Privacy threat analysis

Mina Deng,

Kim Wuyts, Riccardo Scandariato,

Bart Preneel, Wouter Joosen

KATHOLIEKE UNIVERSITEIT LEUVEN

DistriNet

IBBT RESEARCH GROUP

Objectives

- Understanding the nature of privacy requirements and their relationship with anti-requirements
- Method to elicit privacy anti-requirements (LINDDUN)
- Documenting privacy threats

Overview

- Privacy
  - What?
  - Properties
- Privacy methodology
- Example case study
- Project information

Privacy

- what is privacy?
  - Confidentiality
  - Data minimization
  - User empowerment
  - …

What is privacy?

- The right to be let alone *(Warren & Brandeis, 1890)*
- The right of the individual to decide what information about himself should be communicated to others and under what circumstances *(Westin, 1970)*
- Freedom from unreasonable constraints on the construction of one's own identity *(Agre & Rotenberg, 2001)*

People don't care about online privacy?

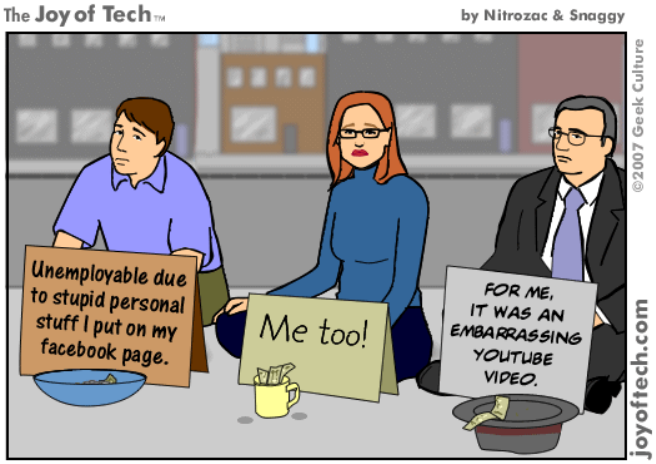
- In the "real world": concerned about information we share
  - Who they tell what
    - You might be willing to tell your best friend that you had an argument with your girlfriend, but you don't want everybody to know about it
  - Concerns over information taken out of context
    - A picture taken at a crazy party being available to a potential employer
  - We value friends who are discreet and keep our secrets
    - We give more information to people we trust
  - The cost of gathering and analyzing information without advanced technologies has guaranteed that we had a rather high level of privacy protection

People don’t care about online privacy?

- Online:
  - less concerned or unaware of privacy violations
- This information is not necessarily secret, but would you want to broadcast it?
  - Identity attributes (Name, age, gender, race, IQ, marital status, place of birth, address, phone number, ID number...)
  - Location (Where you are at a certain point in time, movement patterns)
  - Interests / preferences (Books you read, music you listen, films you like, sports you practice, political affiliation, religious beliefs, sexual orientation)
  - Behavior (Personality type, what you eat, what you shop, how you behave and interact with others)
  - Social network (Who your friends are, who you meet when, your different social circles)
  - Health data (Medical issues, treatments you follow, DNA, health risk factors)
  - Financial data (How much you earn, how you spend your money, credit card number, bank account)
- Combination of them all is even more troublesome

