CSE227 – Graduate Computer Security

Threat Models, Science of Security, A Primer on Security Research

UC San Diego

Housekeeping

General course things to know

- *Due by* **4/10** *at 11:59*
  - Project intention form: https://forms.gle/bgaAtD742X8YSxgw5
  - #FinAid Canvas quiz: https://canvas.ucsd.edu/courses/74259/quizzes/247982
- Start thinking about your teams, the style of project you'd like to do (more on this today), and the topic area
  - I set up a Piazza for group formation, we'll also have some "group time" at the end of the class today for folks to meet people

Some miscellaneous questions I received from last time…

- What's the grading scale?
  - Standard grading without A+ (A: 93 – 100, A-: 90 – 92.99, B+: 87 – 89.99, etc.)
  - **No curve in the class**
- What's the late work policy?
  - Course policy (in the syllabus) is: 10% reduction in overall grade for every 24h late, no exception
- I'm really scared of cold calls. Do I *have* to?
  - Yes, you do. Your life will be filled with scarier people than me asking you for stuff. Best to get used to it now!

Remember…

- Cold calls begin **next class**
  - READ THE PAPERS, the fun begins!

Previously on Graduate Computer Security…

- We talked about **trust**: to have *security*, we must trust something (and for complete *security*, we must trust *everything*)
  - But in today's world, it's hard to trust **anything**, ranging from software to news
- Question: **How do we reason about security in such a fractured trust ecosystem?**

Today's lecture – Security fundamentals, threat models, research

Learning Objectives

- Understand the security mindset, attackers, defenders, what a threat model is, why we have threat models, and get some hands on experience with threat modeling
- Learn about what makes science *science*
- Discuss some paper reading strategies
- Learn about several styles of security research, work through some examples of security research, and work through a potential project idea

The Security Mindset

The adversarial mindset

- The adversarial mindset is: **thinking like an attacker.** How might an attacker break the fundamental assumptions that you (the defender) make about a system?
  - Bruce Schneier: *"Good engineering involves thinking about how things can be made to work; the security mindset involves thinking about how things can be made to fail."*
- Ultimately, the goal of security (and security research!) is simply to explore this type of thinking as it applies to manny areas of computing

The adversarial mindset

- Let’s say you’re playing a game of tag with ten others, and you’re it. What are your strategies for winning as the “attacker?”

The adversarial mindset

- Let's say you're playing a game of tag with ten others, and you're it. What are your strategies for winning as the "attacker?"
- Attacker mentality includes…
  - Looking for the weakest links (find the slowest people)
  - Identifying the **assumptions** that proper functionality depends on. Can you make them false? (it's 10 - 1. Can you make it 9 – 2?)
  - Think outside the box… ignore the limited worldview of the system's designers (turn tag into a stealth mission)

The adversarial mindset

- Let's say you're playing a game of tag with ten others, and you're it. What are your strategies for winning as the "attacker?"
- Attacker mentality includes…
  - Looking for the weakest links (find the slowest people)
  - Identifying the **assumptions** that proper functionality depends on. Can you make them false? (it's 10 - 1. Can you make it 9 – 2?)
  - Think outside the box… ignore the limited worldview of the system's designers (turn tag into a stealth mission)
- You can do this now and all the time. When you interact with a system, think about what that system depends on and how it might be exploited

Adversarial mindset: Break into CSE after hours

- *How might you do it?*

The defender mindset

- The defender mindset is: **thinking like a defender.** What are you trying to protect? What threats are you trying to protect those things from? Who is launching those threats, and what are their capabilities?
  - This type of exercise is known as *threat modeling*

Threat model

Thinking like a defender

- A threat model defines security goals: what are we trying to protect and from whom?
  - Threat models are about *assets* and *attackers*
- How can we think about protecting assets?
  - Three properties: *Confidentiality*, *Integrity*, and *Availability* (C.I.A.) of a particular system
- How can we think about characterizing attackers?
  - Who is the attacker? A curious classmate reading your texts over your shoulder?
  - Two properties: *Capabilities* and *Intent*

Threat model

Thinking like a defender

- A threat model defines security goals: what are we trying to protect and from whom?
  - Threat models are about *assets* and *attackers*
- How can we think about protecting assets?
  - **Three properties: *Confidentiality*, *Integrity*, and *Availability* (C.I.A.) of a particular system**
- How can we think about characterizing attackers?
  - Who is the attacker? A curious classmate reading your texts over your shoulder?
  - Two properties: *Capabilities* and *Intent*

Trifecta of security: Confidentiality, Integrity, Availability

- With a group, come up with definitions of each term, and a few examples of breaches of each property
  - What is confidentiality?
  - What is integrity?
  - What is availability?

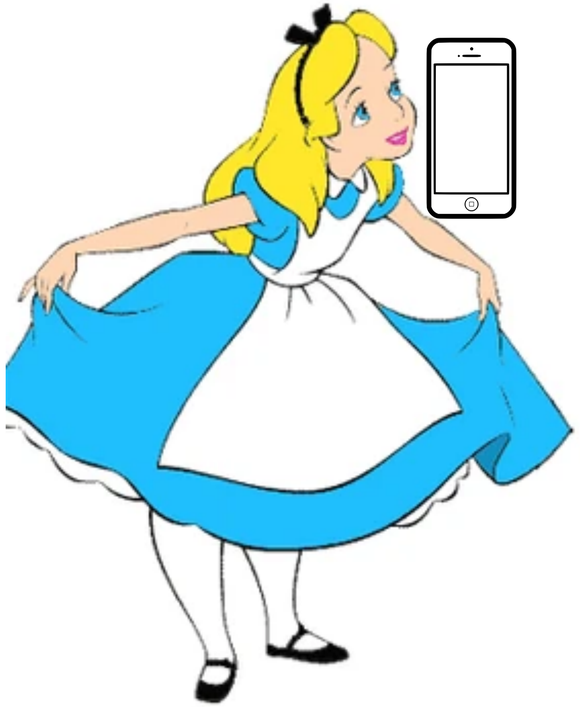
Trifecta of security: Confidentiality, Integrity, Availability

- With a group, come up with definitions of each term, and a few examples of breaches of each property
  - **What is confidentiality?**
  - What is integrity?
  - What is availability?

