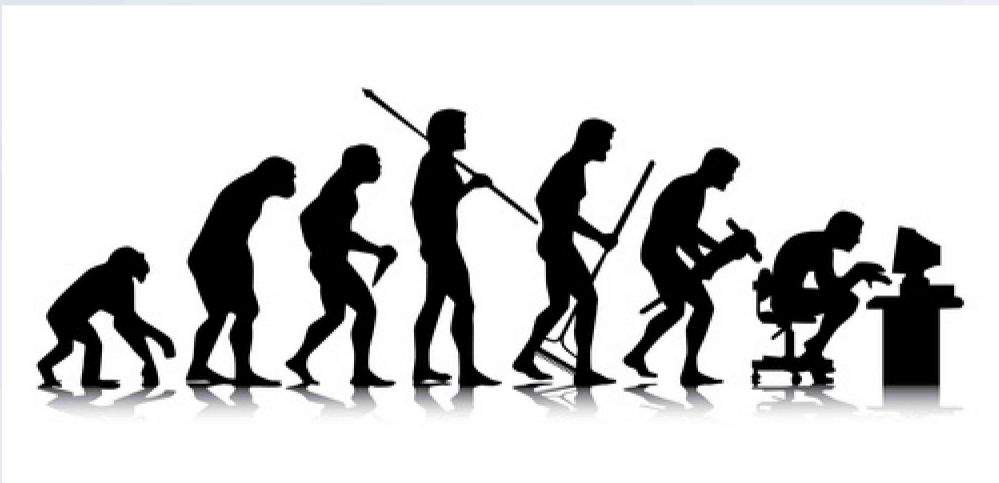


INTRODUCTION TO PRIVACY

CARMELA TRONCOSO*
IMDEA SOFTWARE INSTITUTE
IPICS 2016
5TH JULY 2016

*THANKS TO GEORGE DANEZIS, SEDA GURSES, CLAUDIA DIAZ, AND BART PRENEEL

THE CONTEXT: THE ERA OF DATA



THE CONTEXT: THE ERA OF...

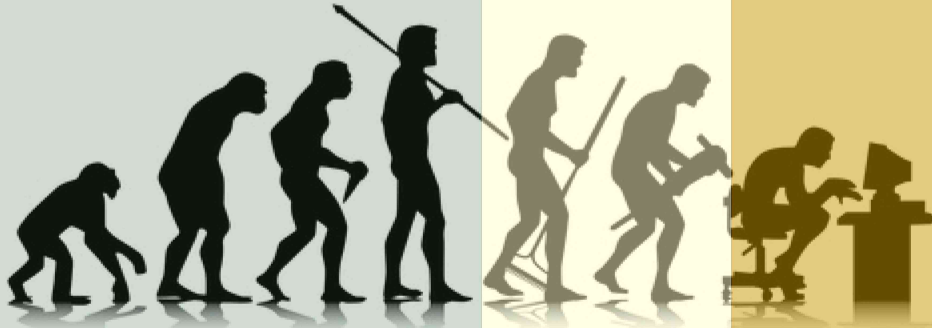


LITTLE

MORE

MUCH MORE

DATA GENERATION

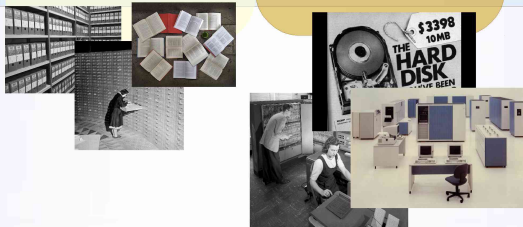


NONE?

