



Welcome to Real-World Cryptography

CPSC 444/544, 2025 Spring

Instructor: Fan Zhang

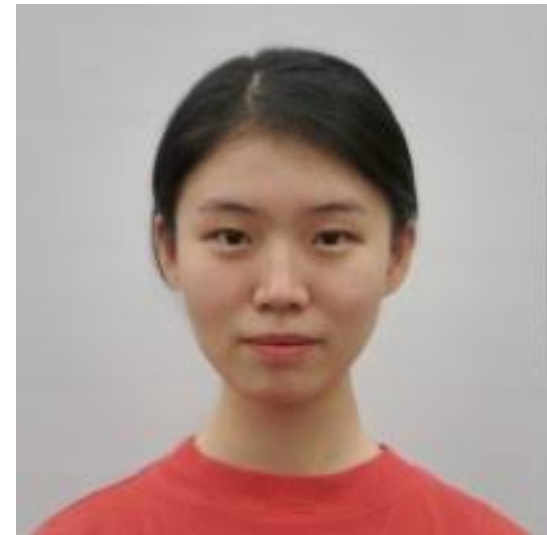
Yale



Course Staff



- Instructor: Fan Zhang
- Assistant Professor
 - PhD from Cornell ('20)
- Research areas:
 - Security
 - Applied Cryptography
 - Blockchains and decentralized systems
- Website: fanzhang.me



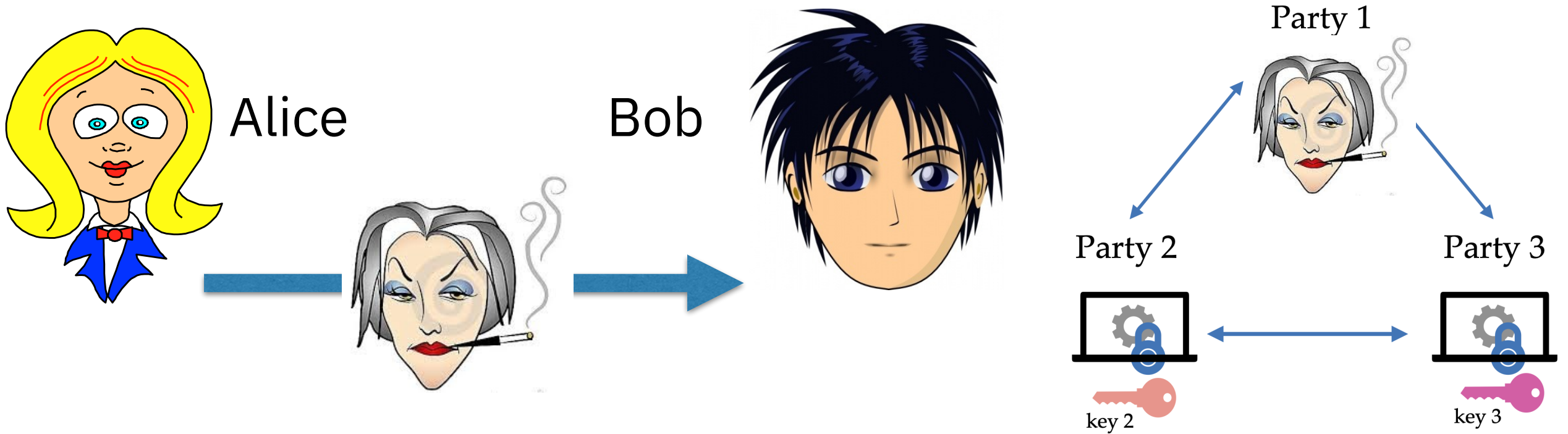
- TA: Yunhao Wang
- 3rd year PhD Student
- OH: 4-5 pm Thursday

This course

- Focuses on application of cryptography in real world systems.
 - TLS — **the green lock in your browsers**
 - Secure messaging (e.g., **Tor, OTR, Signal, WhatsApp**)
 - Anonymous credentials (e.g., **Cloudflare Privacy Pass, Google and Apple's Private Tokens, google age verification**)
 - Digital Identity (**OAuth, SSO, W3C DID** and **BlueSky**)
 - Trustworthy computation (**Intel SGX, TDX, Apple Secure Enclave, Google Titan, Amazon Nitro**)
 - AI systems (e.g., **Phala AI, Steg.ai (watermarking)**)