Privacy problems

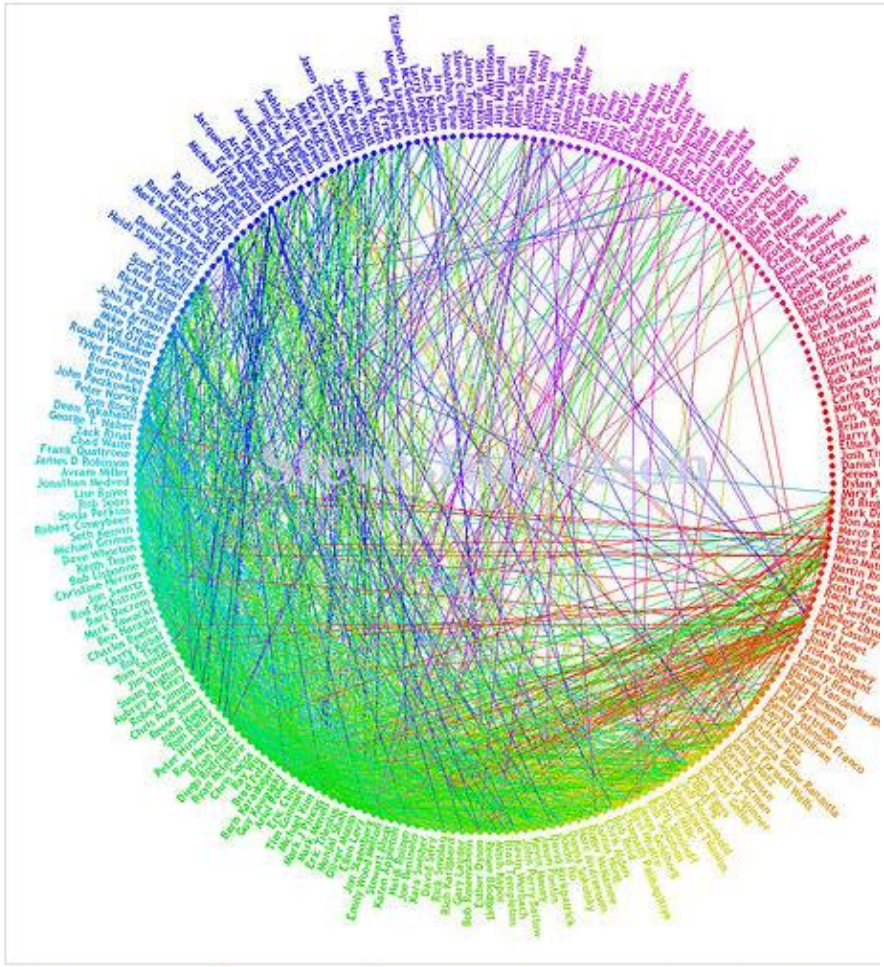
- Identity theft
  - Getting a credit card on your name
- Stalking
- Profiling
  - Find compulsive buyers, ...
- Sensitive information being shared
- Information taken out of context



Signs of the social networking times.

Gaydar Algorithm Outs Facebook Users

By Susannah F. Locke Posted 09.21.2009 at 12:27 pm 12 Comments



What are your friends saying about you? Online social networks like this Facebook one mi more about you than you think *jurvetson (CC licensed)*

A pair of MIT students claim that they have created an algorithm that outs gay members of Facebook by analyzing the sexual orientations of their networks of friends.

The students first analyzed the networks of people who publicized their sexual orientation on Facebook. Turns out that statistically speaking, gay men have more gay friends than straight guys do. So then, they used an algorithm to run the stats on men who kept mum about their sexual orientation on the site. Their computer program was able to correctly identify 10 men whom the students personally knew to be gay in the real world but who hadn't shared that fact on Facebook. (The algorithm didn't work as well with women or with bisexual Facebookers.)

The students completed the project for a class on ethics and the Internet and hope to publish it in a scientific journal.

Their project is far from the first study showing that a simple computer program can sleuth out details you might prefer to keep private by looking at your social network on the Internet. Earlier this year, computer scientists correctly linked 30 percent of anonymous Twitter and Flickr accounts with a simple algorithm that compares who's following who on each site. And other researchers have used Internet social networks to correctly identify peoples' political affiliations or where they live.

It's a good reminder to take a look at your privacy settings. Because you might inadvertently be sharing things you'd rather keep to yourself. Even if you're only declaring to the world that someone's your friend.

'I'm going to destroy America and dig up Marilyn Monroe': British pair arrested in U.S. on terror charges over Twitter jokes

