Authentication

What does it mean for some data (or information, or files) to be **secure**?

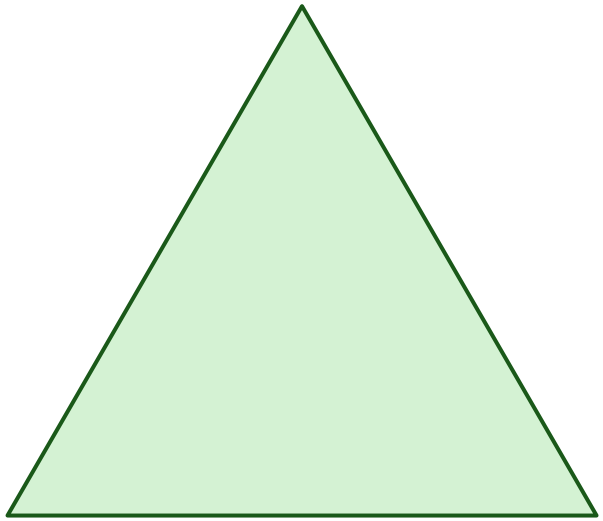
Think of student records (name, address, units & grades …)

CIA

Not this kind of CIA …

CIA

CONFIDENTIALITY



INTEGRITY

AVAILABILITY

CIA

Confidentiality: Restricted / secret data is only available to people who are permitted to see it.

Integrity: Data can only be modified or deleted by people who are permitted to do so.

Availability: People who are *permitted* to read or write data, are also *able* to do so.

Message Authentication Codes

Integrity

"Internet" scenario: attackers might be able to

send a message twice

block a message

modify a message

insert messages of their own

redirect messages

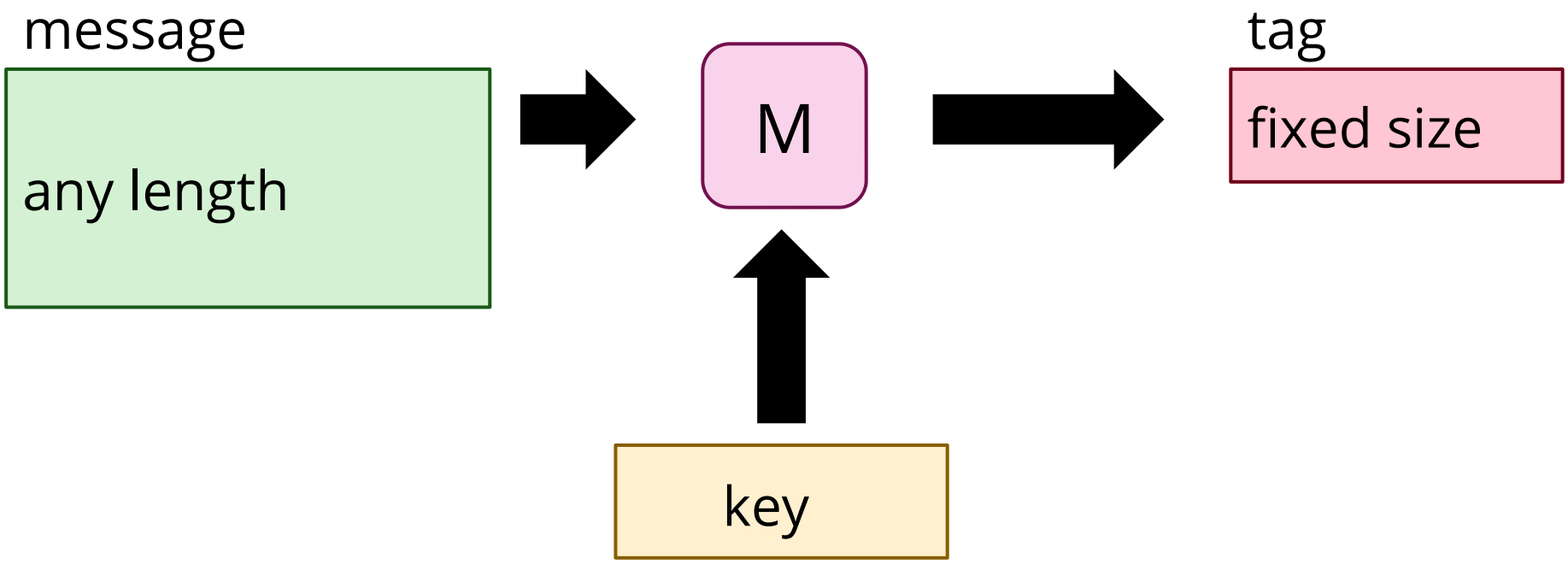
...