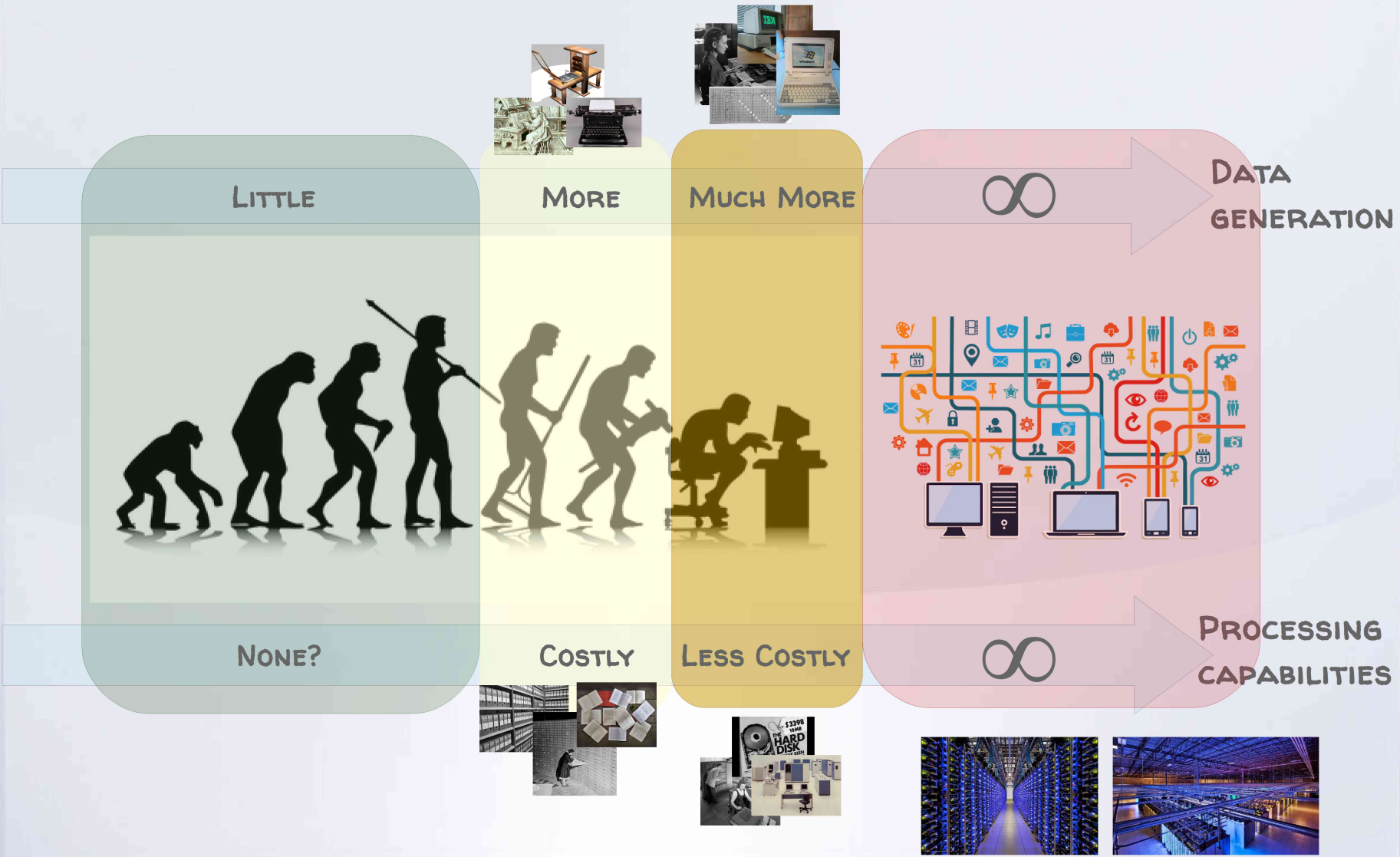
COSTLY

LESS COSTLY

PROCESSING CAPABILITIES



THE CONTEXT: THE ERA OF DATA AND PROCESSING!!



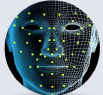
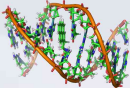


THE CONTEXT: UNIQUENESS OF DATA

USERS



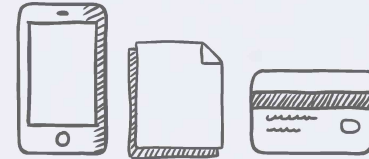
PHYSICAL (BIOMETRICS)

- > fingerprints 
- > iris 
- > face 
- > DNA 


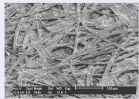
BEHAVIORAL

- > typing 
- > locations 
- > social network 

DEVICES

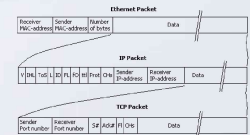


PHYSICAL ("BIOMETRICS")

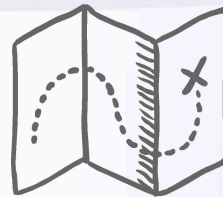
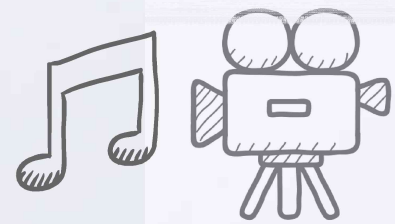
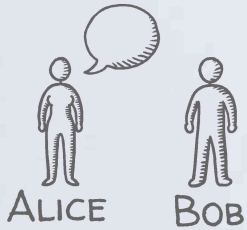
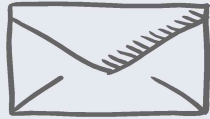
- > radio fingerprinting 
- > paper fiber patterns 
- > magnetic behavior
- > PUFs (Physically Unclonable Functions)

LOGICAL

- > IP, MAC addresses, IMSI
- > Certificates
- > Software, fonts, ...



THE CONTEXT: THE RULES OF THE GAME HAVE CHANGED



THE CONTEXT: THE RULES OF THE GAME HAVE CHANGED



THE CONTEXT: THE RULES OF THE GAME HAVE CHANGED (IS NOT ONLY ACTIVE...)

HOW MANY WAYS HAVE YOU BEEN LOCATED TODAY?

- > cell phone (turned on?)
 - > Bluetooth-enabled devices (printers, headphones, speakers,...)
 - > WiFi AP
- > laptop computer (same)
- > credit card at the gas station or in the ATM machine
- > cameras driving in a monitored intersection or at the supermarket
- > scan badge to enter a building

THE PROBLEM WITH LOCATION

- > Strong inferences!
 - > desk in a building
 - > home location
 - > future locations
- > Highly sensitive!
 - > abortion clinic
 - > business competitor
 - > political headquarters



FROM AVAILABLE (UNIQUE DATA) + PROCESSING CAPABILITIES...

INTELLIGENT DATA-BASED APPLICATIONS

- > Road pricing
- > Health monitoring
- > Children/Elderly trackers
- > Smart metering
- > Intelligent buildings
- > Recommendation systems
 - > Movies (Netflix)
 - > Products (Amazon)
 - > Friends (Social networks)
 - > Music (Spotify, iTunes)
- > Location based services
 - > Friend finders
 - > Maps
 - > Points of interest

INDIVIDUAL APPLICATIONS ARE LEGITIMATE



TOGETHER THEY BECOME A CHEAP
SURVEILLANCE INFRASTRUCTURE

THE PRIVACY DEBATE

1) BUT I HAVE NOTHING TO HIDE!!



Eric Schmidt
Google's CEO
Dec 2009

"I don't care about surveillance because I have nothing to hide"

"If you are so concerned about people/the police/the government knowing what you do, it's because you know you're doing something wrong"

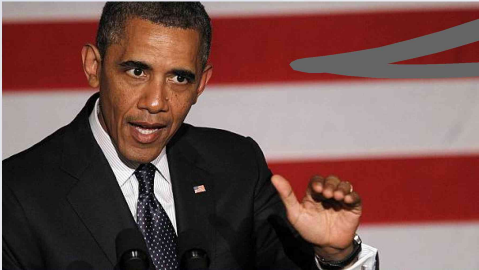
"If you have something that you don't want anyone to know, maybe you shouldn't be doing it in the first place."

THE PRIVACY DEBATE

2) WE NEED A TRADEOFF WITH SECURITY!!

"Nobody is listening to your telephone calls. That's not what this programme is about..."

But I think it's important to recognise that you can't have 100% security and also then have 100% privacy and zero inconvenience"



Barack Obama
US President
Jun 2013

THE PRIVACY DEBATE

3) PEOPLE DO NOT CARE!!

"People have really gotten comfortable not only sharing more information and different kinds, but more openly and with more people"

"That social norm is just something that has evolved over time."



Mark Zuckerberg
Facebook CEO
Jan 2011

THE PRIVACY DEBATE: 1) BUT I HAVE NOTHING TO HIDE!!

