

Kim Wuyts



Owuytski



@kimw@mastodon.social

Privacy Threat Modeling



OUTLINE

PRIVACY

- Why?
- What?
- Hands-on exercise (15min)

PRIVACY THREAT MODELING

- How?
 - LINDDUN / LINDDUN GO
 - Best practices
- Hands-on exercise (15min)

TAKE AWAYS

LINDDUN PRIVACY THREAT MODELING APPROACH

CREATED 10+ YEARS AGO
 AT DISTRINET/COSIC (KU LEUVEN)
 TOPIC OF ONGOING DEVELOPMENT

 INSPIRED BY AND ALIGNED WITH MICROSOFT'S STRIDE

SYSTEMATIC IDENTIFICATION OF PRIVACY ISSUES







WHY PRIVACY MATTERS?

```
01011100101100100011100101011110011
```

I HAVE DONE NOTHING WRONG, SO I HAVE NOTHING TO HIDE

MISCONCEPTION

WHY PRIVACY MATTERS?

```
1001
                 01111
 110101
                101010
 1 1 0 0 1
1110101
               J 1010101
                                1100
                                101
                             110100
                               111001:
      10
                                010110
                                010010
                                110011
                00'
                     100
                10'
                     1 0 1
```

I HAVE DONE NOTHING WRONG, SO I HAVE NOTHING TO HIDE



Roomba testers feel misled after intimate images ended up on Facebook





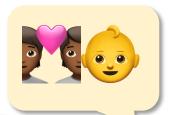
A run a day won't keep the hacker away: privacy in sports apps often subpar

The Strava the End of The US military security policies data shared on soc

bases and patrol rc



WHY PRIVACY MATTERS?



How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did

'Gaydar' on race Feb 2012 **Your Friends Reveal Sexual Orientation?**

ABCnews Sept 2009

Sex toy company admits to recording users' remote sex sessions, calls it a 'minor bug' The Verge, Nov 2017

WHY SHOULD PRIVACY MATTER FOR COMPANIES?



WE DO NOT PROCESS PERSONAL DATA!

MISCONCEPTION

PERSONAL DATA: any information that relates to an identified or identifiable living individual



social security number

name

IP address

Cookie ID

address

date of birth

browser fingerprint

location data





PRIVACY ENGINEERING

TRANSPARENCY
PREDICTABILITY



LINDDUN



LINKING



IDENTIFYING



NON-REPUDIATION



DETECTING



DATA DISCLOSURE



UNAWARENESS



NON-COMPLIANCE





LINKING

PLAYING "GUESS WHO"

Linking multiple properties to the same individual

VS.

IDENTIFYING

WINNING "GUESS WHO"

Reducing the set of individuals to one.

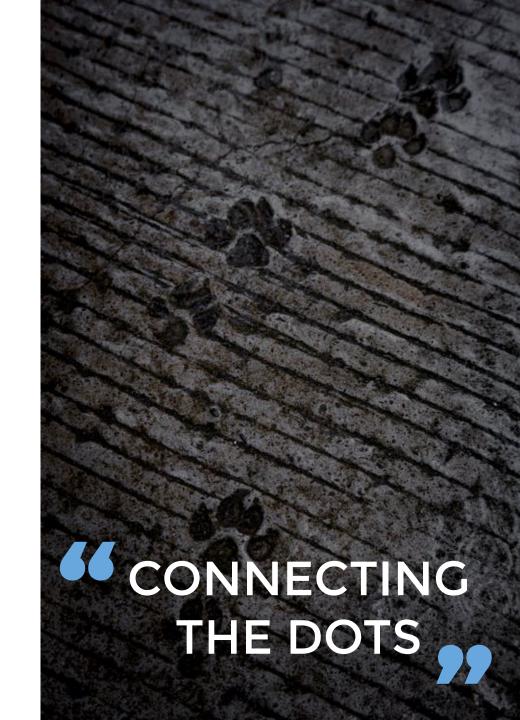


LINKING

LEARNING MORE ABOUT AN INDIVIDUAL (OR GROUP) BY MATCHING DATA ITEMS

TOGETHER

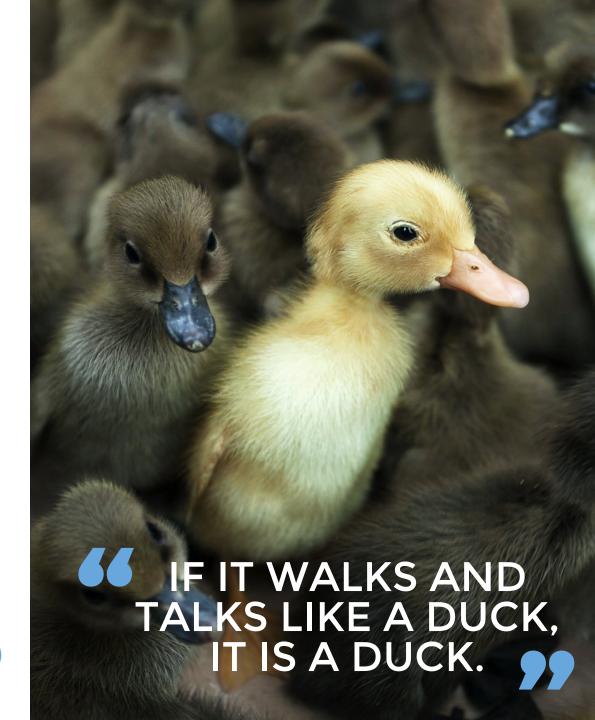
- Through identifiers
- Through combination
- Through profiling/derivation/inference



IDENTIFYING

LEARNING THE IDENTITY

- Through direct identifiers
- Through identifiable information
 - Pseudonyms
 - Revealing content
 - Small anonymity set (set of individuals)



EXAMPLE Identifying

Aol.

