. 1999.

Why Johnny Can’t Encrypt - 1999

Set the stage:

- We have crypto…
- We have crypto tools…
- BUT, they’re **not really being used**…
  (by non-cryptographers)

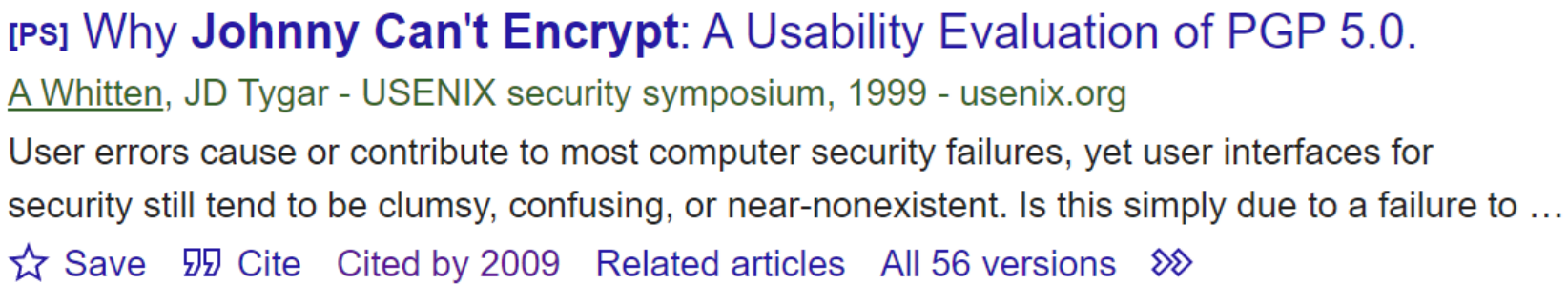
Whitten and Tygar. "Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0." USENIX Security Symposium. 1999.

Why Johnny Can’t Encrypt - 1999

Set the stage:

- We have crypto…
- We have crypto tools…
- BUT, they’re **not really being used**…

  (by non-cryptographers)



Whitten and Tygar

Why Johnny Can’t Encrypt - 1999

Set the stage:

- We have crypto…
- We have crypto tools…
- BUT, they’re not really being used…(by non-cryptographers)

Only a handful of related work…

Whitten and Tygar. "Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0." USENIX Security Symposium. 1999.

Why Johnny Can’t Encrypt - 1999

Set the stage:

- We have crypto…
- We have crypto tools…
- BUT, they’re not really being used…(by non-cryptographers)

Only a handful of related work…

Only one notion of usability across them…

Whitten and Tygar. "Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0." USENIX Security Symposium. 1999.

Why Johnny Can’t Encrypt - 1999

Set the stage:

- We have crypto…
- We have crypto tools…
- BUT, they’re not really being used…(by non-cryptographers)

Only a handful of related work…

Only one notion of usability across them…

“Usability necessarily has different meanings in different contexts”

Whitten and Tygar. "W

Usability - 1999

“Usability necessarily has different meanings in different contexts”

“For some, **efficiency may be a priority**, for others, learnability, for still others, flexibility. In a security context, our priorities must be whatever is needed in order for the security to be used effectively.”

Whitten and Tygar. "Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0." USENIX Security Symposium. 1999.

Usability - 1999

“Usability necessarily has different meanings in different contexts”

“For some, efficiency may be a priority, for others, **learnability**, for still others, flexibility. In a security context, our priorities must be whatever is needed in order for the security to be used effectively.”

Whitten and Tygar. "Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0." USENIX Security Symposium. 1999.

Usability - 1999

“Usability necessarily has different meanings in different contexts”

“For some, efficiency may be a priority, for others, learnability, for still others, **flexibility**. In a security context, our priorities must be whatever is needed in order for the security to be used effectively.”

Whitten and Tygar. "Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0." USENIX Security Symposium. 1999.

Usability - 1999

“Usability necessarily has different meanings in different contexts”

“For some, efficiency may be a priority, for others, learnability, for still others, flexibility. **In** a **security** context, our priorities must be **whatever is needed in order for the security to be used effectively.**”

Whitten and Tygar. "Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0." USENIX Security Symposium. 1999.

Definition (1999)

Security software is usable if the people who are expected to use it:

- are reliably made aware of the security tasks they need to perform
- are able to figure out how to successfully perform those tasks
- don’t make dangerous errors
- are sufficiently comfortable with the interface to continue using it

Whitten and Tygar. "Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0." USENIX Security Symposium. 1999.

Definition (1999)

Security software is usable if the people who are expected to use it:

- are reliably made aware of the security tasks they need to perform
- are able to figure out how to successfully perform those tasks
- don’t make dangerous errors
- are sufficiently comfortable with the interface to continue

How can we improve this?

Whitten and Tygar. "Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0." USENIX Security Symposium. 1999.

Challenges (1999)

Claim: Security has some inherent properties that make it a difficult problem domain for user interface design.

Whitten and Tygar. "Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0." USENIX Security Symposium. 1999.

Challenges (1999)

Claim: Security has some inherent properties that make it a difficult problem domain for user interface design.

What do you think they are (were)?

Whitten and Tygar. "Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0." USENIX Security Symposium. 1999.

Challenges (1999)

Claim: Security has some inherent properties that make it a difficult problem domain for user interface design.

- The unmotivated user property
- The abstraction property
- The lack of feedback property
- The barn door property
- The weakest link property

Whitten and Tygar. "Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0." USENIX Security Symposium. 1999.

Challenges (1999)

Claim: Security has some inherent properties that make it a difficult problem domain for user interface design.

- The unmotivated user property
- The abstraction property
- The lack of feedback property
- The barn door property
- The weakest link property

Task: make computer security usable for people who are not already knowledgeable in that area

Whitten and Tygar. "Why Johnny can't Encrypt: A Usability Evaluation of PGP 5.0." USENIX Security Symposium. 1999.

(Many) Descendents and Branches after Johnny

Finally johnny can encrypt: But does this make him feel more secure?
N Gerber, V Zimmermann, B Henhapl… - Proceedings of the 13th …, 2018 - dl.acm.org
… of E2E **encryption** by non-experts in the email context. An oftenquoted example is the paper '… **Johnny can't encrypt**' [33] as well as subsequent studies on the usability of E2E **encryption** …
☆ Save 99 Cite Cited by 34 Related articles All 4 versions

Teaching Johnny not to fall for phish
P Kumaraguru, S Sheng, A Acquisti, LF Cranor… - ACM Transactions on …, 2010 - dl.acm.org
Phishing attacks, in which criminals lure Internet users to Web sites that spoof legitimate Web sites, are occurring with increasing frequency and are causing considerable harm to victims…
☆ Save 99 Cite Cited by 563 Related ar

Leading Johnny to water: Designing for usability and trust
E Atwater, C Bocovich, U Hengartner, E Lank… - … Symposium On Usable …, 2015 - usenix.org
Although the means and the motivation for securing private messages and emails with strong end-to-end encryption exist, we have yet to see the widespread adoption of existing …
☆ Save 99 Cite Cited by 76 Related articles All 3 versions »

Branches Following Engineering Style Challenges

“PGP 5.0 alerts its users to this compatibility issue…it uses different icons to depict the different key types…”

- NIST (and other) standardization processes
- Tools, libraries, etc…
- Improving intuition of icons (browsers, mobile…)

L. Cranor, Potty Talk at PEPR 2021, https://youtu.be/s4cS5Tqnemo

Branches Following the Visual Metaphors

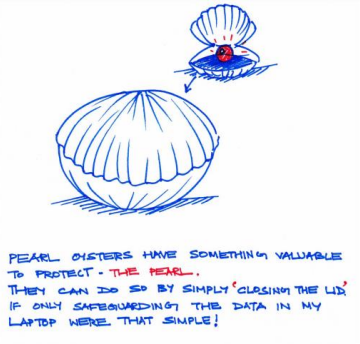


Fig. 62. "Pearl oysters have something valuable to protect - the pearl. They can do so by simply 'closing the lid.' If only safeguarding the data in my laptop were that simple!" By Sharon, age 25.

Fig. 33. "Privacy means that the thoughts in my brain are locked away. What I know does not have to go into the world, which I put an X over." By Thomas, age 19

Fig. 23. "This is me enjoying my privacy. This is the only time during the day, were I am truly alone and nothing bothers me. No man no children no dogs." By Cindy, age 54

Fig. 24. "No one come in when I am in the bathroom!" By Sydney, age 7

M. Oates, et al. Turtles, locks, and bathrooms: Understanding mental models of privacy through illustration." Proceedings on Privacy Enhancing Technologies 2018.

The Branches Towards Usable Cryptography

- Ceremony analysis
- (Novel and Nuanced) threat models
- Human Computer Interaction (HCI) studies
- Software engineering (tooling)

The Principle of Psychological Acceptability

“ It is essential that the human interface be **designed for ease of use**, so that users routinely and automatically **apply the protection mechanisms correctly**.”

- Jerome Saltzer and Michael Schroeder

J. Saltzer and Michael Schroeder. “The Protection of Information in Computer Systems”, Proceedings of the IEEE 63:0. 1975

Important

Theoretical Cryptography?

Applied Cryptography?

Deployable Cryptography?

Question the Assumptions of the Motivation

Private set intersection as “good” for:

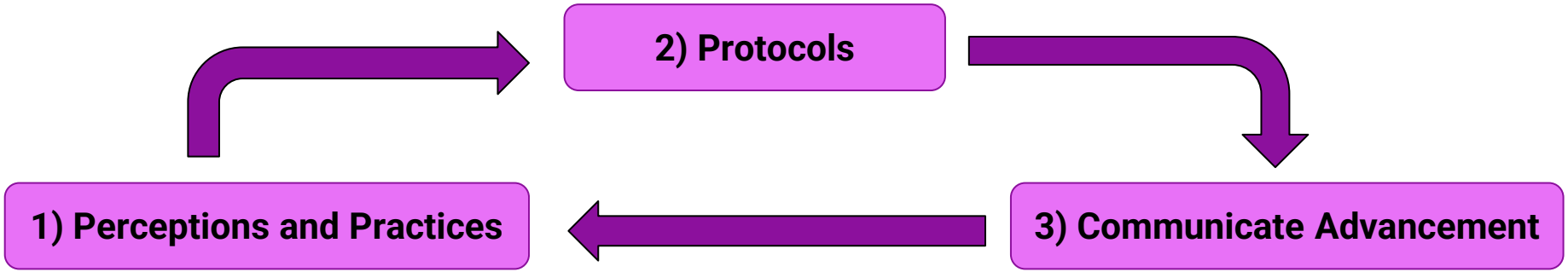
- Ad conversion
- Security incident information sharing
- Contact discovery

Pattern of the claims made:

- Just send it (bad)
- Just hash it (bad)
- Just PSI this (good)

We can do better

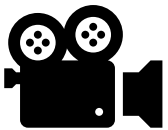
Human-Centered Design



"…that aims to make systems usable and useful by **focusing on the users, their needs and requirements**, … counteracts possible adverse effects of use…" - ISO 9241-210:2019(E)

Usability based Analysis

Mini-Crash Course on some human research methodologies for CS Students



The slides in this crash course section are derived from instructional material from Dr. C. Demmans-Epp

Predominant Methodologies

Quantitative

- Focus on testing theories and hyp.
- Analyzed through math and stats.
  - Descriptive analyses
  - Correlational analyses
  - Inferential analyses (testing)
- (most) Numbers, graphs and tables
- Requires an appropriate # of resp.
  - The number depends on what you are trying to measure or test
- Closed (multiple choice) questions, measures, observation

Qualitative

- Focus on exploring ideas and formulating a theory or hyp.
- Analyzed by summarizing, categorizing, and interpreting
- Mainly expressed in words
  - Rich descriptions are important
  - Alternative representations include graphics and art work (e.g., plays)
- Can require few respondents
- Open-ended questions, observation