Confidentiality

- Prevention of unauthorized access to information
  - Unauthorized parties can’t view protected information…. in other words, *secrecy*
- What are some examples of breaches in confidentiality?

How does confidentiality work in practice?



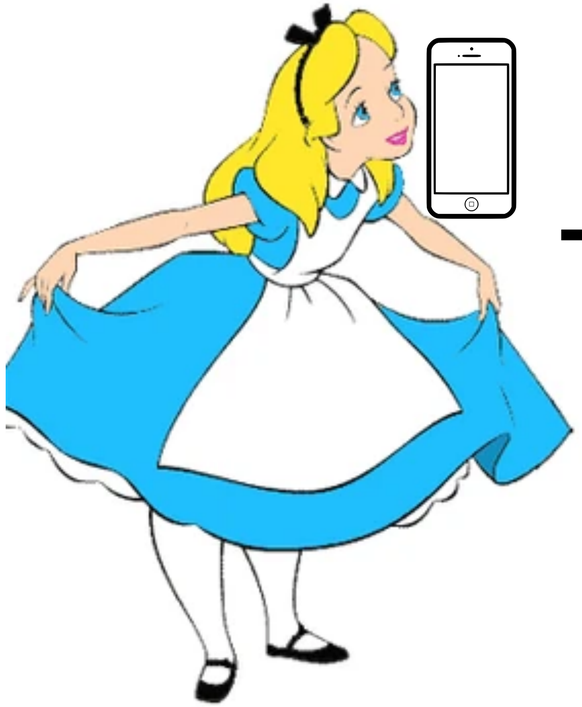
Alice

Bob

AT&T

ISP transmitting message

How does confidentiality work in practice?



message m

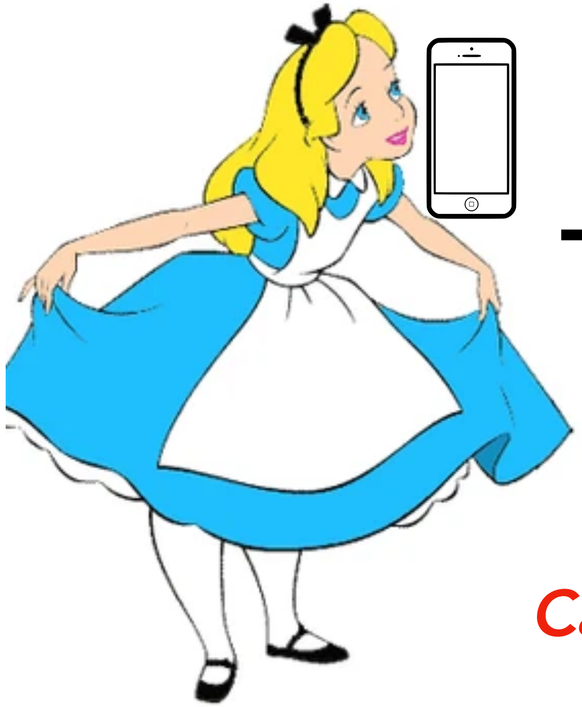
Alice

AT&T

Bob

ISP transmitting message

How does confidentiality work in practice?



message m

Can AT&T read Alice's message to Bob?

Alice

AT&T

Bob

ISP transmitting message

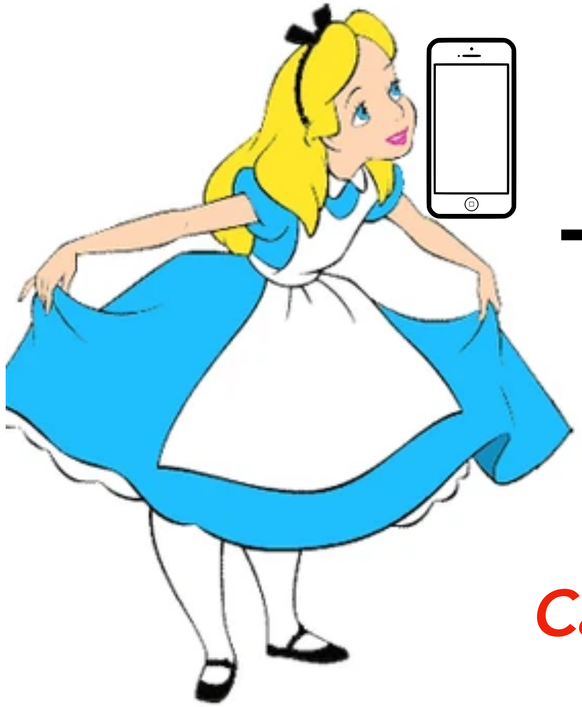
Trifecta of security: Confidentiality, Integrity, Availability

- With a group, come up with definitions of each term, and a few examples of breaches of each property
  - What is confidentiality?
  - **What is integrity?**
  - What is availability?

Integrity (& Authenticity)

- Prevention of unauthorized modification of information, process, or function
  - Unauthorized parties can’t modify private information in flight or at rest
- What are some examples of breaches in integrity?

How does integrity work in practice?



message m

Can AT&T modify Alice's message to Bob?

Alice

AT&T

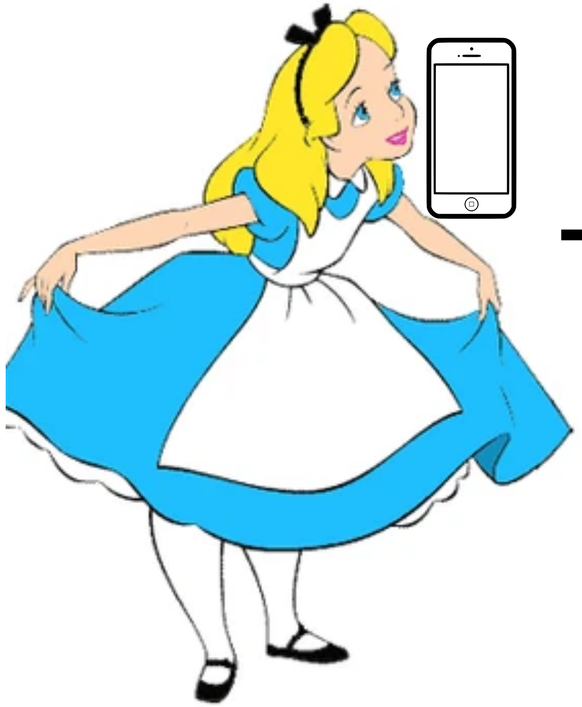
Bob

ISP transmitting message

Integrity **(& Authenticity)**

- Prevention of *impersonation* of another identity… like integrity, but specifically to do with another *actor* (person, system, otherwise)
- What are some examples of breaches in authenticity?

