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 **1/17** at 11:59*
  - Project intention form: https://forms.gle/3efhZJAmfG9Gv4xF8
  - #FinAid Canvas quiz: https://canvas.ucsd.edu/courses/61827/quizzes/199237
- Start thinking about your teams, the style of project you'd like to do (more on this today), and the topic area
- Project specification will be released **1/10,** will provide more details and information about each milestone

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!

News

Meta Says It Will End Its Fact-Checking Program on Social Media Posts

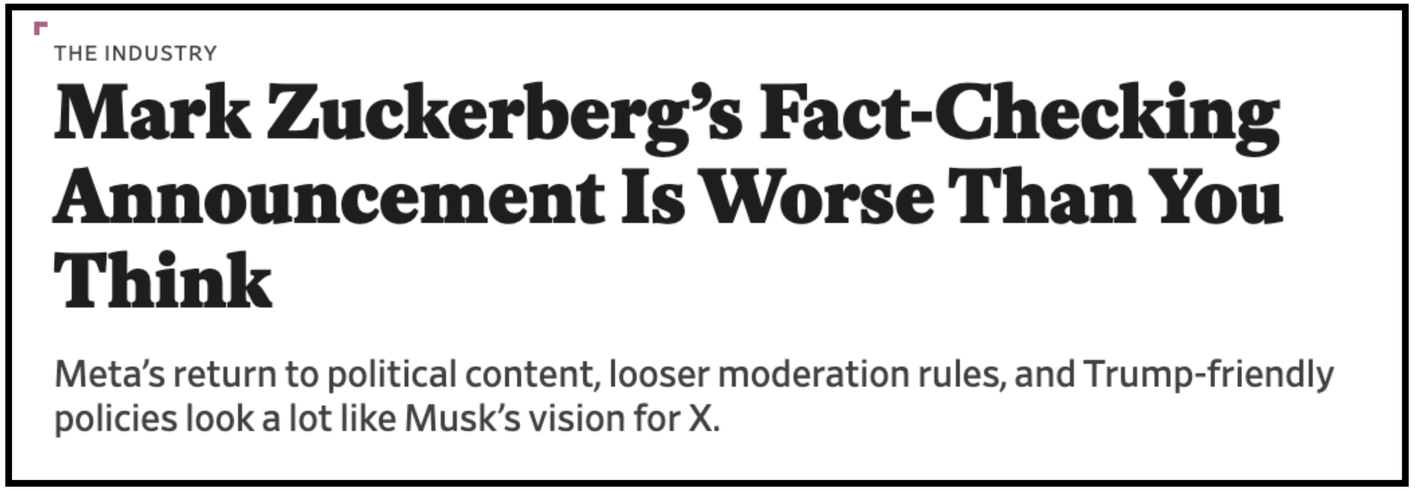
The social networking giant will stop using third-party fact-checkers on Facebook, Threads and Instagram and instead rely on users to add notes to posts. It is likely to please President-elect Trump and his allies.

What is fact checking and how does it work?

- In 2016, among significant concerns about mis/disinformation on their platforms, Meta launched an **independent fact checking program**
    - Outsourcing major stories and facts to independent, verified, and *trusted* organizations
- **Now,** in a reversal, Meta (and Zuck) have decided independent fact checkers come with too much of their own biases, and will end the program
- As a change their approach, they are moving to *Community Notes* (similar to X), which is a crowd-labeled adjudication of truth (because Facebook users can always be trusted 100%)

Why the change?

That’s not the way things played out, especially in the United States. Experts, like everyone else, have their own biases and perspectives. This showed up in the choices some made about what to fact check and how. Over time we ended up with too much content being fact checked that people would understand to be legitimate political speech and debate. Our system then attached real consequences in the form of intrusive labels and reduced distribution. A program intended to inform too often became a tool to censor.


THE INDUSTRY
Mark Zuckerberg's Fact-Checking Announcement Is Worse Than You Think
Meta's return to political content, looser moderation rules, and Trump-friendly policies look a lot like Musk's vision for X.


POLITICS
The danger of Meta's big fact-checking changes
Mark Zuckerberg's efforts to cozy up to Trump have concerning consequences.
by Li Zhou
Jan 8, 2025, 3:45 AM PST


Facebook and Instagram get rid of fact checkers
24 hours ago
Share
Save
Liv McMahon, Zoe Kleinman & Courtney Subramanian
BBC News in Glasgow and Washington

My opinion

Decrying fact-checking as politically biased is… a take

- While it is true that people (and experts) have biases, the **point** is that *experts* are supposed to rise above the bias in their examination
  - Meta believes that's not true; fact-checkers disagree
  - Interestingly, research suggests lay-people may also be suspicious of experts for this reason: https://hci.stanford.edu/publications/2022/ComparingPerceivedLegitimacy.pdf
- In general, we haven't seen **any** evidence to suggest adversarial, political motives in *fact-checking* done by expert review
- However, I *am* somewhat excited about community notes: provides a new avenue for intentional and defensive design against mis / disinformation

