COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
School of Information

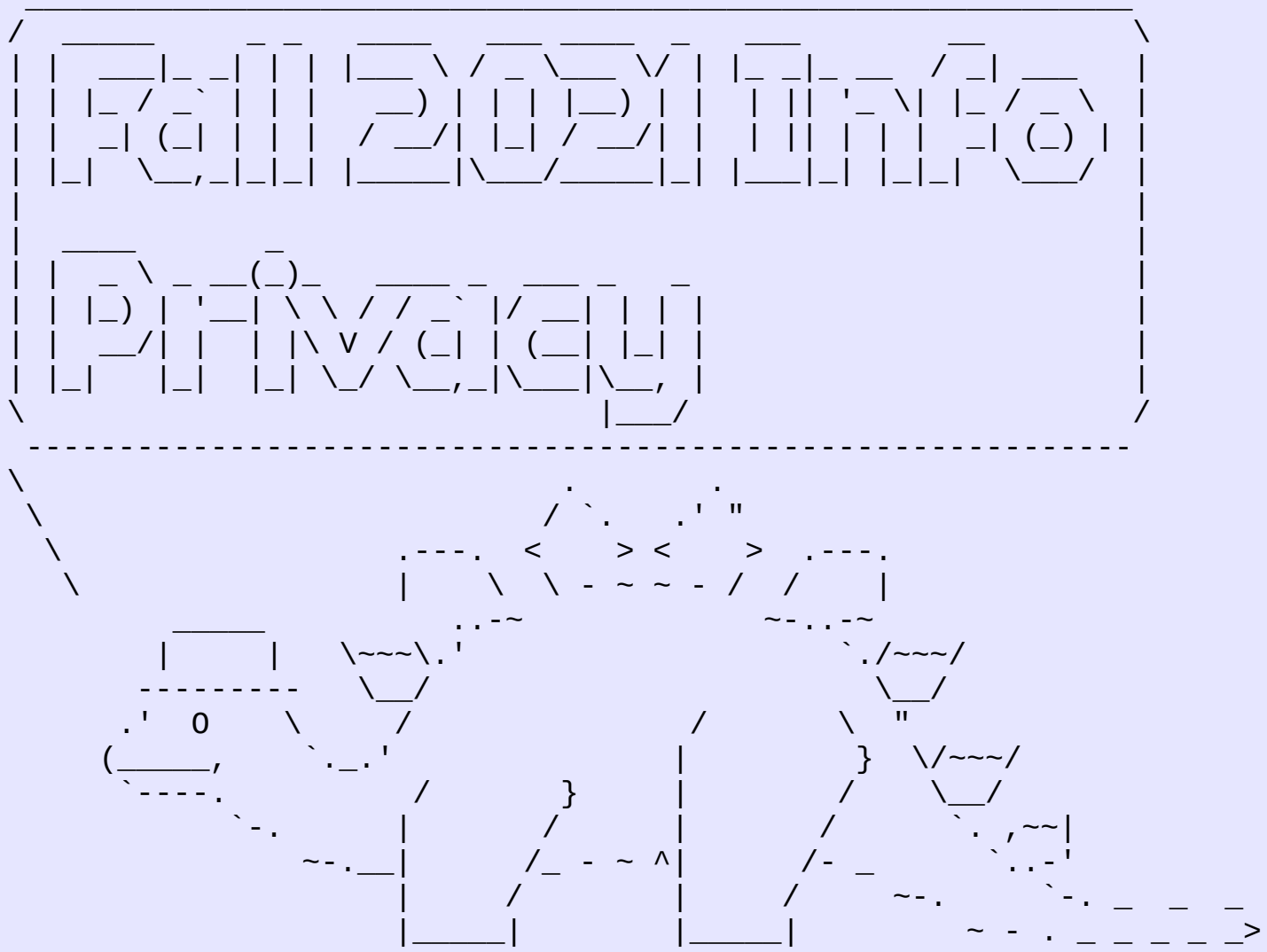
Foundations of Privacy Technology II

Information Privacy with Applications

David Sidi (dsidi@email.arizona.edu)