By RICHARD HARTLEY-PARKINSON

Last updated at 1:08 PM on 31st January 2012

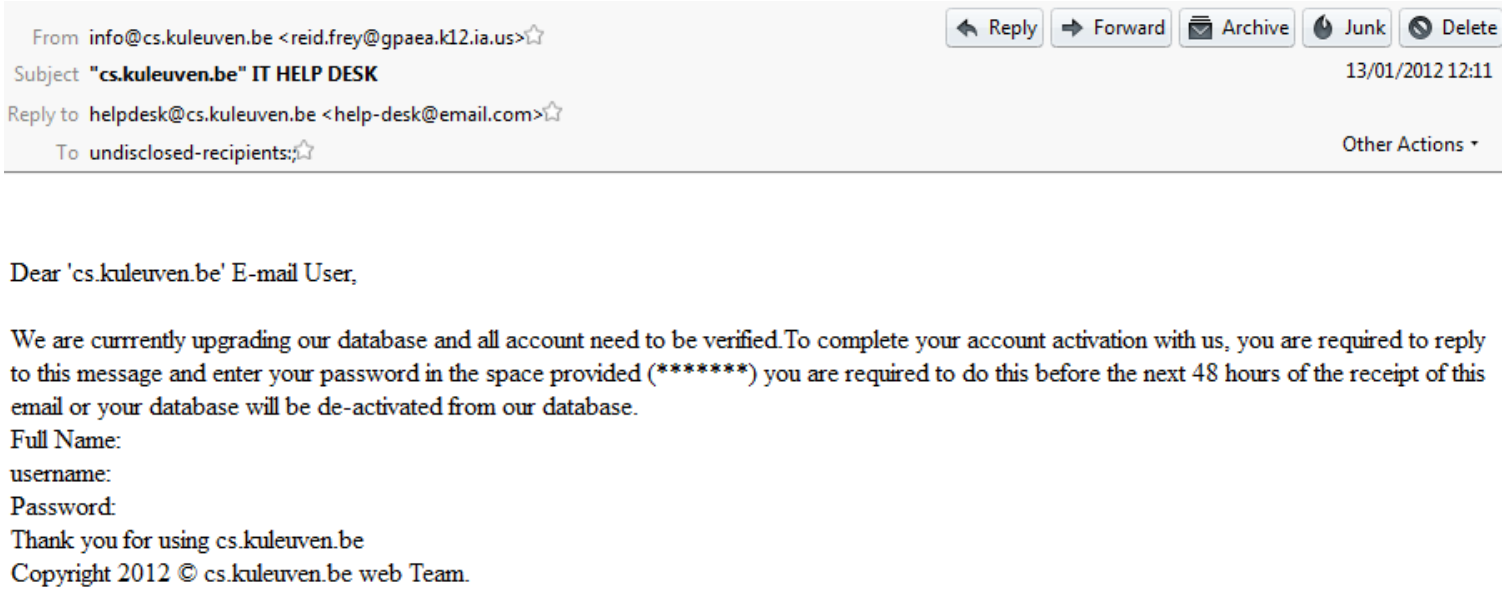
Two British tourists were barred from entering America after joking on Twitter that they were going to 'destroy America' and 'dig up Marilyn Monroe'.

Leigh Van Bryan, 26, was handcuffed and kept under armed guard in a cell with Mexican drug dealers for 12 hours after landing in Los Angeles with pal Emily Bunting.

The Department of Homeland Security flagged him as a potential threat when he posted an excited tweet to his pals about his forthcoming trip to Hollywood which read: 'Free this week, for quick gossip/prep before I go and destroy America?'

Spear-phishing

- Using personal information to make phishing more successful



From info@cs.kuleuven.be <reid.frey@gpaea.k12.ia.us>
Subject **"cs.kuleuven.be" IT HELP DESK**
Reply to helpdesk@cs.kuleuven.be <help-desk@email.com>
To undisclosed-recipients:;

Reply | Forward | Archive | Junk | Delete
13/01/2012 12:11
Other Actions

Dear 'cs.kuleuven.be' E-mail User,

We are currrently upgrading our database and all account need to be verified.To complete your account activation with us, you are required to reply to this message and enter your password in the space provided (*******) you are required to do this before the next 48 hours of the receipt of this email or your database will be de-activated from our database.
Full Name:
username:
Password:
Thank you for using cs.kuleuven.be
Copyright 2012 © cs.kuleuven.be web Team.