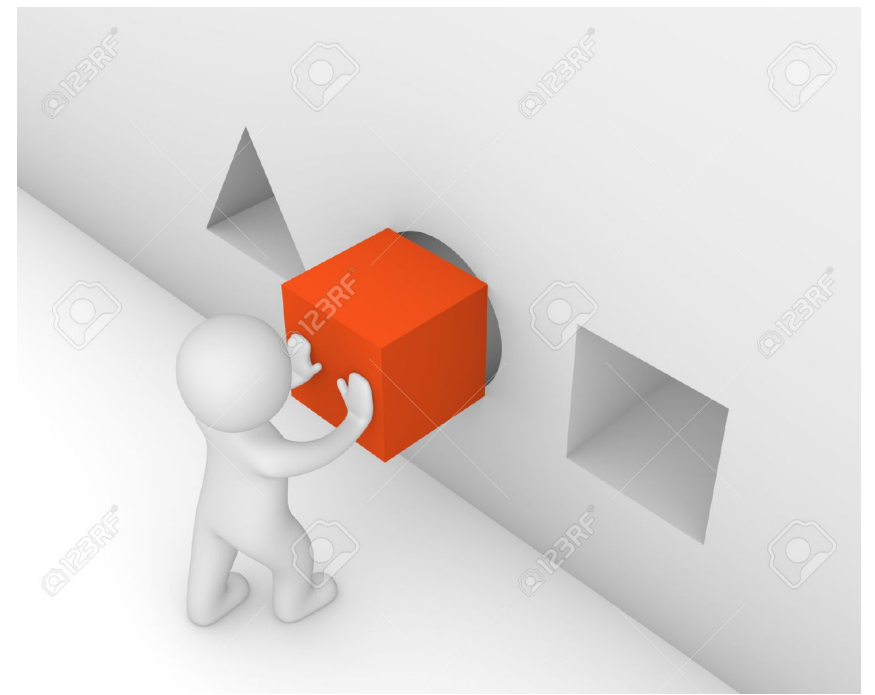
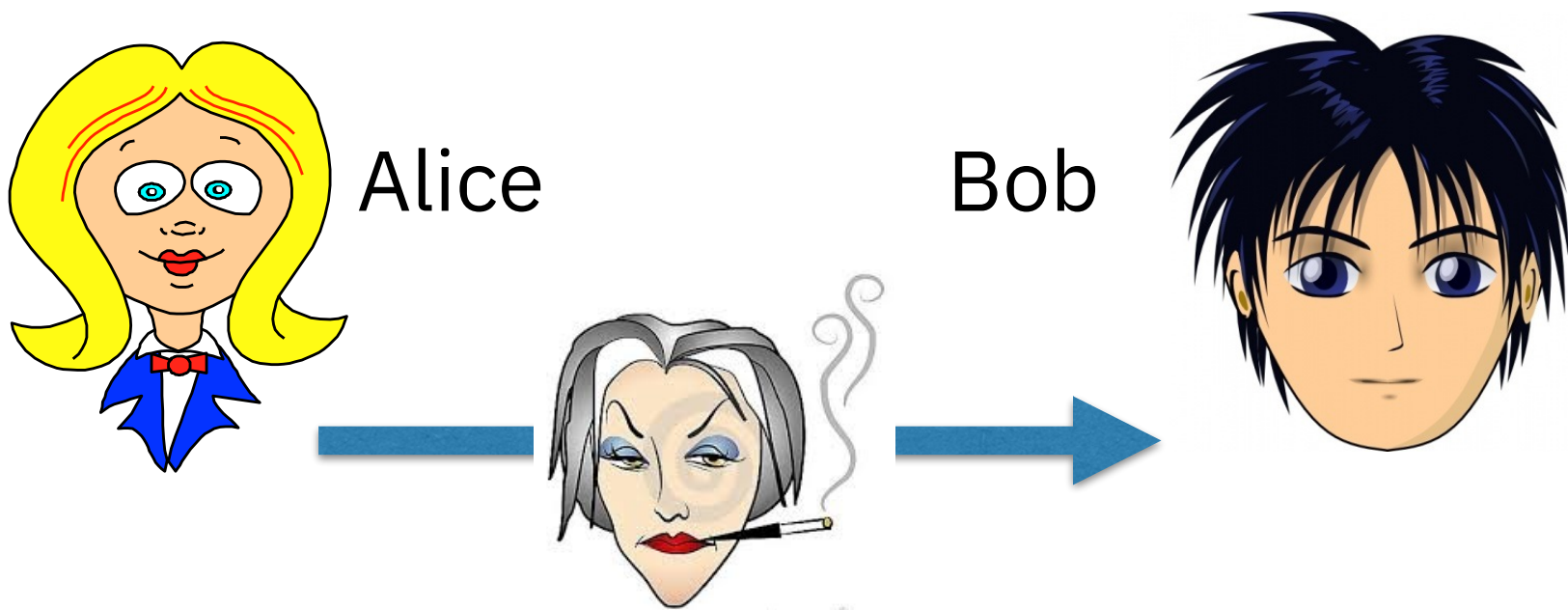
Course work

- Paper reading response
 - 5 papers scheduled; a few questions due before class to facilitate discussion
- Homework / Labs
 - Written problems, and programming in Python
- 5440 only: Independent project & presentation
 - topics will be suggested as we go



Real-World Cryptography

- Techniques for *secure communication* and *computation* in the presence of *adversarial behavior*.
- Basic approaches: Defining and proving security in well-defined mathematical models.