WOULD YOU SHARE WITH THE GROUP...?

IDENTITY ATTRIBUTES: Name, age, gender, race, IQ, marital status, address, phone number, ID number...

HEALTH DATA: Medical issues, treatments you follow, DNA, health risk factors

BEHAVIOR: Personality type, what you eat, what you shop, how you behave and interact with others

LOCATION: Where were you yesterday (or at other point in time), your movement patterns

SOCIAL NETWORK: Who your friends are, who you meet when, your different social circles

FINANCIAL DATA: credit card number, bank account,...
or how much you earn, how you spend your money,

INTERESTS / PREFERENCES: Books you read, music you listen, films you like, sports you practice...
And political affiliation, religious beliefs, or sexual orientation?



Daniel Solove,
Prof. of Law

"The problem with the 'nothing to hide' argument is its underlying assumption that **PRIVACY IS ABOUT HIDING BAD THINGS**."

With whom would you share?

All of it?

Is it about right/wrong?

THE PRIVACY DEBATE: 2) WE NEED A TRADEOFF WITH SECURITY!!

“Surveillance is good and privacy is bad for national security.
We need a tradeoff between privacy and security”

Keywords: Terrorism, Child pornography, Money laundering, Crime

- “we need more surveillance” is a powerful argument
- if attacks increase, you can argue that you need even more
- if attacks decrease, you take credit



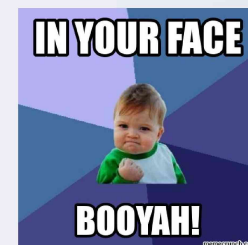
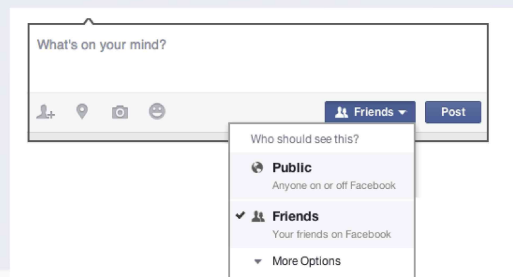
THE PRIVACY DEBATE: 2) WE NEED A TRADEOFF WITH SECURITY!!

(SURVEILLANCE == SECURITY) == TRUE ??

- not **EFFECTIVE**: smart adversaries evade surveillance
 - ISIS uses Telegram, Threema, Signal,...
 - ... but we do not!!
- risk of **ABUSE**: lack of transparency and safeguards
 - Snowden revelations: NSA spying on Americans, companies, ...
 - Spanish Interior ministry spying independentist politicians
- risk of **SUBVERSION** for crime / terrorism
 - Greek Vodafone scandal (2006): "someone" used the legal interception functionalities (backdoors) to monitor 106 key people: Greek PM, ministers, senior military, diplomats, journalists...

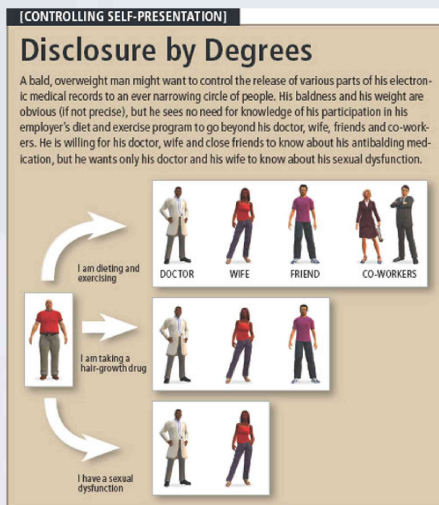
THE PRIVACY DEBATE: 3) PEOPLE DO NOT CARE ABOUT PRIVACY... REALLY??

- In the real world, we want to control information related to us
 - What we tell to whom (aka, who knows what)
 - We value friends who are discreet and keep our secrets
 - We give more information to people we trust



THE PRIVACY DEBATE: 3) PEOPLE DO NOT CARE ABOUT PRIVACY... REALLY??

- We care about Impression management / self-presentation
 - The process through which people try to control the impressions other people form of them
 - Construct an image of ourselves to claim personal identity



Dyson, Esther. "Reflections On Privacy 2.0." *Scientific American* 299.3 (2008): 50-55.
<http://phdcomics.com/> ← THIS IS A MUST!!

THE PRIVACY DEBATE: 3) PEOPLE DO NOT CARE ABOUT PRIVACY... REALLY??

- Concerns over information taken out of context
 - Sharing health data is ok at the hospital but not at your kids' school
 - A picture taken at a crazy party being available to a potential employer
 - **PRIVACY AS CONTEXTUAL INTEGRITY (Nissenbaum)**

- Personal safety
 - Valuable items in an empty house
 - Child alone at home
 - Vulnerability to manipulation (e.g. supermarket that makes you spend more)
 - Identity theft



Nissenbaum, Helen. "Privacy as contextual integrity." Wash. L. Rev. 79 (2004): 119.

<http://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/>

<http://pleaserozme.com/>

http://www.ted.com/talks/alessandro_acquisti_why_privacy_matters

NOT ONLY ABOUT BIG BROTHER

- Solove: “Part of what makes a society a good place in which to live is the extent to which it allows people freedom from the intrusiveness of others. **A SOCIETY WITHOUT PRIVACY PROTECTION WOULD BE SUFFOCATION**”
- Not so much Orwell’s “Big Brother” as Kafka’s “The Trial”:
 - “...a bureaucracy with inscrutable purposes that uses people’s information to make important decisions about them, yet **DENIES THE PEOPLE THE ABILITY TO PARTICIPATE IN HOW THEIR INFORMATION IS USED**”
 - “The problems captured by the Kafka metaphor are of a different sort than the problems caused by surveillance. They often do not result in inhibition or chilling. Instead, they are **PROBLEMS OF INFORMATION PROCESSING—THE STORAGE, USE, OR ANALYSIS OF DATA—RATHER THAN INFORMATION COLLECTION.**”
 - “...not only frustrate the individual by creating a sense of helplessness and powerlessness, but they also **AFFECT SOCIAL STRUCTURE BY ALTERING THE KIND OF RELATIONSHIPS PEOPLE HAVE WITH THE INSTITUTIONS THAT MAKE IMPORTANT DECISIONS ABOUT THEIR LIVES.**”