Active and Passive attacks

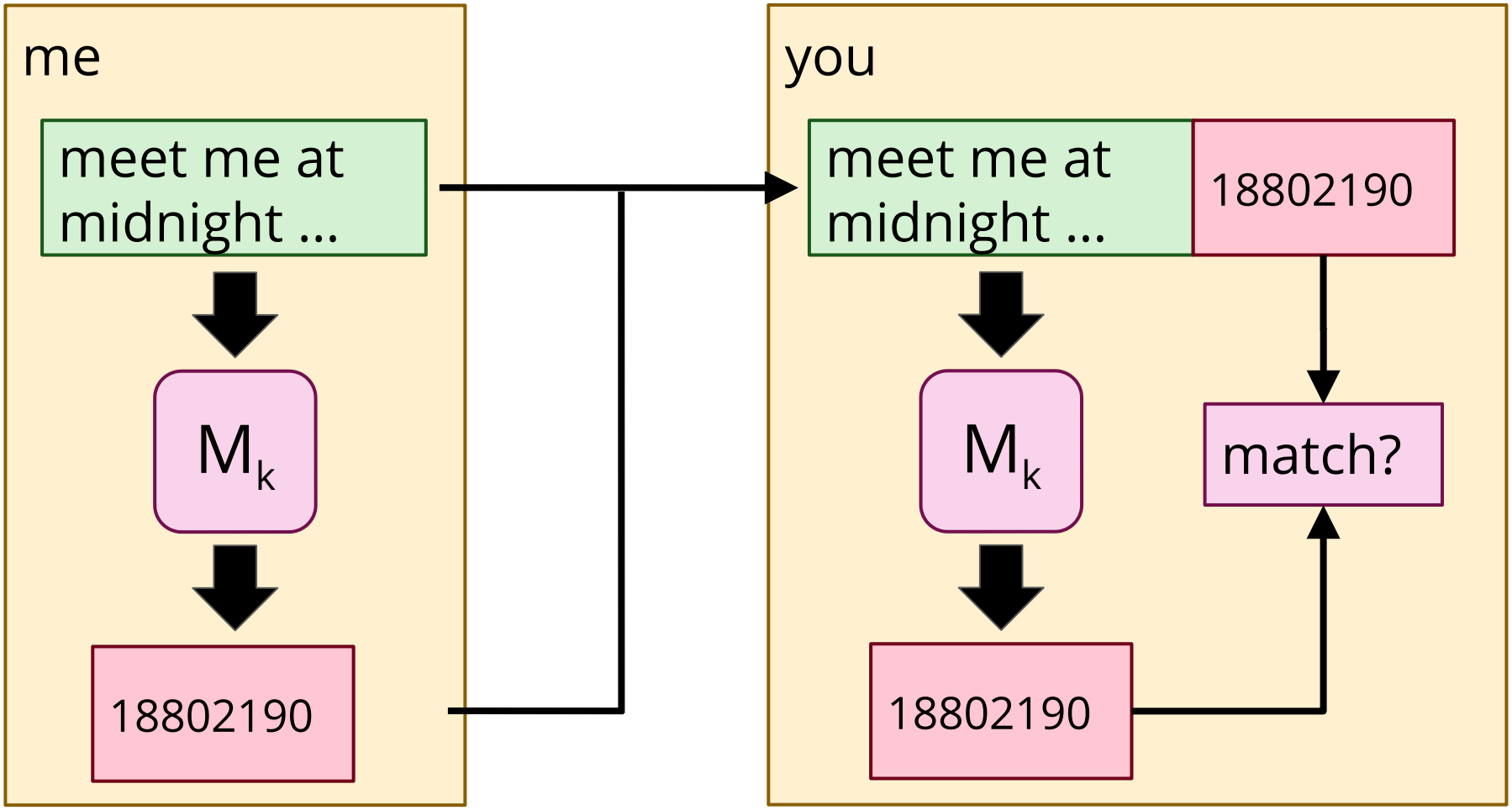
A **passive attack** only observes data, but does not modify anything.

An **active attack** may insert, delete, change ... data.