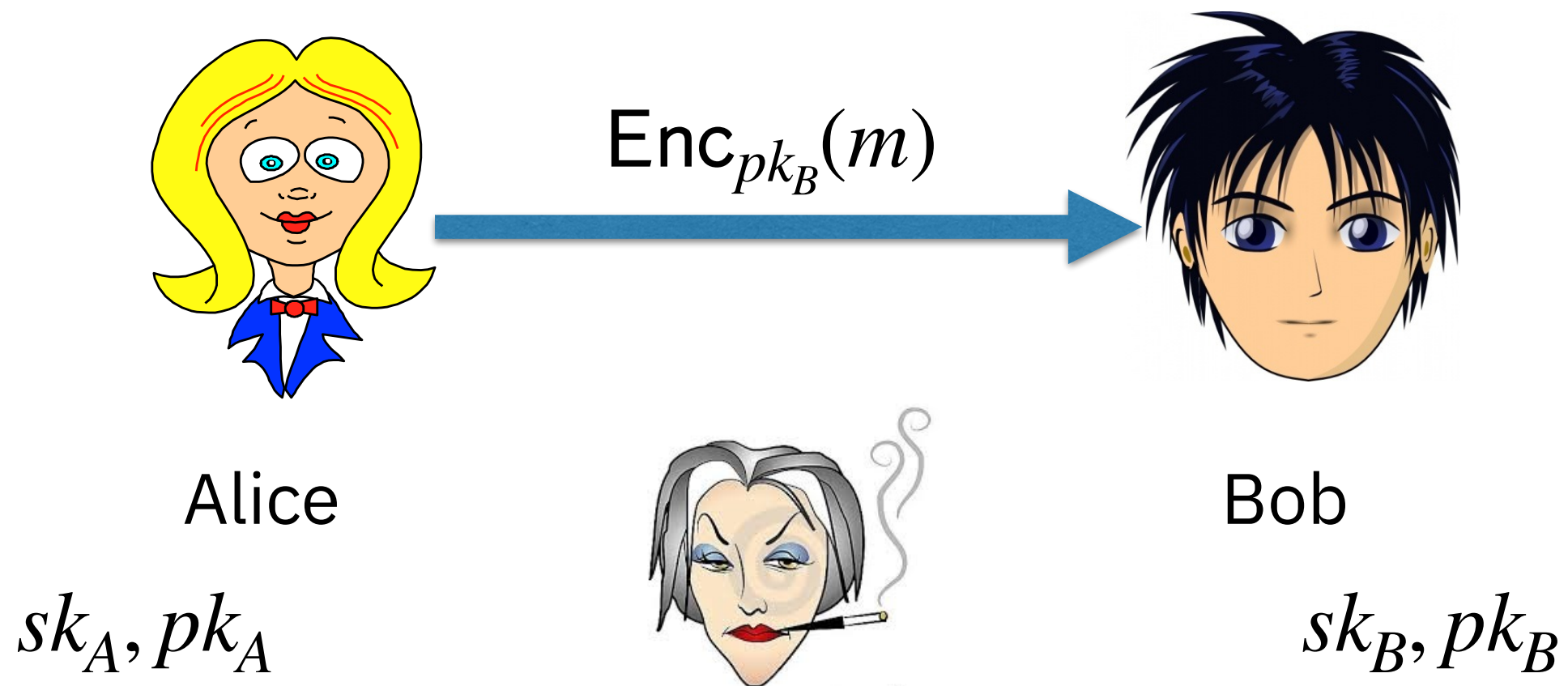


Real-World Cryptography

- Crypto is mathematically perfect, but why do we still see security breach? (Can you name a recent one?)
- We don't live in mathematical models.
- **Real problems: wrong models.** Wrong attacker model, wrong user model (human factors/behaviors)...
- How to get the model right? This is the art part.

Model vs Reality

- What real-world scenarios are captured by this model?
- What are ignored / unprotected?



Model vs Reality



Me

sk_A, pk_A

$Enc_{pk_B} \left(\begin{array}{c} \text{PDF} \\ \text{final_exam.pdf} \end{array} \right)$



Public Printer

sk_B, pk_B

Model vs Reality

- Question: who might be the adversary in this case?



Commander

sk_A, pk_A

Enc_{pk_B} (“Hey drone,
strike Bin Laden at
noon”)