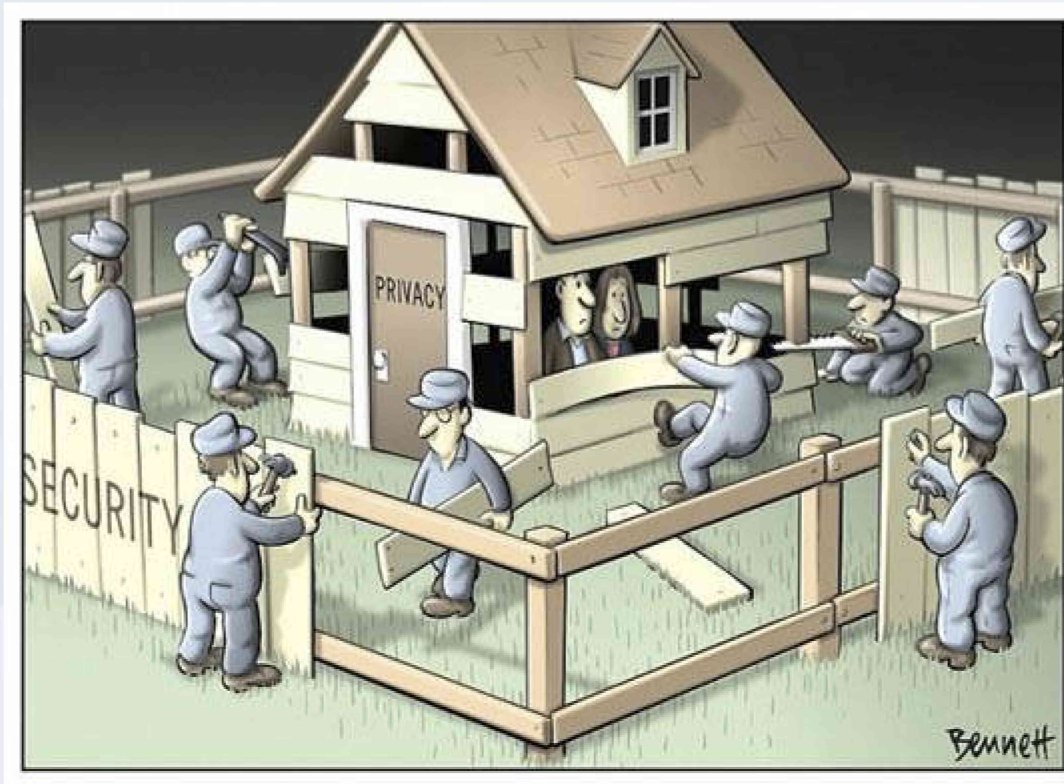
THE CHEAP SURVEILLANCE
INFRASTRUCTURE

MAY BECOME



ONE RING TO RULE THEM ALL

IS THERE A TRADEOFF BETWEEN SECURITY AND PRIVACY?



By Clay Bennet
From <http://www.roymckenzie.me>

PRIVACY IS A SECURITY PROPERTY

➤ INDIVIDUALS

- freedom from intrusion, profiling and manipulation, protection against crime / identity theft, flexibility to access and use content and services, control over one's information

➤ COMPANIES

- protection of trade secrets, business strategy, internal operations, access to patents

➤ GOVERNMENTS / MILITARY

- protection of national secrets, confidentiality of law enforcement investigations, diplomatic activities, political negotiations

➤ SHARED INFRASTRUCTURE

- despite varying capabilities infrastructure is shared
- telecommunications, operating systems, search engines, on-line shops, software, ...
- denying security to some, means denying it to all: crypto wars redux?

WE LOVE IT

WE NEED IT

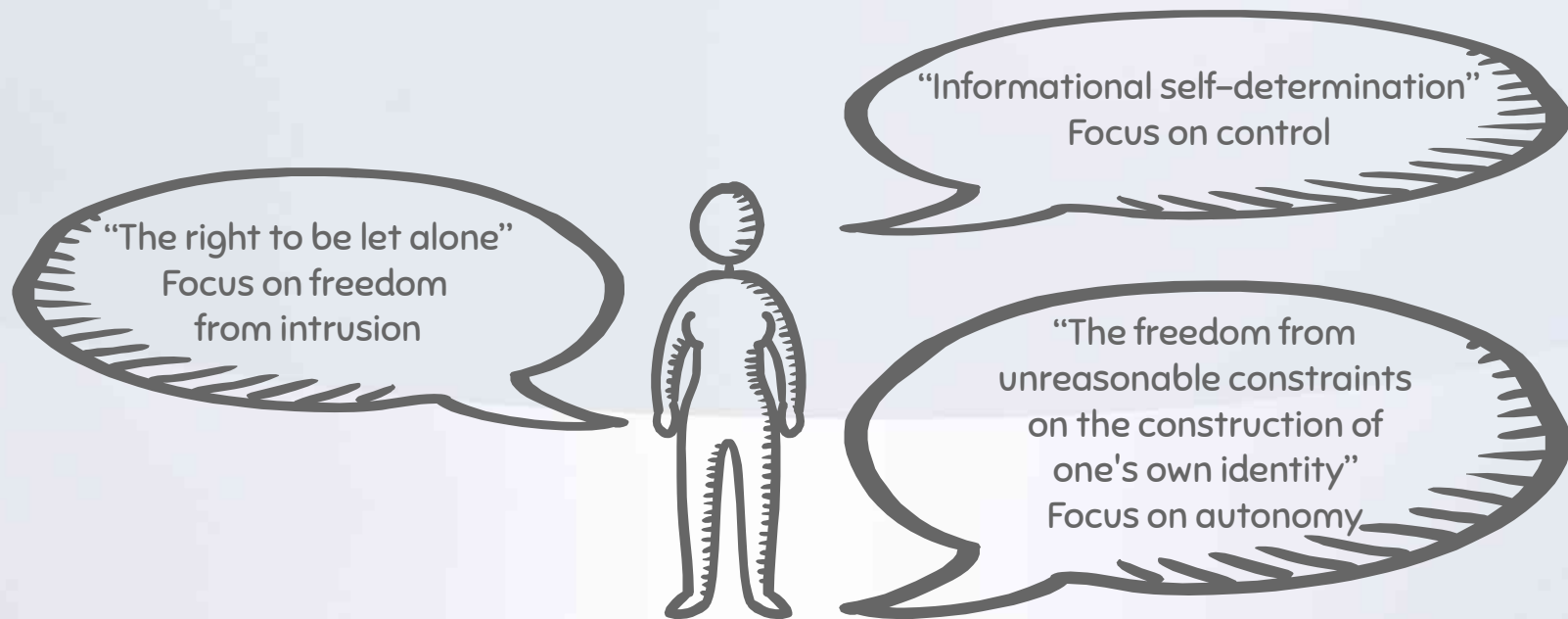
WE MUST HAVE IT

BUT....