- Clothes for age 60
- 60 single men
- Best retirement city
- Jarrett arnold
- Jack t. arnold
- Jaylene and jarrett arnold
- Gwinnett county yellow pages
- Rescue of older dogs
- Movies for dogs
- Sinus infection



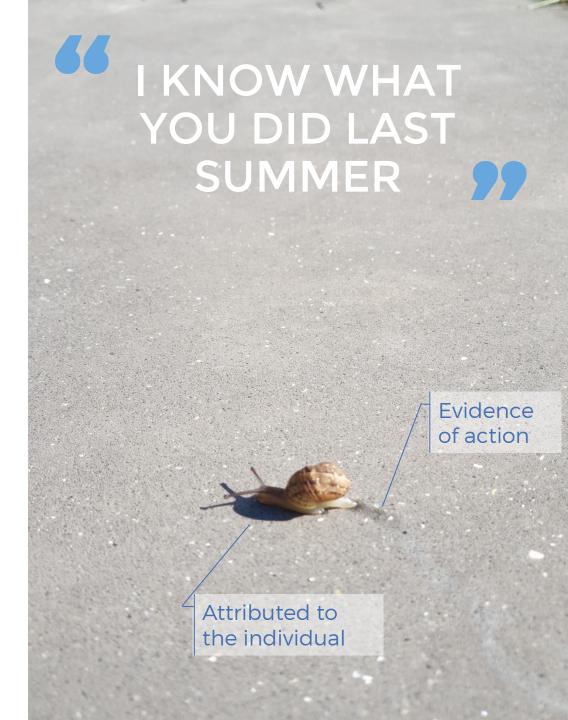
Thelma Arnold 62 year old widow Lilburn, Georgia



NON-REPUDIATION

PROOF OF A CLAIM
ABOUT AN INDIVIDUAL

- Evidence of the claim / action
- Attribution to the individual





DETECTING

DEDUCING SUBJECT INVOLVEMENT
BY OBSERVING EXISTENCE OF
RELEVANT INFORMATION

- Observed communication
- Application side-effects
- System responses



DATA DISCLOSURE

UNNECESSARY USE

OF DATA

- Excessive data types
- Excessive volume
- Excessive processing
- Excessive exposure

- collection
- storage
- processing
- sharing



- Unawareness of data subject
- Unawareness of user sharing personal data (about others or themselves)



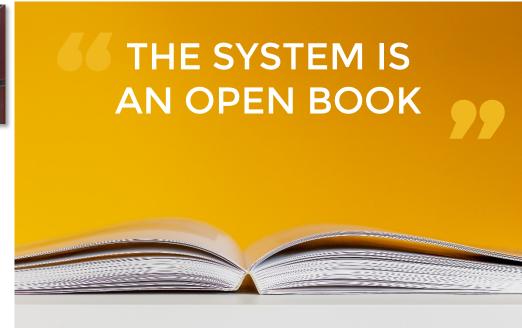
INSUFFICIENTLY INFORMING ABOUT

THE PROCESSING OF PERSONAL DATA





- Lack of preferences control
- Lack of access
- Lack of rectification/erasure





DARK PATTERNS

HIDING THE PRIVACY INFORMATION

EXAMPLEUnawareness

WANT TO KNOW MORE?

and therefor do you allow us to collect and process all your personal data?

YES!
SHOW ME THE COOL STUFF

No. I don't consent

NUDGING TO SELECT THE LESS PRIVACY-FRIENDLY BUTTON

HIDING THE OPT-OUT

NON-COMPLIANCE

LACK OF ADHERENCE TO LEGISLATION,
REGULATION, STANDARDS AND BEST
PRACTICES

- Lawfulness
- Data lifecycle management
- Cybersecurity risk management



TRUTH PRIVACY REQUIRES A DIFFERENT MINDSET

000

SECURITY

- Protecting data
- Company assets
- (External) attacker

PRIVACY

- Protecting personal data
- Data subject assets
- Attacker + (internal)'misbehavior'

SECURITY AND PRIVACY



PRIVACY DOESN'T NEED TO CONFLICT SECURITY

HANDS-ON #1

- Threat Modeling Manifesto value

Doing threat modeling over talking about it.





Privacy Threat Modeling



OUTLINE

PRIVACY

- Why?
- What?
- Hands-on exercise (15min)

PRIVACY THREAT MODELING

- How?
 - LINDDUN / LINDDUN GO
 - Best practices
- Hands-on exercise (15min)

TAKE AWAYS

THREAT MODELING



2. WHAT CAN GO WRONG?

3. WHAT TO DO ABOUT IT?

4. WAIT A MINUTE?!

HOW TO THREAT MODEL?



1. MODEL THE SYSTEM

Create DFD / white board sketch / ...

2. ELICIT THREATS

- Map model components
- Identify threats

3. MITIGATE THREATS

- Assess & prioritize
- Mitigate

4. REFLECT

• Reflect & repeat

All models are wrong, some are useful - G. Box

1. MODEL THE SYSTEM

Create DFD / white board sketch / ...

2. ELICIT THREATS

- Map model components
- Identify threats

3. MITIGATE THREATS

- Assess & prioritize
- Mitigate

4. REFLECT

• Reflect & repeat

7 • • • • • •

1. MODEL THE SYSTEM

- 2. ELICIT THREATS
- 3. MITIGATE THREATS
- 4. REFLECT



PROCESS

1. MODEL THE SYSTEM

Create DFD / white board sketch / ...

NOW WHAT?



2. ELICIT THREATS

- Map model components
- Identify threats

3. MITIGATE THREATS

- Assess & prioritize
- Mitigate

4. REFLECT

• Reflect & repeat

REUSABLE KNOWLEDGE

STRIDE

SPOOFING

TAMPERING

REPUDIATION

INFORMATION DISCLOSURE

DENIAL OF SERVICE

ELEVATION OF PRIVILEGE

LINDDUN

LINKING

IDENTIFYING

NON-REPUDIATION

DETECTING

DATA DISCLOSURE

UNAWARENESS

NON-COMPLIANCE

PROCESS

1. MODEL THE SYSTEM

Create DFD / white board sketch / ...

2. ELICIT THREATS

- Map model components
- Identify threats

3. MITIGATE THREATS

- Assess & prioritize
- Mitigate

4. REFLECT

• Reflect & repeat

REUSABLE KNOWLEDGE

PROCESS

Create DFD / white board sketch / ...