Recap

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

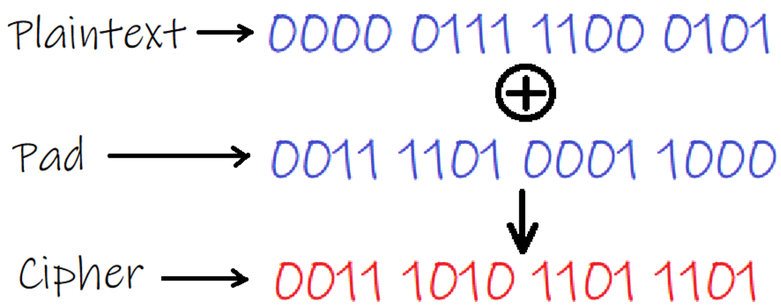
- Understand what computer security *is*, and different types of security models
- Understand what a threat model is, why we have threat models, and get some hands on experience with threat modeling
  - Get experience with the security mindset
- Learn about what makes science *science*
- Learn about several styles of security research, work through some examples of security research, and work through a potential project idea

Security Models

Two competing philosophies for security

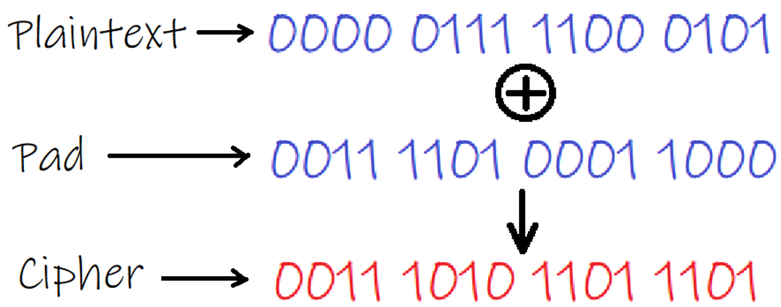
- **Binary** model [secure vs. insecure]
  - Traditional cryptography and trustworthy systems
  - Assume adversary limitations X and define security policy as Y
  - if Y cannot be violated without needing X then system is secure, else insecure
  - Code words: “Proof of security,” “Secure by design,” “Trustworthy systems”
- **Risk management** model [more secure vs. less secure]
  - Most commercial software development (and real-world security… e.g., terrorism)
  - Try to minimize biggest risks and threats
  - Improve security where most cost effective
  - Code words: “Risk,” “Mitigation,” “Defenses,” “Resilience”

Binary model example: Perfect substitution cipher



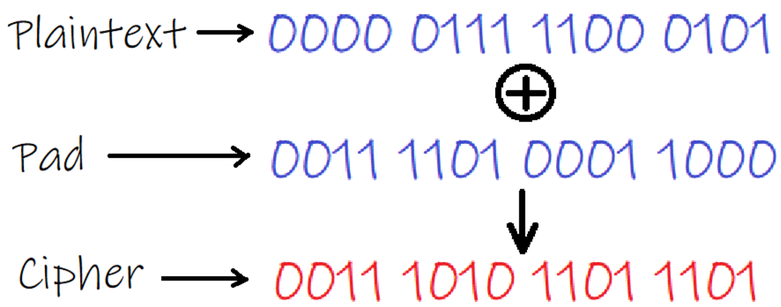
- For a given plaintext, choose a string of **random** bits the same length of the plaintext, XOR them to obtain the ciphertext

Binary model example: Perfect substitution cipher



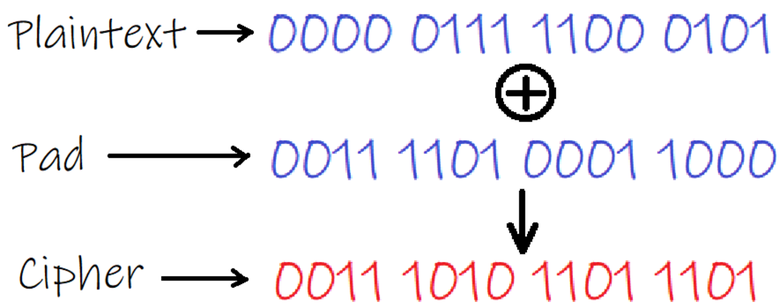
- For a given plaintext, choose a string of **random** bits the same length of the plaintext, XOR them to obtain the ciphertext
  - *Why is this considered perfect?*

Binary model example: Perfect substitution cipher



- For a given plaintext, choose a string of **random** bits the same length of the plaintext, XOR them to obtain the ciphertext
  - *Why is this considered perfect?*
- **Perfect secrecy** – probability that a given message is encoded in the ciphertext is unaltered by knowledge of the ciphertext
- **Forward secrecy** – Future messages encrypted in this scheme will not reveal information about previous plaintext messages

Binary model example: Perfect substitution cipher



What are some assumptions about perfect substitution ciphers that might make this not as perfect as it seems?