Administration

- Assignment I, part I due today

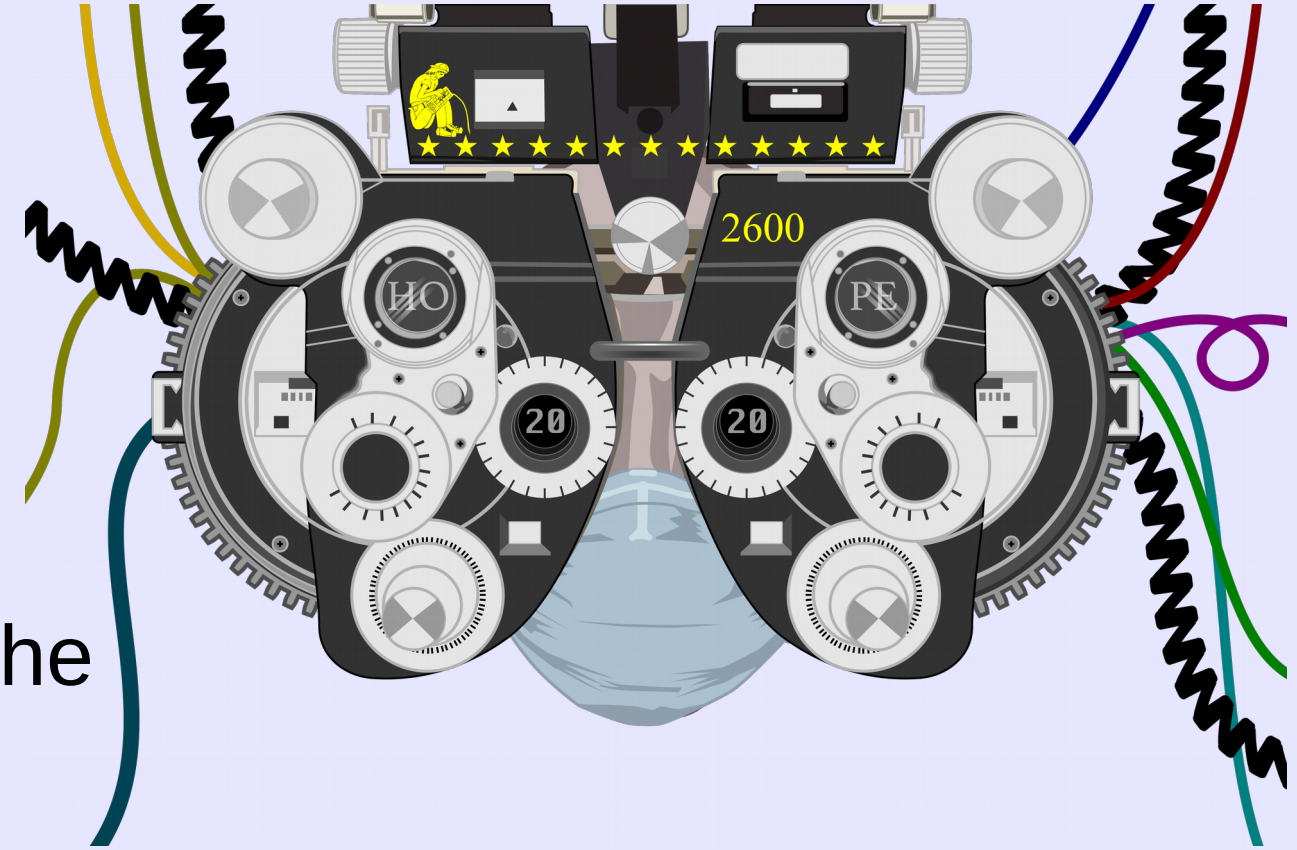


```
WELCOME TO THE FALL 2021 SANDBOX! ENJOY YOURSELF...
Last login: Tue Aug 24 16:59:00 2021
fa21-course-1vcpu-1gb-sfo1-01:dsidi $
```

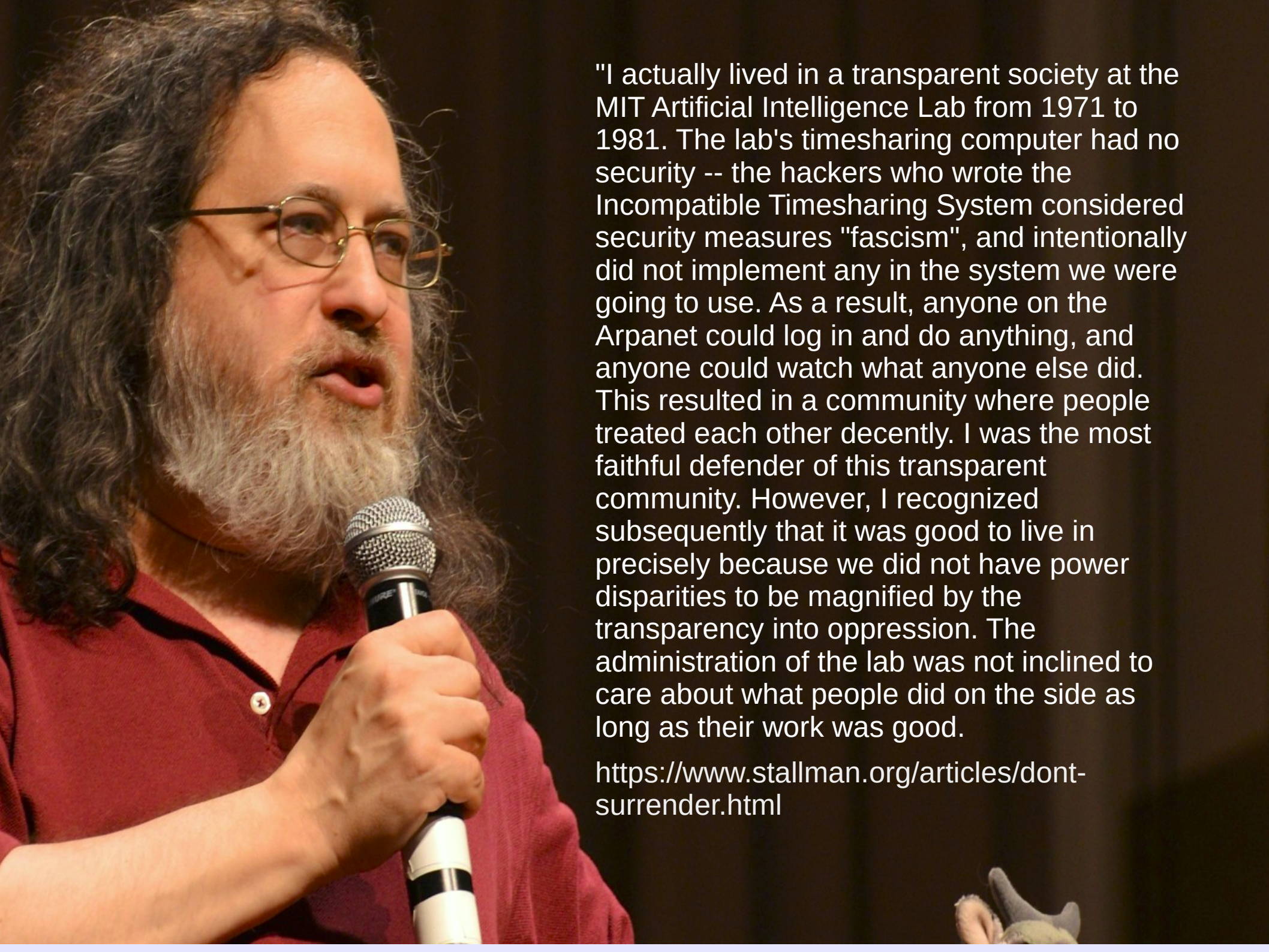
Let’s get down to it

- recording what you’re doing: `script`
- getting help: `man`
- getting out of trouble: `^C`, `^D`, `^|`, `kill`
  - kill: processes, how to find the PID. Related: pkill, top
- getting around: `ls`, `cd`, `pwd`, `find`
  - ls: hidden files (including the special ones in every directory), globs
- Open a terminal. Which directory are you in?
- Change to `/var/log`, and list all files starting with ‘s’
- List exactly the names of the hidden directories in the current directory, *using only ls*
- List all the files in any subdirectory below the current one
- Start vim, then kill it from a different window