1. MODEL THE SYSTEM



The data sent to the system can be used to identify the user (with a sufficient degree of likelihood).

- (i.e. identified data, data that can be linked to already obtained identified data, or data that, when combined, become identifie
- Data subject anonymously shares his preferences in a feedback form (of his employer, school, ...). When these preferences are unique, they can identify the user.
- Data subject can be identified by linking data to previously obtained data (from same or other
- Likelihood depends on previous knowledge of

13

- The data subject is not necessarily the sender.
- Combining several data items can lead to identification.
- Identifying credentials (II) and actions (I2) are subtypes of this threat.

LINDDUN

2. ELICIT THREATS

- Map model components
- Identify threats

3. MITIGATE THREATS

- Assess & prioritize
- Mitigate

4. REFLECT

Reflect & repeat

WWW.LINDDUN.ORG

Identifiability of

transactional data (transmitted data)

(future) receive

untrusted

Data available to

untrusted party

Data flow not fully

protected

Disclosure

🔼 LINDDUN

LINDDUN - privacy threat trees

dentifiability of data flow

Identifishility

of content

(weak anonymizatio

Identifiability of

contextual data

Non-anonymous

Based on

Insecure anonymit

system deployed

Passive attacks

Traffic analysis possible

Based on behavioral patterns (time,

frequency, location

Active attack possible

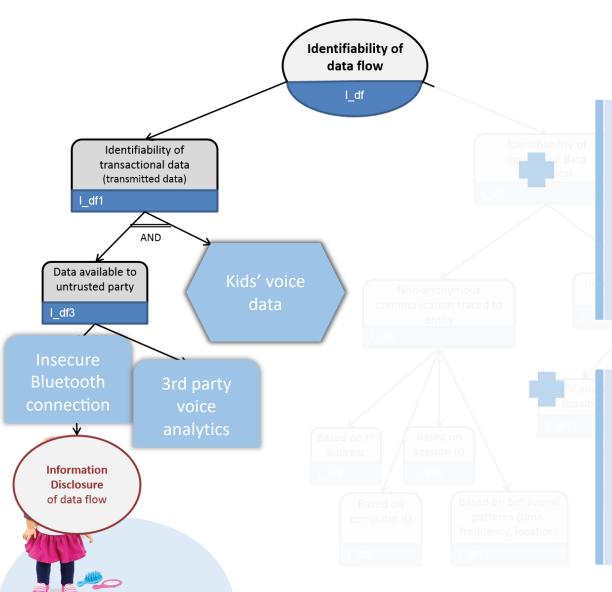
LINDDUN GO cards

1. MODEL THE SYSTEM

2. ELICIT THREATS

3. MITIGATE THREATS

4. REFLECT

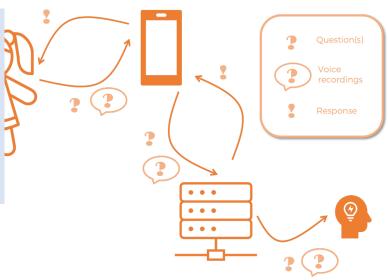


THREAT 01

Identifiable kids'
voice data is being
sent over an insecure
communication
channel

THREAT 02

Identifiable kids'
voice data is
being shared with
an untrusted 3rd party





LINDDUN GO

LEAN APPROACH TO PRIVACY THREAT MODELING

DECK OF 34 PRIVACY THREAT CARDS
DESCRIBING MOST COMMON THREATS
FOR EACH LINDDUN CATEGORY

FACILITATE AND STRUCTURE
DISCUSSION

COLLABORATIVE APPROACH
TO ENGAGE ALL PARTICIPANTS

- **Prioritize** threats
 - assess risk (impact & likelihood)

- Mitigate threats
 - Tactics & strategies
 - Privacy patterns
 - PETs



PROCESS

1. MODEL THE SYSTEM

Create DFD / white board sketch / ...

2. ELICIT THREATS

- Map model components
- Identify threats

3. MITIGATE THREATS

- Assess & prioritize
- Mitigate

4. REFLECT

Reflect & repeat

- 1. MODEL THE SYSTEM
- 2. ELICIT THREATS
- 3. MITIGATE THREATS

4. REFLECT

THREAT 01

Identifiable kids' voice data is being sent over an insecure communication channel

THREAT 02

Identifiable kids'
voice data is being
shared with
an untrusted 3rd party

Before sharing

- Hide Restrict access. Secure communication between doll and phone.
- Separate Distribute processing. Local speech to text translation (no sharing of voice to the back-end).

When shared to back-end

- Abstract summarize/group/perturb recordings.
 When share to external party, aggregate data, scramble recordings, etc.
- Minimize select/exclude/strip/destroy data. Don't store recordings. Delete once speech is translated to text. Don't link questions to user profiles.



Privacy design strategies

(Hoepman)



Jaap-Henk Hoepman, Privacy design strategies (little blue book) https://www.cs.ru.nl/~jhh/publications/pds-booklet.pdf

MITIGATING THREATS

Minimise

 Limit as much as possible the processing of personal data.

Separate

Separate the processing of personal data as much as possible.

Abstract

• Limit as much as possible the detail in which personal data is processed.

Hide

Protect personal data, or make it unlinkable or unobservable.
 Make sure it does not become public or known.

Inform

 Inform data subjects about the processing of their personal data in a timely and adequate manner.

Control

 Provide data subjects adequate control over the processing of their personal data.

Enforce

 Commit to processing personal data in a privacy-friendly way, and adequately enforce this.

Demonstrate

 Demonstrate you are processing personal data in a privacyfriendly way.

MITIGATE THREATS

- LINDDUN privacy mitigation strategies and privacy solutions www.linddun.org/mitigation-strategies-and-solutions
- NIST SP 800-53 security and privacy controls
 https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-53r5.pdf

Privacy design strategies (Hoepman)
 https://www.cs.ru.nl/~jhh/publications/pds-booklet.pdf

Patterns

Security: https://securitypatterns.distrinet-research.be/

• Privacy: <u>privacypatterns.org</u> <u>privacypatterns.eu</u>

 OWASP cheat sheets: https://cheatsheetseries.owasp.org/

• ...