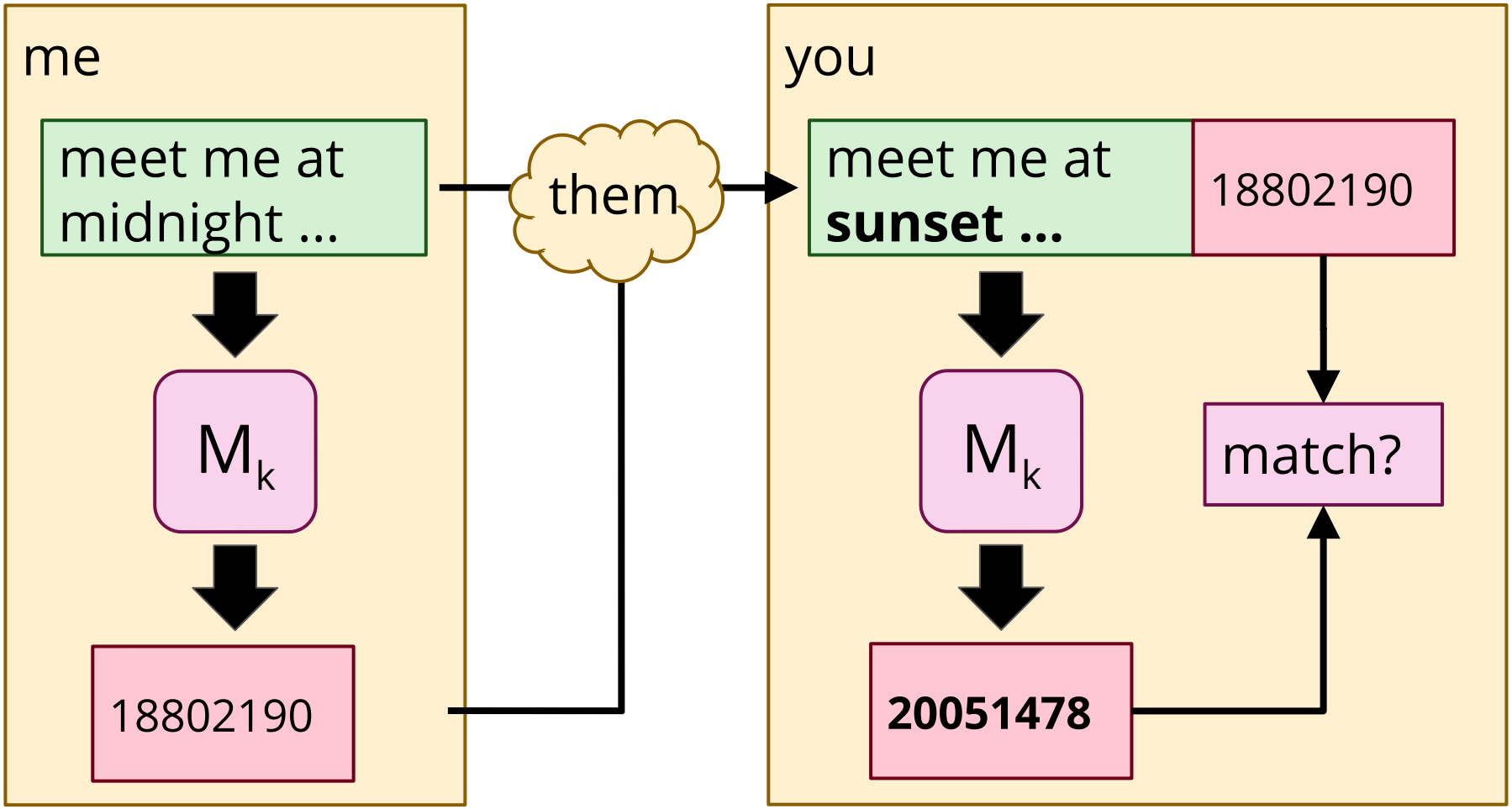
MAC = message authentication code



The idea


me
meet me at midnight ...
M_k
18802190
you
meet me at midnight ...
18802190
M_k
match?
18802190

The idea


me
meet me at midnight …
M_k
18802190
them
you
meet me at **sunset …**
18802190
M_k
match?
20051478

Integrity

You can detect if the message received was the same one that was sent, or whether it has been modified in transit.

A **MAC** (message authentication code) is a single, deterministic function M(k, m).

Without k, it should be hard to find M(k, m) for any m.

Things that are not MACs

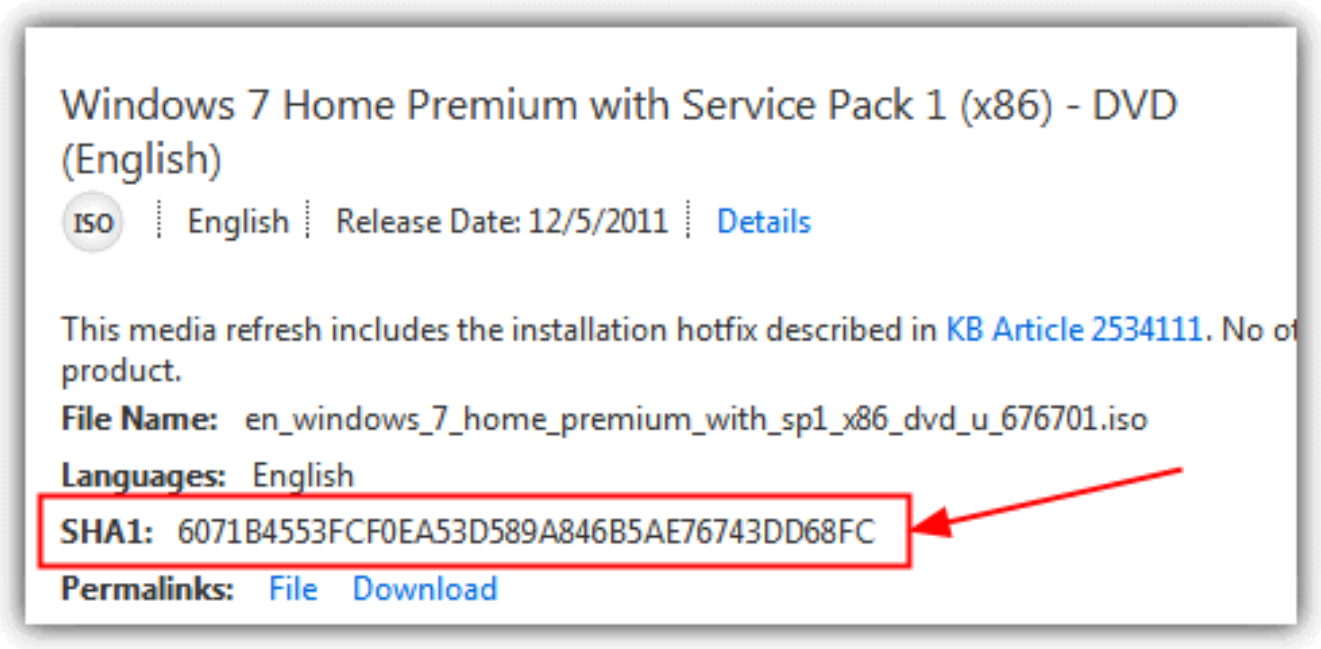
Checksums - as used in barcodes, QR codes, ISBN numbers … (but not bank account numbers!)

Hash functions (on their own).

None of these provide any *integrity* - an attacker can just recompute the checksum / hash of the new message.

A MAC needs a *key*.