Design-based Research

- A form of (mostly)qualitative research that aims to iteratively improve processes/artefacts

Predominant Methodologies – Key Terms

Quantitative

- Objectivity
- Testing
- Measurement
  - Central tendency (e.g., M, Mdn)
  - Variability (e.g., SD, IQR, Min, Max)
- Validity
- Replicability: someone can do the same thing themselves
- Reproducibility: someone gets the same results using the original researcher's data and analysis procedures

Qualitative

- Subjectivity
  - Positionality
- Understanding
- Complexity
- Context
  - Thick descriptions
  - Common views
  - Dissenting or other views
- Replicability
  - Some prefer to call it methodological accounting

Mixed Methods

- The world requires more complex views that combine approaches from qualitative and quantitative methodologies
- Will be biased towards either a qualitative or a quantitative methodology
    - Methods or techniques from the sub-ordinate methodology will be used to support the dominant one.
    - e.g., qualitative methods can be used to explain quantitative results (mixed-methods explanatory design)
- Not all fields agree that mixed methods are real

All Methods Are Limited and Provide Opportunities

- Methods enable and limit evidence
- All are valuable when used appropriately
- All have weaknesses or limitations
- You can combine multiple methods
  - to offset or mitigate their weaknesses
  - select them so that the strength of one method will address the weakness of another method
  - e.g., log files only tell you what a user did and cannot tell you why so you can combine their analysis with questionnaire, interview, or think-aloud data to understand why certain actions were taken

Things to Consider when Reading Research

- Are the methods appropriate to what is being studied?
  - What strengths or weaknesses exist?
  - Have they met the major quality criteria for the method chosen?
- Does the paper acknowledge the strengths and weaknesses of the methods employed?
- Is the research evidence based on only a single evaluation method?

Beliefs About Evidence

"Credible empirical knowledge requires convergence of evidence across studies based on different methods."

To enhance credibility, we try to maximize:

- Evidence **generalizability**
- Measurement **precision**
- **Control over extraneous factors** that are not under investigation
- **Realism** of the situation or context within which we gather evidence

Large samples do not give you generalizability – Generalizability comes from study design and sampling procedures

Methods for Learning About Users & Designing

- Interviews
- Observation
- Questionnaires
- Analyse their tasks
- Research

- Have them help you design the software
- Have them try to use early prototypes
  - See if they can complete specific tasks
  - Have them “think aloud” while using the system

Questionnaires & Scales

- Use these to quickly collect
  - Perceptual data
  - Demographic data
- Often quantifiable
- Reuse others' instruments where possible
  - They may need adjustment
  - They may not apply to your context, in which case they need additional validation
  - Report measured reliability
- Give non-response, "other", N/A response options
  - Sensitive topics: Gender, ethnicity, race, …
  - Things people may not have done or used
- Rating scale selection
  - Forced Choice
  - Neutral response: 5 or 7 items
- Include at least one open-ended item

Interviews

- Unstructured
  - Scriptless
  - Open-ended
  - Rich but not replicable
- Structured
  - Tightly scripted
  - Often like a questionnaire
  - Replicable but may lack richness
  - Cognitive interviewing
- Semi-structured
  - Guided by a script
  - Interesting issues can be explored in more depth
  - Balance b/w richness and replicability
- Focus on their EXPERIENCES
  - Ask them for examples
  - Ask them to tell you a story of when they…

Semi-Structured Interviews: Example

CONTEXT: A study of mobile use for supporting learning English as an additional language (EAL) and supporting EAL learner communication.

GOALS: General approach and use of a specific mobile application (i.e., MyVoice)

- What has your experience with learning languages been like? — **General Background**
- What has your experience with technology been like?
- Before using MyVoice, what was your experience with using computer programs and mobile devices (e.g., iPhone, Android, iPad) for language learning like? — **Specific Background**
- What would you like to see added to these programs and technologies to make learning easier for you?
- What has your experience with using MyVoice for language learning been like? — **The Application**

Semi-Structured Interviews: Example (cont.)

- What has your experience with learning languages been like?
  - Which languages have you tried to learn? Why?
  - What is a typical day like for you in that language?
  - How did you go about learning the language?
  - What types of things help you with learning languages? Why?
  - What tools and strategies did you use? Why?
  - How did they help/frustrate you?
  - Was there anything that you felt was missing that might have been helpful to you?
- What has your experience with technology been like?
  - What technologies have you used? (computer, mobile phone, VCR, TV, robot, ...)
  - Where did you use that technology? (home, the library, work, ...)
  - How did you use that technology?
  - What does that technology let you do that you couldn't do before?
  - What does that technology prevent/stop you from doing?
  - What does it make easier/harder?
  - Why do you keep using that technology?
  - Do you have an example of when you liked using it? What happened?
  - Do you have an example of when you hated using it? What happened?

Interviews

- Take detailed notes
  - Possibly check them with participant
- Record and transcribe
  - Member-checking: check with the participant later to make sure you interpreted things properly
- Make your participant comfortable
- Do not judge
- Ask them to
  - Provide examples
  - Tell you a story about when it happened

Interviews – General Guidelines

- Show your gratitude — and be clear about what is (and isn't) being tested
- Assume your interviewee is in an uncomfortable situation
  - Develop a bit of a relationship with the user: this means SHARING and listening
- Pay attention to their behaviour and reflect it back to them
- Prioritize open-ended questions
  - Be Socratic: pretend you know nothing and have them explain it to you
- Be quiet
- Confirm interpretations
- Save demographics for the end or collect them well in advance

Qualitative Methods - Saturation

- A key component of rigor in qualitative work
- Basically, when new data is expected to add no new insights
  - When little in your code book changes following the addition of data from one more unit
  - When you start to only see things that you have seen before
  - When the amount of insight gained by each new unit starts to decline
- It can be reached in
  - As little as 3-6 interviews
  - Often reached within 12 interviews or 4-8 focus groups

Resources for Methods and Statistics

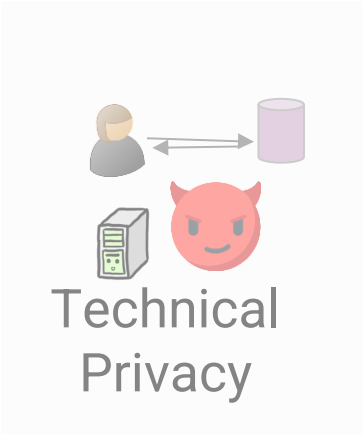
- Stats: http://yatani.jp/HCIstats/HomePage and https://online.stat.psu.edu/stat501/
- Reporting
  - Twining, P., Heller, R. S., Nussbaum, M., & Tsai, C.-C. (2017). Some guidance on conducting and reporting qualitative studies. *Computers & Education*, *106*(Supplement C), A1–A9. https://doi.org/10.1016/j.compedu.2016.12.002
  - López, X., Valenzuela, J., Nussbaum, M., & Tsai, C.-C. (2015). Some recommendations for the reporting of quantitative studies. *Computers & Education*, *91*(Supplement C), 106–110. https://doi.org/10.1016/j.compedu.2015.09.010
  - Joelle Pineau's Checklist: https://www.cs.mcgill.ca/~jpineau/ReproducibilityChecklist.pdf
- Mixed Methods
  - Leech, N. L., & Onwuegbuzie, A. J. (2007). A Typology of Mixed Methods Research Designs. *Quality & Quantity*, *43*(2), 265–275. https://doi.org/10.1007/s11135-007-9105-3 (relatively accessible)
  - Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed). SAGE Publications.

Using analysis towards cryptography

Another example: finding design failures -- HCI and PSI

Kacsmar, Duddu, Tilbury, Ur, and Kerschbaum. Comprehension from Chaos: Towards Informed Consent for Private Computation. 2*023 ACM SIGSAC Conference on Computer and Communications Security* (CCS).

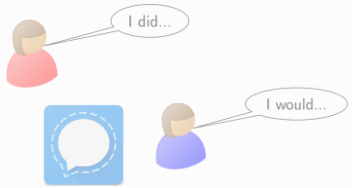
A Wider View of Technical Privacy



Technical Privacy

Conceptual Privacy

Legal Privacy



Usable Privacy

Understanding privacy notions and behaviours, **right to privacy**, and privacy expectations

M. Oates, et al. Turtles, locks, and bathrooms: Understanding mental models of privacy through illustration." Proceedings on Privacy Enhancing Technologies 2018.

Cryptography from Research Papers to Products

- What **steps** are involved in adopting cryptography, and who are the **relevant stakeholders**?
- What are the **key obstacles** hindering the widespread **adoption** and **correct use** of cryptography?
- What are potential ways to **overcome** these obstacles?

K. Fischer, I. Trummová, P. Gajland, Y. Acar, S. Fahl, & A. Sasse. “The Challenges of Bringing Cryptography from Research Papers to Products: Results from an Interview Study with Experts”. Usenix Security Symposium 2024

A Path from Research Papers to Products

1. Algorithm and Protocol Development
2. Standardization
3. Secure Implementation (Cryptography Libraries)
4. Product Development
5. Adoption and Use of Cryptographic Products

K. Fischer, I. Trummová, P. Gajland, Y. Acar, S. Fahl, & A. Sasse. “The Challenges of Bringing Cryptography from Research Papers to Products: Results from an Interview Study with Experts”. Usenix Security Symposium 2024

A Visualization of the Cryptography Ecosystem

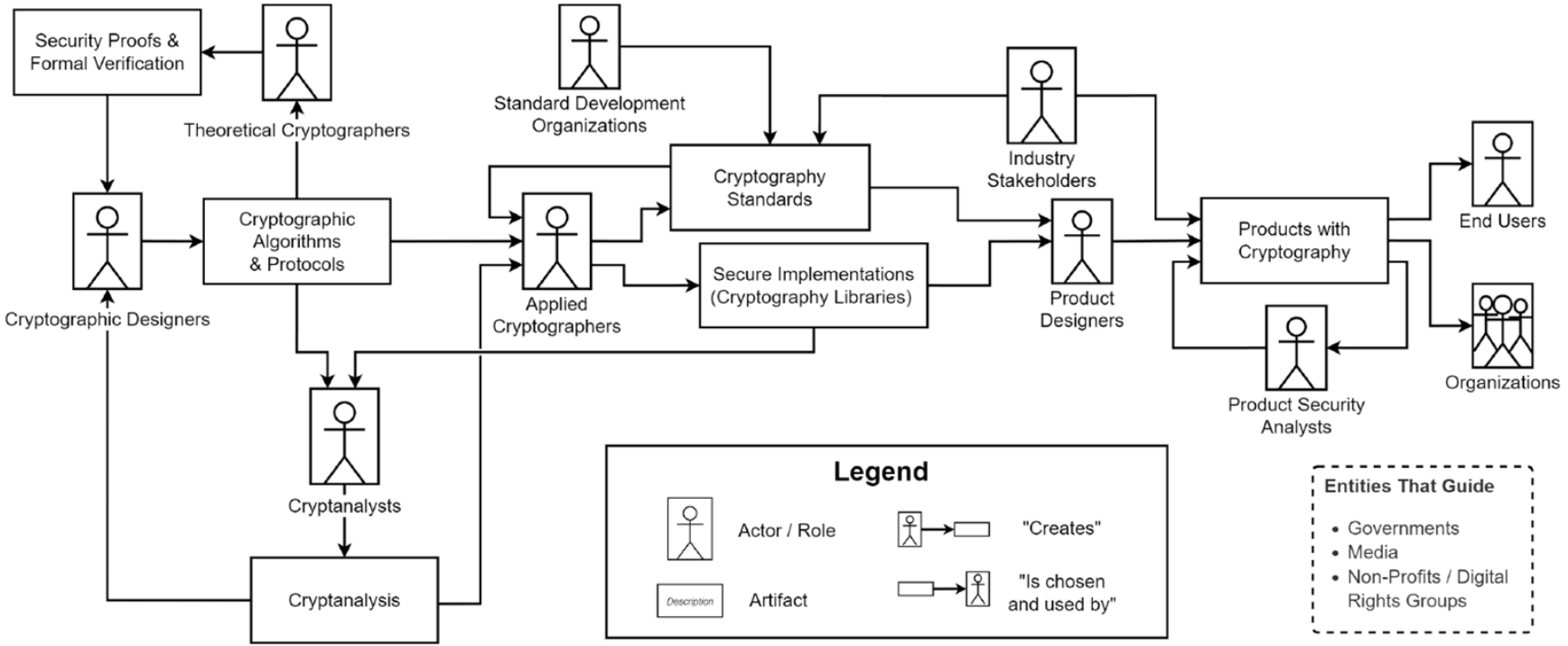


Figure 2 from: K. Fischer, I. Trummová, P. Gajland, Y. Acar, S. Fahl, & A. Sasse. "The Challenges of Bringing Cryptography from Research Papers to Products: Results from an Interview Study with Experts". Usenix Security Symposium 2024