- Using Facebook data

Freddi Staur

- 41% agreed to be friends with Freddi which (often) led to access to
  - Email address
  - Full date of birth
  - Details on education and workplace
  - Current address
  - Pictures of family and friends
  - Name of their partner / relatives

Privacy properties

- Unlinkability
- Anonymity/ pseudonymity
- Plausible deniability
- Undetectability

Hard privacy

- Confidentiality

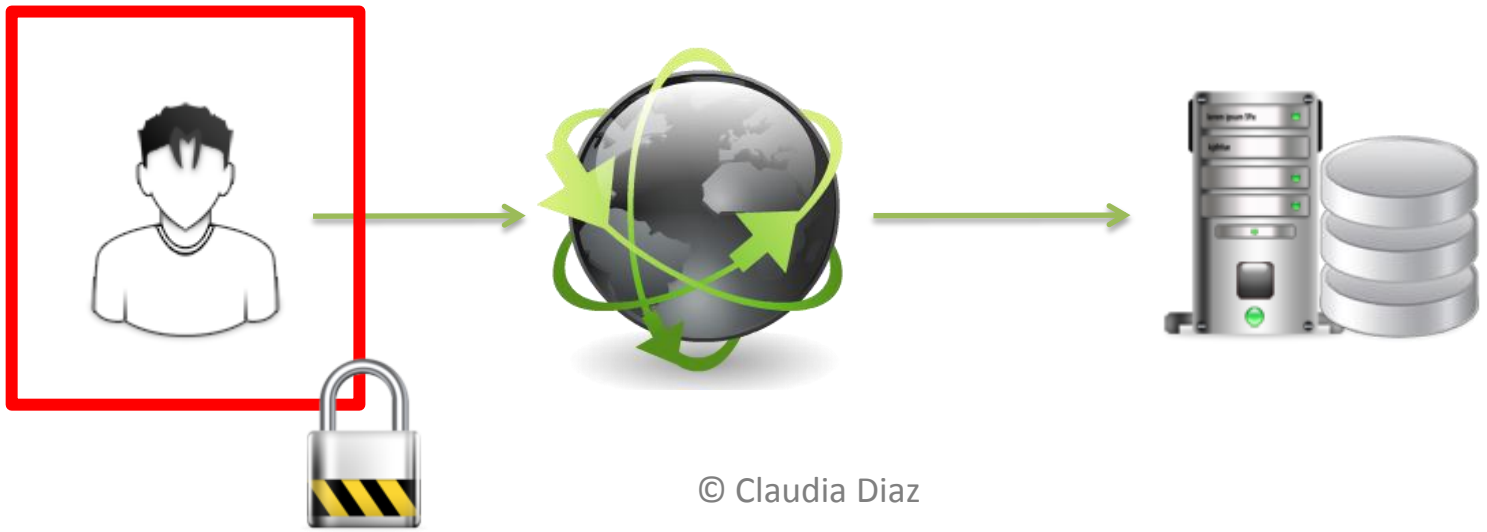
Security

- Content awareness
- Policy and consent compliance

Soft privacy

Hard privacy

- Data minimization
  - Subject provides as little data as possible
  - Reduce as much as possible the need to “trust” other entities



Soft privacy

- Data subject has already lost control of her data
  - In practice, very difficult for data subject to verify how her data are collected and processed

Soft privacy

- Need to trust data controllers (honesty, competence)