“we focus on one key aspect of these technologies, namely, the kind of trust they can provide” (396)



- 2600 meetings near Harvard in the early 1990s, recounted in the cDc book



"I actually lived in a transparent society at the MIT Artificial Intelligence Lab from 1971 to 1981. The lab's timesharing computer had no security -- the hackers who wrote the Incompatible Timesharing System considered security measures "fascism", and intentionally did not implement any in the system we were going to use. As a result, anyone on the Arpanet could log in and do anything, and anyone could watch what anyone else did. This resulted in a community where people treated each other decently. I was the most faithful defender of this transparent community. However, I recognized subsequently that it was good to live in precisely because we did not have power disparities to be magnified by the transparency into oppression. The administration of the lab was not inclined to care about what people did on the side as long as their work was good.

https://www.stallman.org/articles/dont-surrender.html

- can bad security be good for privacy? How else might allowing access to be delegated, and loosely, improve anonymity, for example? How might it build a community?

“we focus on one key aspect of these technologies, namely, the kind of trust they can provide” (396)

- “Trust” here is ersatz, from security, making up for a deficit in social trust
- technologies “providing trust” means: making trust moot

Not all privacy technologies reduce the perimeter of trust

- privacy vs security

- **Question: how can social trust be built with technology?**

A division among safeguards offered by privacy technologies (Le Métayer)

- Minimization of disclosure of personal data
- Enforcement of rights when personal data is disclosed
- each has distinctive trade-offs and design challenges

Minimization is hard, since data is useful

- Data is not just good for the individual; it’s good for society---you can’t just restrict its use without cost
  - open data in science, government
  - “Tragedy of the Data Commons”

Enforcing rights requires people as well as technology, so it’s hard too

- Data is already disclosed, out of technology’s hands and into people’s
  - “Information does not just want to be free, it longs to be free. ... Information is Rumor's younger, stronger cousin; Information is fleeter of foot, has more eyes, knows more, and understands less than Rumor..” Eric Hughes

Policy enforcement and technology

Rawat and Saxena (2008). “Practical Data Protection,” Journal of Craptology, 5.

Practical Data Protection

Sanjay Rawat *and Amitabh Saxena
<rawat, amitabh>@dit.unitn.it
Dept. of Information and Communication Technology
University of Trento
38050 Trento, Italy

April 23, 2008

Abstract

We present a very easy and practical method to send the information in a secure manner such that its disclosure to unintended recipient is not possible. Our method does not require the distribution of shared key at all. Our idea is inspired by the popularity of a very recent phenomenon of "Disclaimer Statement" in corporate emails.

2 Our Method

As mentioned above, our method is based on the a popular phenomenon of putting a "disclaimer" (a similar method was used for creating a very deadly virus [1]). This disclaimer is appended at the end of the mail. We propose that instead of putting the disclaimer at the end of the mail/message, it should be inserted at the very beginning of the mail. In this way, the receiver will first read the disclaimer and if he is not the intended recipient, he must not read that message and must delete that. These "MUST" properties are the characteristics of the disclaimer method and are well accepted in practice [5]. Following is an example of such a disclaimer:

> "*This message is being sent from University of Trento (Italy) and may contain information which is confidential. If you receive this message but you are not the intended recipient, stop reading a single line after this disclaimer onwards, advise the sender immediately by replying this e-mail and delete this message and any attachments without retaining a copy (don't forget to delete the mail from your "Sent Mails" folder). We appreciate your cooperation, otherwise you will be in big trouble.*"

We can see that the above disclaimer provides very tight confidentiality. Our method is well protected by the law [7, 8] and is highly flexible in the sense that you can design very creative disclaimers based on your security requirements and level. Few disclaimers are available online [6][1]. Furthermore, the disclaimers can be inserted via the outgoing mail server so that individuals don't have to worry about that.

As a fine tuning parameter, we advise you not to mention the subject of the message in the "Subject" field as it may give an adversary some information leakage to do further cryptanalysis. Instead, write the Subject after the disclaimer, as a part of message body.

Finally, to counter the powerful cryptanalysis method of Knudsen and Mirza [3], we recommend that the very first line of the disclaimer must be "*Do not remove this disclaimer.*" This makes our method resistant to "deletion cryptanalysis."

The distinction from today's reading can be compared to others

- Minimizing disclosure of personal data
- Enforcing rights when personal data is disclosed

The distinction from today's reading can be compared to others

- Minimizing disclosure of personal data
- Enforcing rights when personal data is disclosed

- **Question: What distinction that we've seen before among kinds of privacy technology is very close to the above?**

From Diaz and Gürses last time

- Privacy as control: a matter of policy, which **controls data use**. Does not try to minimize trust in a third party for linkable data
  - example: privacy settings
- Privacy as confidentiality: a matter of applied mathematics, which obviates policy, and **minimizes disclosure**. Tries to minimize trust in a third party with linkable data.
  - example: PIR

From Danezis

Soft Privacy Technologies

- Focus on compliance.
- Focus on “internal controls”.
- **Assumption: a third party is entrusted with the user data.**
- Threat model: third party is trusted to process user data according to user wishes.
- Examples technologies:
  - Access control, tunnel encryption (SSL/TLS)
- “Keeping honest services safe from insiders / employees”.