A Visualization of the Cryptography Ecosystem

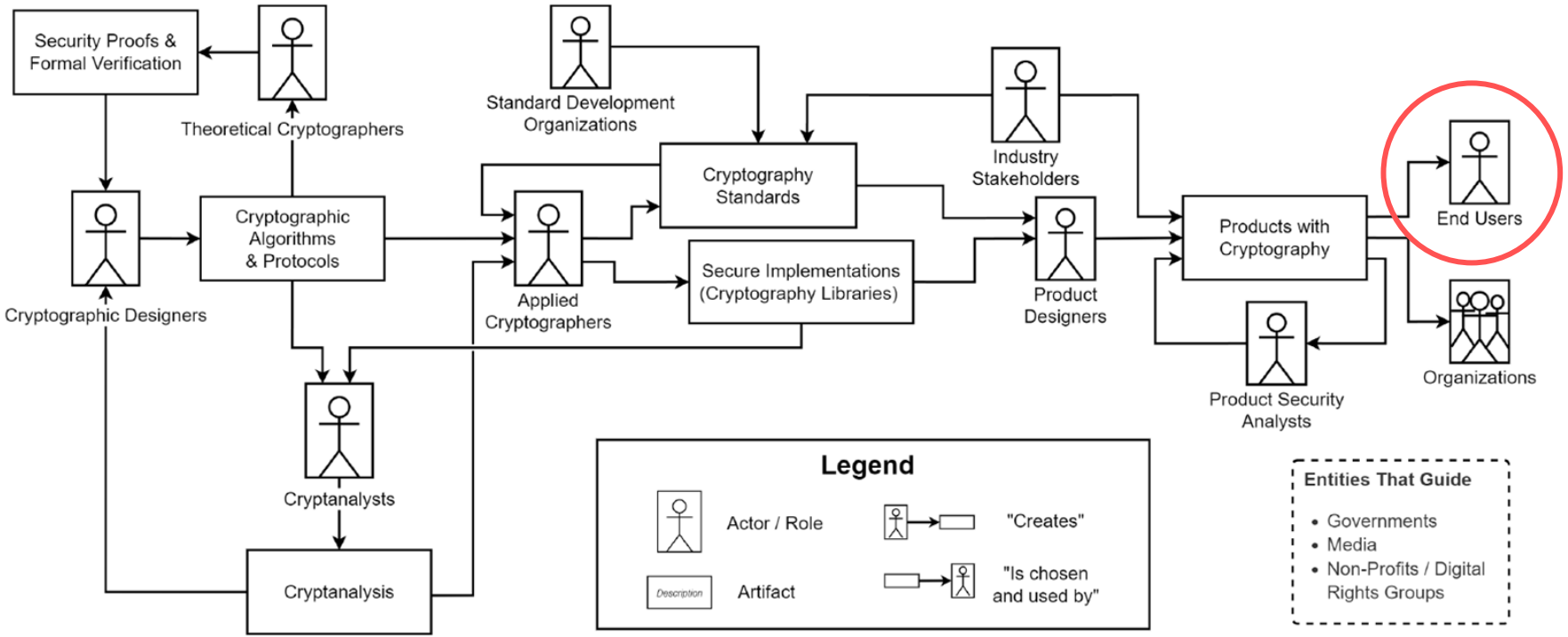


Figure 2 from: K. Fischer, I. Trummová, P. Gajland, Y. Acar, S. Fahl, & A. Sasse. "The Challenges of Bringing Cryptography from Research Papers to Products: Results from an Interview Study with Experts". Usenix Security Symposium 2024

A Visualization of the Cryptography Ecosystem

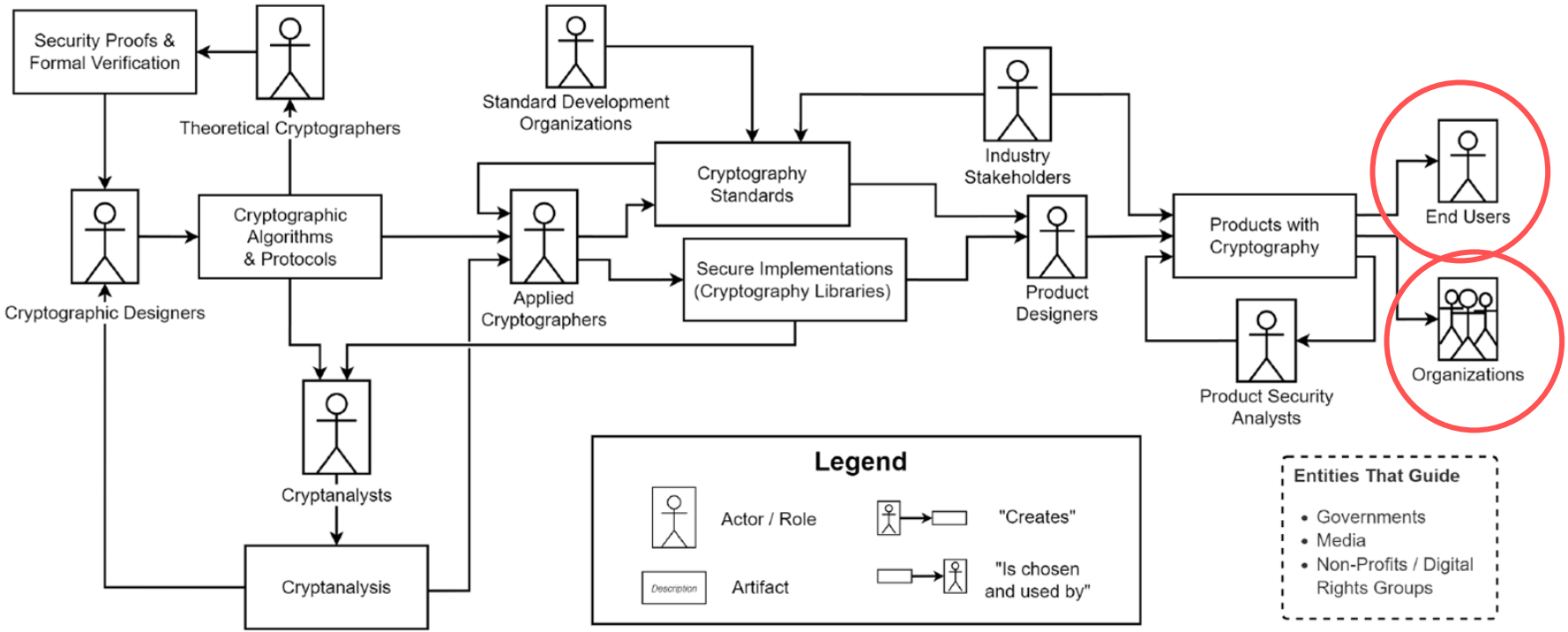


Figure 2 from: K. Fischer, I. Trummová, P. Gajland, Y. Acar, S. Fahl, & A. Sasse. "The Challenges of Bringing Cryptography from Research Papers to Products: Results from an Interview Study with Experts". Usenix Security Symposium 2024

A Visualization of the Cryptography Ecosystem

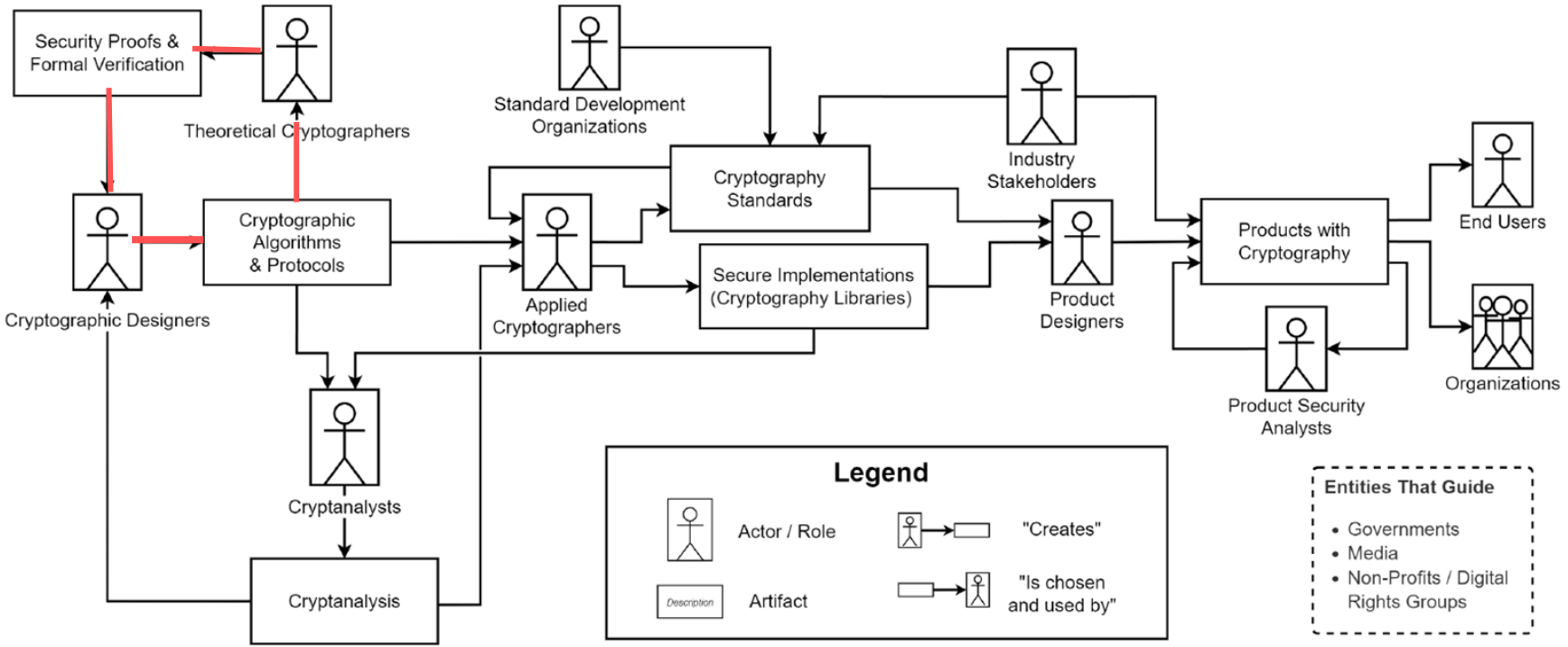


Figure 2 from: K. Fischer, I. Trummová, P. Gajland, Y. Acar, S. Fahl, & A. Sasse. "The Challenges of Bringing Cryptography from Research Papers to Products: Results from an Interview Study with Experts". Usenix Security Symposium 2024