How does authenticity work in practice?



Alice

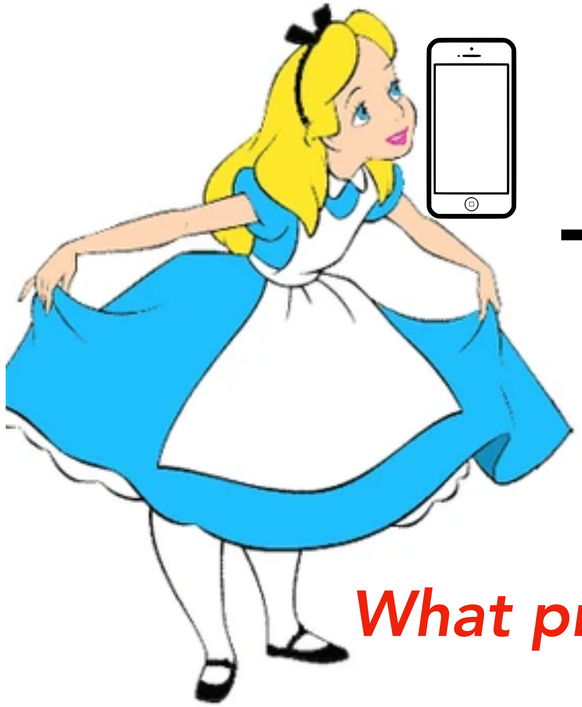
message m

Bob

AT&T

Totally 100% Alice

How does authenticity work in practice?



message m

What prevents AT&T from masquerading as Alice?

Alice

AT&T

Bob

Totally 100% Alice

Trifecta of security: Confidentiality, Integrity, Availability

- With a group, come up with definitions of each term, and a few examples of breaches of each property
  - What is confidentiality?
  - What is integrity?
  - **What is availability?**

Availability

- Prevention of unauthorized *denial of service* to others
  - Unauthorized parties can’t prevent authorized users from accessing a system
- What are some examples of breaching *availability*?

How does availability work in practice?

message m

Can AT&T block Alice's message to Bob?

Alice

AT&T

Bob

ISP transmitting message

C.I.A. + Privacy

- **Privacy:** A person's right or expectation to control the disclosure of their personal information, **including** activity metadata
- What is the difference between privacy and secrecy?

C.I.A. + Privacy

- **Privacy:** A person's right or expectation to control the disclosure of their personal information, **including** activity metadata
- What is the difference between privacy and secrecy?
  - Secrecy is about explicitly hiding information from third-parties
  - Privacy is about not being observed / monitored, including public data
- Activity metadata
  - What can you figure out about a person just from their location history?

C.I.A. + Privacy

- What security property is violated if someone…
  - Unplugs your alarm clock while sleeping?

C.I.A. + Privacy

- What security property is violated if someone…
  - Unplugs your alarm clock while sleeping?
  - Changes the time on your alarm clock?

C.I.A. + Privacy

- What security property is violated if someone…
  - Unplugs your alarm clock while sleeping?
  - Changes the time on your alarm clock?
  - Watches you through your window via a telescope?

Side note: Some lingo

- ***A vulnerability*** is something that can be ***exploited*** (made use of) to cause damage to ***assets*** (usually in the form of a violation of C.I.A. + Privacy)
  - Default passwords (e.g., “admin123”)
  - Bad passwords (e.g., “password123”)
  - Implementation flaws in software
  - Old software left open to the network
  - Cryptography based on weak keys
- Lots of security is organized around vulnerabilities (e.g., National Vulnerability Database, run by NIST), always buzz when a **0-day vuln** is released

Side note 2: More lingo

- **Security Boundary**
  - Perimeter around components of the same trust level
  - Any data or signals coming in from outside is untrusted and potentially malicious (e.g., a bouncer)
- **Attack Surface**
  - Set of interaction points across a security boundary
  - Parts of your system handling input from or otherwise interacting with less trusted and potentially malicious entities
  - Some *highly sensitive* systems are even “air-gapped” to minimize the attack surface

Threat modeling exercise: defending CSE

How do you protect yourself from the undergrads?

- Assets?
  - What are you trying to protect? What's at risk?
- Attackers?
  - What are their capabilities?
  - What are their motivations?
- What's your security boundary? What's the attack surface?

Threat model

Thinking attacker and defender

- A threat model defines security goals: what are we trying to protect and from whom?
  - Threat models are about *assets* and *attackers*
- How can we think about protecting assets?
  - Three properties: *Confidentiality*, *Integrity*, and *Availability* (C.I.A.) of a particular system
- How can we think about characterizing attackers?
  - **Who is the attacker? A curious classmate reading your texts over your shoulder?**
  - **Two properties: *Capabilities* and *Intent***

Attackers / Adversaries

Know thine enemy

- Understand who the attacker is
  - Individual? Are they an outsider, insider, or privileged insider?
  - Group? Are they ad hoc? Established hacking group?
  - Organization? Are they a competitor? A supplier? A customer?
  - Nation state? How powerful is the nation state?

Capabilities

- Clearly understanding the capabilities and scope of the attacker is crucial to proper security defense
- What are some other capabilities of attacker that might be useful to know about?

Motivation / Intent

- Motivation plays a big role in our lives
  - Would you rather fight a motivated 5-year old or an unmotivated 35-year old?
- What are some motivations attackers might have?