Problems with Binary: Assumptions often fail in practice

- Many assumptions are **brittle** in real systems
  - Real artifacts are fragile, imperfect, have bugs/limitations
  - *How can you ensure you always generate a truly random one-time pad?*
    - Turns out this is *really hard to do* – we'll read a paper on this in **week 5**
- Huge gap between abstraction and implementation
  - **Deepak's version:** *The real world is hard.*

Problems with Binary: Security evolution

- As engineers, we like to pretend like we understand our own creations, or that we can create complex systems that only do what they're meant to do…
  - This is a lie, nobody *really* knows how these systems work
- Complex systems co-evolve with attacks against them
  - Systems deemed secure today may not be resilient to new threats: e.g., quantum computers

Risk-mitigation model example: Antivirus

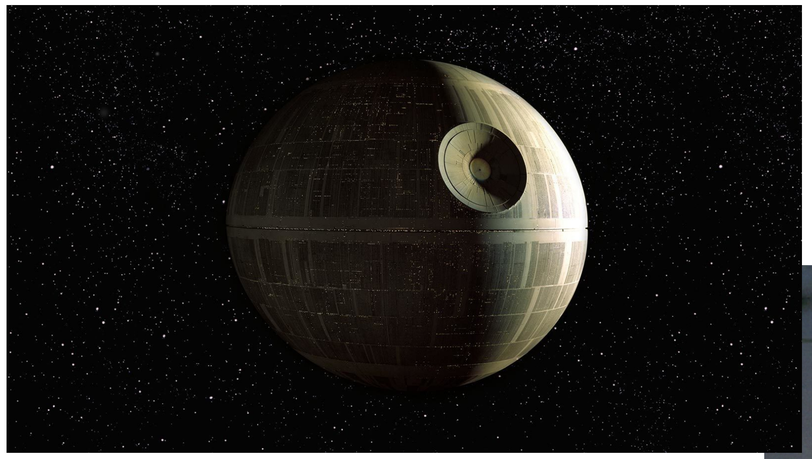
- Antivirus is software that you install on your machine that monitors your machine to detect + remove **malware** or other bad software
- *Question: What's the difference between different anti-virus software?*

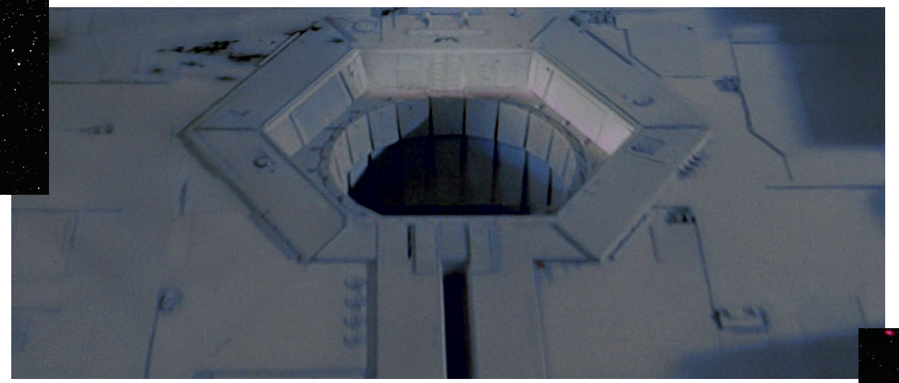
Risk-mitigation model example: Antivirus

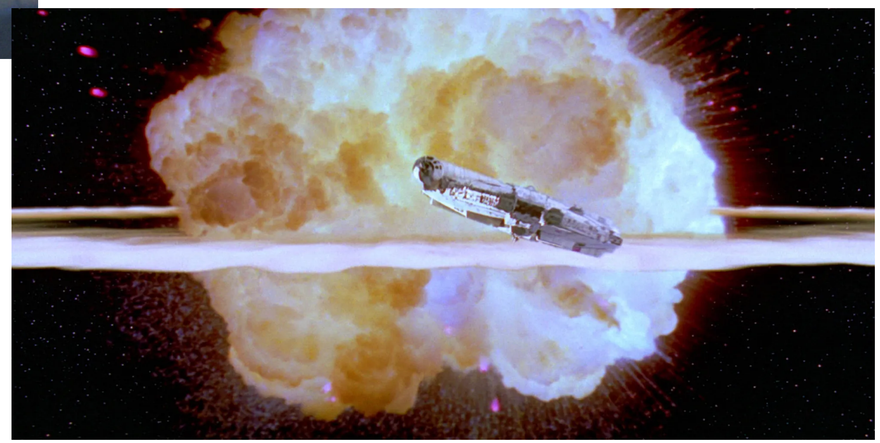
- Antivirus is software that you install on your machine that monitors your machine to detect + remove **malware** or other bad software
- *Question: What's the difference between different anti-virus software?*
- Answer: ¯_(ツ)_/¯, could make for a good research project!