A Visualization of the Cryptography Ecosystem

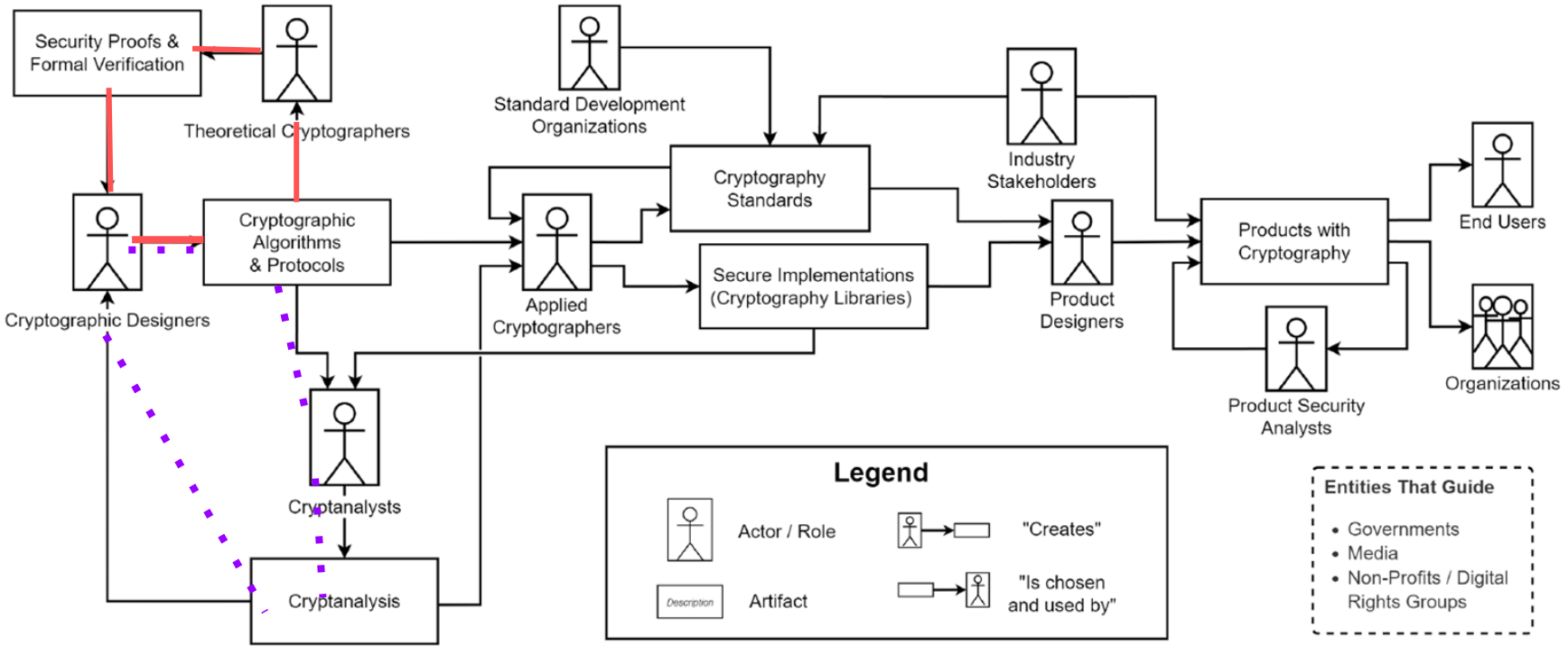


Figure 2 from: K. Fischer, I. Trummová, P. Gajland, Y. Acar, S. Fahl, & A. Sasse. "The Challenges of Bringing Cryptography from Research Papers to Products: Results from an Interview Study with Experts". Usenix Security Symposium 2024

A Visualization of the Cryptography Ecosystem

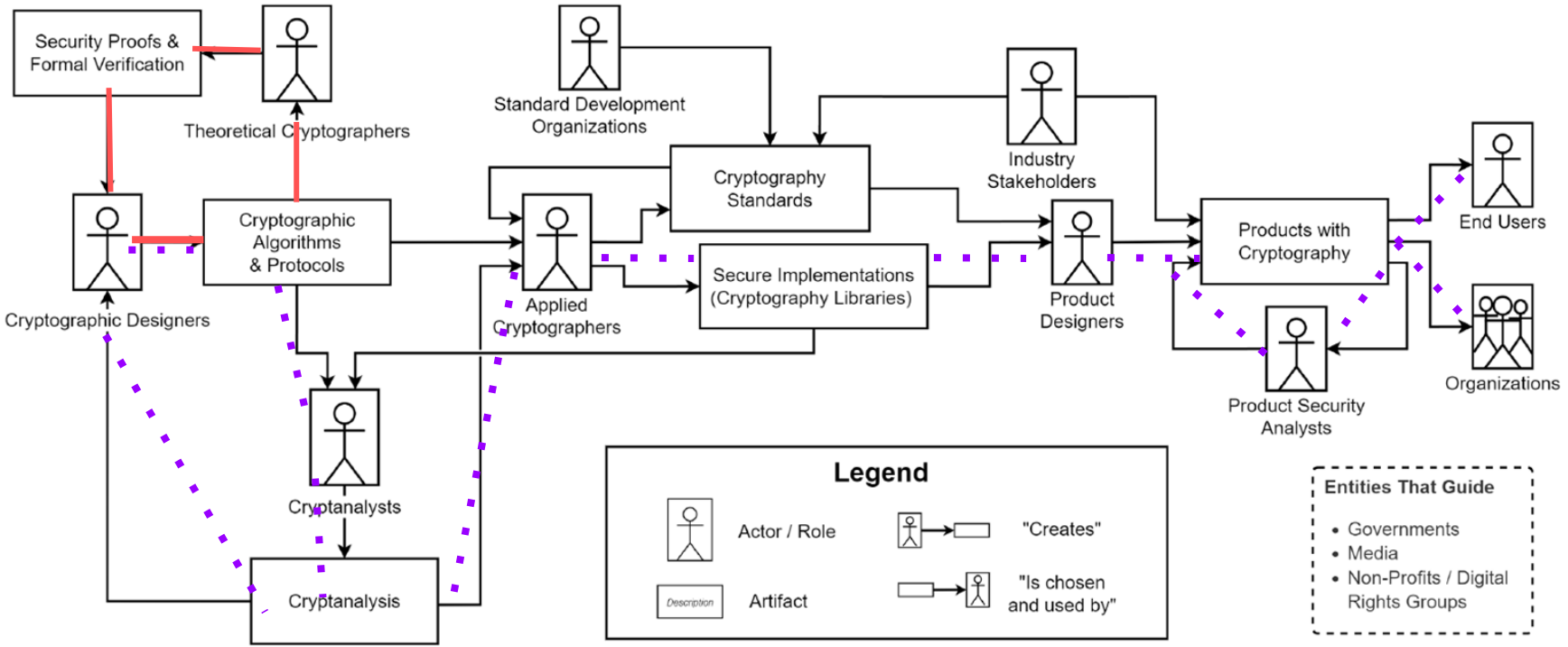
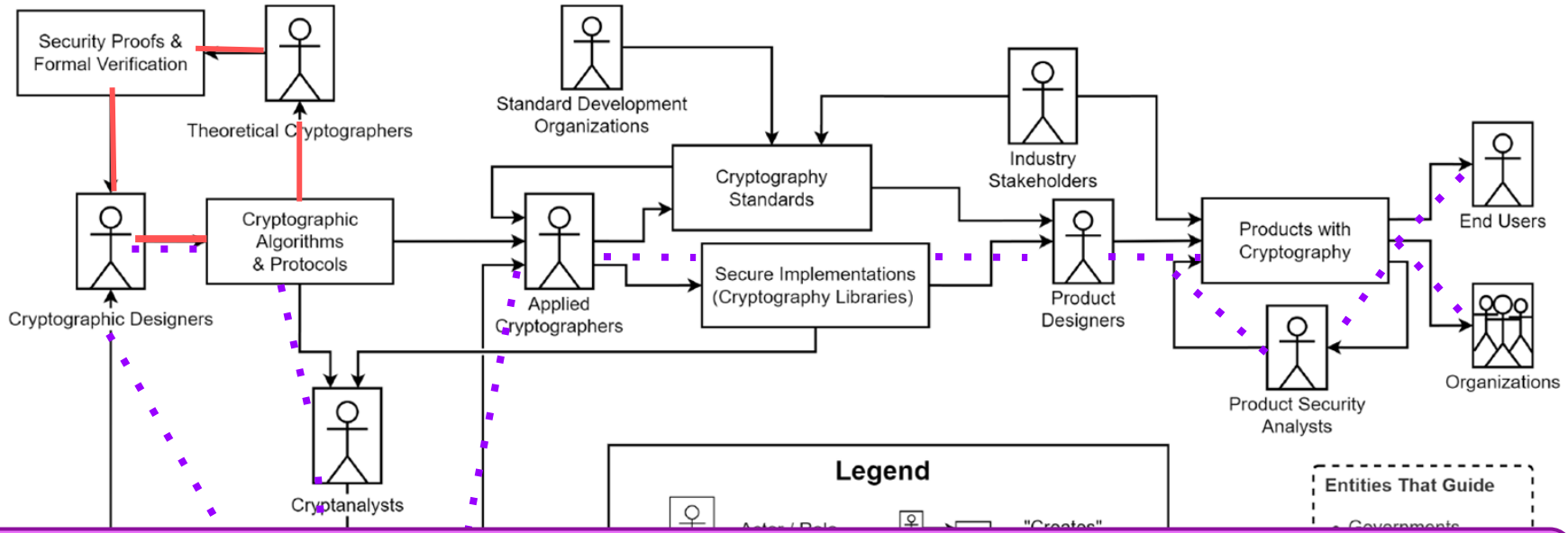


Figure 2 from: K. Fischer, I. Trummová, P. Gajland, Y. Acar, S. Fahl, & A. Sasse. "The Challenges of Bringing Cryptography from Research Papers to Products: Results from an Interview Study with Experts". Usenix Security Symposium 2024

A Visualization of the Cryptography Ecosystem



Question: Can we agree this is a problem?

Figure 2 from: K. Fischer, I. Trummova, P. Gajland, Y. Acar, S. Fahl, & A. Sasse. "The Challenges of Bringing Cryptography from Research Papers to Products: Results from an Interview Study with Experts". Usenix Security Symposium 2024

Diverging (Expert) Views

K. Fischer, I. Trummová, P. Gajland, Y. Acar, S. Fahl, & A. Sasse. “The Challenges of Bringing Cryptography from Research Papers to Products: Results from an Interview Study with Experts”. Usenix Security Symposium 2024

Diverging (Expert) Views

"[RWC] is actually a wonderful place where **industry and academia come together**. [. . .] The community is growing and a lot of papers that analyse a crypto standard will now actually appear at the security conferences." (P3)

K. Fischer, I. Trummová, P. Gajland, Y. Acar, S. Fahl, & A. Sasse. "The Challenges of Bringing Cryptography from Research Papers to Products: Results from an Interview Study with Experts". Usenix Security Symposium 2024

Diverging (Expert) Views

"[RWC] is actually a wonderful place where **industry and academia come together**. [. . .] The community is growing and a lot of papers that analyse a crypto standard will now actually appear at the security conferences." (P3)

"RWC, even by it's name, it conveys what the message is: '**Don't bring your theoretical nonsense here**. We don't want to hear about it!'" (P13).