Back to threat models…

- Your very first question in any security discussion should be

What's the threat model?

- You can't argue about attacks or defenses without understanding the threat model
- The threat model is what *defines* the problem to be solved
  - And if there's no consensus on the problem, there's going to be no consensus on the solution

A word of caution about threat models

- Your threat model is your problem scope… attackers do not care about them
- Just because an attacker doesn't exist in your threat model doesn't mean they don't exist :)
  - It just means you have explicitly decided you will not address them in your solution
- *"All models are wrong, but some are useful" – George E.P. Box*

Doing Research in Cybersecurity

Why do we do research at all?

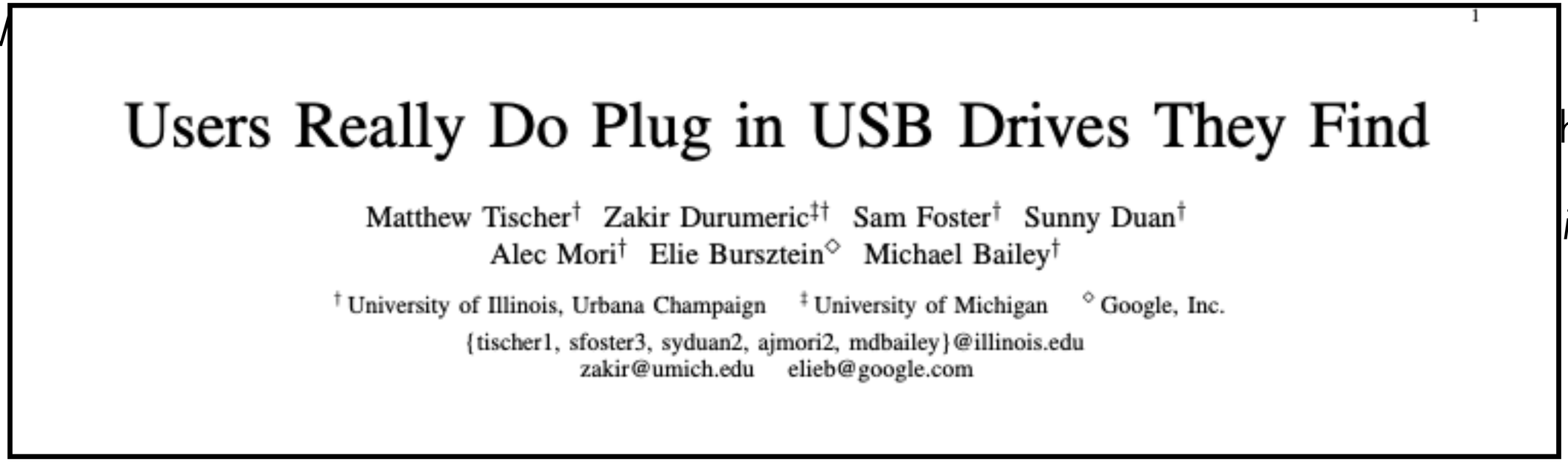
Isn't security just breaking stuff?

- Want to make the world a safer, more secure place – **but we don't know how**
  - With research, we can provide evidence to claims and help ascertain their truth
  - E.g., *Do people pick up USBs they find on the ground and plug them into their computers?* Answer: **Yes**

Why do we do research at all?

Isn't security just breaking stuff?

- W
  - h
  - *ir*



1

Users Really Do Plug in USB Drives They Find

Matthew Tischer† Zakir Durumeric‡† Sam Foster† Sunny Duan†
Alec Mori† Elie Bursztein◇ Michael Bailey†

† University of Illinois, Urbana Champaign ‡ University of Michigan ◇ Google, Inc.

{tischer1, sfoster3, syduan2, ajmori2, mdbailey}@illinois.edu
zakir@umich.edu elieb@google.com

Why do we do research at all?

Isn't security just breaking stuff?

- Want to make the world a safer, more secure place – **but we don't know how**
  - With research, we can provide evidence to claims and help ascertain their truth
  - E.g., *Do people pick up USBs they find on the ground and plug them into their computers?* Answer: **Yes**
- Want to hold companies, products, services to task in protecting people
  - E.g., finding vulnerabilities in popular services

Why do we do research at all?

Isn't security just breaking stuff?

- Want to make the world a safer, more secure place – **but we don't know how**
  - With research, we can provide evidence to claims and help ascertain their truth
  - E.g., *Do people pick up USBs they find on the ground and plug them into their computers?* Answer: **Yes**
- Want to hold companies, products, services to task in protecting people
  - E.g., finding vulnerabilities in popular services
- Want to build the next generation of defenses against new, evolving threats
  - E.g., How do we defend against online harassment? Cyberstalking?

What is science?

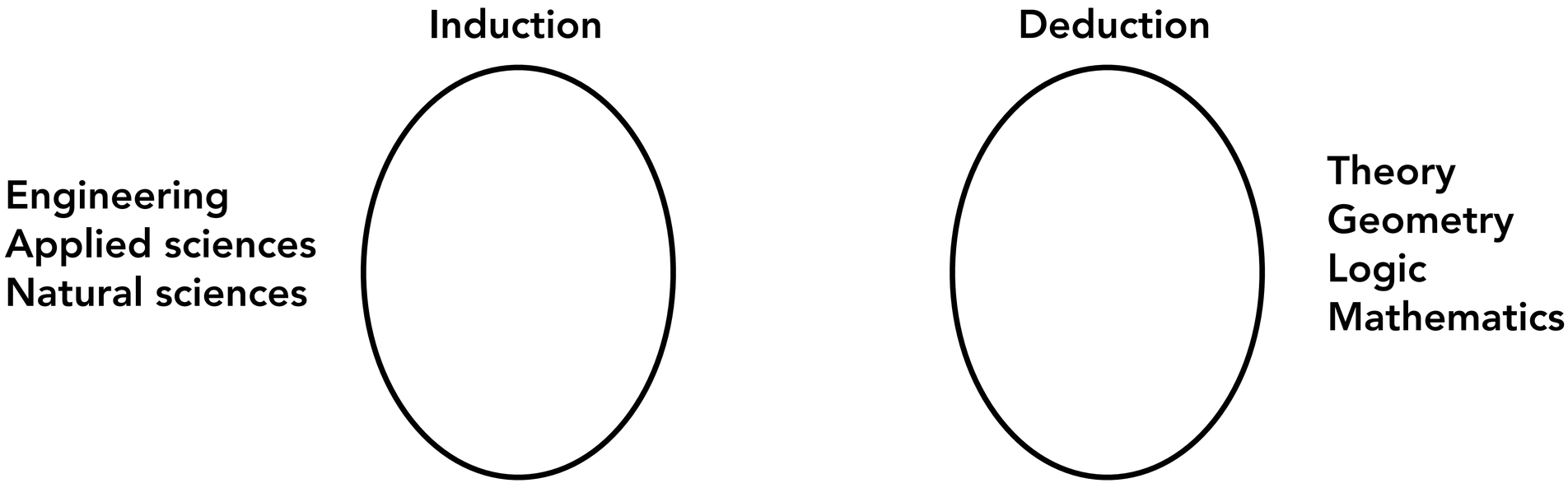
What is science?