Problems with Risk-mitigation

*One unforeseen vulnerability can matter **a lot***



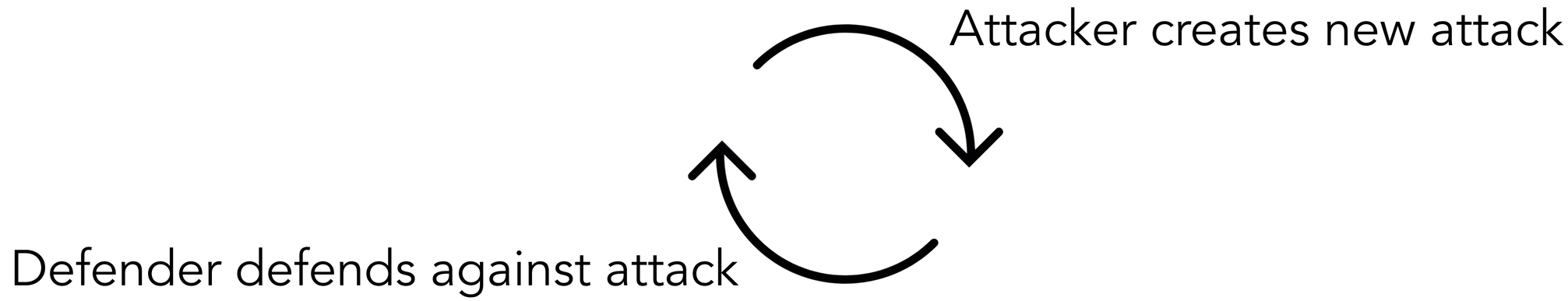




Problems with Risk-mitigation

You never win

- Created arms-race – forced co-evolution



Problems with Risk-mitigation

How do you know if you're making progress?

- How do you **evaluate** risk or reward?
  - How many points of security does antivirus give you? How do you measure those points?
- Big, existential question for the field: how do we measure security?
  - How do we do this in other fields? Are those strategies applicable here too?

Key meta-issues in security

Some areas to consider as you think about projects

- Policy – what makes a thing bad?
- Assets, Risks, Threats – what do I care about protecting, against what?
- Value – what's the cost if the bad thing happens? how much does it cost to prevent?
- Protection – *how* do I defend against threats? (e.g., technology)
- Deterrence – how might I *deter* the bad thing from happening in the first place?

Threat Models

The Security Mindset

Thinking attacker and defender

- To evaluate security, you need to think *both* like an attacker and a defender
- Thinking like an attacker
  - Understand techniques for circumventing security
  - Look for ways security *can* break, not reasons why it won't work
- Thinking like a defender
  - Know what you're defending and against whom
  - Weigh benefits vs. costs: **no system is ever "completely" secure**
  - Rational paranoia

Thinking like an attacker

Thinking attacker and defender

- First step: **weak links** – these are the easiest to attack
  - In real systems, this often requires important technical knowledge that comes from deep study and consideration
- Identify *assumptions* that security depends on. Are those assumptions always true? Under what conditions?
- Think outside the box – not constrained by the system designer's worldview

Exercise: Breaking into CSE after hours

How might you do it?

Thinking like an defender

Thinking attacker and defender

- Security Policy
  - What assets are we trying to protect, and what properties are we trying to enforce?
- Threat model
  - Who are the attackers? What are their capabilities? Motivations?
  - What kind of attack are we trying to prevent? What kinds of attacks should we ignore?
- Risk assessment
  - Rational paranoia, likelihood of risks
- Countermeasures
  - Costs vs. benefits?
  - Technical vs. nontechnical?

Threat modeling exercise: Should I lock your door?

- Assets?
- Adversaries?
- Risk assessment?
- Countermeasures?
- Costs/benefits?

Threat modeling exercise: Should I enter my CC into this website?

- Assets?
- Adversaries?
- Risk assessment?
- Countermeasures?
- Costs/benefits?

Break Time

Codeword:
Pumpkin-Spice

https://tinyurl.com/cse291attendance

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*

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?

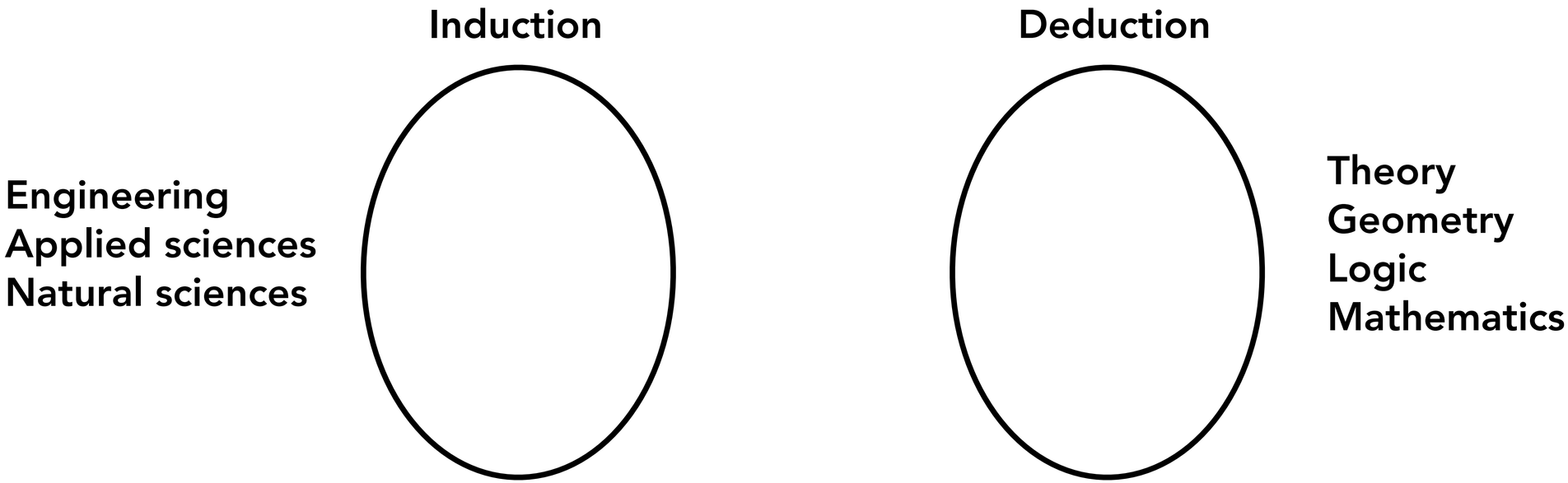
What is science?