K. Fischer, I. Trummová, P. Gajland, Y. Acar, S. Fahl, & A. Sasse. "The Challenges of Bringing Cryptography from Research Papers to Products: Results from an Interview Study with Experts". Usenix Security Symposium 2024

Diverging (Expert) Views

"[RWC] is actually a wonderful place where **industry and academia come together**. [. . .] The community is growing and a lot of papers that analyse a crypto standard will now actually appear at the security conferences." (P3)

"RWC, even by it's name, it conveys what the message is: '**Don't bring your theoretical**

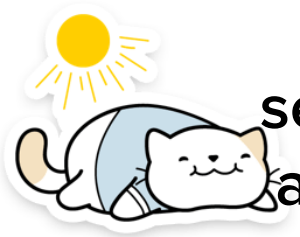
Posits: Motivators/Rewards are the issue

K. Fischer, I. Trummová, P. Gajland, Y. Acar, S. Fahl, & A. Sasse. "The Challenges of Bringing Cryptography from Research Papers to Products: Results from an Interview Study with Experts". USENIX Security Symposium 2024

More Diverging (Expert) Views

K. Fischer, I. Trummová, P. Gajland, Y. Acar, S. Fahl, & A. Sasse. “The Challenges of Bringing Cryptography from Research Papers to Products: Results from an Interview Study with Experts”. Usenix Security Symposium 2024

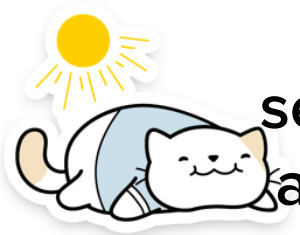
More Diverging (Expert) Views



“[Engineers] have a system and they want to make it secure. And so you indeed have to **translate** your scheme and explain them what you want to do, what you want to achieve and **why these properties are important.**” (P7)

K. Fischer, I. Trummová, P. Gajland, Y. Acar, S. Fahl, & A. Sasse. “The Challenges of Bringing Cryptography from Research Papers to Products: Results from an Interview Study with Experts”. Usenix Security Symposium 2024

More Diverging (Expert) Views

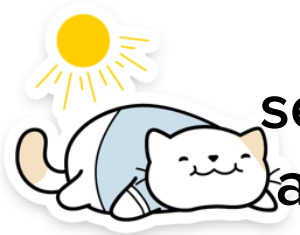


"[Engineers] have a system and they want to make it secure. And so you indeed have to **translate** your scheme and explain them what you want to do, what you want to achieve and **why these properties are important.**" (P7)

"**No! I don't want to understand the problem with the application. That's your job! My job is just the design and mathematics!**" (P10)

K. Fischer, I. Trummová, P. Gajland, Y. Acar, S. Fahl, & A. Sasse. "The Challenges of Bringing Cryptography from Research Papers to Products: Results from an Interview Study with Experts". Usenix Security Symposium 2024

More Diverging (Expert) Views



"[Engineers] have a system and they want to make it secure. And so you indeed have to **translate** your scheme and explain them what you want to do, what you want to achieve and **why these properties are important.**" (P7)

"No! I don't want to understand the problem with the application. That's your job! **My job is just the**

Posits: Lack of translators is the issue

K. Fischer, I. Trummova, P. Gajland, Y. Acar, S. Fahl, & A. Sasse. "The Challenges of Bringing Cryptography from Research Papers to Products: Results from an Interview Study with Experts". USENIX Security Symposium 2024

All together now

"Of course, **not everyone needs to be an expert in multiple areas**. However, our interviews have shown that the role of a translator, "a crypto plumber", or a person in the middle is often poorly rewarded and insufficiently incentivized. Our results suggest that there is certainly a need for people to step into this role." - Fischer et al. 2024

K. Fischer, I. Trummová, P. Gajland, Y. Acar, S. Fahl, & A. Sasse. "The Challenges of Bringing Cryptography from Research Papers to Products: Results from an Interview Study with Experts". Usenix Security Symposium 2024

All together now

"Of course, not everyone needs to be an expert in multiple areas. **However**, our interviews have shown that the role of **a translator, "a crypto plumber", or a person in the middle** is **often poorly rewarded and insufficiently incentivized**. Our results suggest that there is certainly a need for people to step into this role." - Fischer et al. 2024

K. Fischer, I. Trummová, P. Gajland, Y. Acar, S. Fahl, & A. Sasse. "The Challenges of Bringing Cryptography from Research Papers to Products: Results from an Interview Study with Experts". Usenix Security Symposium 2024

All together now

"Of course, not everyone needs to be an expert in multiple areas. However, our interviews have shown that the role of **a translator, "a crypto plumber", or a person in the middle** is often poorly rewarded and insufficiently incentivized. Our results suggest **that there is certainly a need for people to step into this role**." - Fischer et al. 2024

K. Fischer, I. Trummová, P. Gajland, Y. Acar, S. Fahl, & A. Sasse. "The Challenges of Bringing Cryptography from Research Papers to Products: Results from an Interview Study with Experts". Usenix Security Symposium 2024

“So what?” - The Audience

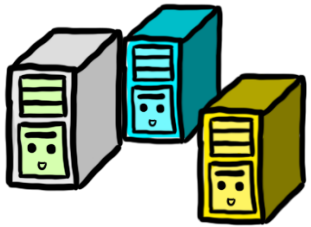
“In general users don’t care very much: I mean good cryptography is cryptography that users don’t see, right?” (P7).

Then what do we need to tell them? Do we need to?

What cryptography do we need to make? How do we know?

K. Fischer, I. Tkachenko, P. Gajland, T. Acar, S. Fahl, & A. Sasse. “The Challenges of Bringing Cryptography from Research Papers to Products: Results from an Interview Study with Experts.” Usenix Security Symposium 2024

Return: Why Private Computation?



A company wants to analyze data

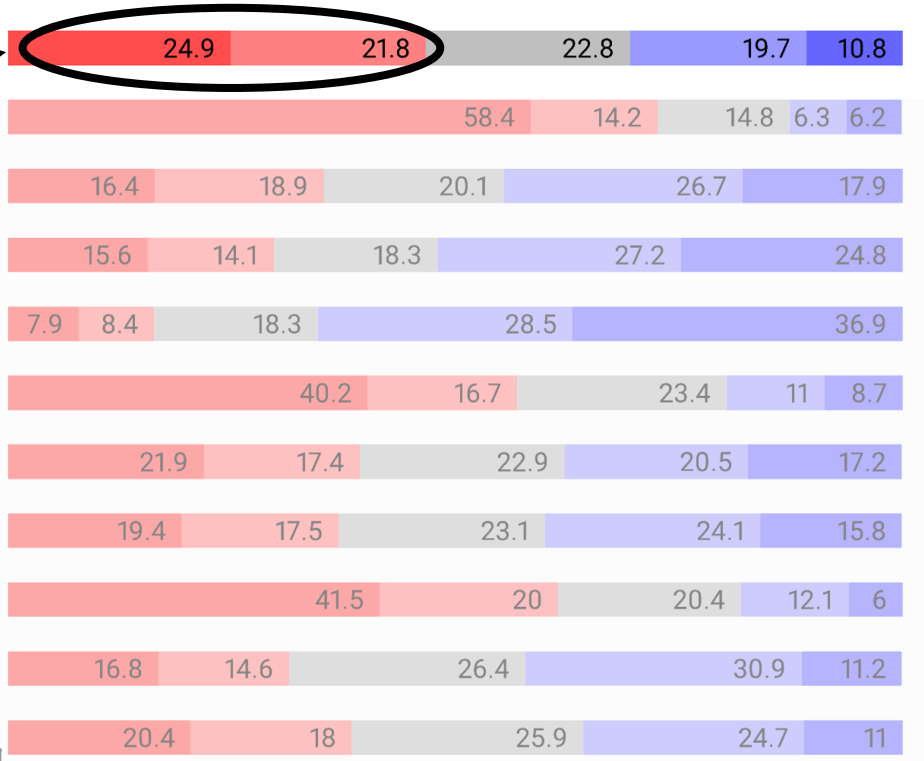
But the data has privacy implications for the data subjects

Researchers develop technical solutions

In what ways does private computation matter to people?

Overall Acceptability Across Scenarios

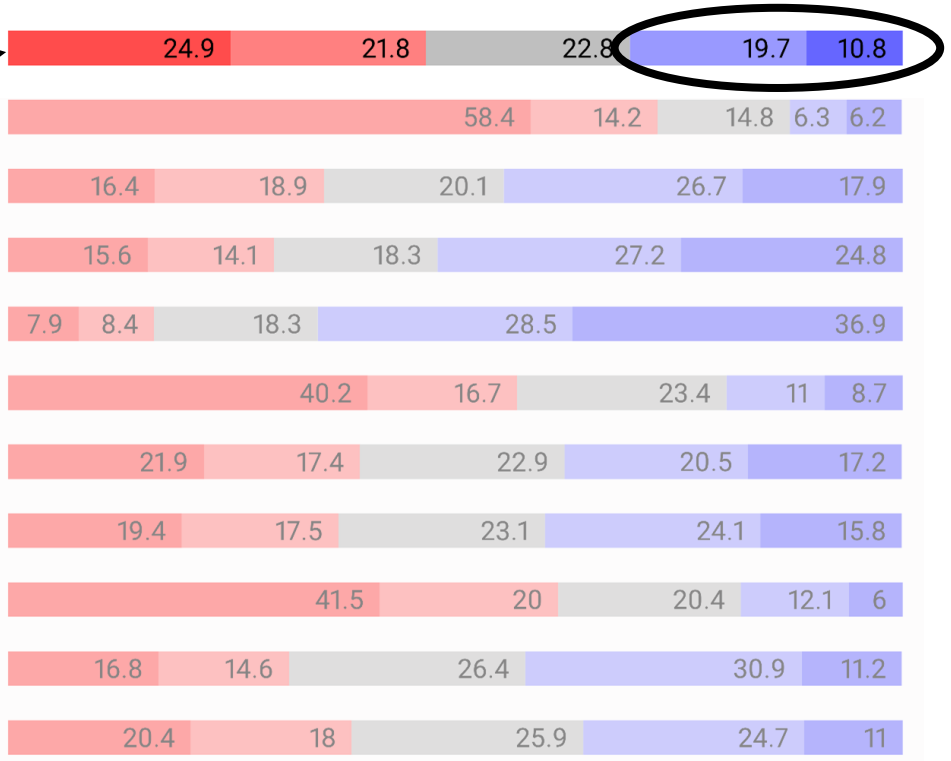
General Scenario Acceptability?



Kacsmar, Tilbury, Mazmudar, Kerschbaum. Caring about Sharing: User Perceptio

Overall Acceptability Across Scenarios

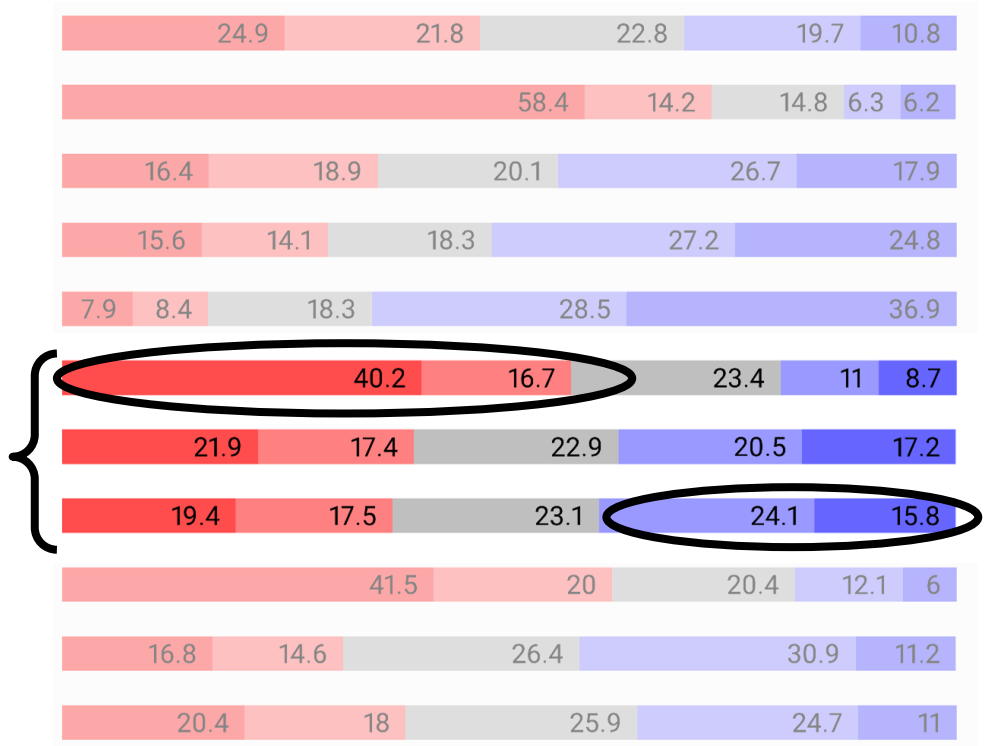
General Scenario Acceptability?