DID I DO A GOOD ENOUGH JOB?

PROCESS

1. MODEL THE SYSTEM

Create DFD / white board sketch / ...

2. ELICIT THREATS

- Map model components
- Identify threats

3. MITIGATE THREATS

- Assess & prioritize
- Mitigate

4. REFLECT

• Reflect & repeat

USEFUL RESOURCES

- Threat modeling. Designing for security. By Adam Shostack, 2014.
- Threat Modeling A Practical **Guide for Development Teams** by Izar Tarandach & Matthew J. Coles, 2020
- Securing systems. Applied security architectures and threat models by Brook Schoenfield, 2015.
- Threat Modeling Manifesto www.threatmodelingmanifesto.org
- Threat Modeling Connect community www.threatmodelingconnect.com

THREAT MODELING **APPROACHES**

- STRIDE
- LINDDUN PRIVACY

- EoP
- PASTA
- TRIKE
- TARA
- Continuous Threat Modeling

- INCLUDES NO DIRT PRIVACY
- PLOT4AI PRIVACY
- TRIM PRIVACY
- STRIPED PRIVACY

PRIVACY THREAT MODELING

BEST PRACTICES



PRIVACY THREAT MODELING

What you need?

Understanding of the system & data (dfd, whiteboard sketch, ...)

What you do?

Analyze the privacy threats that are posed by the different threat sources at play to the individuals associated with the data that are being processed by the system

PRIVACY THREAT KNOWLEDGE

FOCUS AREAS



DATA PERSPECTIVES





PRIVACY THREAT KNOWLEDGE

WHAT CAN GO WRONG?

LINDDUN

- Threat trees
- GO cards
- Threat categories

Use **knowledge** as

- Structured guidance
- ☐ Facilitation of discussion
- Gap analysis
- **.**...

Privacy threat categories

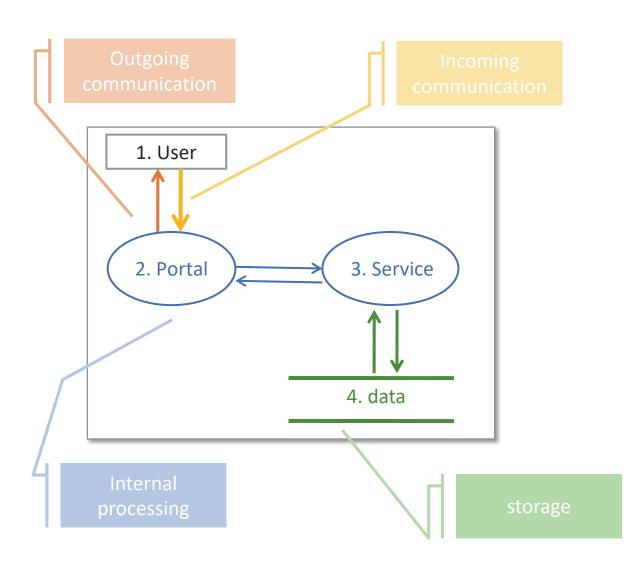
Classification of known privacy issues.

- Data disclosure
- Linking
- Identifying
- Detecting *
- Non-repudiation *
- Unawareness
- Non-compliance *

1



FOCUS AREAS



WHAT PART OF THE SYSTEM ARE WE ASSESSING?

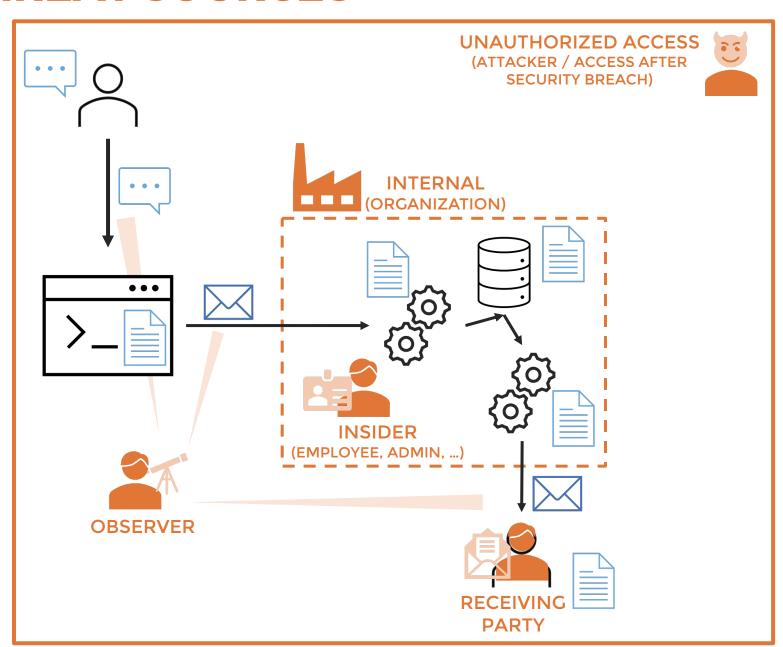
Focus areas

Parts of the system model that require a detailed privacy analysis.

- Incoming communication
- Outgoing communication
- Storage
- Internal processing *