Hard Privacy Technologies

- Stronger focus on data minimization.
- **Assumption: there exists no single third party that may be trusted with user data.**
- Threat model: a service is in the hands of the adversary; may be coerced; may be hacked.
- Common assumption: k-out-of-n honest third parties.
- May relay on service integrity if auditing is possible.
- **Challenge: achieve functionality without revealing data!**

Slide credit: George Danezis

Minimization technologies

Communication services

- email
- online social networks
- blogs
- web pages
- instant messaging
- (storage services)
- ...

Two properties of minimization technology for communication services

- (payload) confidentiality
- three related properties: unobservability, unlinkability, anonymity

Trusted relays and semitrusted relays

- Following convention, call the intermediaries ‘relays’
- there are approaches with trusted and semitrusted relays
- we’ll do this in more detail in the anonymity lectures; this will be a superficial introduction to follow the reading

Trusted relays

- Example: Type-0 Remailers.
  - a server keeps a dictionary between real and pseudonymous emails
  - request comes to the remailer, which forwards it, gets the response, and returns it to the user
- Example: VPNs

Question: problems with this?

Semi-trusted relays

- Example: Mix-nets (Chaum, 1980s). Routing protocol with a chain of servers called ‘mixes’ that shuffle (blocks from) messages received from multiple senders, and pass them to the next node, which could be another mix. Mixes only know their neighbors.

- sidenote: inexplicably, David Chaum is not cited in the reading. He is the originator of not just mix-nets but many of the ideas we are discussing.

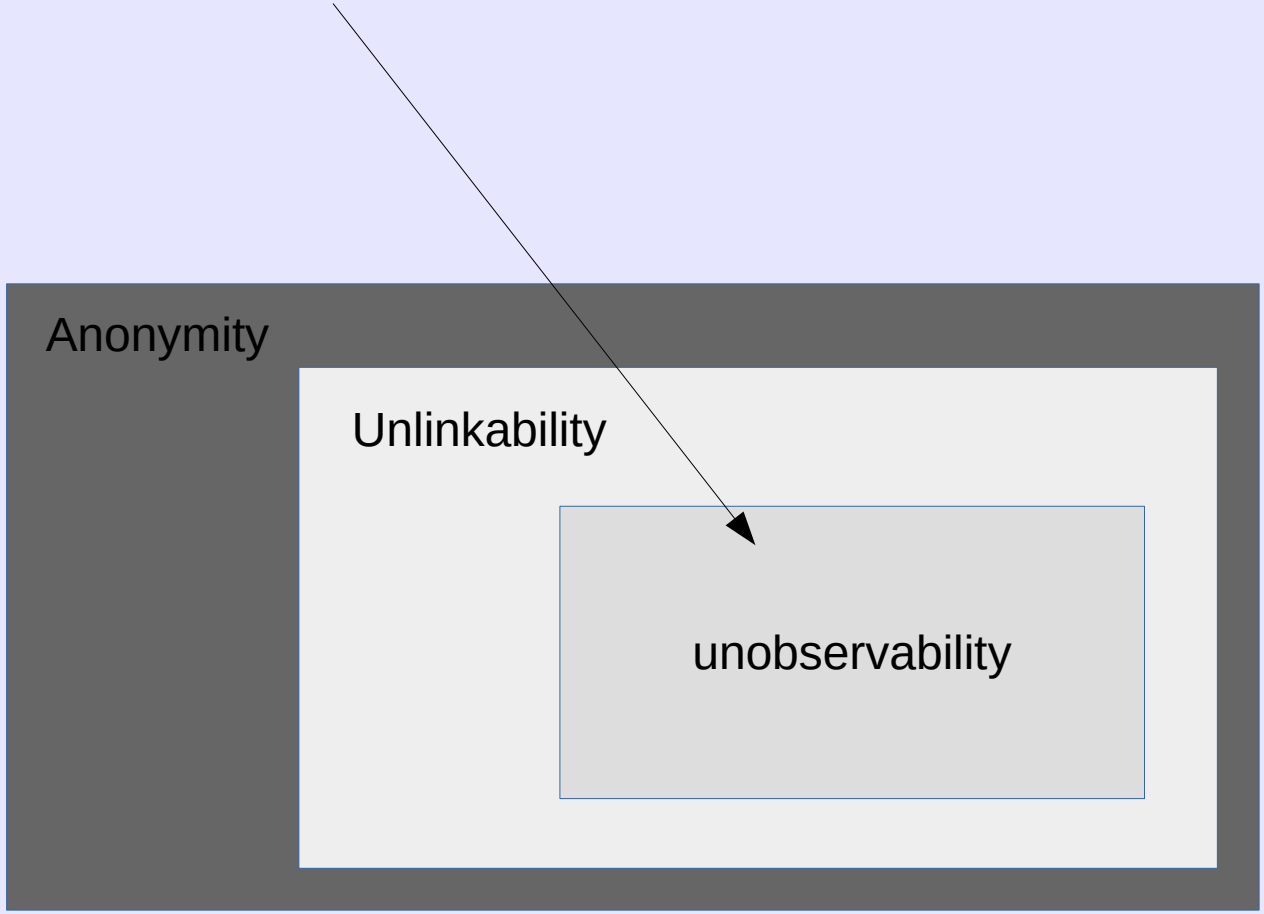
Two properties of minimization technology for communication services

- (payload) confidentiality
- three related properties: unobservability, unlinkability, anonymity
  - getting them involves an intermediary