Privacy

Anonymity & Pseudonymity

Anonymity

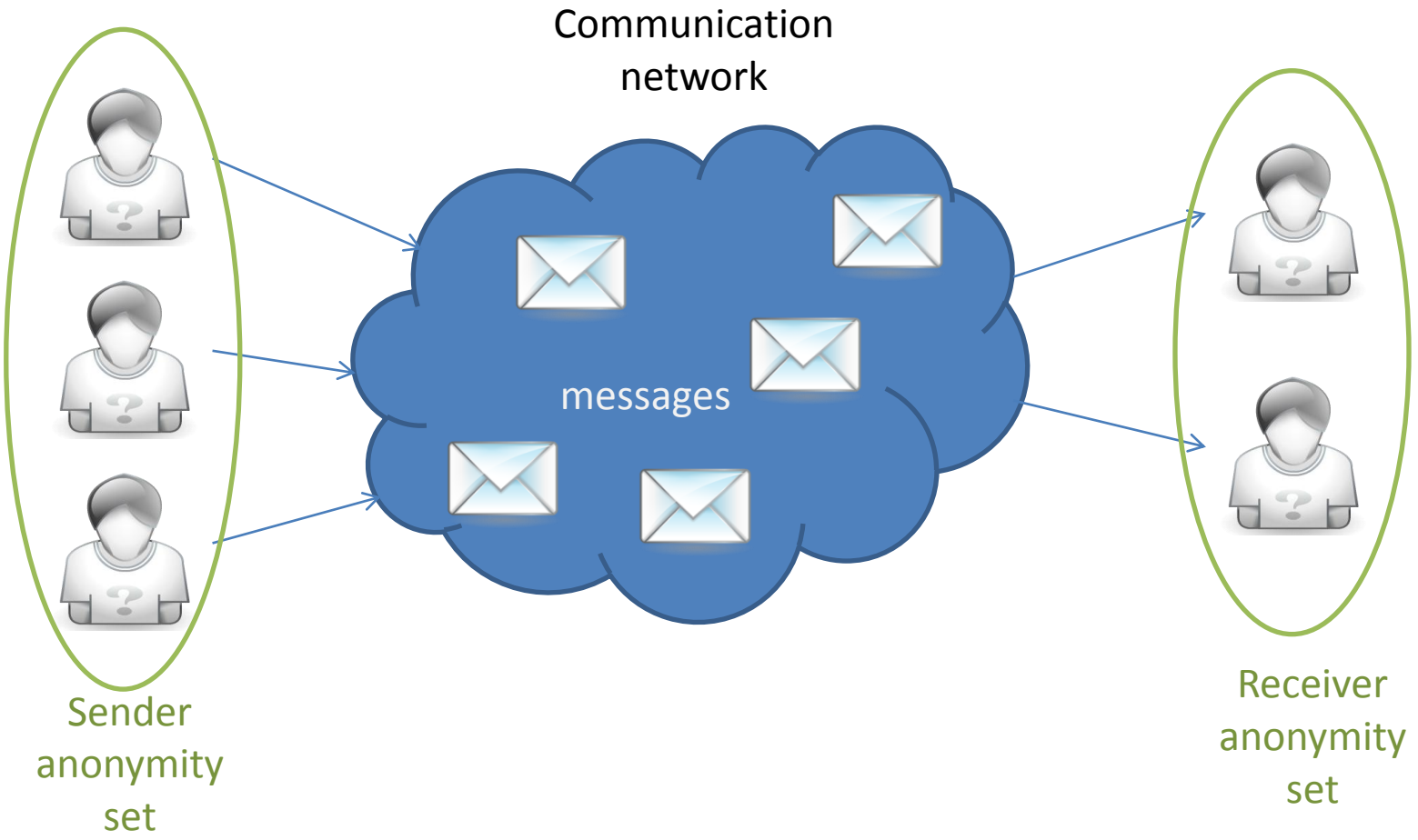
An attacker cannot sufficiently identify the subject within a set of subjects, the anonymity set