THREAT SOURCES







RECEIVING PARTY

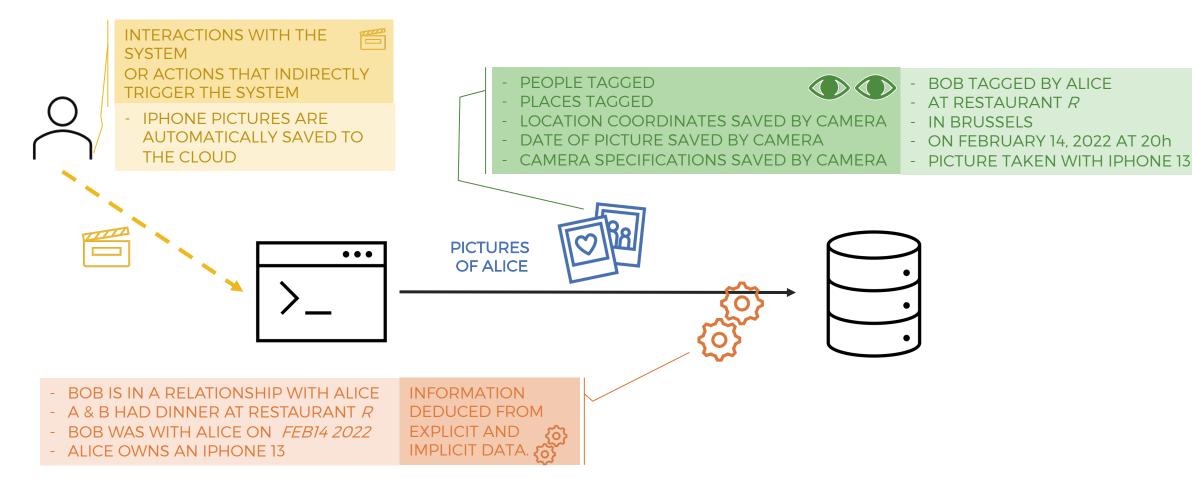


INSIDER **
(EMPLOYEE, ADMIN, ...)





UNAUTHORIZED ACCESS **
(ATTACKER / ACCESS AFTER
SECURITY BREACH)



DATA PERSPECTIVES



NOT A BLACK/WHITE CONCEPT



INSPIRATION

not LIMITATION

"Allow for creativity by including both craft and science."

- Threat Modeling Manifesto pattern Informed Creativity

HANDS-ON #2

- Threat Modeling Manifesto value

Doing threat modeling over talking about it.

EXERCISE 2: LINDDUN GO

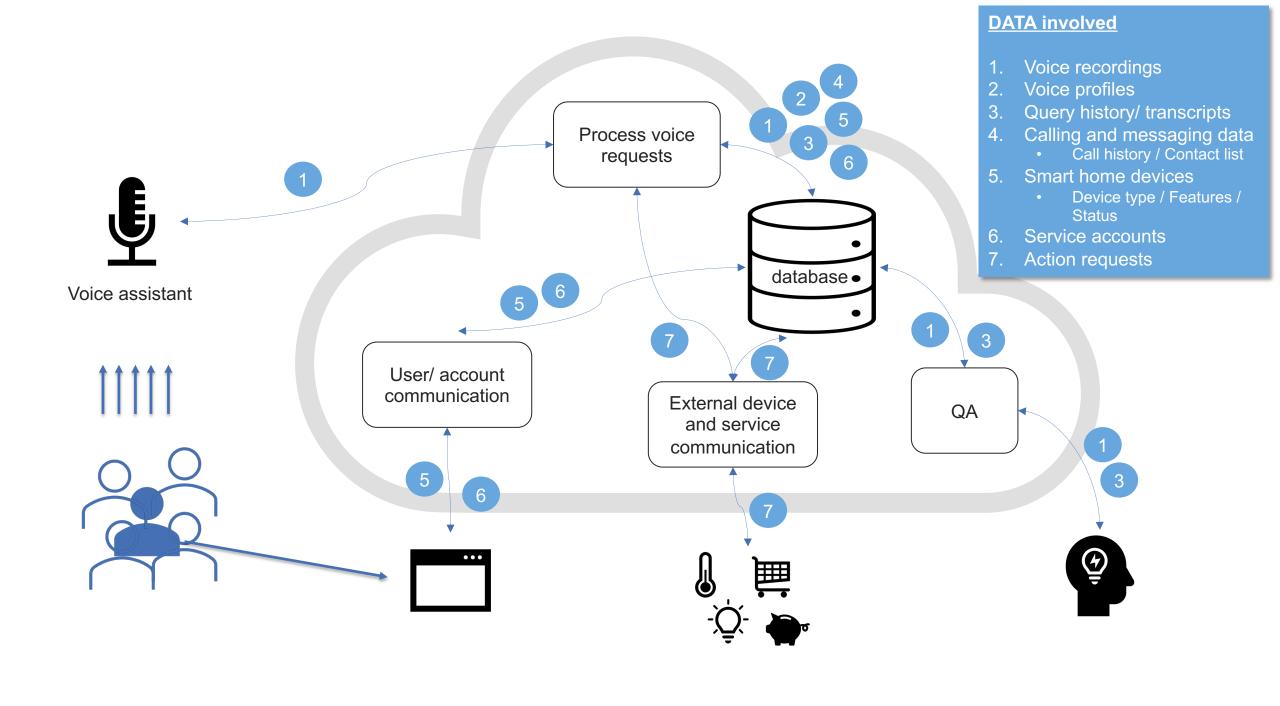
VOICE ASSISTANTS



Siri Privacy Whistleblower Unmasks to **Urge Stricter Voice Assistant Privacy** Regulation



'Alexa, are you invading my privacy?' the dark side of our voice assistants



LINKABILITY OF INBOUND DATA

- Hotspot

INBOUND FLOW CONTAINING PERSONAL DATA



- Threat source —

ORGANIZATIONAL

The data sent to the system are linked to already collected data of the same or other data subjects (from same or other source).

- 1. Does the flow contain personal data?
 - 2. Does (or can) the system link these data (i.e. are data items sufficiently unique to link to each other) in a privacy-violating way?

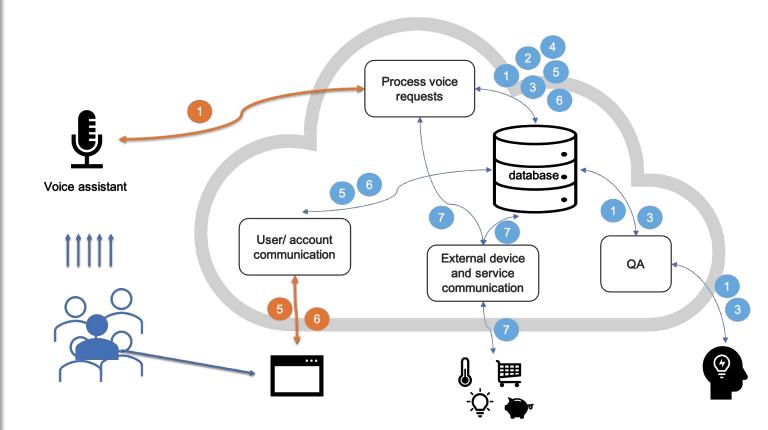


- Data subject shares minimal set of information (e.g. city instead of full address), yet given the information already available (e.g. only 1 person of that city in the system), the data can be easily linked.
- Information can be deduced based on the linked data (inference).
- Threat depends on the knowledge of the organization.
- The data subject is not necessarily the sender of the data.
- Linkability of credentials (L1) and actions (L2) are subtypes of this threat.