Retention: Acceptability Across All Scenarios

Data Retention?

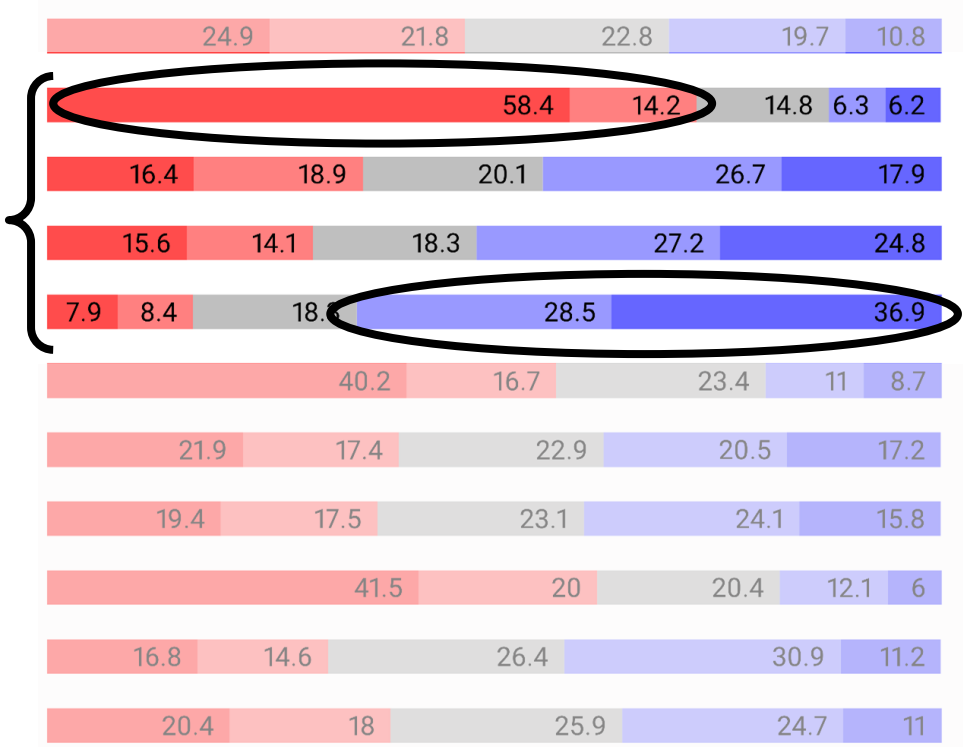
- **Indefinitely**
- **While in use**
- **For set time**



Consent: Acceptability Across All Scenarios

Informed Consent?

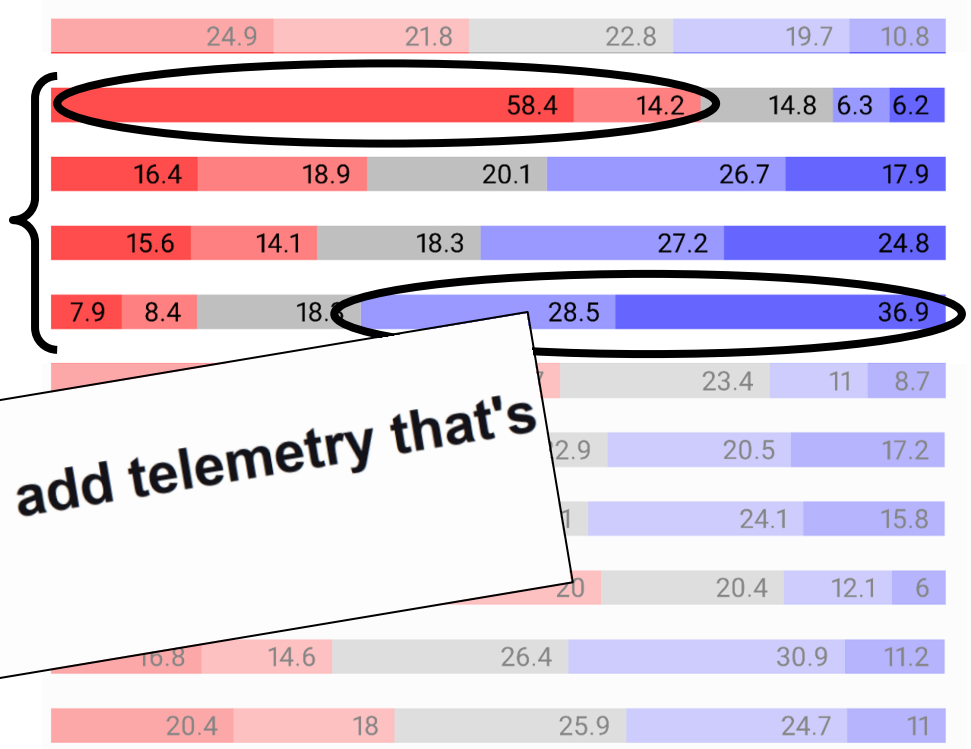
- **Concealed** 
- **Assumed**
- **Opt-out**
- **Opt-in** 



Consent: Acceptability Across All Scenarios

Informed Consent?

- **Concealed** 
- **Assumed**
- **Opt-out**
- **Opt-in** 

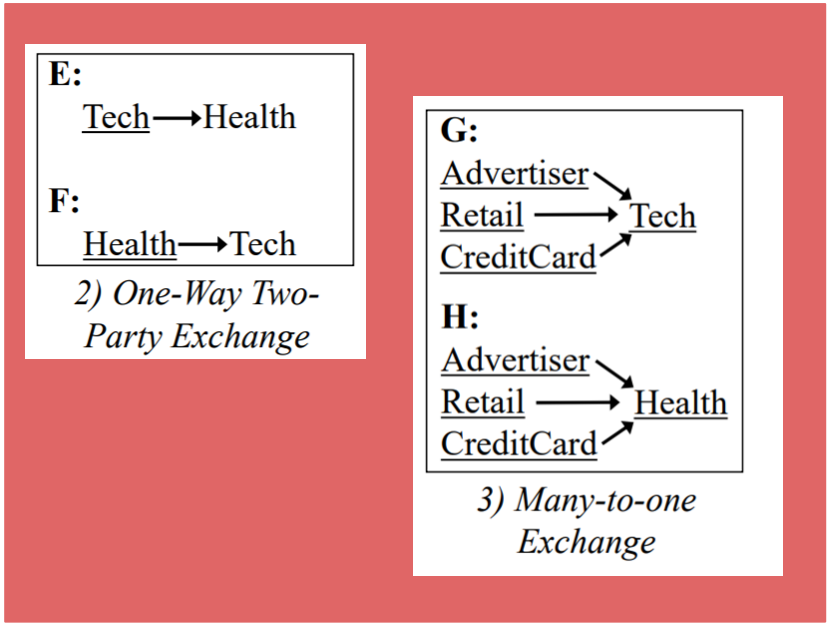


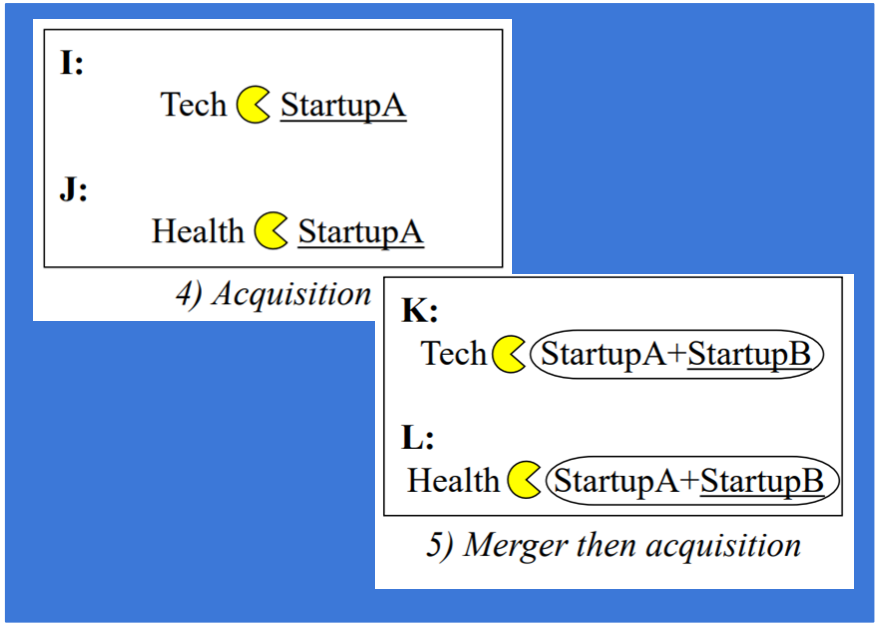
theregister.com

Google's Go may add telemetry that's on by default

Thomas Claburn

Sharing Type Impact on Overall Acceptability

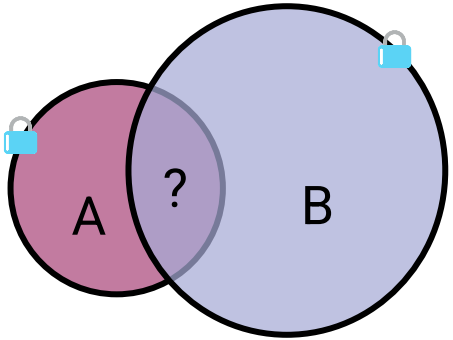




General acceptability is statistically different between types.