WHAT IS PRIVACY??????

WHAT IS PRIVACY

- Abstract and subjective concept, hard to define
 - Dependent on cultural issues, study discipline, stakeholder, context
- Popular definitions:



WHAT IS PRIVACY IN PRIVACY ENHANCING TECHNOLOGIES

- 3 different flavors
 - the concept of “privacy” they embed
 - their goals
 - their challenges and limitations

Gürses, Seda, and Claudia Diaz. "Two tales of privacy in online social networks." *IEEE Security & Privacy* 11.3 (2013): 29–37.

Diaz, Claudia, and Seda Gürses. "Understanding the landscape of privacy technologies." *Information Security Summit* (2012): 58–63.

Danezis, George, and Seda Gürses. "A critical review of 10 years of privacy technology." *Surveillance cultures: a global surveillance society* (2010): 1–16.

"SOCIAL PRIVACY": GOALS

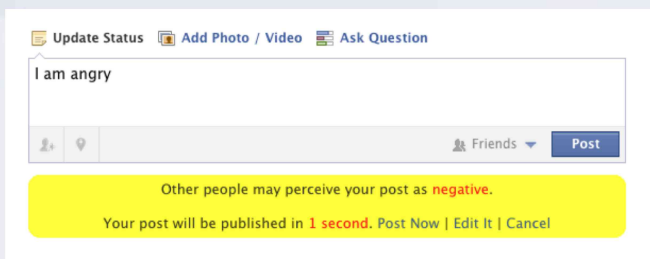
- Meet privacy expectations: **"DON'T SURPRISE THE USER"**
- TWO MAIN APPROACHES
 - Support decision making
 - Privacy controls visible and easy to use
 - Predict actions to avoid regret
 - Help users develop appropriate privacy practices
 - e.g. use bcc

APPROPRIATE DEFAULTS: only friends

EASY CONFIGURATION: automated grouping

CONTEXTUAL FEEDBACK: "how X sees my profile"

PRIVACY NUDGES: force to reconsider
Audience, time, sentiment,...

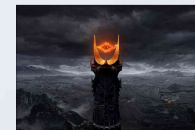


"SOCIAL PRIVACY": LIMITATIONS

- Focus on concerns directly related to actions (and implicit?)
- Front-end oriented
 - No info about the server, only privacy towards third parties
- Limited by user understanding
 - As much as policies can do...
 - Based on average consumer
- Based on privacy expectations
 - What if expectations are null....



AWESOME FOR INDUSTRY!!
MAKE USERS COMFORTABLE



“INSTITUTIONAL PRIVACY”: CONCERNS

- Data **COLLECTED** without users' awareness or *informed consent*
- Data **PROCESSED** for illegitimate purposes

- Data **SECURITY**
 - correctness, integrity, deletion
 - Information not becoming public
 - Safety (crime protection, stalking,...)

WHO DEFINES THE PRIVACY PROBLEM: **LEGISLATION**

"INSTITUTIONAL PRIVACY": GOALS

- Ensure compliance with data protection principles:

- informed consent
- purpose limitation
- data minimization
- subject access rights

APPROPRIATE DEFAULTS: towards organization!

EASY CONFIGURATION: policy negotiation with organization

ACCESS CONTROL: limit and log who accesses what

"PRIVATE" DATA PUBLISHING: anonymization, differential privacy

- **Data SECURITY**

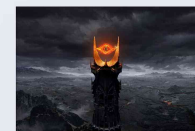
- Prevent (or mitigate) data breaches

- **Auditability and accountability**

"INSTITUTIONAL PRIVACY": LIMITATIONS

- Assumes:
 - collection and processing by organizations is necessary
 - organizations are (semi)-trusted and honest
 - Reliance on punishment
 - No technical protection of the data
- Focuses on limiting misuse, **NOT** collection
 - Easy to circumvent minimization to collect in bulk
 - Auditing may require more data!
 - The danger of *informed consent*: if compliant is ok!
- Limited
 - Scope (personal data != all data)
 - transparency (proprietary sw and algorithms)

AWESOME FOR INDUSTRY!!
MAKE USERS COMFORTABLE
+ LEGAL COMPLIANCE!!



"ANTI-SURVEILLANCE PRIVACY": CONCERNS

- Data disclosure **BY DEFAULT** through ICT infrastructure
- Threat model **ANYBODY** that may see the data
 - ISP
 - Service provider
 - Government
- Concerned about
 - Surveillance
 - Censorship
 - Other democratic values:
 - Freedom speech
 - Freedom association
 - Democracy itself!



WHO DEFINES THE PRIVACY PROBLEM: SECURITY EXPERTS

"ANTI-SURVEILLANCE PRIVACY": GOALS

- Prevent/minimize default disclosure of personal information anyone:
 - Only information explicitly disclosed is made available to intended recipients (confidentiality)
 - Both user-generated and implicit!