Two properties of minimization technology for communication services

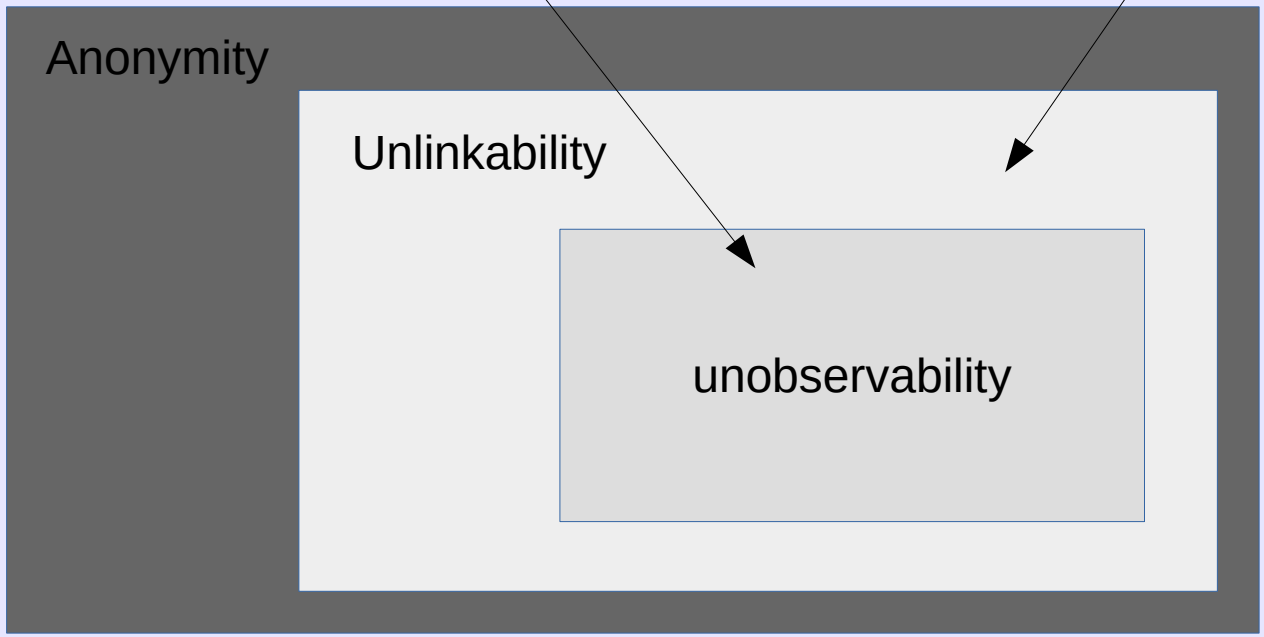
- (payload) confidentiality
- three related properties: unobservability, unlinkability, anonymity
  - getting them involves an intermediary
  - orderable by strength

can’t even tell that a message is being sent



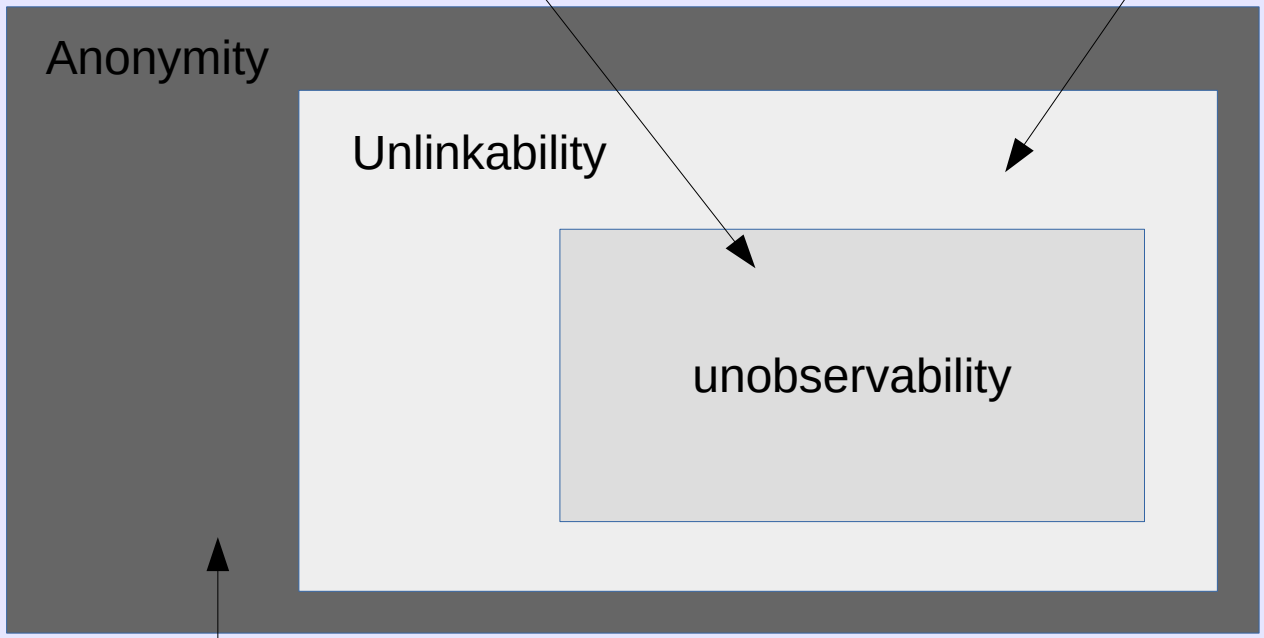
can’t even tell that a message is being sent

can tell that messages are sent, but can’t tell if messages are from the same source or different, going to the same place or different



can’t even tell that a message is being sent

can tell that messages are sent, but can’t tell if messages are from the same source or different, going to the same place or different



Can group messages by sender (receiver) but can’t identify the sender (receiver)

Question: What is it to identify a sender or receiver?

Anonymity set

- Anonymity is relative to a subset, called the anonymity set.
  - Think of it as answering “who might you be?”
- Can also consider the complement, “who is definitely not you?”