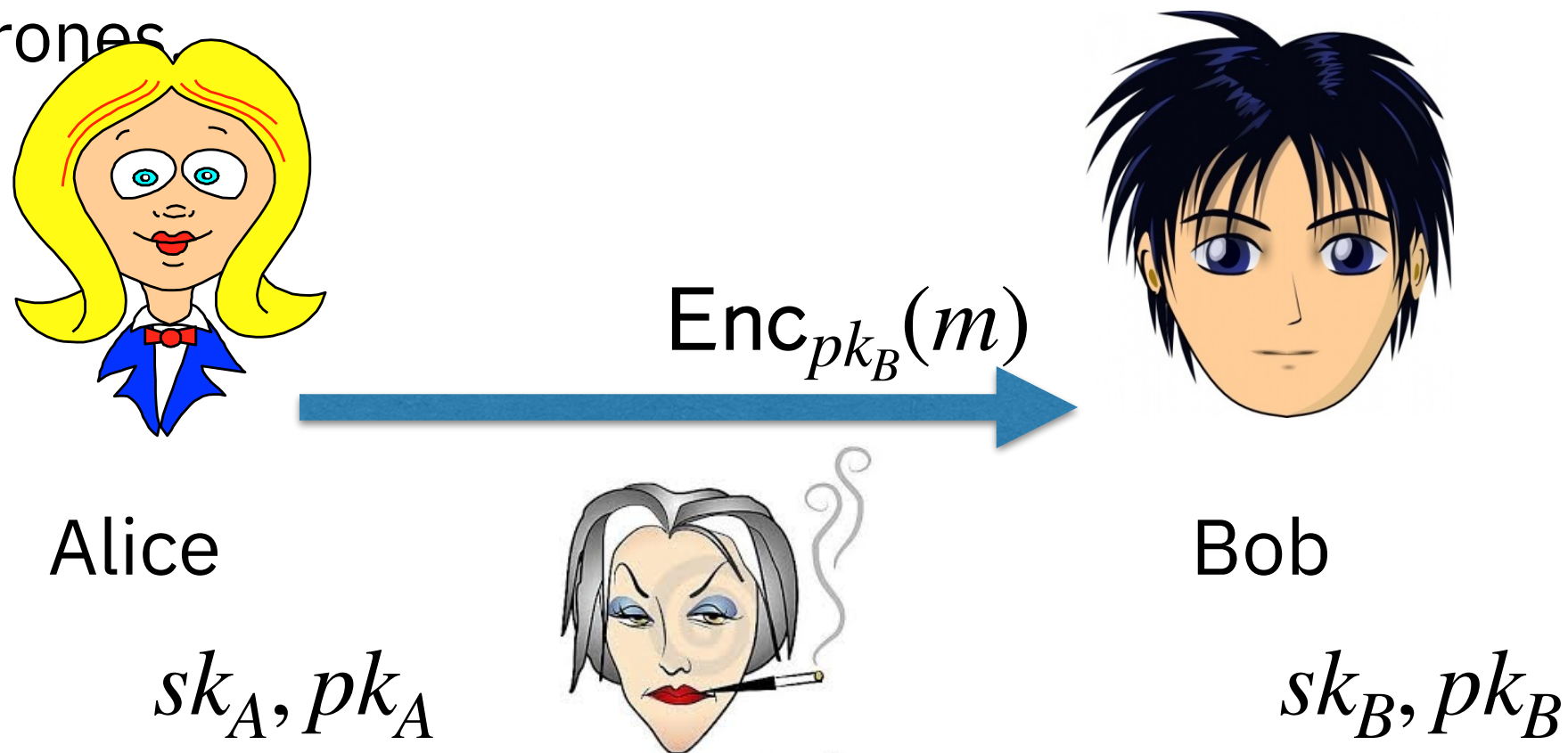


Drone

sk_B, pk_B

Model vs Reality

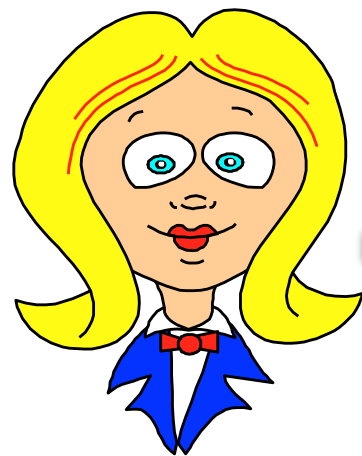
- It captures all kinds of secret messaging (goals) over insecure channel (adversary model).
- Amazing fact: cryptography is **commodity** now. In this model, my final exam gets the same protection as the war commands sent to drones



Model vs Reality

Can Alice get Bob's key correctly?

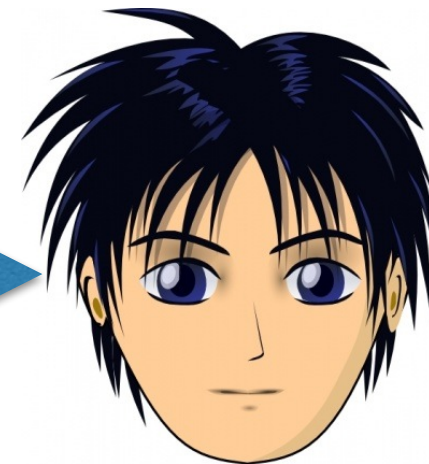
Who wrote the code?



Alice

sk_A, pk_A

$Enc_{pk_B}(m)$



Bob

sk_B, pk_B



Adv learns quite a lot!