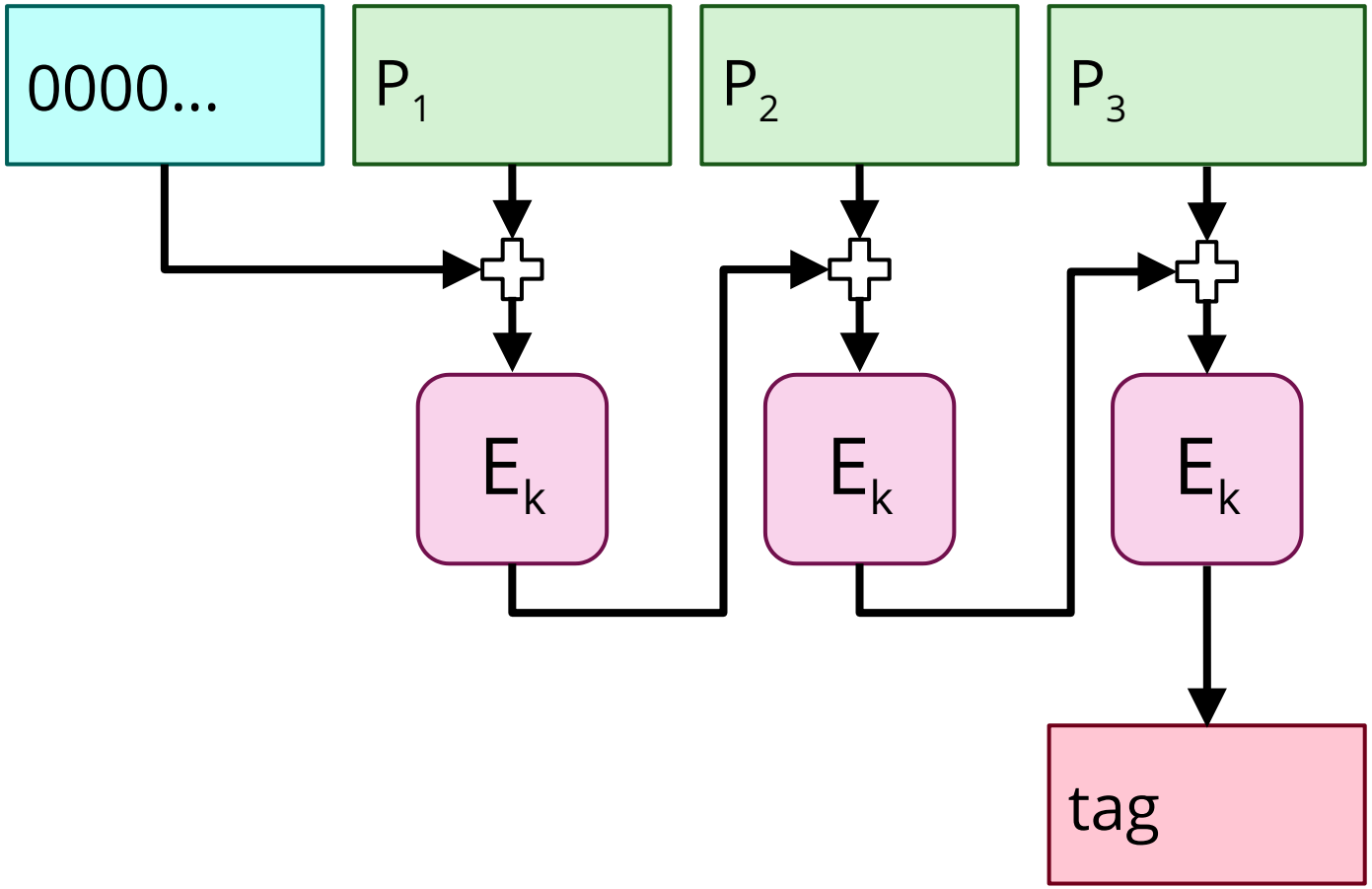
Things that are not MACs



source: raymond.cc
(comparing several file hash utilities)