Observation, experimentation, test

Two styles of research

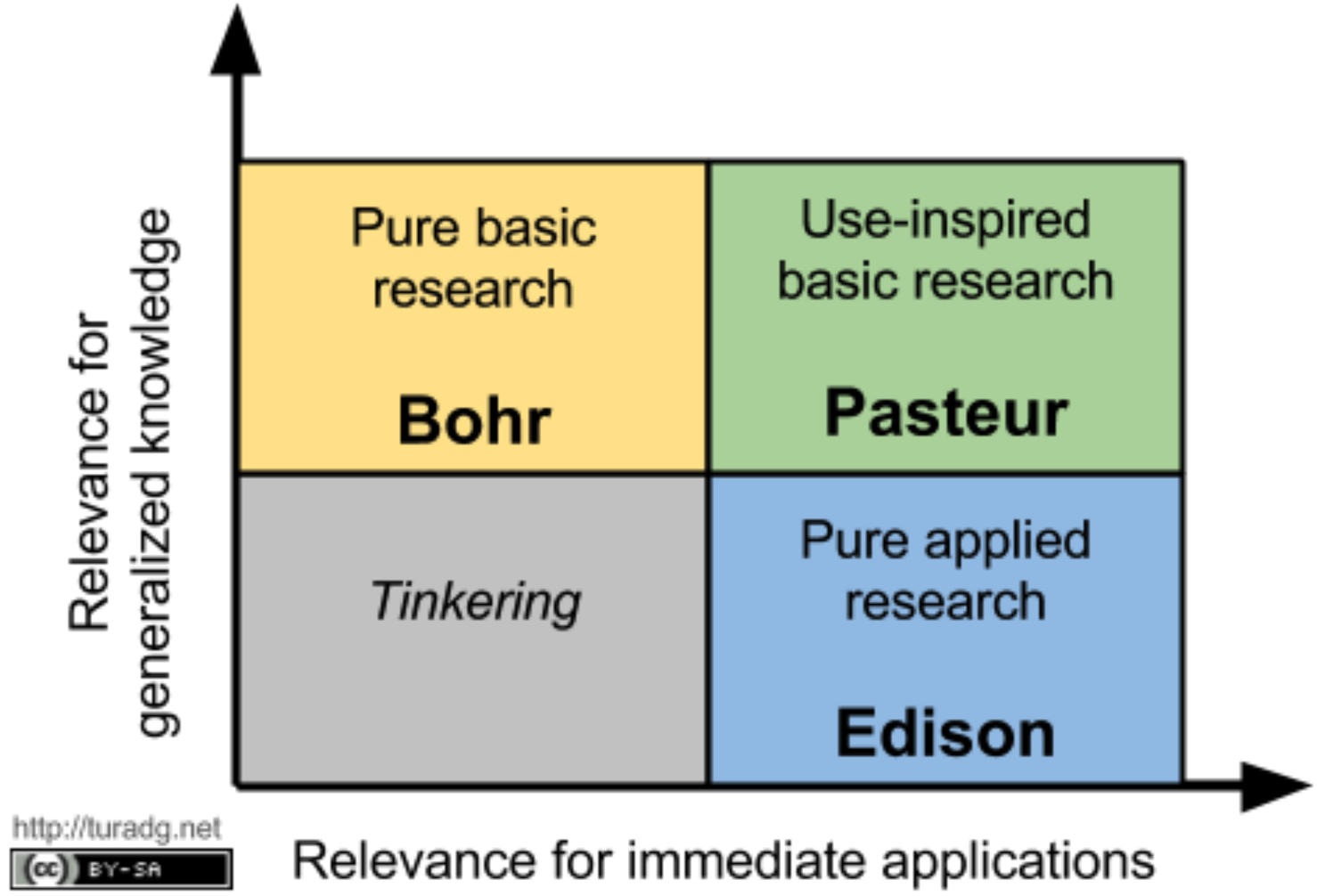
Induction vs. Deduction

- **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?"