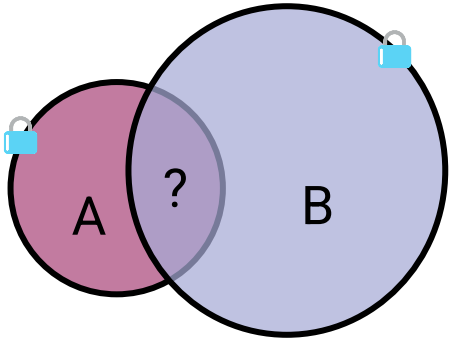
Kacsmar, Tilbury, Mazmudar, Kerschbaum. Caring about Sharing: User Perceptions of Multiparty Data Sharing. *USENIX Security 2022*

Private Set Intersections



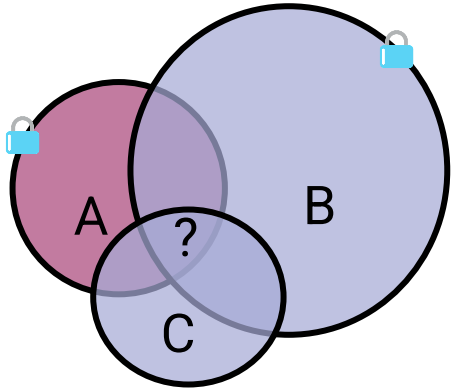
2-Party, One-Way PSI

A → B



2-Party, Two-Way PSI

A ↔ B

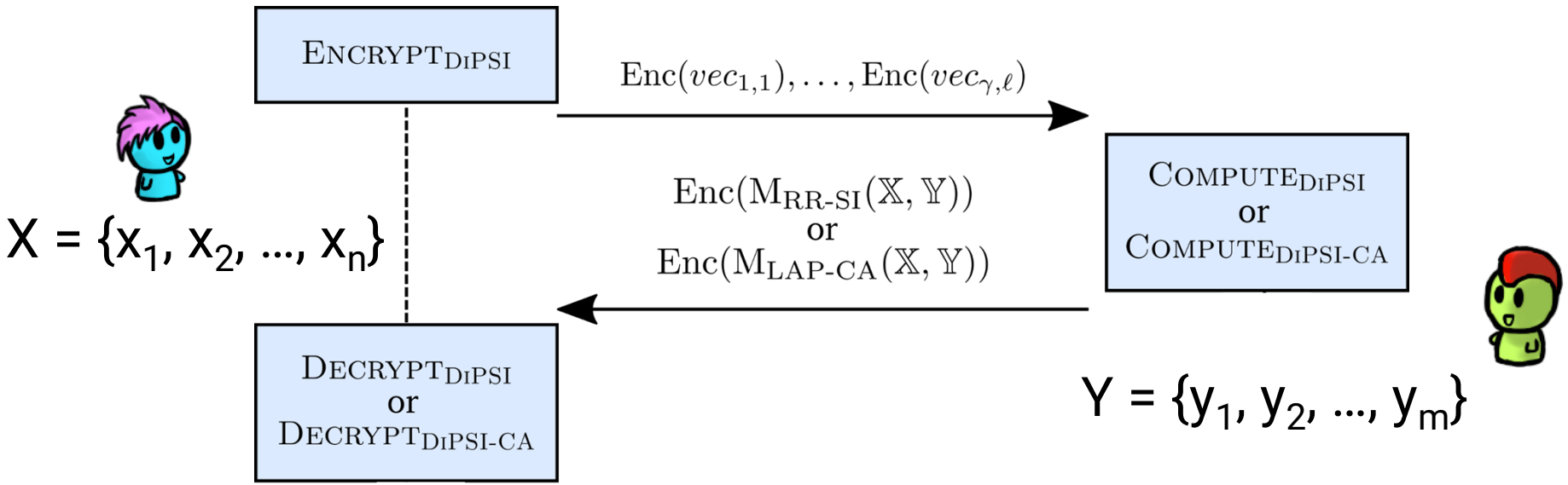


n-Party PSI

Directionality | **Reducing Information** | **Multi-party** | **Varying Guarantees**

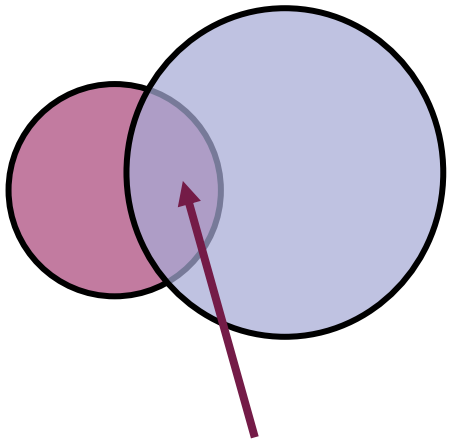
Throw some differential privacy at it.

Private Set Intersection



Kacsmar Khurram, Lukas, Norton, et al. "Differentially private two-party set operations." In 2020 IEEE European Symposium on Security and Privacy (EuroS&P), pp. 390-404. IEEE, 2020.

Why Differentially Private Set Intersection?



Individuals with transactions at **R** who saw ads for **R**

1. Let **s** be the sum of matched credit card transactions
2. Ads for **R** are very specific, if only one individual is at the match, **s** reveals purchase history for them
3. The goal of a DP-sum for this intersection is to prevent such revelations.

B. Kacsmar, B. Khurram, N. Lukas, A. Norton, et al. "Differentially private two-party set operations." In 2020 IEEE European Symposium on Security and Privacy (EuroS&P), pp. 390-404. IEEE, 2020.

Perceptions and Expectations

- What do data subjects understand?
- How is a data subject's willingness to share impacted?
- How do data subjects perceive the risks?

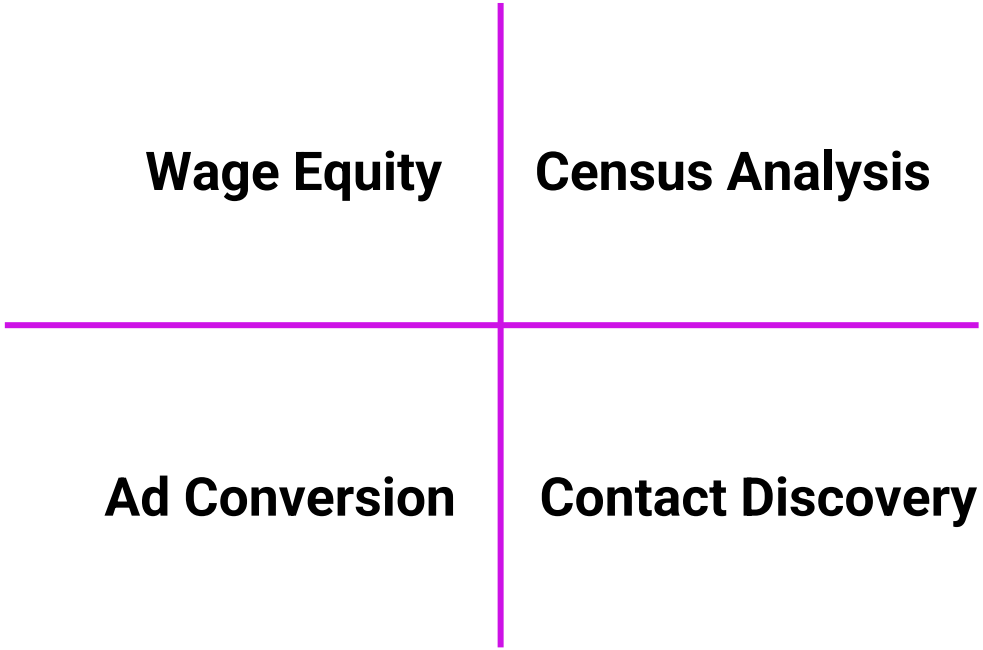
What they "want"

What they "need"

Build towards those attributes

Kacsmar, Duddu, Tilbury, Ur, and Kerschbaum. Comprehension from Chaos: Towards Informed Consent for Private Computation. 2*023 ACM SIGSAC Conference on Computer and Communications Security* (CCS).

The Scenarios



Kacsmar, Duddu, Tilbury, Ur, and Kerschbaum. Comprehension from Chaos: Towards Informed Consent for Private Computation. 2*023 ACM SIGSAC Conference on Computer and Communications Security* (CCS).

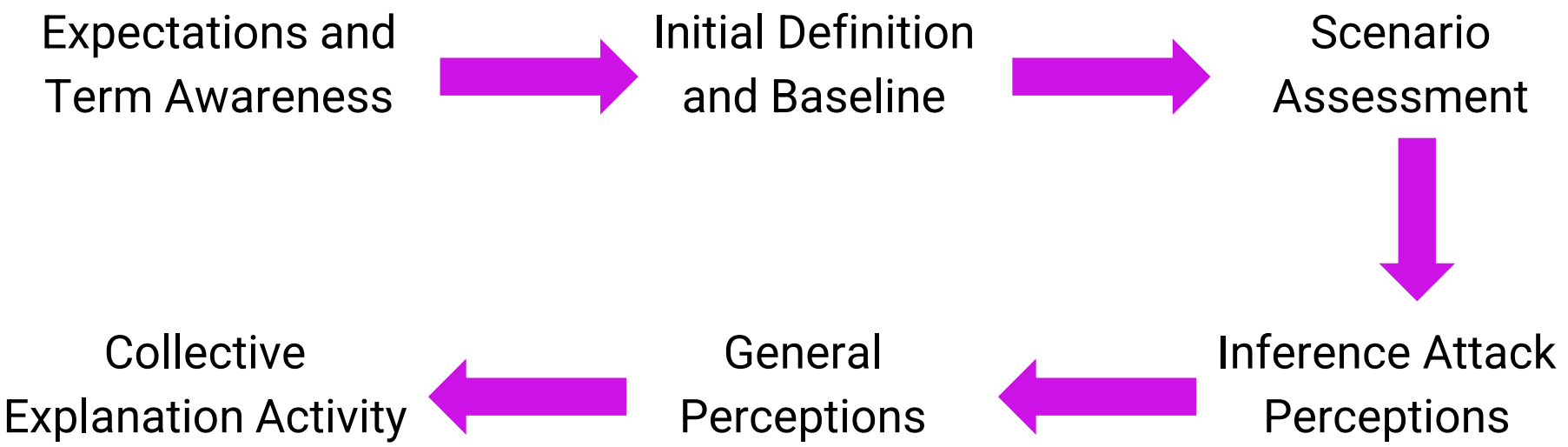
Contact Discovery Conceptual Example

The app wants to **determine the common contacts** between the new user and the existing users via…

1. …the new user shares all their contact information with the social media app.
2. … the new user shares **a modified version** of their contact information…**such that** the social media app does not learn non-users…thus, **this means**…

Kacsmar, Duddu, Tilbury, Ur, and Kerschbaum. Comprehension from Chaos: Towards Informed Consent for Private Computation. 2*023 ACM SIGSAC Conference on Computer and Communications Security* (CCS).

The Interview



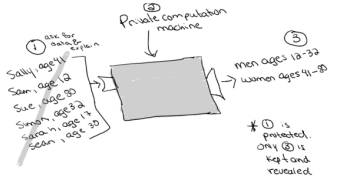
Kacsmar, Duddu, Tilbury, Ur, and Kerschbaum. Comprehension from Chaos: Towards Informed Consent for Private Computation. 2*023 ACM SIGSAC Conference on Computer and Communications Security* (CCS).

Participant Comprehension and Expectations

First Attempt

Second Attempt

Secure computation is a way that a company analyzes your data. The final analysis will be made public [at access location]. However, your specific data is protected and cannot be traced back to you nor can your specific data points be traced back to you. The analysis will be specifically [example], and this is being done because [purpose].



This is the information we're getting from you, but, rest assured, only Part Three will be shown. You can trust us to keep your information private. <If true>This information will only be used for this project and nothing else in the future.

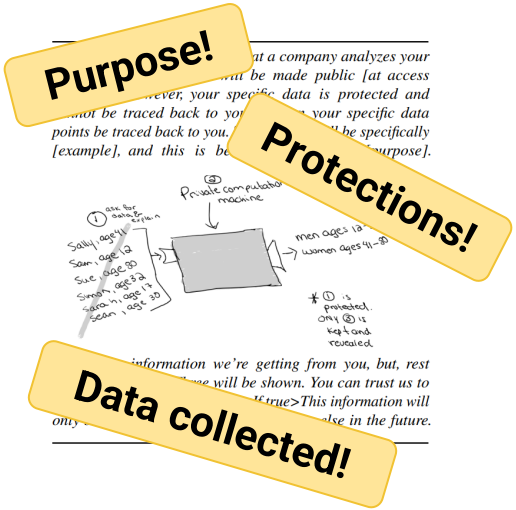
Final Consensus

Kacsmar, Duddu, Tilbury, Ur, and Kerschbaum. Comprehension from Chaos: Towards Informed Consent for Private Computation. 2*023 ACM SIGSAC Conference on Computer and Communications Security* (CCS).