Observation, experimentation, test

Two styles of research

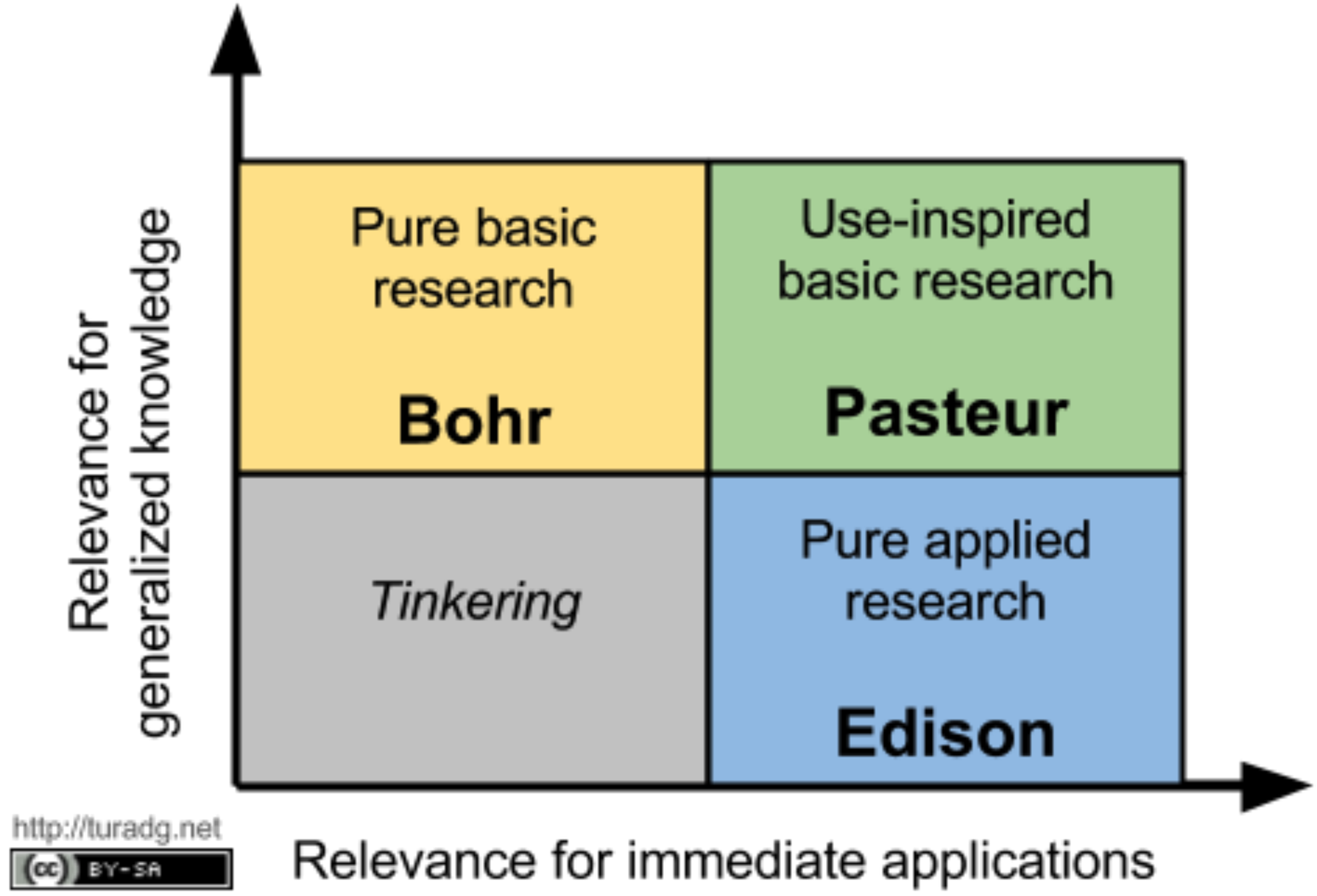
Induction vs. Deduction

- **Induction:** statements about real world (always uncertain) based on observation
- **Deduction:** proved-true statement from axioms



What is science?

Pasteur's Quadrant



What does this have to do with security?

Is security "science?"