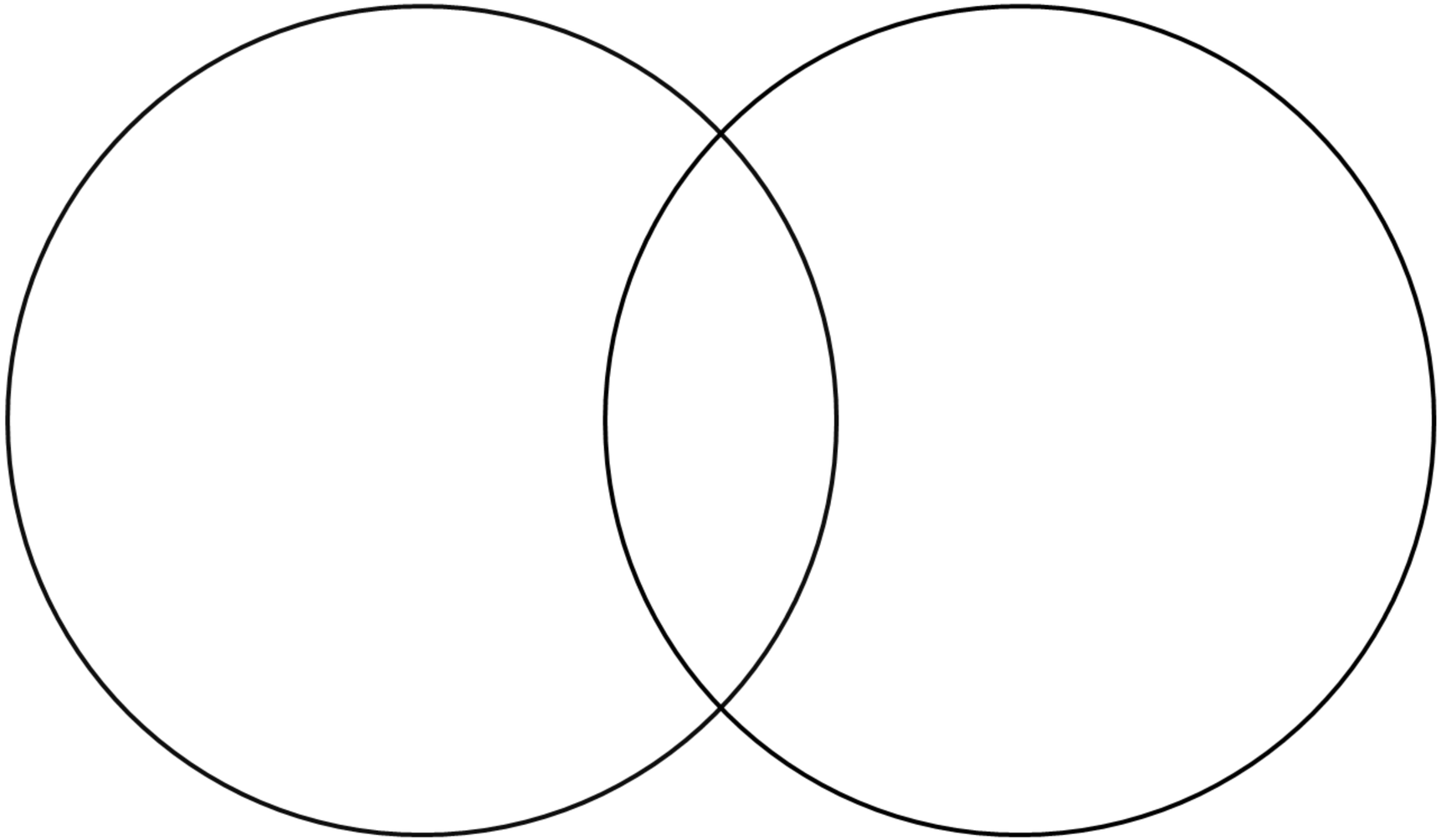
- In academia, we do *research;* science and scientific processes help us to make **sense** of our work and evaluate our claims
- Claim: *Changing passwords every 90 days is critical to protecting user accounts.*
  - **How might we study this?**