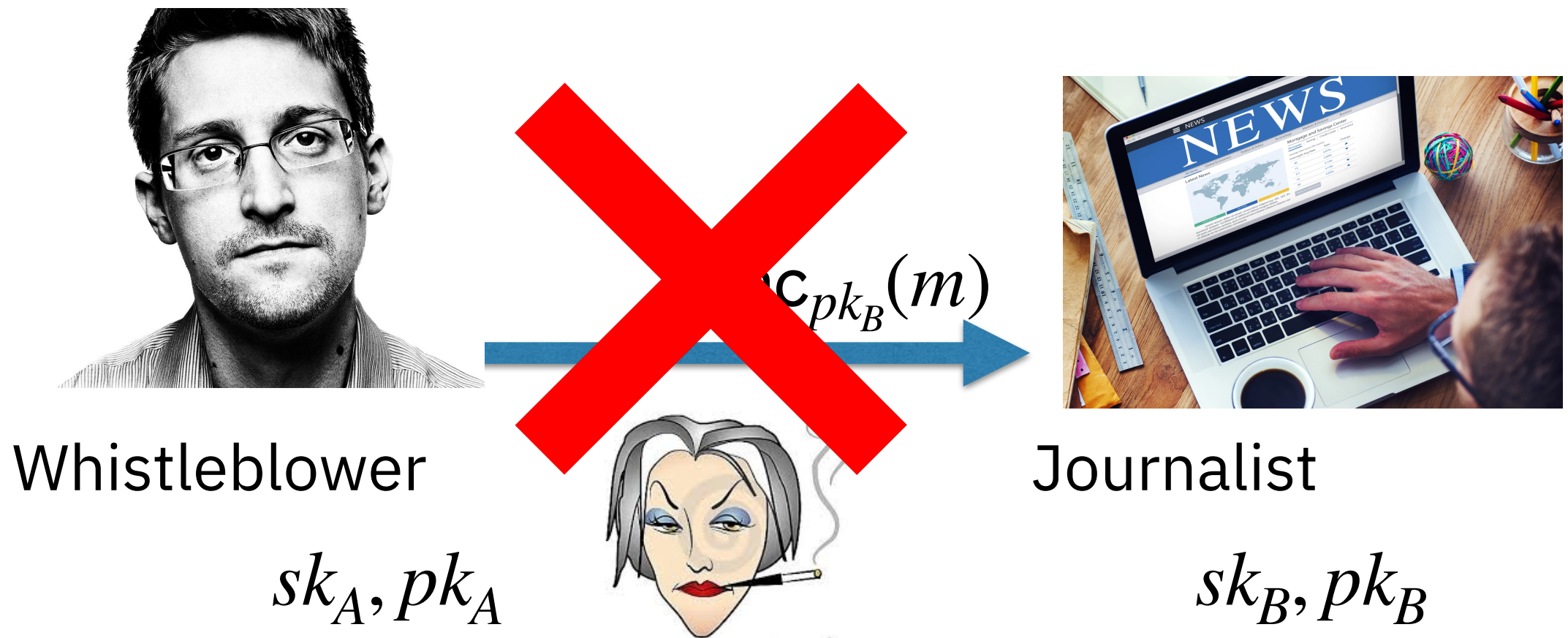
What if Bob loses his key?

Model vs Reality

- Alice may get ***malicious PK key***
 - E.g., browser trusting bogus certs
- Attacker may steal Alice's or Bob's ***secret key?***
 - E.g., Virus, trojans, lost devices
- What if Alice downloaded ***malicious software?***
 - E.g., supply chain attack
- Even none of the above happens, adversary still learns **important metadata**
 - E.g., the IP address of two parties

Model vs Reality

- Whistleblowers probably want to hide his IP address from adversary.



Quiz

- Is cryptography is a subfield of
 - (A) theory, or
 - (B) computer security?
- Explanation: The purpose of cryptography is to solve *security problem*.

Content of this course

- Learn the ***mindset*** and ***techniques*** that secure the ***substrate of digital society***
 - 1) adversarial mindset
 - 2) cryptographic building blocks
 - 3) socially important applications

Goals of the course

- More concretely, given a system design, you should understand:
 1. Potential security and privacy vulnerabilities and attacks, i.e., how things might break.
 2. The implications and cost of security and privacy failures, i.e., why/if should we care.
 3. Tools, techniques, and principles to use for defense (e.g., cryptography and more)

What's an adversarial
mindset?

“Security requires a particular mindset. Security professionals -- at least the good ones -- see the world differently. They can't walk into a store without noticing how they might shoplift...They can't vote without trying to figure out how to vote twice. They just can't help it.”

—Bruce Schneier



- Can card readers for this building be skimmed / cloned.
- Can the printer's billing systems be bypassed?
-

The adversarial mindset:

Four key questions

1. **Security goal:** What policy or good state is meant to be enforced?
2. **Adversarial model:** Who is the adversary? What is the adversary's space of possible actions?
3. **Mechanisms:** Are the right security mechanisms in place to achieve the security goal given the adversarial model?
4. **Incentives:** Will human factors and economics favor or disfavor the security goal?

Four key security goals

Confidentiality: Data not leaked

Integrity: Data or resource not tampered with

Availability: Data or resource accessible when needed

Authenticity: Correct belief in data or resource origin

(CIA + Authenticity)

Example: Adversarial Air Travel

— A bug in airport security (fixed circa 2007-2010)

Step 1: Home



Step 2: Security



Step 3: Gate



FRI, MAR 30, 2012		DELTA	
Diamond Testacct GT9549 / SKY PRIORITY		SkyMiles #XXXXXX9718 BOARDING DOCUMENT DIAMOND/ELITEPLUS/SKY CLUB	
JFK ▶ LAX		Alice	
NYC-KENNEDY (JFK) ▶ Los Angeles (LAX) FLIGHT DL120	BOARDING 8:20am	GATE* -	ZONE Sky SEAT 24C Economy (H) Depart Fri, 9:00am Arrive Fri, 12:20pm
*Gates may change. Check airport monitors.		Fly Paperless: www.delta.com, pp	
Ticket#: 006 2144236059			

FRI, MAR 30, 2012		DELTA	
Diamond Testacct GT9549 / SKY PRIORITY		SkyMiles #XXXXXX9718 BOARDING DOCUMENT DIAMOND/ELITEPLUS/SKY CLUB	
JFK ▶ LAX		Alice	
			

FRI, MAR 30, 2012		DELTA	
Diamond Testacct GT9549 / SKY PRIORITY		SkyMiles #XXXXXX9718 BOARDING DOCUMENT DIAMOND/ELITEPLUS/SKY CLUB	
JFK ▶ LAX		Alice	
NYC-KENNEDY (JFK) ▶ Los Angeles (LAX) FLIGHT DL120	BOARDING 8:20am	GATE* -	ZONE Sky SEAT 24C Economy (H) Depart Fri, 9:00am Arrive Fri, 12:20pm
*Gates may change. Check airport monitors.		Fly Paperless: www.delta.com, pp	
Ticket#: 006 2144236059			

What's the **security goal** for passport / ID checking?