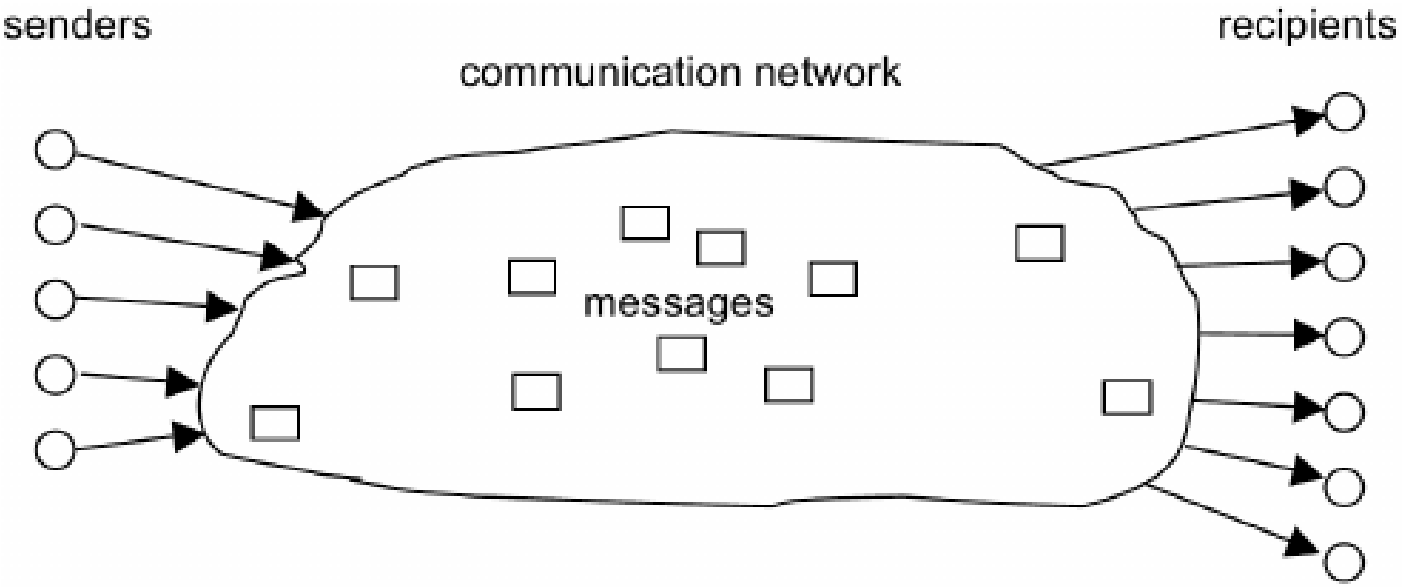
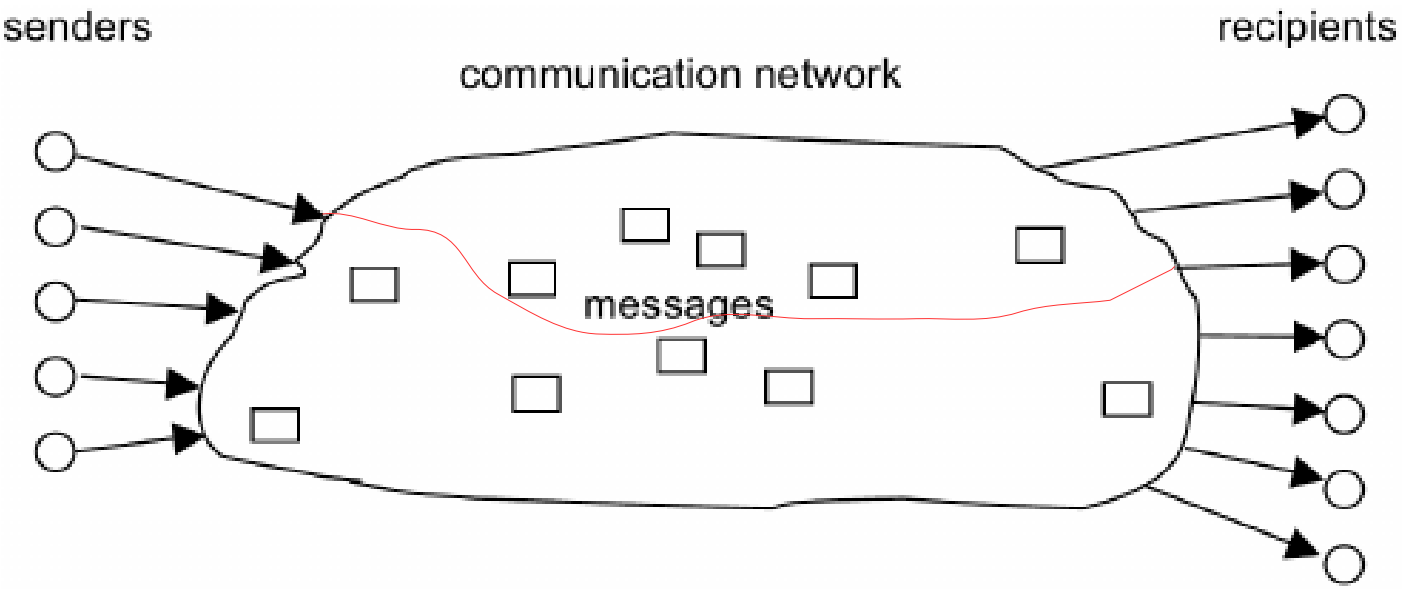
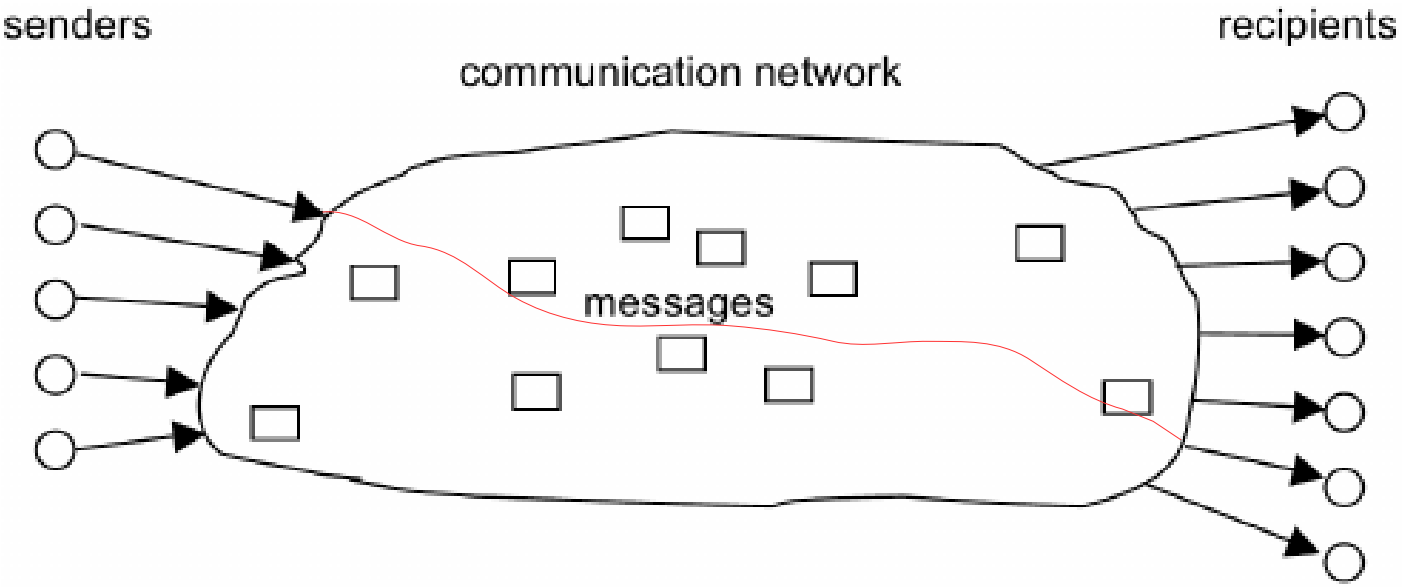