Anonymity

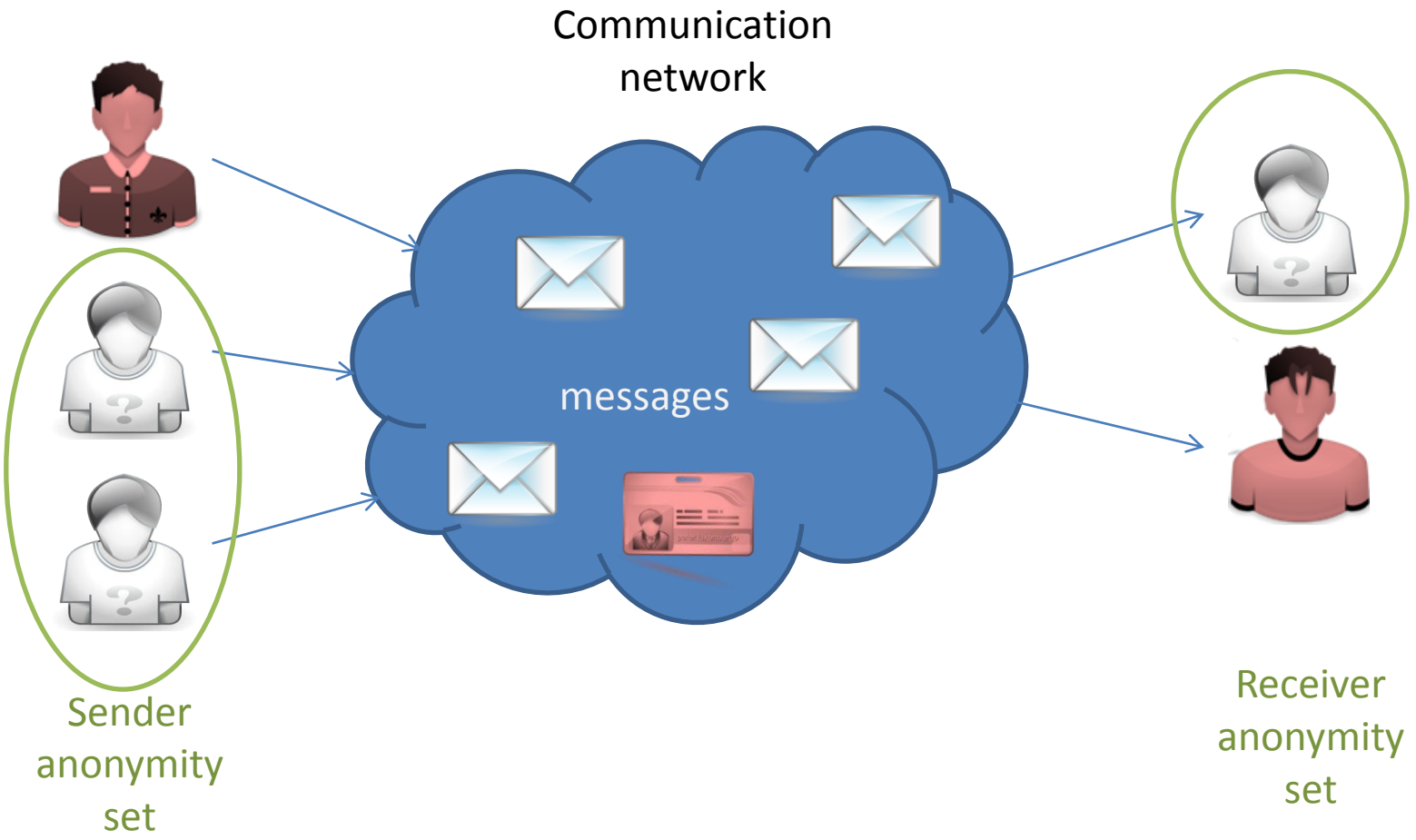
- *An attacker cannot sufficiently identify the subject within a set of subjects, the anonymity set (Pfitzmann)*
- Hiding link between identity and action / piece of information.

- Examples:
  - Reader of a web page, person accessing a service
  - Sender of an email, writer of a text
  - Person to whom an entry in a database relates
  - Person present in a physical location