- In academia, we do *research;* science and scientific processes help us to make **sense** of our work and evaluate our claims
- Claim: *Changing passwords every 90 days is critical to protecting user accounts.*
  - **How might we study this?**

Practical advice for your research projects

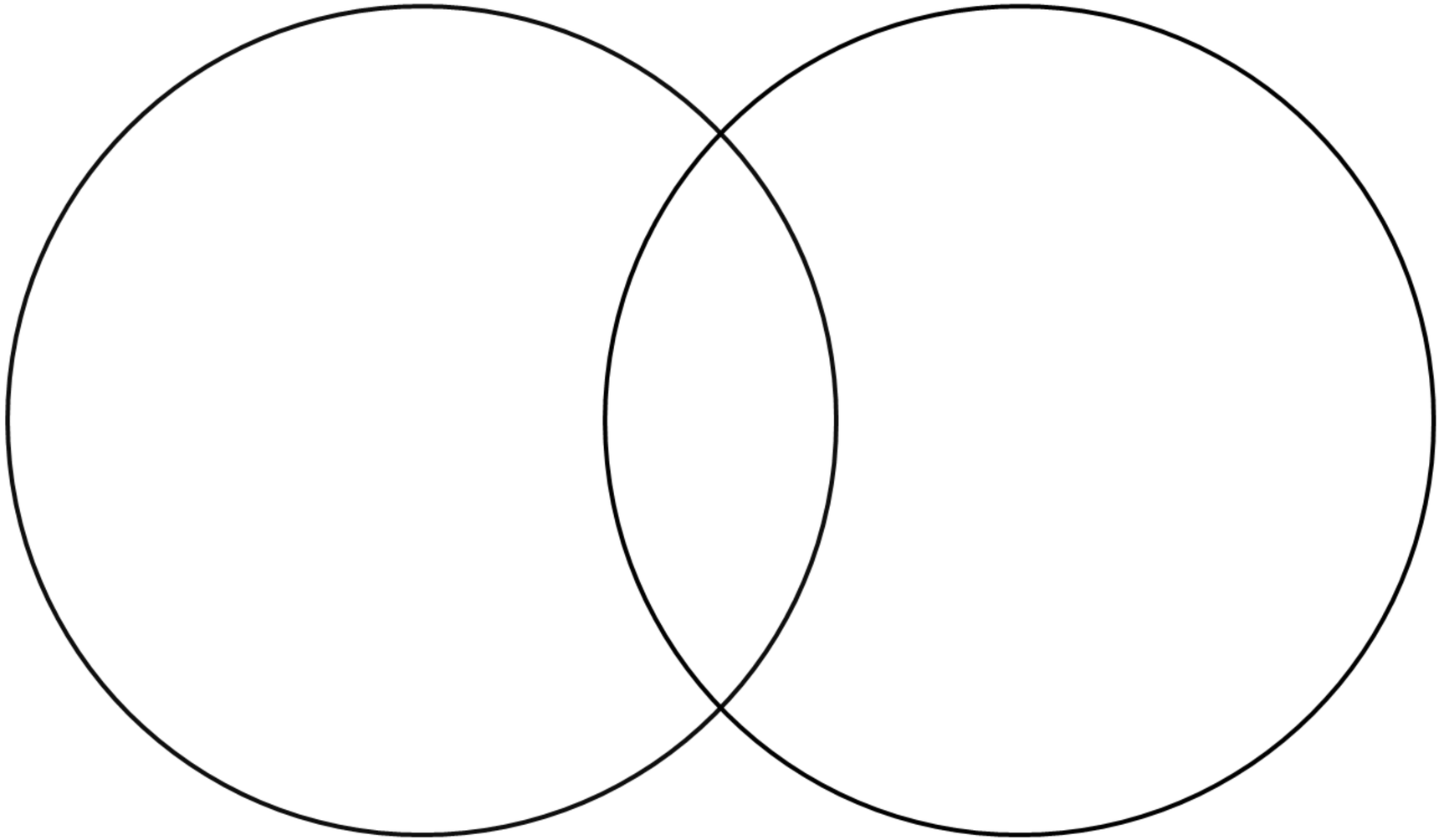
Exercises to help think about projects

- Scanning / skimming papers in a related area
  - Say you're excited about the web… what parts of the web are you interested in? Why are you interested in them?
- Mesearch: What services do I use **in my life** that are critical to me? How can I evaluate whether they are secure?
- General "cool" factor
  - *It would be very cool if someone could break Zoom background blur and identify objects based on images*

Side note: How to read papers

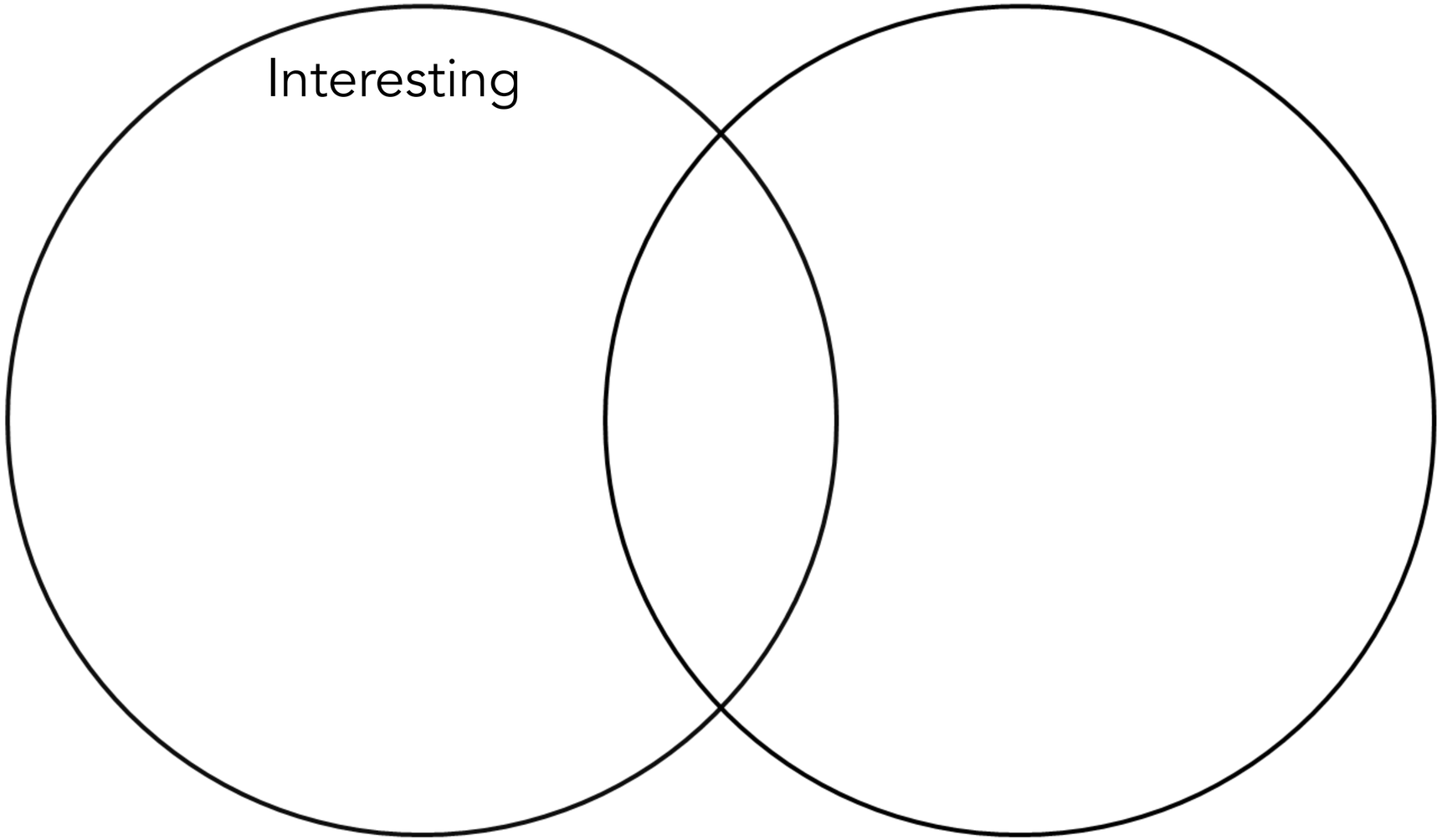
- Reading papers can be a drag
- My general process:
  - Read through the *abstract + introduction* first; figure out what the high level takeaways of the paper are
  - Read through the *figures + captions* — this will tell you the concrete results the authors thought were worth highlighting
  - Read through the *methods* — this might point you to something interesting you can use in your own research
  - Read the rest
- Papers might take 2h to get through in the beginning…. by the end you'll be able to digest the main concepts and ideas of a paper in ~30 – 45 mins (it just takes practice)