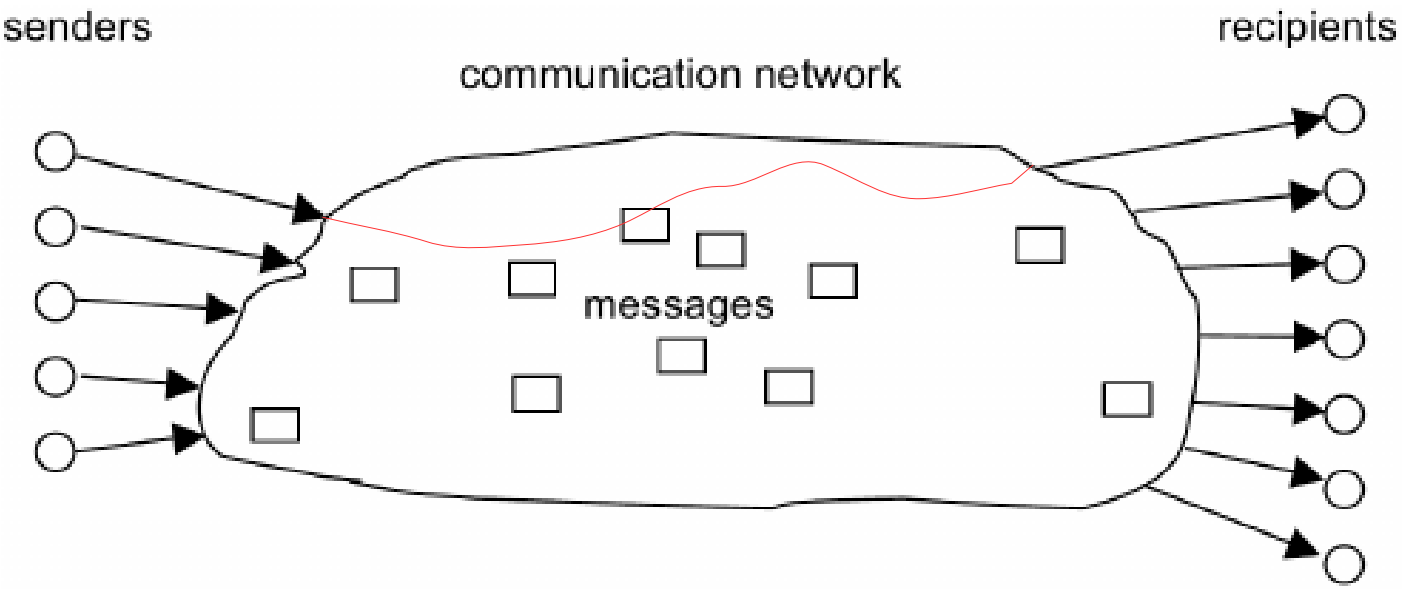
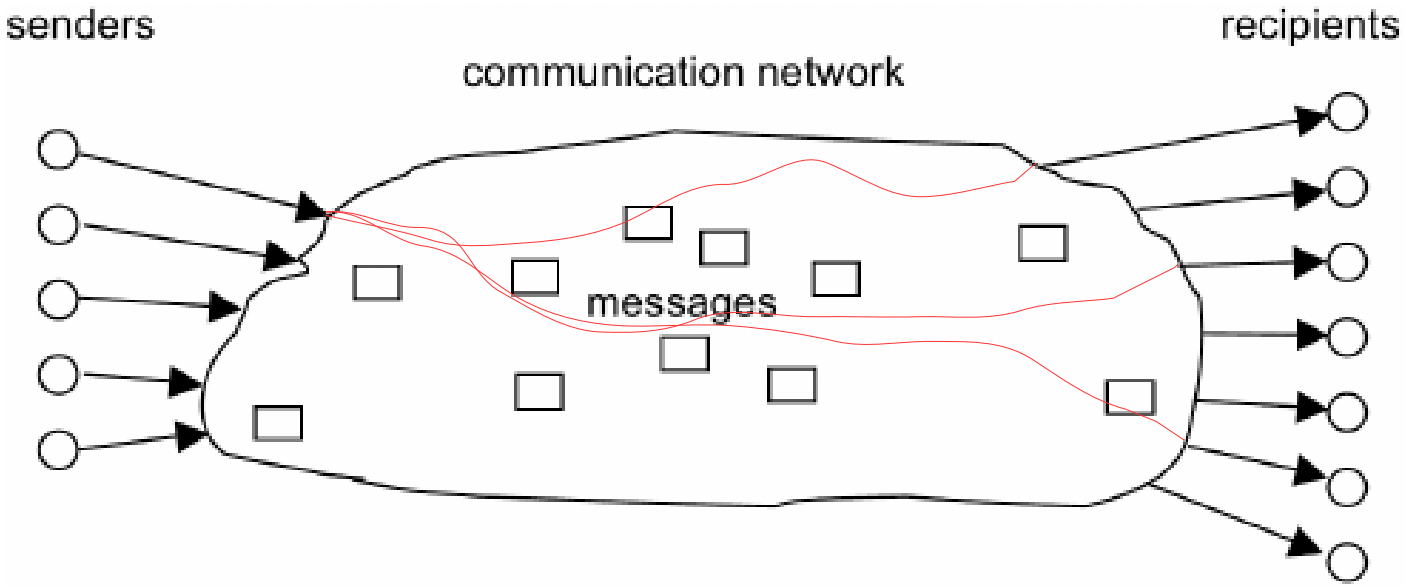