Practical advice for your research projects

Exercises to help think about projects

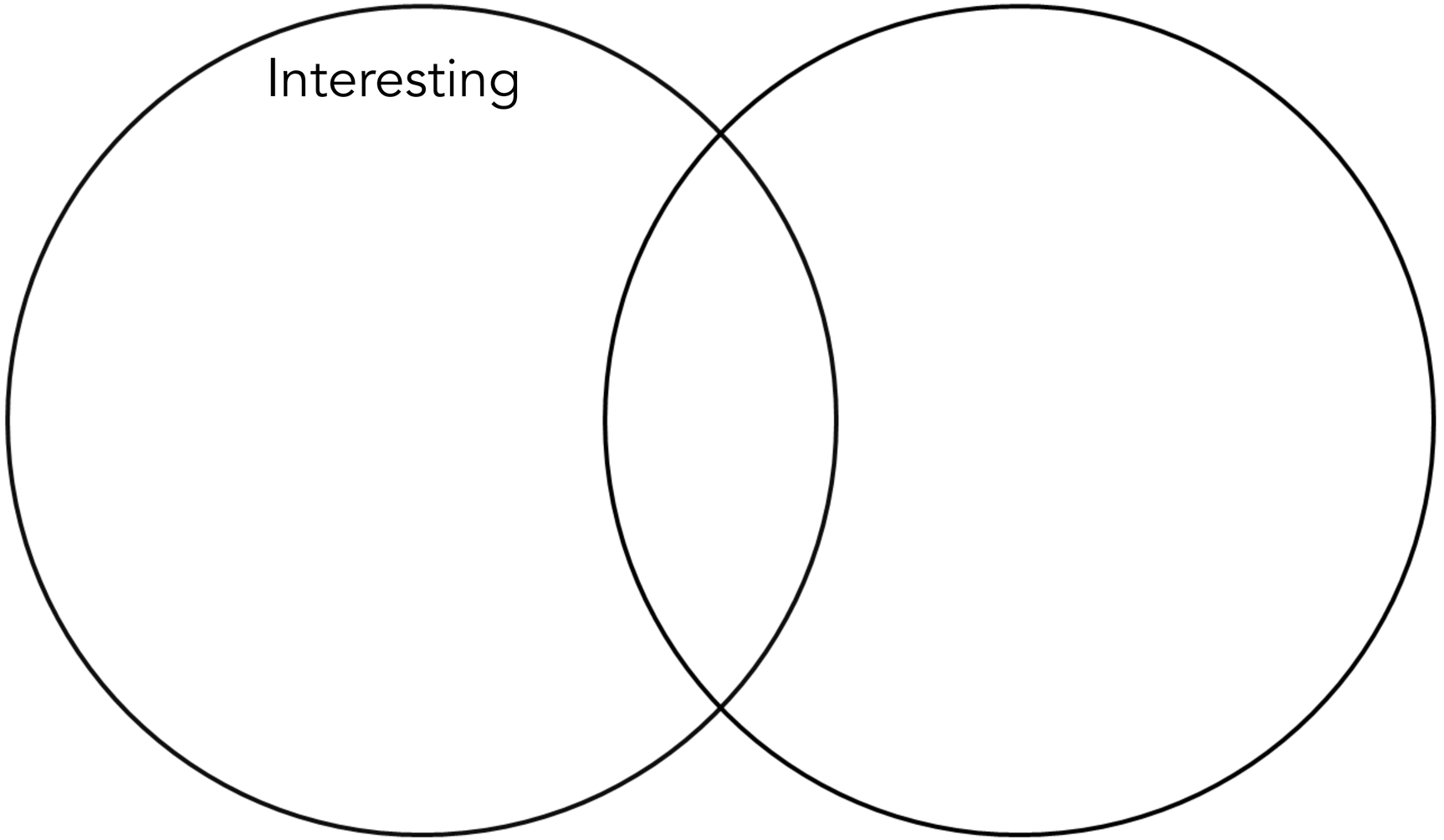
- Scanning / skimming papers in a related area
  - Say you're excited about the web… what parts of the web are you interested in? Why are you interested in them?
- Mesearch: What services do I use **in my life** that are critical to me? How can I evaluate whether they are secure?
- General "cool" factor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