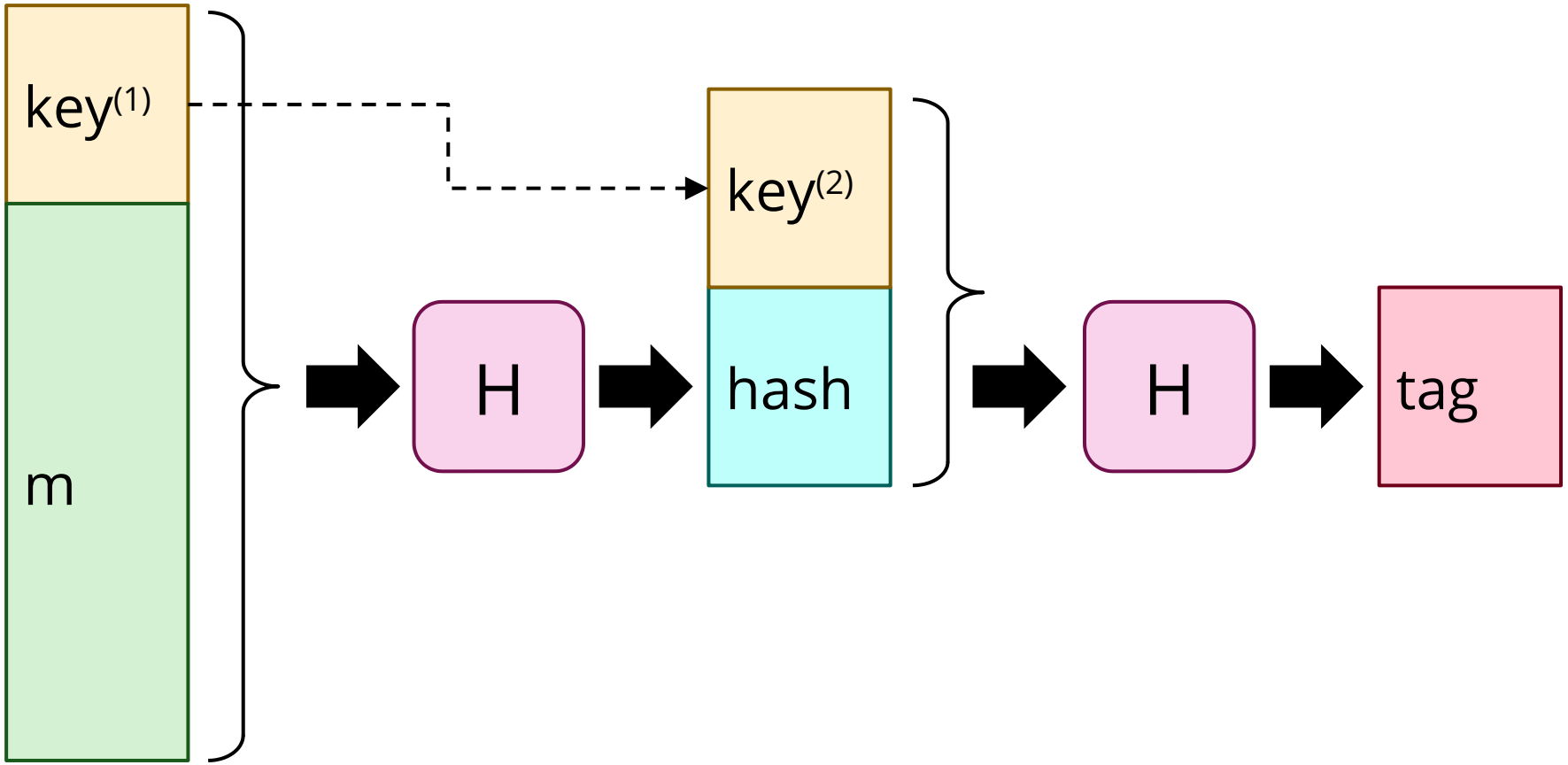
CBC MAC

MAC = last block of CBC with IV = 0.



HMAC

Hash-MAC: prepend (twiddled) key and hash, then repeat.



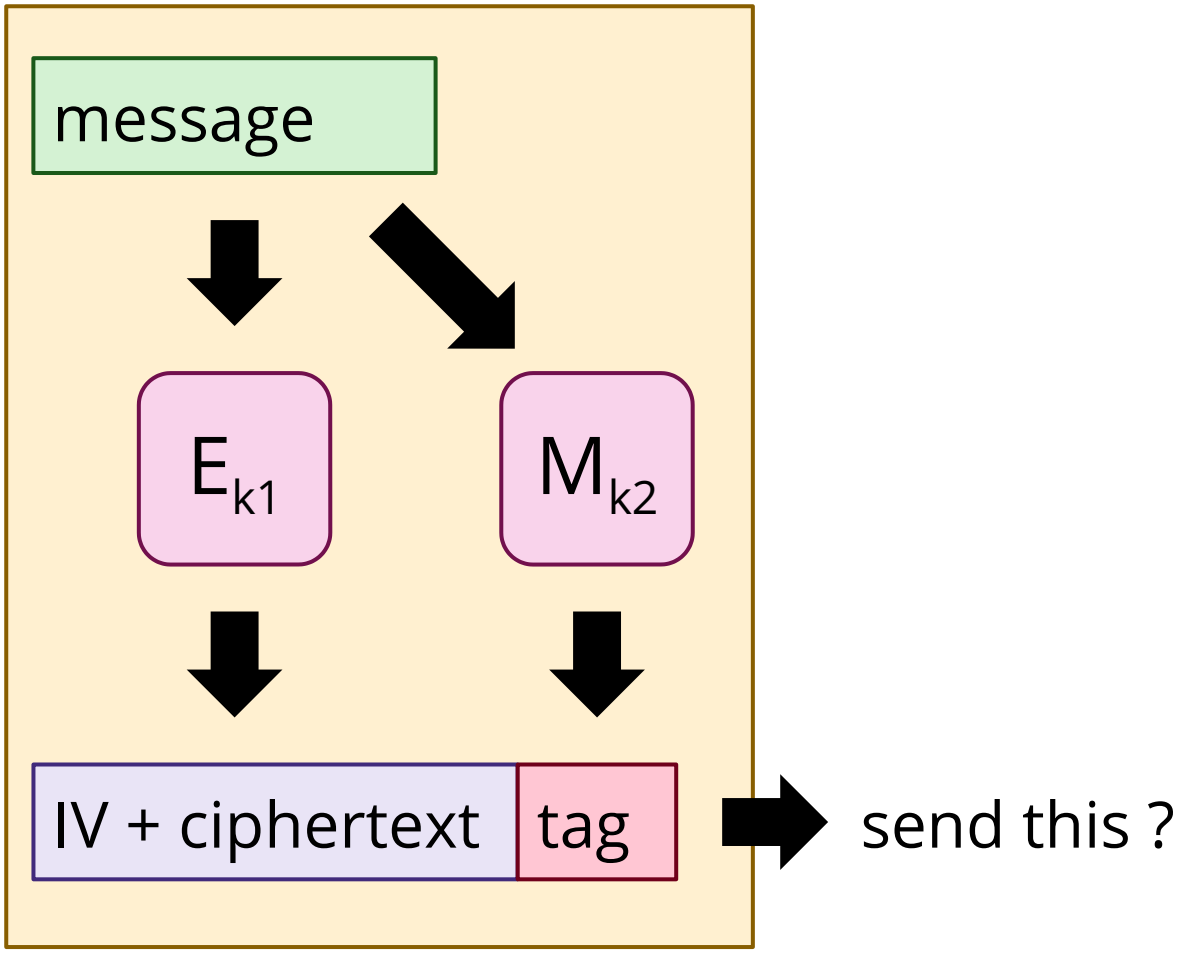
short summary

Encryption (e.g. block cipher in CBC mode) provides *confidentiality*.