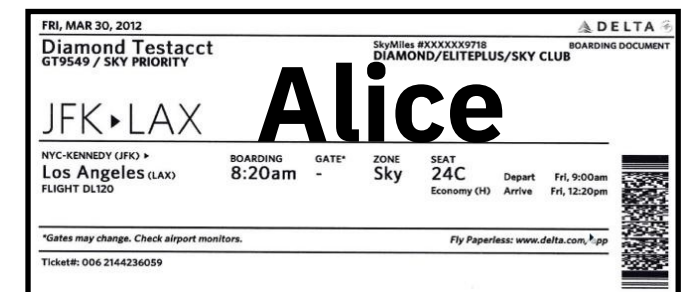
- Passengers on **no-fly lists** can't fly.
- Done in two steps
 1. Security check: verify passenger matches the name on boarding pass
 2. Gate check: verify the name on boarding pass is not on no-fly list.
- **Is this design secure?**

(Evil) Eve wants to get on a plane (she's on a no-fly list)

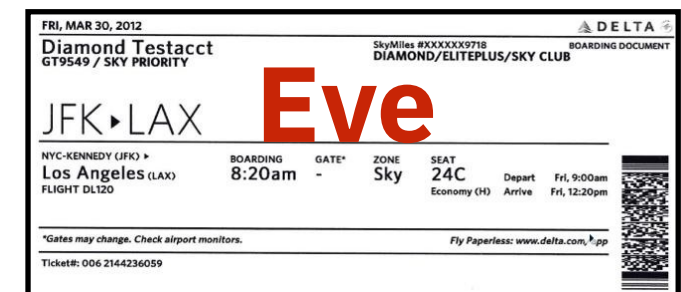


Eve

1. She buys a ticket in Alice's name (e.g. from eBay), and prints a boarding pass for Alice.



2. She also forges a boarding pass with name of Eve.



Example: Adversarial Air Travel

— A bug in airport security (fixed circa 2007-2010)

Step 1: Home



Step 2: Security



Step 3: Gate



FRI, MAR 30, 2012

Diamond Testacct
GT9549 / SKY PRIORITY

SkyMiles #XXXXXX9718
DIAMOND/ELITEPLUS/SKY CLUB

Alice

JFK ▶ LAX

NYC-KENNEDY (JFK) ▶
Los Angeles (LAX)
FLIGHT DL120

BOARDING 8:20am GATE* - ZONE Sky SEAT 24C
Economy (H) Depart Fri, 9:00am
Arrive Fri, 12:20pm

*Gates may change. Check airport monitors. Fly Paperless: www.delta.com

Ticket#: 006 2144236059

FRI, MAR 30, 2012

Diamond Testacct
GT9549 / SKY PRIORITY

SkyMiles #XXXXXX9718
DIAMOND/ELITEPLUS/SKY CLUB

Eve

JFK ▶ LAX

NYC-KENNEDY (JFK) ▶
Los Angeles (LAX)
FLIGHT DL120

BOARDING 8:20am GATE* - ZONE Sky SEAT 24C
Economy (H) Depart Fri, 9:00am
Arrive Fri, 12:20pm

*Gates may change. Check airport monitors. Fly Paperless: www.delta.com

Ticket#: 006 2144236059

FRI, MAR 30, 2012

Diamond Testacct
GT9549 / SKY PRIORITY

SkyMiles #XXXXXX9718
DIAMOND/ELITEPLUS/SKY CLUB

Eve

JFK ▶ LAX

NYC-K
Los
FLIGHT

NEW YORK STATE
Commissioner of Motor Vehicles

ENHANCED DRIVER LICENSE

ID: 012 345 678 CLASS D

DOCUMENT
SAMPLE, LICENSE
2345 ANYPLACE AVE
ANYTOWN NY 12345

Eve

DOB: 06-09-85
SEX: F EYES: BR HT: 5-09
E: NONE
R: NONE
ISSUED: 09-30-08 EXPIRES: 10-01-16

8AJ120T521

FRI, MAR 30, 2012

Diamond Testacct
GT9549 / SKY PRIORITY

SkyMiles #XXXXXX9718
DIAMOND/ELITEPLUS/SKY CLUB

Alice

JFK ▶ LAX

NYC-KENNEDY (JFK) ▶
Los Angeles (LAX)
FLIGHT DL120

BOARDING 8:20am GATE* - ZONE Sky SEAT 24C
Economy (H) Depart Fri, 9:00am
Arrive Fri, 12:20pm

*Gates may change. Check airport monitors. Fly Paperless: www.delta.com

Ticket#: 006 2144236059

Where's the mistake?

- The **adversarial model** should include boarding pass tampering, but doesn't.
- Assumption: pass that's *issued* is pass that's *presented*
- I.e., The boarding pass lacks **integrity**...