- Circumvent censorship

- Minimize the need to trust others
 - Distribute trust by avoiding single points of failure



END-TO-END ENCRYPTION: PGP, OTR

ANONYMOUS COMMS: Tor

OBFUSSION:

- geo-indistinguishability
- dummy actions
- hiding
- generalization

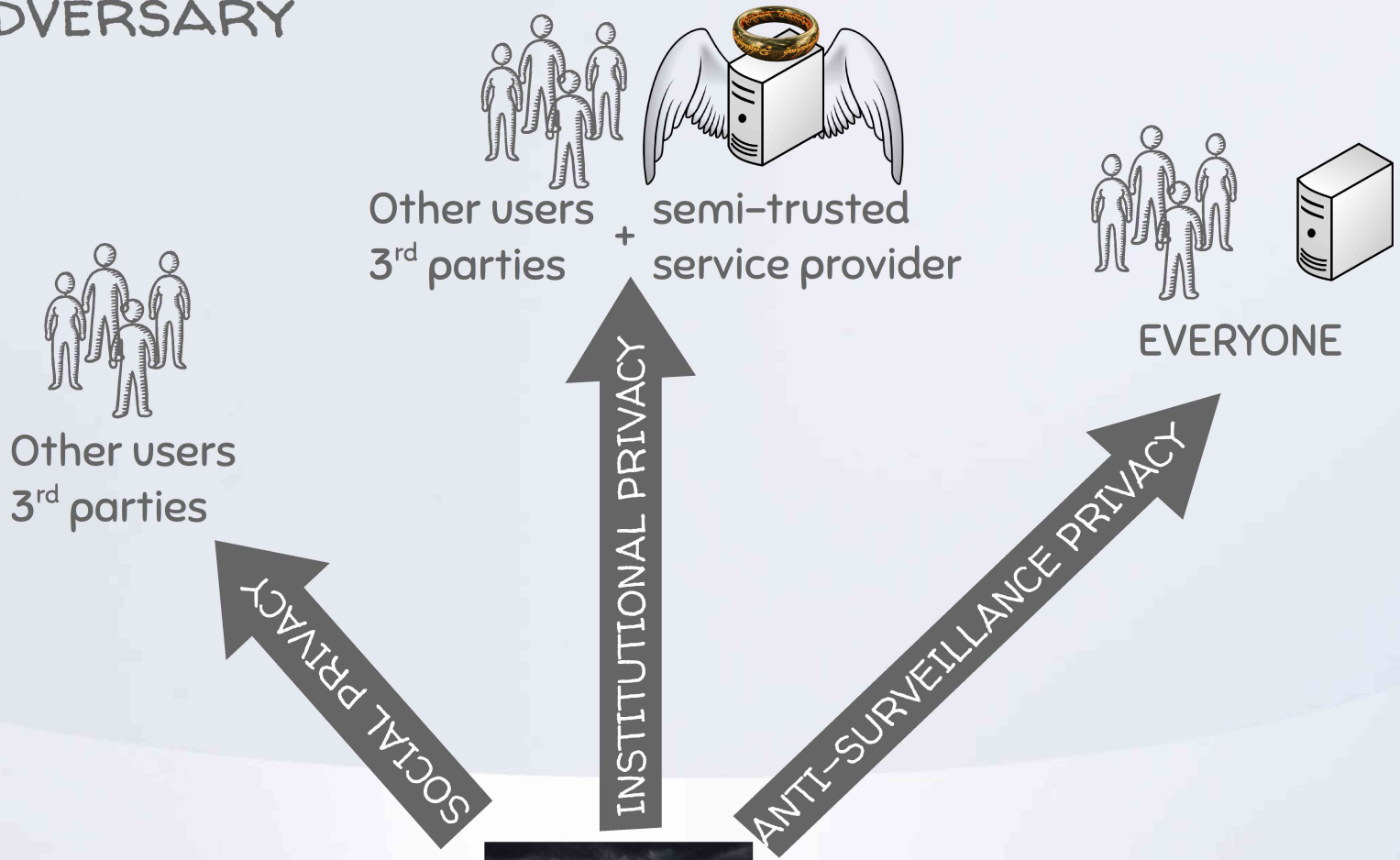
ADVANCED CRYPTO:

- Private information retrieval
- Anonymous authentication
- Multiparty computation
- Oblivious transfer
- Cryptographic commitments

"ANTI-SURVEILLANCE PRIVACY": LIMITATIONS

- Making secure private designs is hard
 - "Narrow" tools
 - Difficult to combine
- Usability problems
 - For developers:
 - how the @\$%&#\$Ŷ& do I program this?
 - performance
 - For users:
 - Unintuitive
- Incentives are low
 - For providers: they lose the data!
 - For governments: national security, fraud detection, surveillance & control

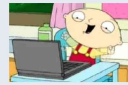
THE ADVERSARY



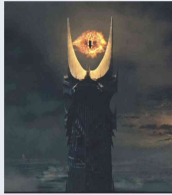
THE ADVERSARY IS ANYONE AND VERY POWERFUL