A *MAC* provides *integrity*.

What if we want both?

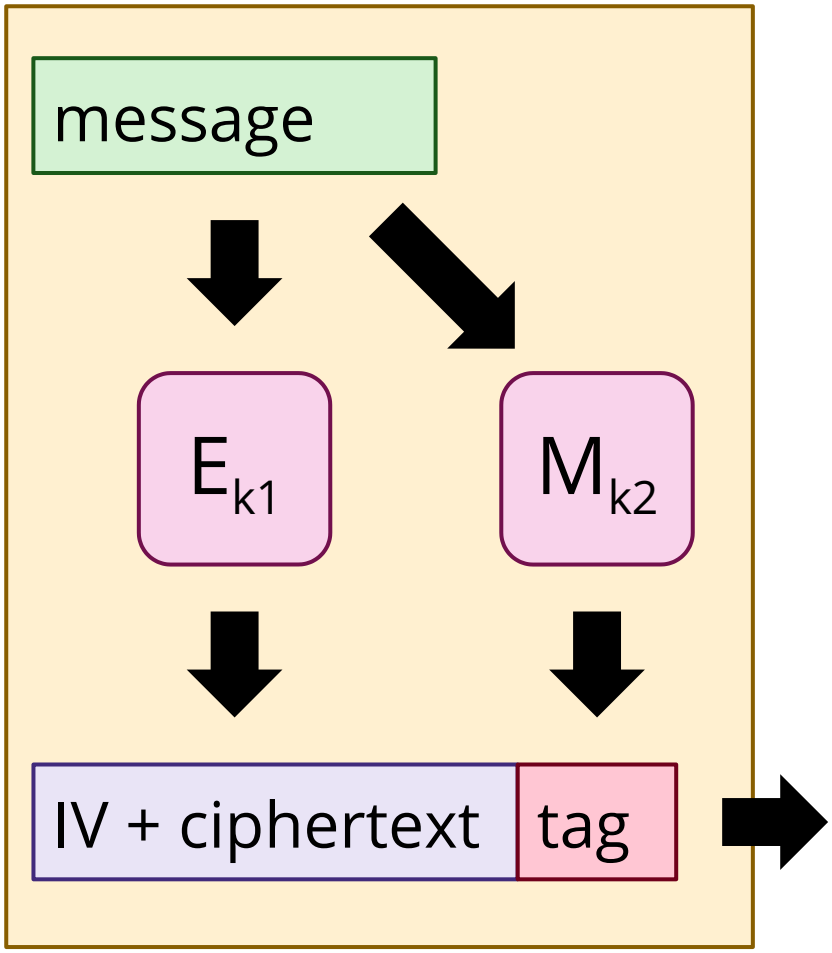
How not to do it



Cryptographic protocols do *exactly* what they were designed for, and no more.

Lots of mistakes happen when people combine features without understanding the implications.

How not to do it



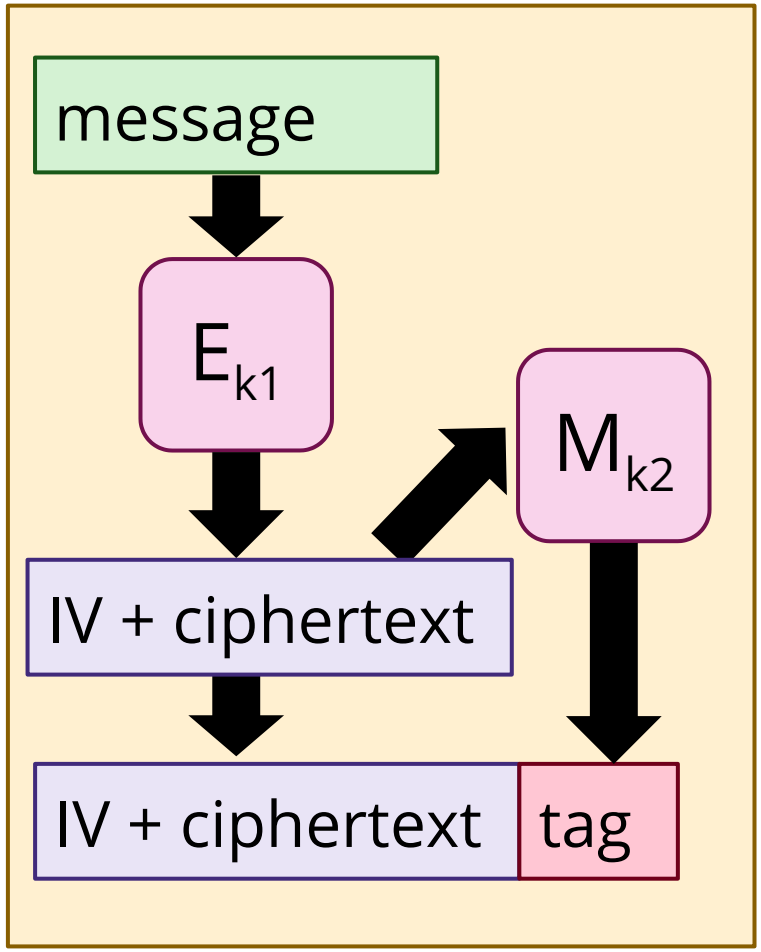
This is called encrypt-and-MAC.