The adversarial model used to
be different

- Alaska Airlines introduced home-printable boarding passes in 1999.
- Before that time, boarding passes were printed on special card stock.
- Security mechanism to protect integrity—passes were harder to modify



The real story

- Such security vulnerabilities had been widely known
- In 2006, IU PhD student Christopher Soghoian created a website that generated fake airline boarding passes.
- Website taken down. Arrested by FBI.
- Next day, DHS hired him to fix boarding pass security flaws.
- Long story short: **security flaw fixed after being made embarrassingly obvious (think incentives)**



Integrity forgotten in adversarial model in many, many other places!

- Cookies: states stored on a client's browser.
- E.g., user is shopping at an e-commerce site.



Simple cookies lack integrity

- Clients can *tamper with* cookies (“cookie poisoning”).

E.g., Edit Cookies Firefox extension

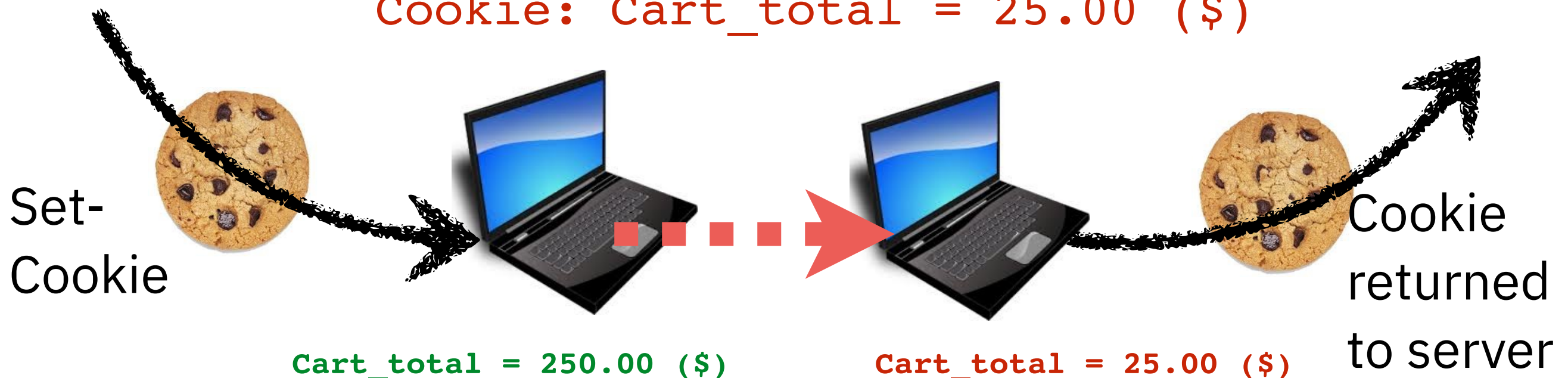
- Example:

E-commerce site executes

Set-Cookie: Cart_total = 250.00 (\$)

Before paying, user substitutes

Cookie: Cart_total = 25.00 (\$)



Quiz (incentives)

- What did Richard Clarke's famous quote say: “**many company spend more on _____ than on IT security**” (paraphrased)
- a) employees insurance
- b) utility bills
- c) coffee

Who is the adversary?

It depends on who you are

“Hackers”

- Breaking into your **own devices** or **the services you use** to steal secret data.
 - E.g., Equifax data breach impacted almost half of US population
- Hard to eliminate all bugs: implementation errors, social engineering, misused cryptography
- But, better security design can help. We will talk about:
 - Proper password hashing can *slow down* extraction
 - Honey objects can *detect* breaches.
 - E2EE services (e.g. Signal) means there is nothing to steal

Privacy-prying companies



- Economic [e.g. ads] and functional [e.g., authentication] requirements led to extensive tracking on Internet
- We will talk about how cryptography can balance privacy and functionality, e.g.,
 - Privacy-preserving authentication: proving you have access without revealing who you are (e.g, Cloudflare, Google, Apple)
 - privacy-focused advertisement: measuring ads efficacy without individual-level tracking (e.g., Meta)