Participant Comprehension and Expectations

First Attempt

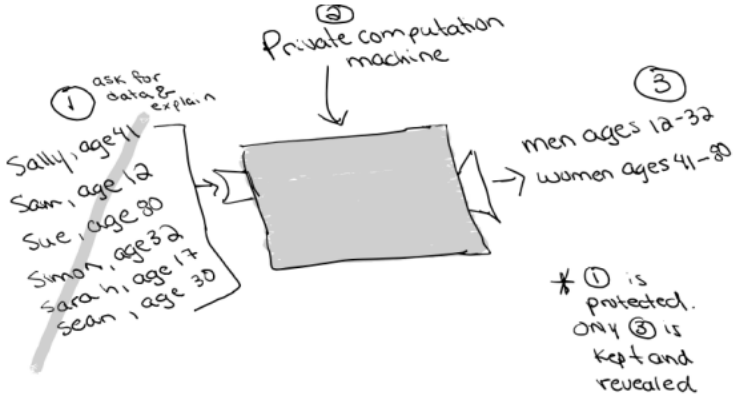
Second Attempt



Final Explanation

Unconcerned with details of the mechanism, **impact** matters

Secure computation is a way that a company analyzes your data. The final analysis will be made public [at access location]. However, your specific data is protected and cannot be traced back to you nor can your specific data points be traced back to you. The analysis will be specifically [example], and this is being done because [purpose].



This is the information we're getting from you, but, rest assured, only Part Three will be shown. You can trust us to keep your information private. <If true>This information will only be used for this project and nothing else in the future.

Kacsmar, Duddu, Tilbury, Ur, and Kerschbaum. Comprehension from Chaos: Towards Informed Consent for Private Computation. 2*023 ACM SIGSAC Conference on Computer and Communications Security* (CCS).

Impact of Private Computation

“...they’re trying to make it sound a little bit better” (P19).

“...it feels a little bit more protected that way” (P12)

Kacsmar, Duddu, Tilbury, Ur, and Kerschbaum. Comprehension from Chaos: Towards Informed Consent for Private Computation. 2*023 ACM SIGSAC Conference on Computer and Communications Security* (CCS).

Bounded Impact of Private Computation

Intentions Matter

Divulge the Details

Regulate the Restrictions

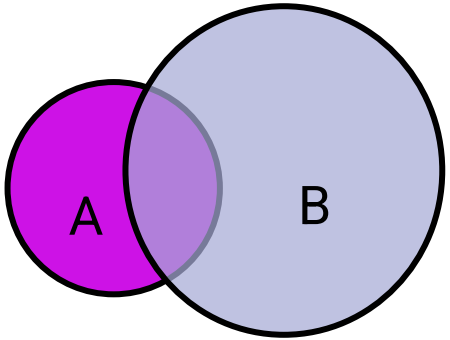
Consent Above All

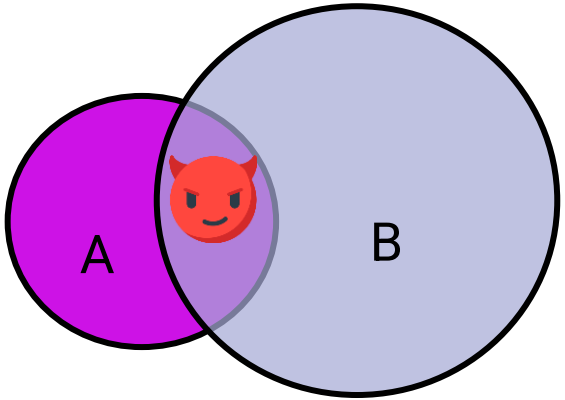
“At the end of the day,
they’re still like learning specific things about me” (P7)

-So what - in technical design terms

Awareness of Unique Threat Models

Alice





Joins Social App

Contact Discovery

Real Identity Connected

There exist, and will continue to exist risks that cannot be regulated by technology

How can we modify PSI for Alice?

Do we understand the problem?

Not just consent, what is the attack?

Not just consent, what is the attack?

Consider:

- Alice joins the app and signs up with her phone number and “E(contact list)”, not shared with other users
- The app, uses contact discovery, but does so with PSI

Not just consent, what is the attack?

Consider:

- Alice joins the app and signs up with her phone number and “E(contact list)”, not shared with other users
- The app, uses contact discovery, but does so with PSI
- **Mallory**, joins the app

Not just consent, what is the attack?

Consider:

- Alice joins the app and signs up with her phone number and “E(contact list)”, not shared with other users
- The app, uses contact discovery, but does so with PSI
- **Mallory**, joins the app
- **Mallory**, has Alice’s number in her contact list

Not just consent, what is the attack?

Consider:

- Alice joins the app and signs up with her phone number and “E(contact list)”, not shared with other users
- The app, uses contact discovery, but does so with PSI
- **Mallory**, joins the app
- **Mallory**, has Alice’s number in her contact list
- The app connects **Mallory** and Alice

Not just consent, what is the attack?

Consider:

- Alice joins the app and signs up with her phone number and “E(contact list)”, not shared with other users
- The app, uses contact discovery, but does so with PSI
- **Mallory**, joins the app

Easy fix you say?

Alice should just get a new number you say?

Variant: Not just consent, what is the attack?

Consider **Alice got a new number**:

- Alice joins the app and signs up with her phone number and “E(contact list)”, not shared with other users
- The app, uses contact discovery, but does so with PSI
- **Mallory**, joins the app

Variant: Not just consent, what is the attack?

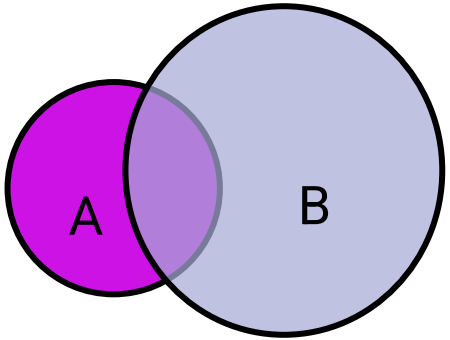
Consider:

- Alice joins the app and signs up with her phone number and “E(contact list)”, not shared with other users
- The app, uses contact discovery, but does so with PSI
- **Mallory**, joins the app

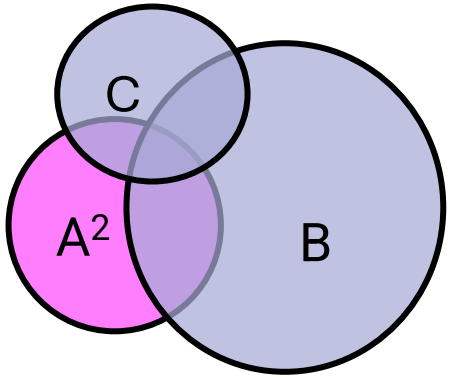
- **Mallory**, tries a set of numbers for Alice’s area code, excluding known non-Alice’s as her contact list
- The app connects **Mallory** and Alice

How can we modify PSI for Alice?

Attempt Fix 1



Alice’s #’s ∩ App users

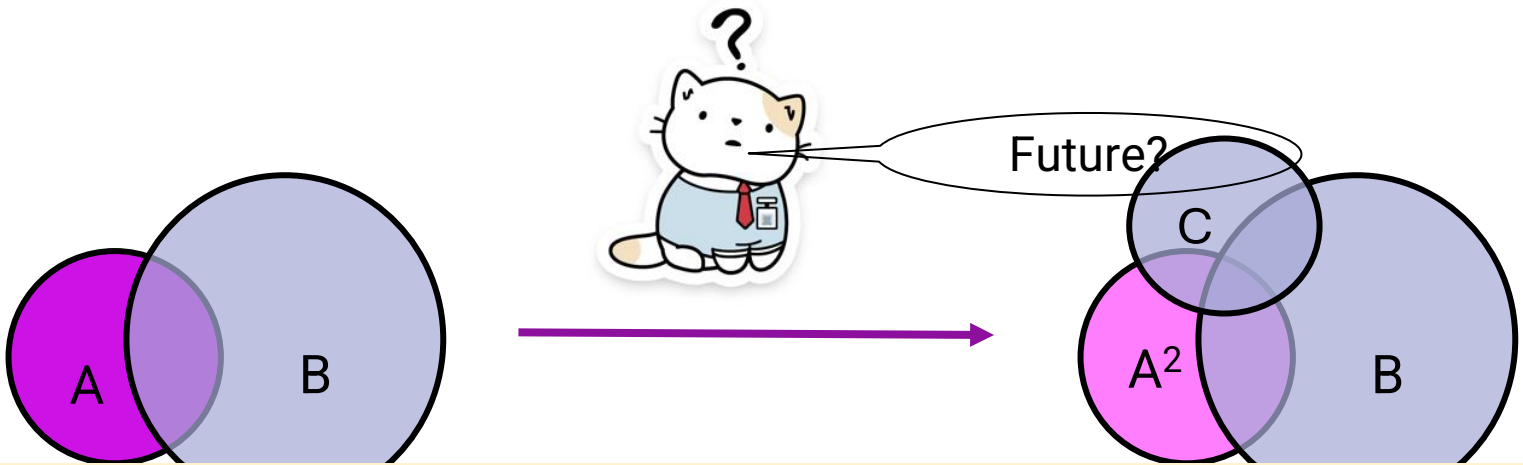


$A^2 \subseteq$ A #’s ∩ App users

And

Match iff $A^2 \cap B \cap C$

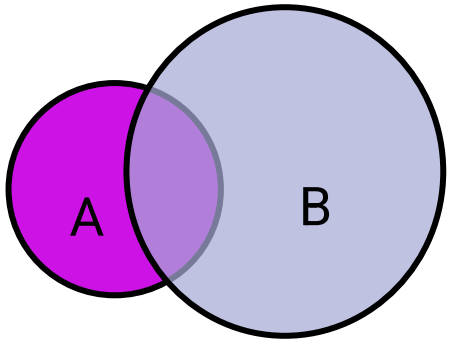
Attempt Fix 1



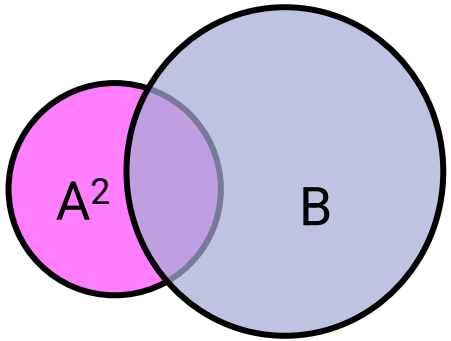
Problem: 3 Party PSI where server will need to find the third party for every element in the primary client set.

Match in $A^2 \cap B \cap C$

Attempt Fix 2



Alice’s #’s ∩ App users



For all $a \in A^2$, $a \leftarrow a + A\#$

$A^2 \subseteq$ A #’s ∩ App users

Take this: Usability is Critical for Privacy

We need usability to support:

- **Accessibility** of secure systems for organizations big and small, used by individuals and populations
- **Enforceability** from legaslaters
- **Verifiability** for those implementing and deploying
- **Meaningful privacy** from applied cryptography for privacy

Module 1 Exercise

- Form groups of 2-4 people (one of you needs a mobile phone that they’re willing to use for this)
- Go to the devices app store
- Search “Math”
- Someone take notes, and the device user narrate decisions:
    - Pick one of the apps. (how did you pick them, tell the others, they should ask you questions)
    - Go to install page
    - Initiate install
    - Open the app

Module 1 Exercise Part 2

- Answer the following (without going back):
  - What permissions did it ask for?
  - How frequently are they used?
  - What are they used for?
  - (other questions generated by group)
- Repeat before, pay attention to privacy nutrition labels and permission requests. Someone take notes, and the device user narrate decisions:
  - Uninstall the app and start over. (how did you pick them, tell the others, group ask questions)
  - Go to install page
  - Initiate install
  - Open the app

Module 1 - Exercise Part 3

- Report on the processes for both part 1 and 2
  - Did either take longer than the other?
  - How did your approach change for these? Did it?
- Report on how effective do you think the original process was at conveying to you the information about permissions/privacy?
  - Was it efficient
  - Was it clear
  - What terms were there? What did they mean? Were any confusing?
- Propose: how could you improve the conveyance of the privacy/permission information?