The problem is, the MAC does not need to keep anything confidential.

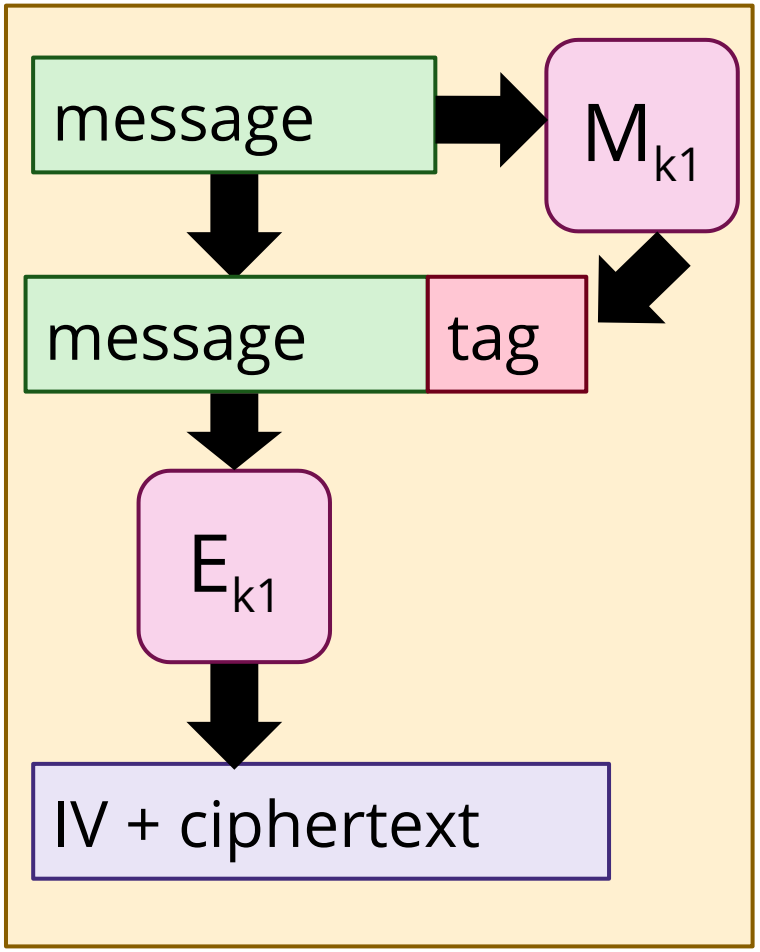
This is not secure - you lose confidentiality!

Combining MAC and encryption

Encrypt-then-MAC



MAC-then-Encrypt



Authenticated Encryption

Provides both confidentiality and integrity: without k, it should be hard to

- learn anything about m from c = E(k, m).
- compute E(k, m) for any m.

How do you authenticate someone?

How do you prove that you …

are old enough to buy alcohol?

have paid to ride on a train?

are a student at the University of Bristol?

are allowed to drive a car?

are *(insert your full name here)*?

Authentication

Identification: who is this?

Authorisation: are they allowed to do this?

Authentication: identification + authorisation

Why would a card just saying “this person is at least 18 years old” be a bad idea?

Is this a problem for train tickets? Which ones?

Authentication

Traditionally: authenticate with something you

have (passport, ID card, key)

know (password)

are (biometrics)

Multi-factor authentication (often “2FA”) is becoming more and more common.

Authentication

passport, ID card: pyhiscal object

typically gives you much more information than you need to sell someone a bottle of beer ...

... but it would be too costly to use it to track people every time they bought something.
(And they would notice.)

Online / Offline

Offline authentication: you hand out a **credential**, which can be misused / stolen / forged.

Online authentication means you also check against a **database** under your own control.

passwords

Passwords


Sorry but your password must contain an uppercase letter, a number, a haiku, a gang sign, a hieroglyph, and the blood of a virgin.
someecards
user card

Top passwords (Adobe 2013 breach)