Choosing a research problem



Choosing a research problem

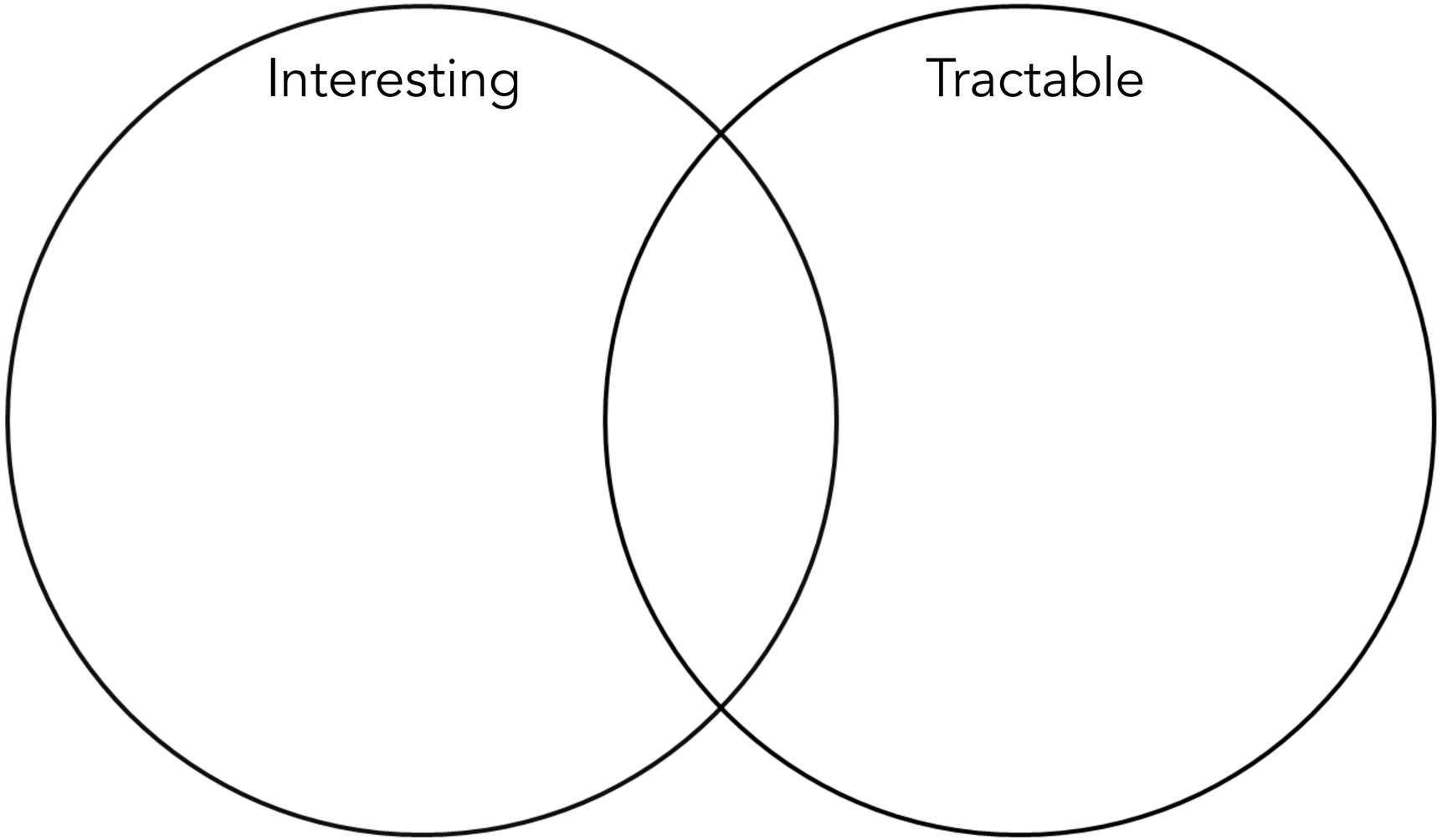
Choose Interesting Problems



- 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

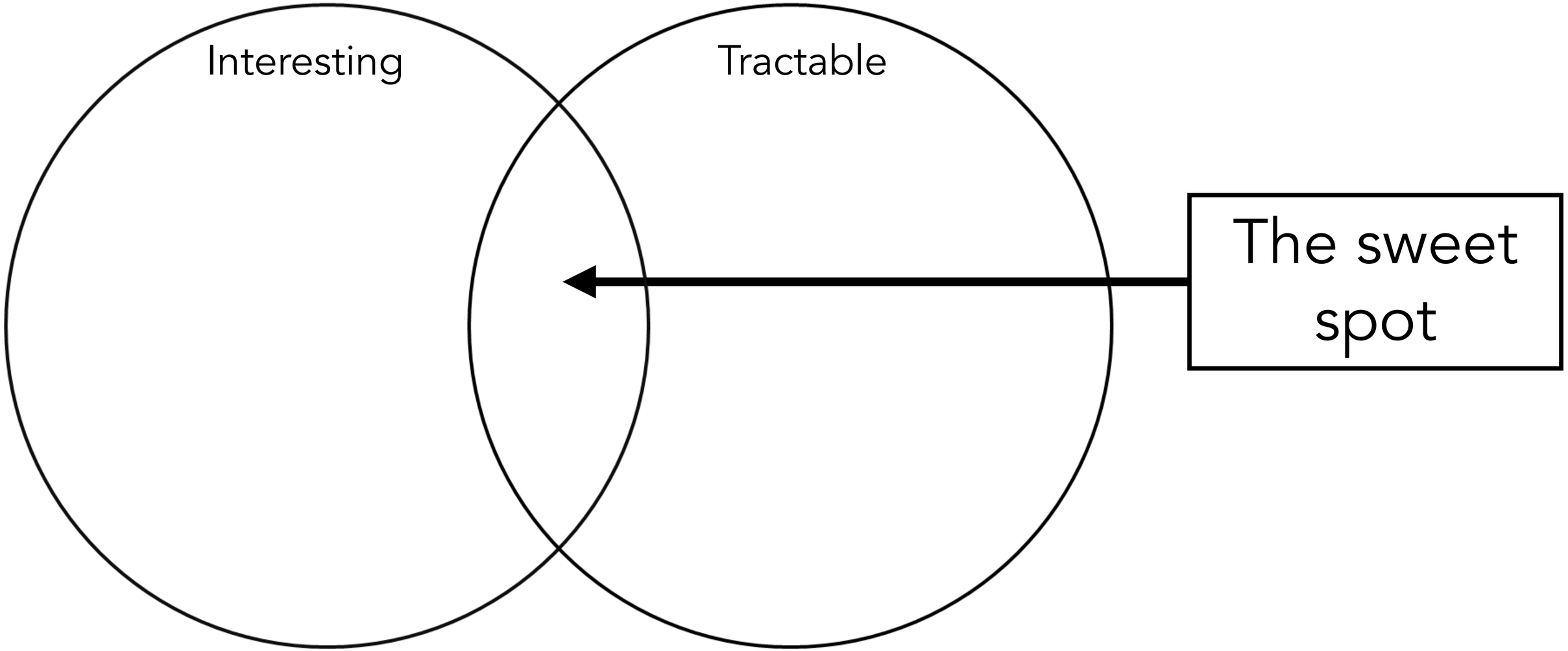
Choose Tractable Problems



- 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 this function. 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
  - 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)

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 collected 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)

Some random project ideas…

From my brain and my interests

- How do different antivirus providers behave with specific pieces of malware? When they're different, *why* might they be different?
- Demystifying the ad ecosystem on misinformation websites – who are the major advertisers?
- Studying Snakeoil ads on YouTube: LOSE FAT USING THESE THREE TRICKS!
- Building a system to ingest, index, and analyze the Digital Services Act data
- Do accounts on Reddit that engage in personal attacks *repeat offense* after being moderated? Why, where, and how?

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