Anonymity set



Anonymity set wrt attacker

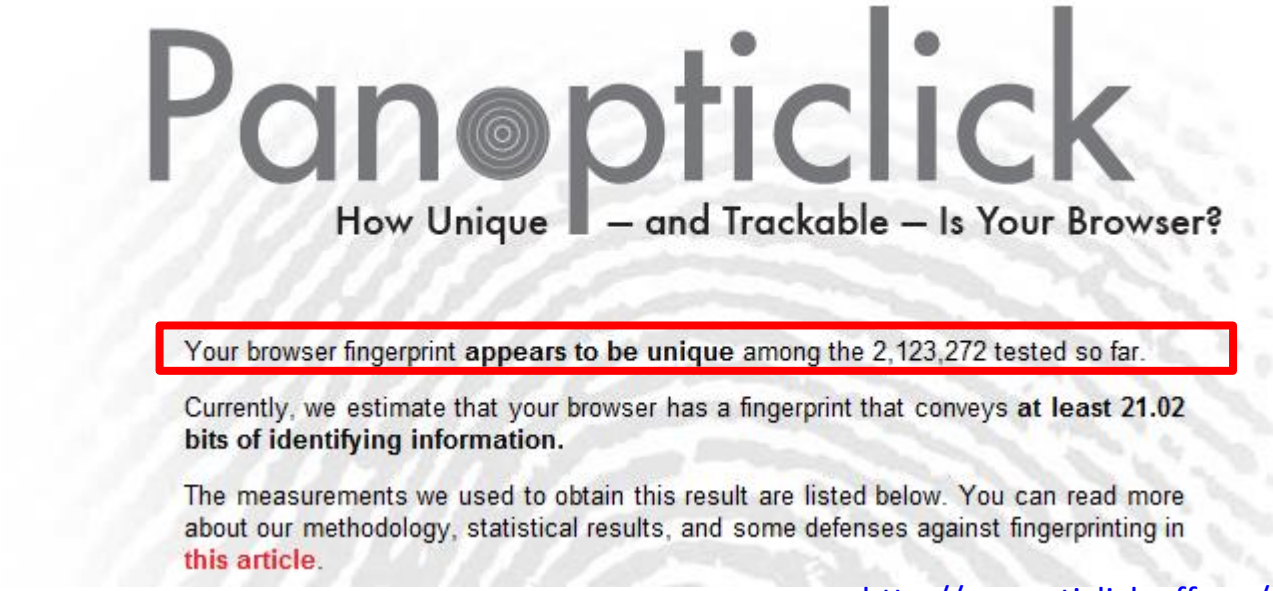


Identifiability

- *The attacker can sufficiently identify the subject within a set of subjects, the identifiability set (pfitzmann)*
- A identity is any subset of attribute values of an individual personal which sufficiently identifies this individual person with any set of persons.
  - There can thus be many “identities”

Identifiability example

- Browser uniqueness



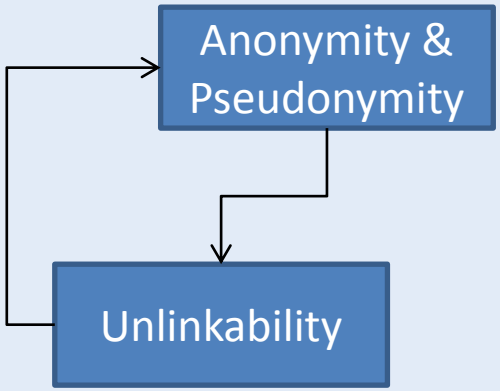
http://panopticlick.eff.org/

- Possible to track "anonymous" visitors

Pseudonymity

- *A pseudonym is an identifier of a subject other than one of the subjects real names. Pseudonymity is the use of pseudonyms as identifiers. (Pfitzmann)*
- Pseudonymity is the entire field between anonymity and identifiability

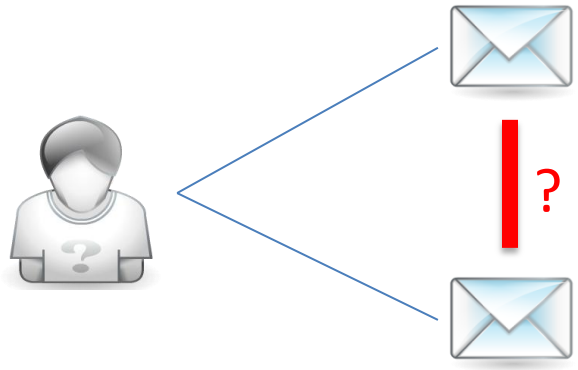
Privacy



⟶ requires

Unlinkability

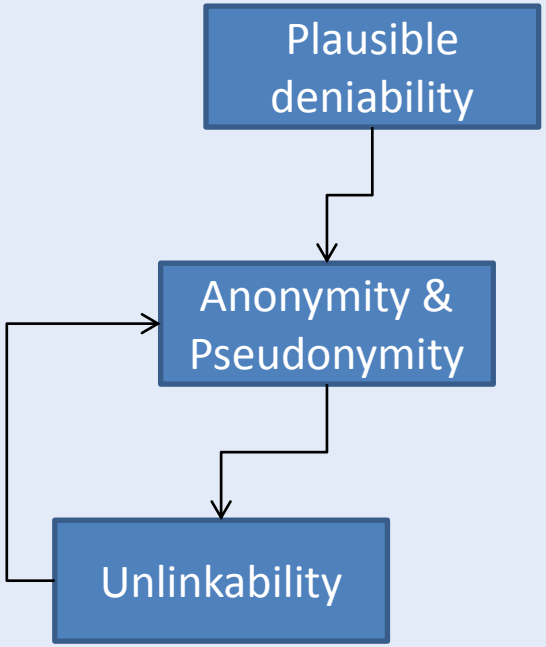
Within a system, the attacker cannot sufficiently distinguish whether two or more items of interest (IOI) are related or not