Choosing a research problem



Choosing a research problem

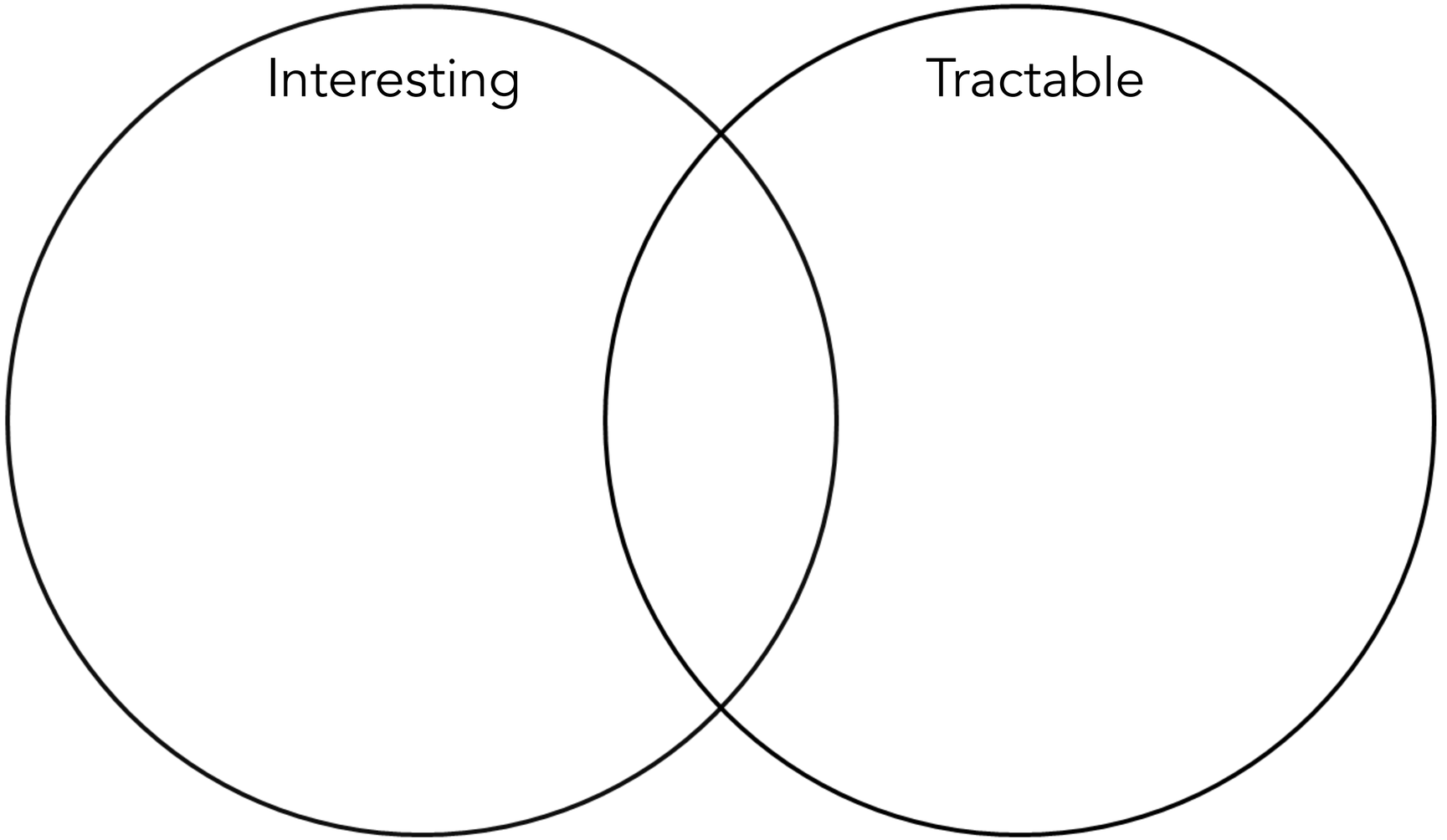
Choose Interesting Problems



- Is this project interesting?
  - Are there technical contributions?
  - Are there societal contributions?
  - Am I doing something new and exciting?
  - Who will care if I do this project?