- 1. Voice recordings
- 2. Voice profiles
- 3. Query history/ transcripts
- . Calling and messaging data
 - Call history / Contact list
- 5. Smart home devices
 - Device type / Features / Status
- 6. Service accounts
- 7. Action requests







EXERCISE 2: LINDDUN GO

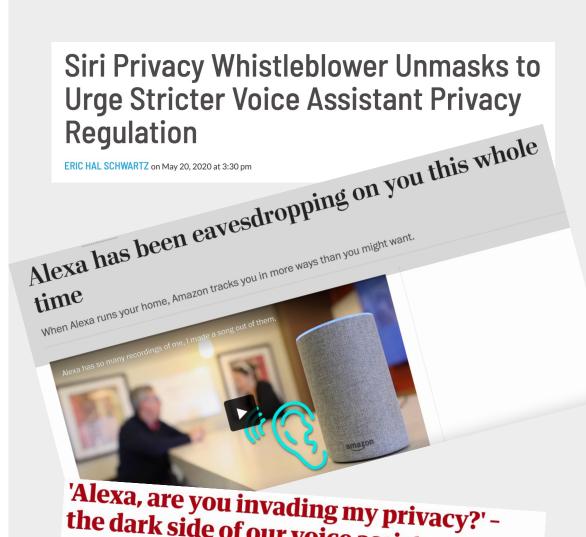
VOICE ASSISTANTS



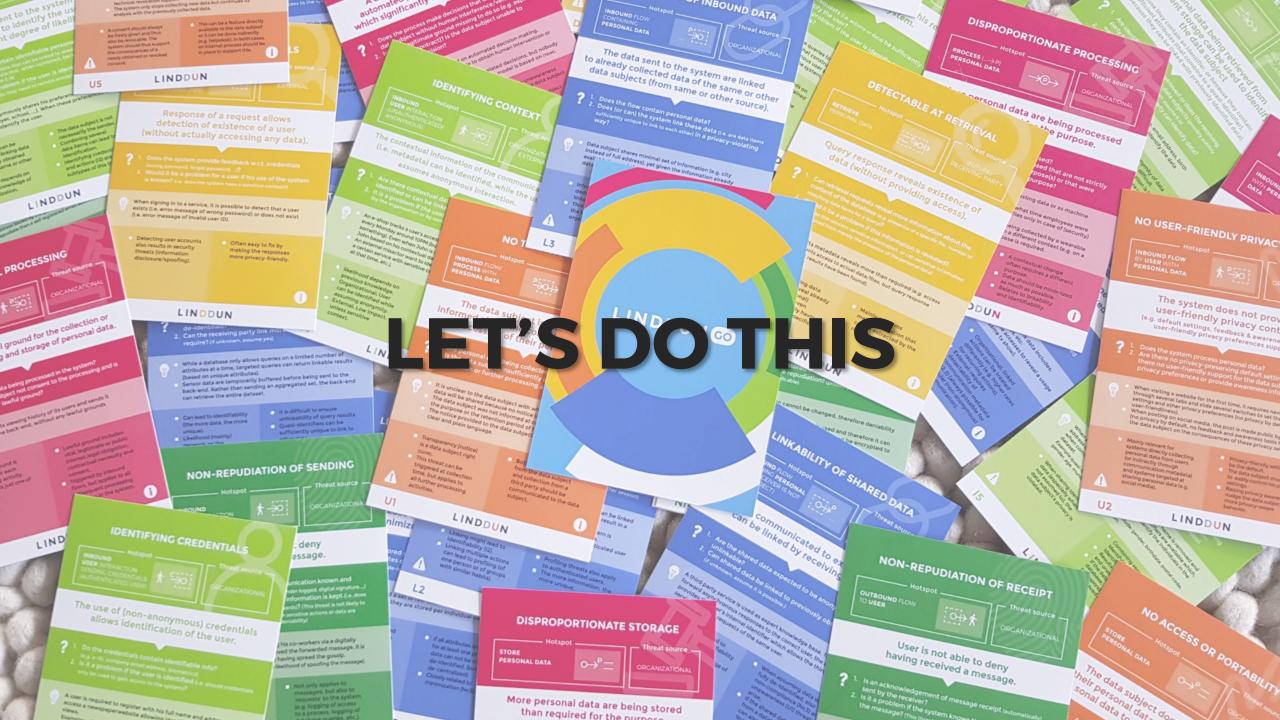
ITERATE OVER THE LINDDUN GO CARDS (SUBSET IN SLIDES). DISCUSS AND LIST ALL POTENTIAL PRIVACY THREATS RELATED TO THE LINDDUN GO CARD(S) FOR THE SYSTEM COMPONENTS INVOLVED.

START WITH THE # OF YOUR BREAKOUT GROUP

TAKE TURNS DISCUSSING POTENTIAL THREATS



the dark side of our voice assistants



LINKABILITY OF INBOUND DATA

Hotspot

INBOUND FLOW CONTAINING PERSONAL DATA



Threat source —

ORGANIZATIONAL

The data sent to the system are linked to already collected data of the same or other data subjects (from same or other source).

- 1. Does the flow contain personal data?
 - 2. Does (or can) the system link these data (i.e. are data items sufficiently unique to link to each other) in a privacy-violating way?