Your Parents



Your Children



The Boss



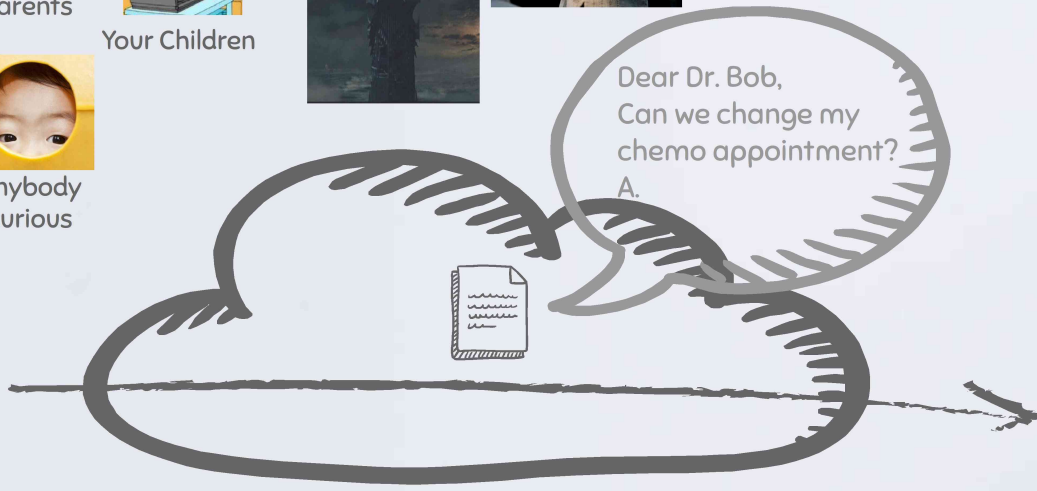
Anybody curious



ISPs



ALICE

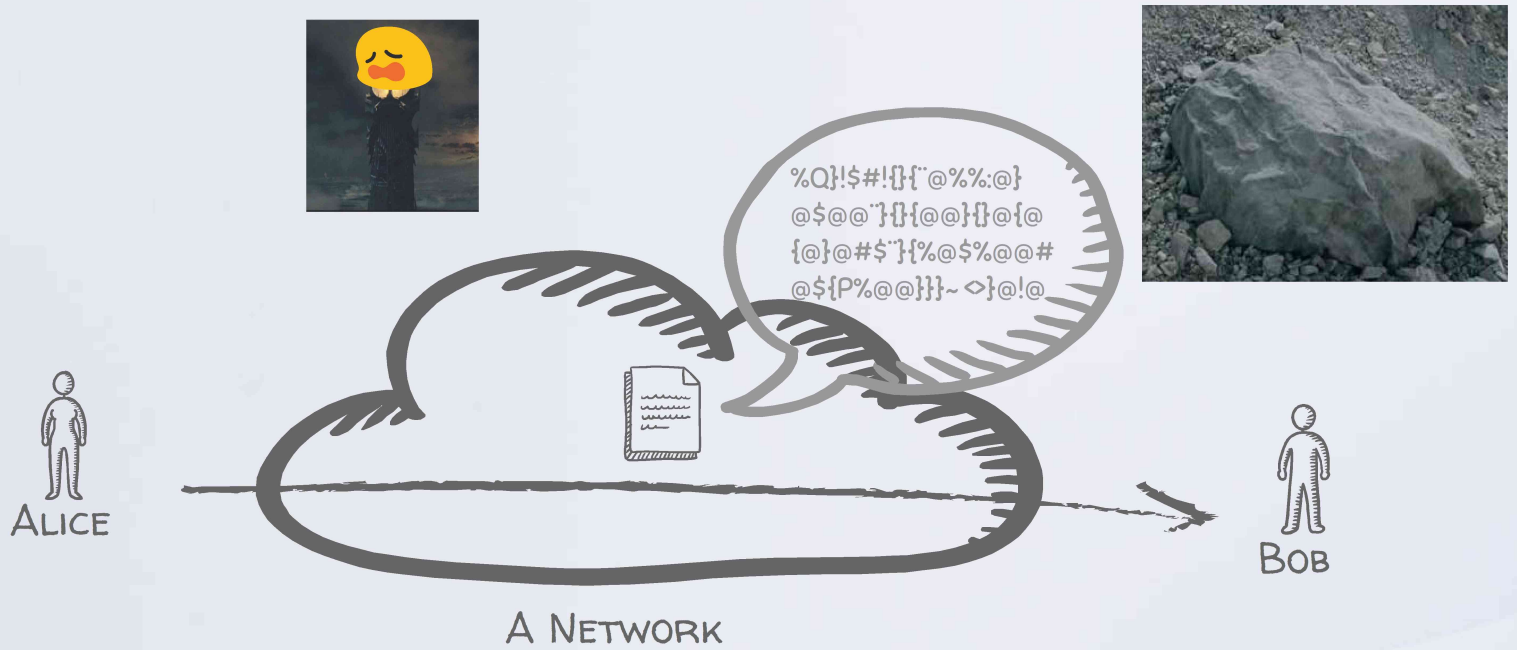


BOB

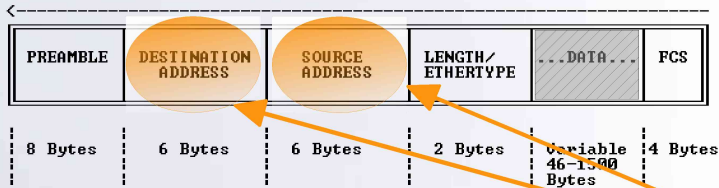
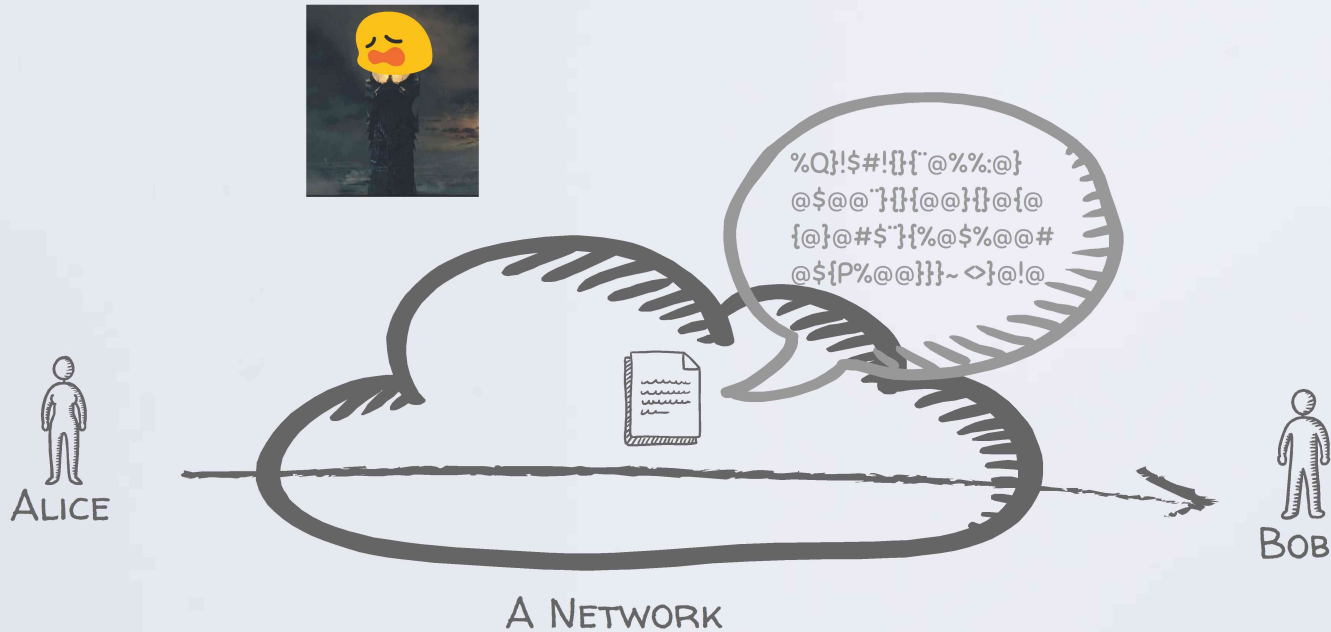
A NETWORK



BUT WE CAN ENCRYPT! WHAT IS THE PROBLEM?