Choosing a research problem

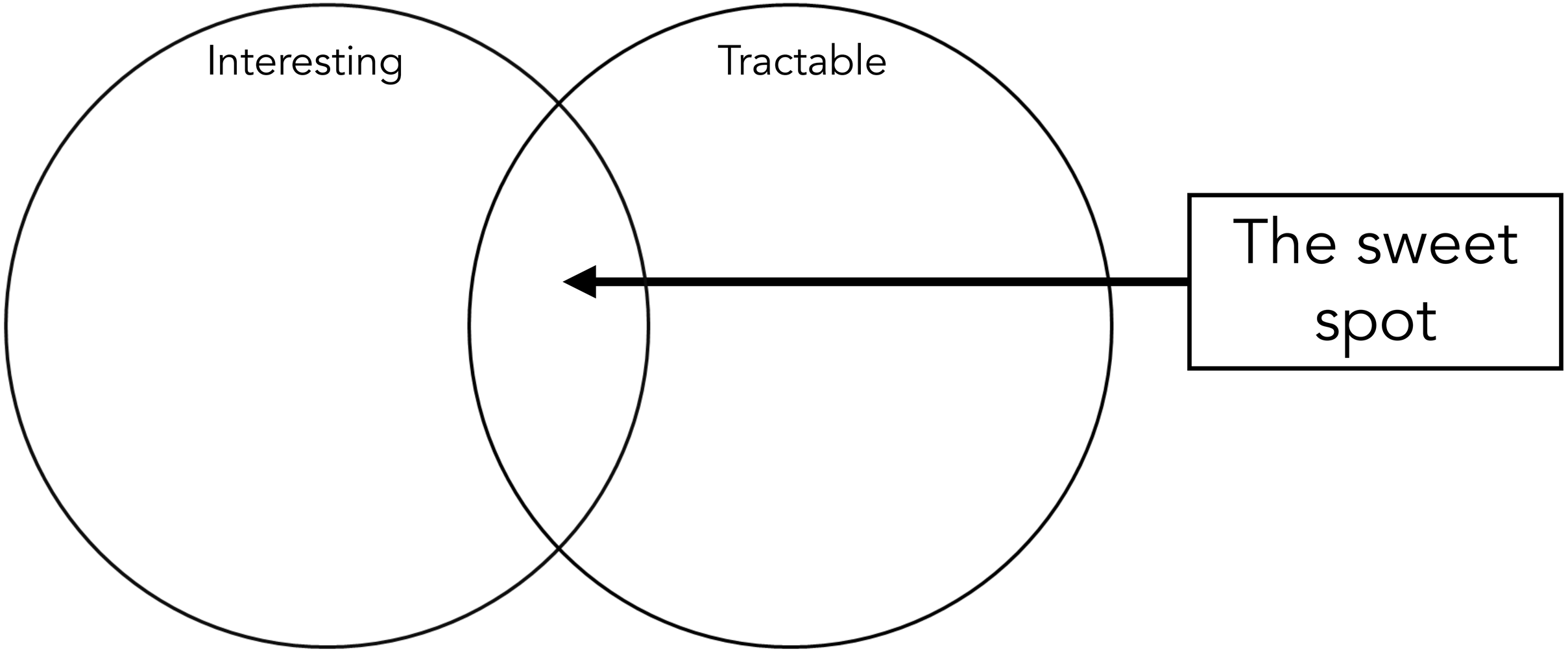
Choose Tractable Problems



- Is this project tractable?
  - How would I go about answering the interesting research questions?
  - Do I have the data to answer this question?
  - Do I have the *time* to answer this question?

Choosing a research problem

Right in the middle



Nonexhaustive List of Research Styles in Security

4 main types of research

- Offensive security research
- Defensive security research
- Measurement / Empirical research
- Human subject research* (you won't do this in this class)

Offensive Security Research

Breaking systems

- Typical construction: “X is a system that exists in the wild that performs some function F. I am an adversary with Y capabilities. Through Z series of technical fu, I have broken X such that it no longer performs function F properly.”
- Pros
  - Clear success criteria (the attack either works or it doesn’t)
  - Coolness factor is very high (e.g., *We hacked a car*)
- Cons
  - Deep technical knowledge needed to understand where to even look in the first place
  - Most attacks end up being very convoluted and hard to conduct

Defensive Security Research

Defending against existing attacks

- Typical construction: “Someone has invented attack A with adversary X. There is currently no defense for A. We design defense D against attack A that works like this. We evaluated defense D in these N scenarios, and demonstrate the defense is robust against adversary X”
- Pros
  - You finish with an end-product – something you have actually built
  - Forces you to think about practicalities in building the thing (software issues, performance, scale, etc.)
- Cons
  - Framing is everything — defending against an attack that no one cares about is usually worthless
  - Defenses are *hard* (and if you reviewer breaks your defense, you’re hosed!)

Measurement Research

How do we measure Internet problems?

- Typical construction: “I have a security related question about X ecosystem. I have devised a system to collect data to measure that ecosystem. I make a lot of assumptions about how the data ought to look. I analyze the data and check my understanding”
- Pros
  - Construction is similar across different projects – similar techniques but vastly different areas
  - Provides more basic understanding of problems
- Cons
  - Takeaways are not always obvious (framing is important)
  - Often requires a lot of assumptions and a lot of data (and you never have perfect data)

Human Subjects Research

Surveys, interviews, and many more

- Typical construction: “I want to understand how people experience X harm/phenomena, but it’s hard to measure with existing metrics. I carefully design a survey or interview experiment Y, pilot that experiment with test participants, iterate on my research instrument, and then deploy it to the world. I analyze the data and try to understand what’s going on.”
- Pros
  - Grounds your work in lived experience – one of the most important and overlooked aspects in computer science
  - Results are often much more nuanced + complex and match the reality of how people behave on the Internet
- Cons
  - Humans are messy: small effect sizes with modeling
  - Very hard to do well, and they take a longer time (hence, not in these 10 weeks)

Exercise: Ideate on a project

- Three questions:
  - What is a technology that I have used recently that I am interested in?
  - What kinds of security or privacy considerations are there for this technology?
  - Is there any way to attack, defend, or measure those considerations?

Group Time