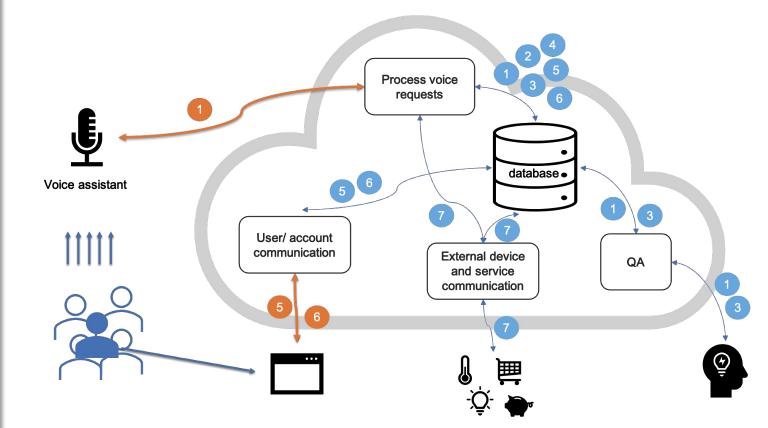


- Data subject shares minimal set of information (e.g. city instead of full address), yet given the information already available (e.g. only 1 person of that city in the system), the data can be easily linked.
- Information can be deduced based on the linked data (inference).
- Threat depends on the knowledge of the organization.
- The data subject is not necessarily the sender of the data.
- Linkability of credentials (L1) and actions (L2) are subtypes of this threat.





- 1. Voice recordings
- 2. Voice profiles
- 3. Query history/ transcripts
- I. Calling and messaging data
 - Call history / Contact list
- 5. Smart home devices
 - Device type / Features / Status
- 6. Service accounts
- 7. Action requests







NON-REPUDIATION OF SENDING

- Hotspot

INBOUND
USER INTERACTION



Threat source —

ORGANIZATIONAL

The user cannot deny having sent a message.

- Is the origin of incoming communication known and traceable to the sender? (e.g. sender logged, digital signature,...)
 - 2. Is it a problem if a trace of this information is kept (i.e. does the sender require deniability afterwards)? (This threat is not likely to be applicable as it only applies when sensitive actions or data are being communicated that require deniability)
- An employee shares gossip among his co-workers via a digitally signed email. When his boss received the forwarded message, it is difficult for the employee to deny having spread the gossip.

 (level of deniability depends on the likelihood of spoofing the message).
 - Mainly applies when the receiving end requires a proof of authenticity during communication, but the sender wants to be able to deny to external parties (afterwards).
- Not only applies to messages, but also to 'requests' to the system (e.g. logging of access to a process, logging of database queries, etc.).
- Credential non-repudiation (Nr1) is a subtype of this threat.

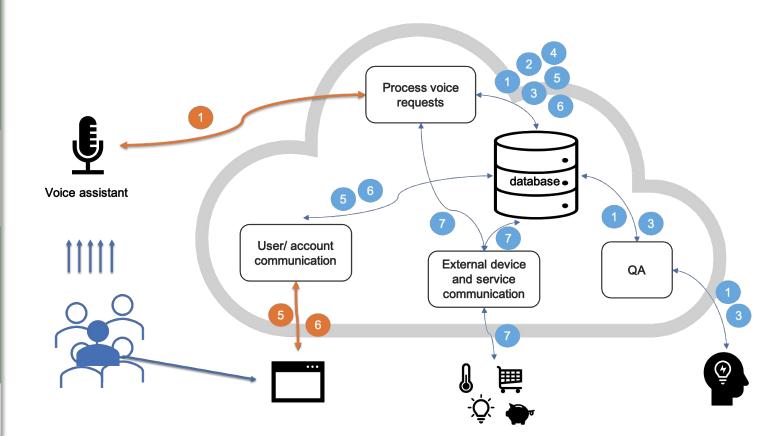


Nr2





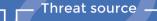
- 1. Voice recordings
- 2. Voice profiles
- 3. Query history/ transcripts
- 4. Calling and messaging data
 - Call history / Contact list
- 5. Smart home devices
 - Device type / Features / Status
- 6. Service accounts
- 7. Action requests



LINKABILITY OF SHARED DATA

- Hotspot

OUTBOUND FLOW



RECEIVING PARTY

CONTAINING PERSONAL **DATA (RECEIVER IS NOT** DATA SUBJECT)

Content communicated to external party can be linked by receiving party.

- 1. Are the shared data expected to be anonymous or unlinkable?
 - 2. Can shared data be linked to previously obtained data? (if unknown, assume it is possible.)

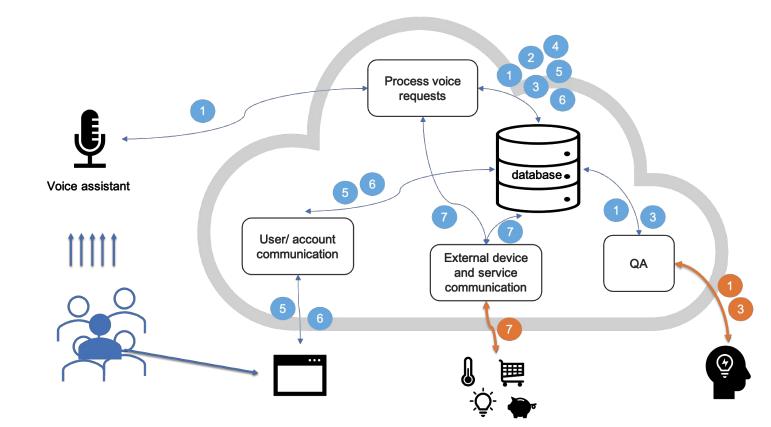


A third party service is used as expert knowledge base. To easily forward asynchronous responses to the correct user, the system provides the user's internal identifier which allows the third party service to link all requests of the same user.

- Linkability can lead to profiling and identifiability
- Depends on knowledge of the receiving party.
- The more shared attributes, the higher the risk.
- When assuming data were fully de-identified, also non-compliance (Nc3) and unawareness (U1) threats will arise.
- If the shared data originate from a database, the threat can also be categorized as 'linkability of retrieved data (L7)'.