BUT WE CAN ENCRYPT! WHAT IS THE PROBLEM?

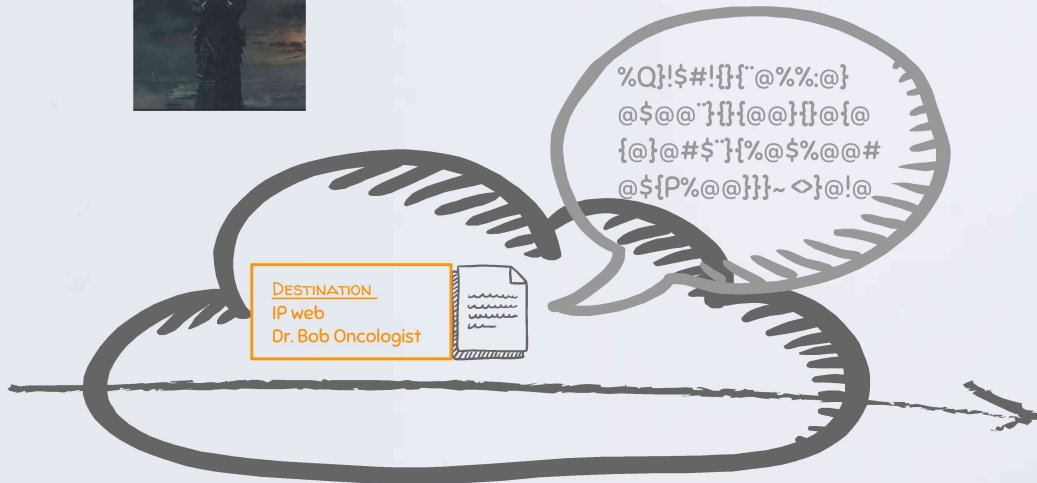
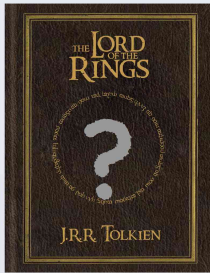


ETHERNET
(IEEE 802.3, 1997)

*Same for IP, TCP, SMTP, IRC,
HTTP, ...*



OMG!! META DATA IS ALSO SENSITIVE!!



A NETWORK

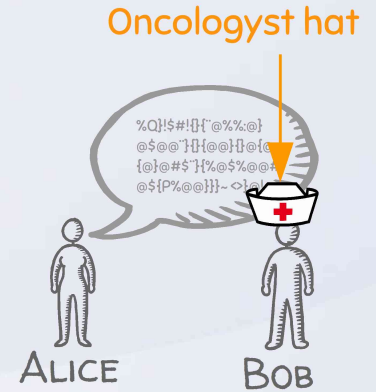
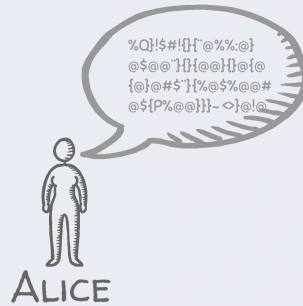
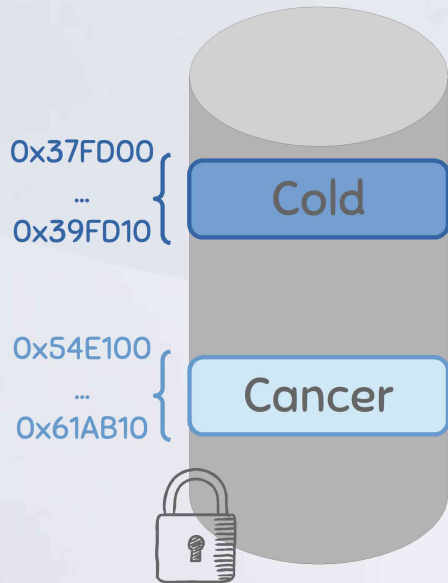
PREAMBLE	DESTINATION ADDRESS	SOURCE ADDRESS	LENGTH/ ETHERTYPE	...DATA...	FCS
8 Bytes	6 Bytes	6 Bytes	2 Bytes	Variable 46-1500 Bytes	4 Bytes

ETHERNET
(IEEE 802.3, 1997)

*Same for IP, TCP, SMTP, IRC,
HTTP, ...*



OMG!! META DATA IS ALSO SENSITIVE!!



ADDRESS
Dr. Bob Oncologist

WHAT CAN WE DO ABOUT IT?

PRIVACY PROPERTIES

PSEUDONYMITY: pseudonymous as ID (personal data!)

ANONYMITY: decoupling identity and action

UNLINKABILITY: hiding link between actions

UNOBSERVABILITY: hiding the very existence of actions

PLAUSIBLE DENIABILITY: not possible to prove a link between identity and action

PRIVACY ENHANCING TECHNOLOGIES (PETS)

THE PROBLEM WITH ANONYMIZATION

(Art 29 Opinion)

- 1- No singling out of individuals but Metadata are unique!
- 2- No linking data from one individual, but sadly is possible
- 3- No inference about individuals, difficult to avoid

TAKEAWAYS

- The Lord of The Rings is a great timeless book
- Privacy IS a security property
 - It is not about hiding bad behaviors, should not be traded off and we do care!
- There are different flavors of privacy
 - Who sets the problem?
 - Who is the adversary?
- The adversary is powerful
 - Cryptography alone does not guarantee privacy
- Privacy is formalized as properties that are achieved through PETs
 - How to measure privacy??

THANKS!

ANY QUESTIONS?

More about privacy: <https://www.petsymposium.org/>



carmela.troncoso@imdea.org
<https://software.imdea.org/~carmela.troncoso/>
(these slides will be there soon)

Template: <http://www.brainybetty.com/>
Figures: [SlidesCarnival](#)