Privacy-prying companies

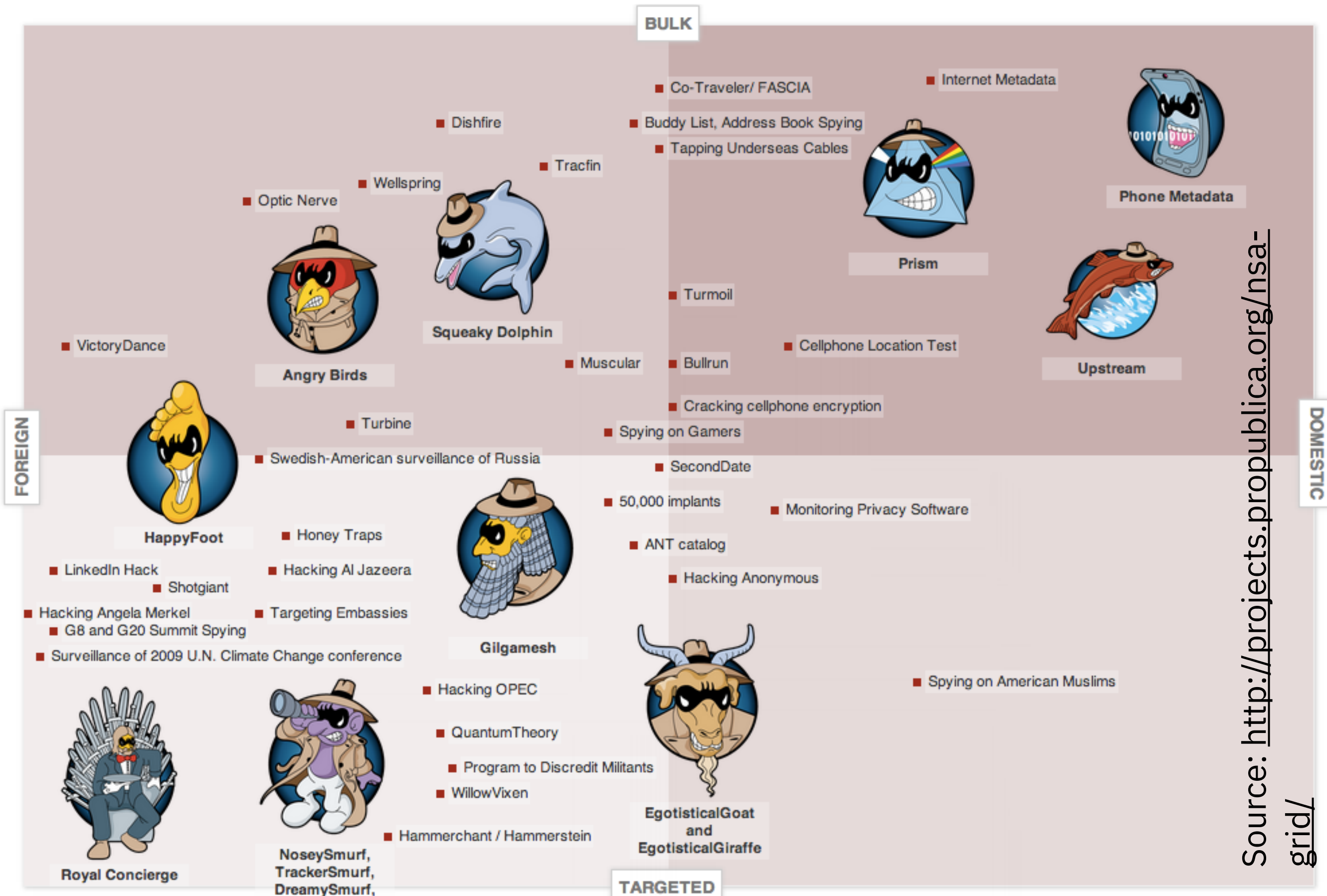


- Centralized social media platforms (e.g., X, YouTube, TikTok) are increasingly *influential* by controlling user identity, moderation, and ranking algorithms.
 - E.g., what X promotes or demotes can impact US election.
- We will talk about new architecture called **Decentralized social media** (e.g., BlueSky)
 - Goals: enables portable accounts, user-chosen feeds and moderation, and competition among service providers.

Rogue Governments

- Governments typically have intelligence agency for *national security*, but they increasingly want to know more about *their citizens*.
- Thanks to Snowden we now know a little bit more about what NSA is up to.

NSA Grid



Source: <http://projects.propublica.org/nsa-grid/>

What can a *state adversary* do?

- Block certain network traffic (censorship)
- Mandate bogus root TLS certificates in browser (MITM attack)
- Mandate installation of spyware on personal devices.
- Examine electronic devices at air ports or raid home.
- Instrument major ISP to monitor network traffic
- Produce misinformation (e.g., using gen AI)
- **Cryptography can prevent *some* of these powerful attacks, as we will see.**

Adversary evolves

- genAI posts new threats, e.g.,
 - Deepfake (e.g., Finance worker pays out \$25 million after video call with deepfake ‘chief financial officer’ - CNN)
 - LLM-crafted phishing bypassing spam filters
 - 45% of AI code has known security vulnerabilities (source)
 - LLMs access & store highly sensitive user data
- Where cryptography *may* help
 - Provenance (e.g. AI watermarking)
 - Confidentiality and Integrity of AI services (e.g. with TEEs)

Administrativa

Resources

- Schedule at: <https://fanzhang.me/teaching/444>
- References, links, lecture notes/videos, etc will be posted there
- We use Canvas for homework

Etiquette

- On-time class attendance is expected
 - Let me know if you can't
 - Matter of respect for instructor, especially guest lecturers
 - Pop-up quiz will be inserted randomly
- Academic Integrity violation will be prosecuted to the fullest extend possible.
 - AI policy next page
- To get H (or A), I expect not only great grades, but also active engagement throughout.

AI Policy

- What AI policy should we enforce?
 - Also an adversarial game?
- Rule of thumb: Think of AI as your classmate
 - If it would be inappropriate to ask a classmate to do, it's inappropriate to ask AI to do.
 - Help received from classmates should be cited and disclosed, which also applies to AI.
- Homework comes with more precise instructions.

Ethics

- Ideally, you'll come out *thinking like a criminal mastermind*, but *behaving like a gentlewoman/gentleman*.
- What we build matters.
 - Cryptography is not neutral math, and cryptographers are morally responsible for how their work affects society [*The Moral Character of Cryptographic Work* by Phillip Rogaway]

ANY QUESTIONS?



makeameme.org