Of 153 million leaked passwords,

# users	password	freq.
1 911 938	123456	1.2%
446 162	123456789	0.3%
345 834	password	0.29%
5.9M	*(in the top 100)*	3.89%

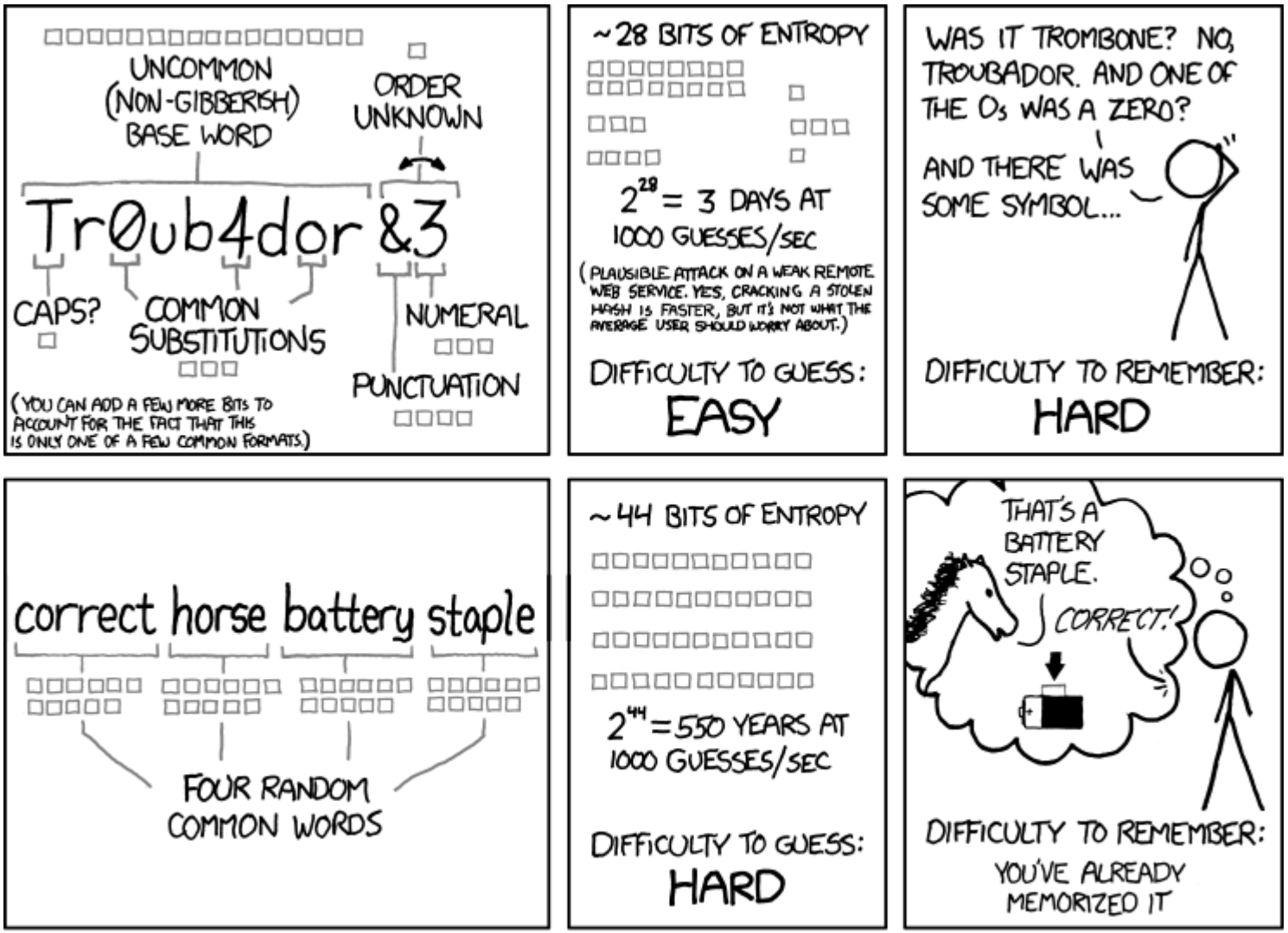
Passwords

A good password is

easy for a human to remember

hard for a computer to guess

Most of the time we do the exact opposite!


UNCOMMON (NON-GIBBERISH) BASE WORD
ORDER UNKNOWN
Tr0ub4dor&3
CAPS?
COMMON SUBSTITUTIONS
NUMERAL
PUNCTUATION
(YOU CAN ADD A FEW MORE BITS TO ACCOUNT FOR THE FACT THAT THIS IS ONLY ONE OF A FEW COMMON FORMATS.)
~28 BITS OF ENTROPY
2^28 = 3 DAYS AT 1000 GUESSES/SEC
(PLAUSIBLE ATTACK ON A WEAK REMOTE WEB SERVICE. YES, CRACKING A STOLEN HASH IS FASTER, BUT IT'S NOT WHAT THE AVERAGE USER SHOULD WORRY ABOUT.)
DIFFICULTY TO GUESS:
EASY
WAS IT TROMBONE? NO, TROUBADOR. AND ONE OF THE Os WAS A ZERO?
AND THERE WAS SOME SYMBOL...
DIFFICULTY TO REMEMBER:
HARD
correct horse battery staple
FOUR RANDOM COMMON WORDS
~44 BITS OF ENTROPY
2^44 = 550 YEARS AT 1000 GUESSES/SEC
DIFFICULTY TO GUESS:
HARD
THAT'S A BATTERY STAPLE.
CORRECT!
DIFFICULTY TO REMEMBER:
YOU'VE ALREADY MEMORIZED IT
THROUGH 20 YEARS OF EFFORT, WE'VE SUCCESSFULLY TRAINED EVERYONE TO USE PASSWORDS THAT ARE HARD FOR HUMANS TO REMEMBER, BUT EASY FOR COMPUTERS TO GUESS.

xkcd.com/936

The bad news

If passwords are not stored properly and a site gets hacked, it hardly matters how good they are.

Last.fm hack (2012): passwords were hashed (md5) but that's not good enough - passwords up to **46** characters were recovered:

```
perseusandtheseamonsterperseusandtheseamonster
pnukgirl5121990pnukgirl5121990pnukgirl5121990
WthAmIDoingHereWthAmIDoingHereWthAmIDoingHere
paracoccidioidomicosisproctitissarcomucosis
lllllllllllllllllllllllllllllllllllllllll
lorenaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
```

Passwords

Can we …

- write them down?
- reuse them for different sites?
- share them with friends and family?
- choose a password for a site and never change it?

Passwords

If someone knows your **university password**, they can cause you a lot of trouble.

Choose one that you've never used, and never will use anywhere else. Make it strong.

Your e-mail account(s) are the next most valuable asset to protect, as anyone who gets in can find all linked sites and usually break in through the *password reset* feature.

Password breaches

haveibeenpwned.com :

Over 5 billion breached accounts from 75’408 breaches, and counting …

check your e-mail addresses there - if it appears assume the password for that service is compromised