- Voice recordings
- Voice profiles
- Query history/ transcripts
- Calling and messaging data
 - Call history / Contact list
- Smart home devices
 - Device type / Features / Status
- Service accounts
- Action requests





IDENTIFYING CREDENTIALS

- Hotspot

INBOUND
USER INTERACTION
SENDING CREDENTIALS
(AUTHENTICATED USER)



Threat source

ORGANIZATIONAL

The use of (non-anonymous) credentials allows identification of the user.

- 1. Do the credentials contain identifiable info? (e.g. e-ID, company email address, biometrics)
 - 2. Is it a problem if the user is identified (i.e. should credentials only be used to gain access to the system)?



A user is required to register with his full name and address to access a newspaperwebsite allowing identification of webpage views.

Examples of identifying credentials include: email address with full name, e-ID, biometrics, too specific attributes of (anonymous) credentials, etc.

- When data are identified rather than identifiable, stronger security measures need to be in place.
- Relates to non-compliance (Nc1) and unawareness (U1).

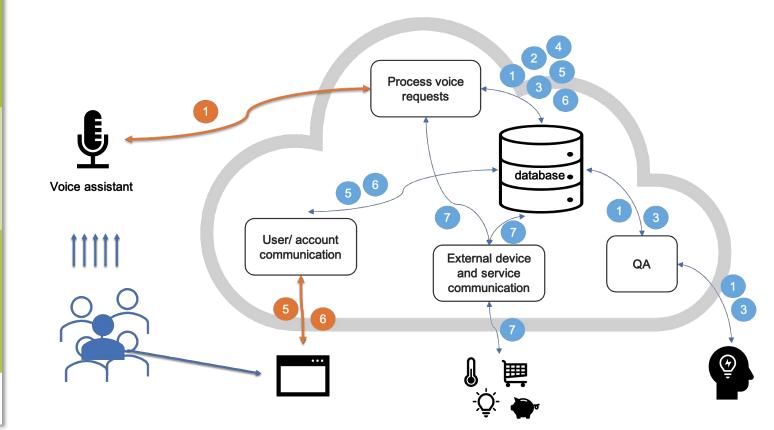








- 1. Voice recordings
- 2. Voice profiles
- 3. Query history/ transcripts
- 1. Calling and messaging data
 - Call history / Contact list
- 5. Smart home devices
 - Device type / Features / Status
- 6. Service accounts
- 7. Action requests



NO TRANSPARENCY

Hotspot

INBOUND FLOW/ ..
PROCESS WITH
PERSONAL DATA



ORGANIZATIONAL

The data subject is insufficiently informed about the collection and further processing of their personal data.

- 1. Are personal data being collected and/or processed?
 - 2. Is the data subject insufficiently informed about this collection or further processing activities?

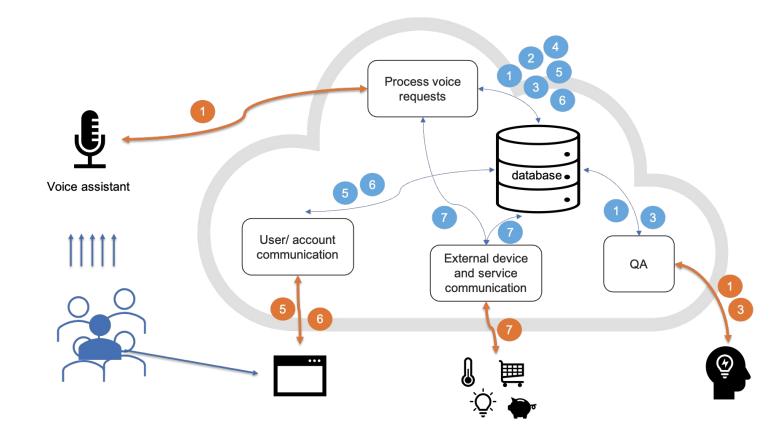


- It is unclear to the data subject with which third parties their data will be shared because no notice is provided.
- The data subject was not informed at collection time about the purpose or the retention period of their personal data.
- The notice provided to the data subject was not written in clear and plain language.
- Transparency (notice) is a data subject right [GDPR].
- This threat can be triggered at collection time, but applies to all further processing activities.
- Both collection directly from the data subject and collection from a third party should be communicated to the data subject.





- 1. Voice recordings
- 2. Voice profiles
- 3. Query history/ transcripts
- L. Calling and messaging data
 - Call history / Contact list
- 5. Smart home devices
 - Device type / Features / Status
- 6. Service accounts
- 7. Action requests







EXERCISE 2: READOUT

LINDDUN GO – VOICE ASSISTANTS

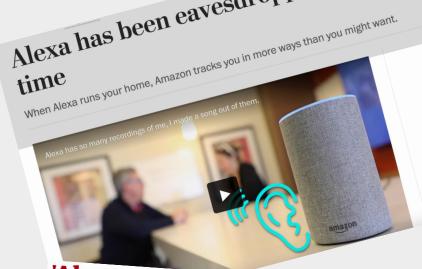


SHARE 1-2 SCENARIOS THAT YOU IDENTIFIED.

HOW WAS THE INDIVIDUAL'S PRIVACY VIOLATED?

HOW WOULD YOU MITIGATE THIS?

Siri Privacy Whistleblower Unmasks to Urge Stricter Voice Assistant Privacy Regulation ERIC HAL SCHWARTZ on May 20, 2020 at 3:30 pm Alexa has been eavesdropping on you this whole time.



'Alexa, are you invading my privacy?' the dark side of our voice assistants



Take aways

- Privacy matters
- Security practices, such as threat modeling, are equally applicable to privacy
- Privacy requires a different mindset than security
- LINDDUN threat trees or LINDDUN GO cards can facilitate privacy threat discussions

INDDUN FEEDBACK? LET US KNOW!



Kim Wuyts

Privacy engineering researcher | Threat modeling enthusiast | privacy-by-design advocate | LINDDUN privacy threat modeling designer



- PhD in privacy engineering
- Researcher at imec-DistriNet, KU Leuven, Belgium











Kim Wuyts



Owuytski



@kimw@mastodon.social