Image credit (before modification):
Christina Pöpper
Ruhr-University Bochum


senders
recipients
communication network
messages


senders
recipients
communication network
messages


senders
recipients
communication network
messages


senders
communication network
recipients
messages

Anonymity set

- Can you clearly describe the limiting cases for the anonymity set?

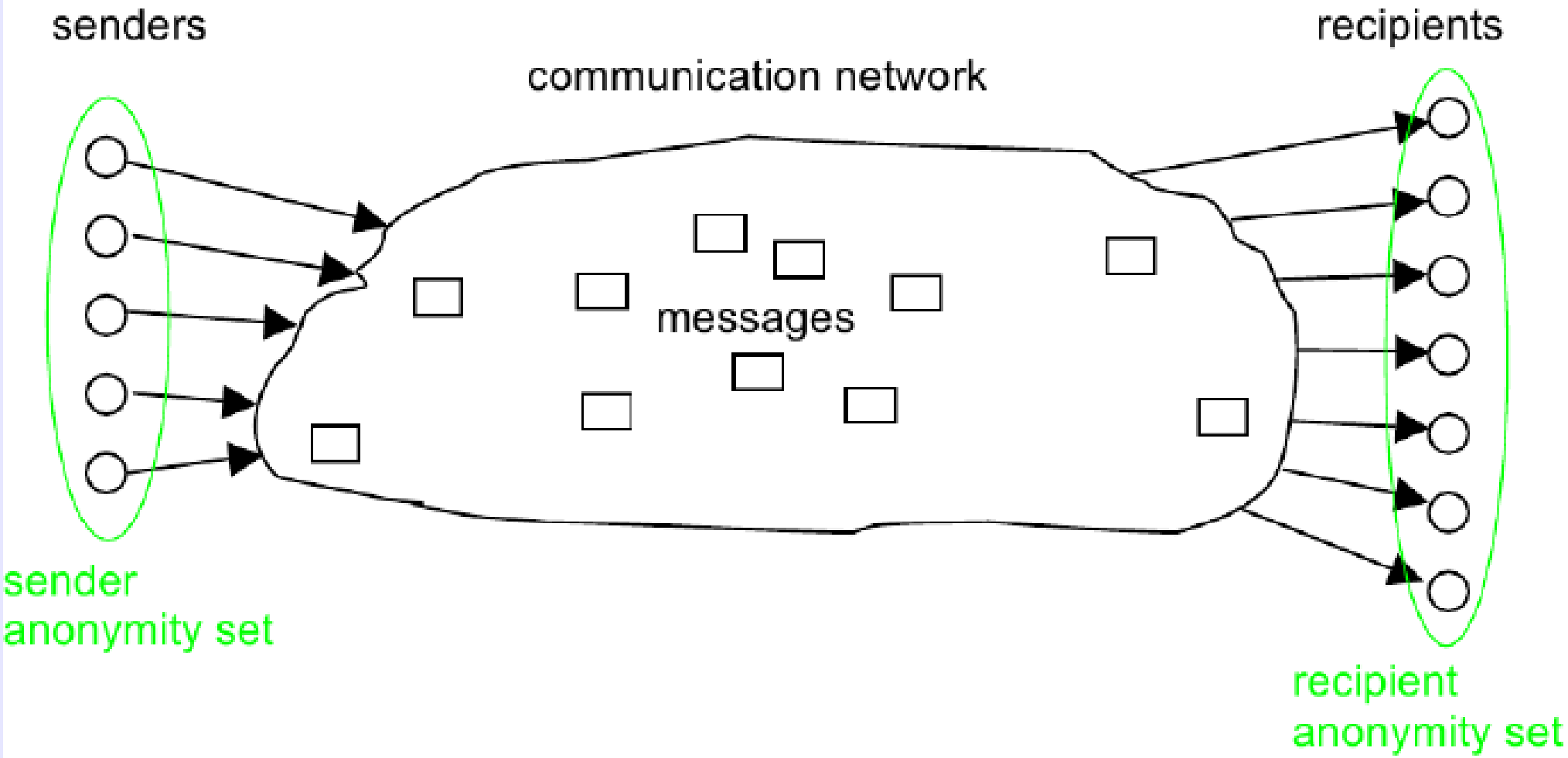


Image credit Christina Pöpper
Ruhr-University Bochum

On the 9th: Integrated class

- Remember to surveil yourself over time (at least 8 hours) in preparation for our integrated class
  - if you don’t feel comfortable doing so, just let me know well in advance
  - Meet here as usual