Unlinkability

- *Within a system, the attacker cannot sufficiently distinguish whether two or more items of interest (IOI) are related or not (Pfitzman)*
- Hiding link between two or more actions / identities /pieces of information
- Examples:
  - Two anonymous letters written by the same person
  - Two web page visits by the same user
  - Entries in two databases related to the same person
  - Two people related by a friendship link
  - Same person spotted in two locations at different points in time

Privacy



Plausible deniability

Not possible to prove user knows, has done or has said something

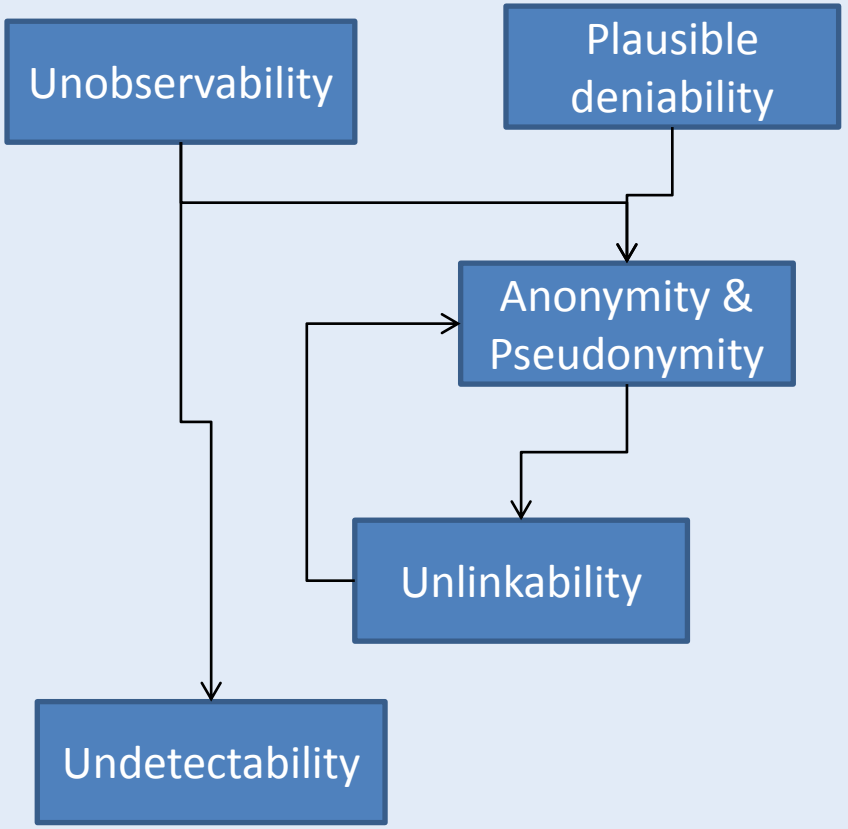
⟶ requires

Plausible deniability

- Not possible to prove user knows, has done or has said something
- Examples:
  - Resistance to coercion:
    - Not possible to prove that a person has hidden information in a computer
    - Not possible to know that someone has the combination of a safe
  - Possibility to deny having been in a place at a certain point in time
  - Possibility to deny that a database record belongs to a person
  - Off-the-record conversations

Privacy

Hard privacy



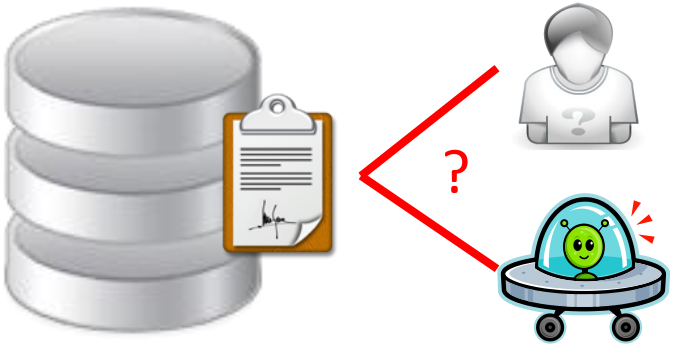
⟶ requires

Undetectability

The attacker cannot sufficiently distinguish whether it exists or not

Unobservability

undetectability + anonymity of subjects involved in the IOI even against the other subjects involved in that IOI



Undetectability & Unobservability

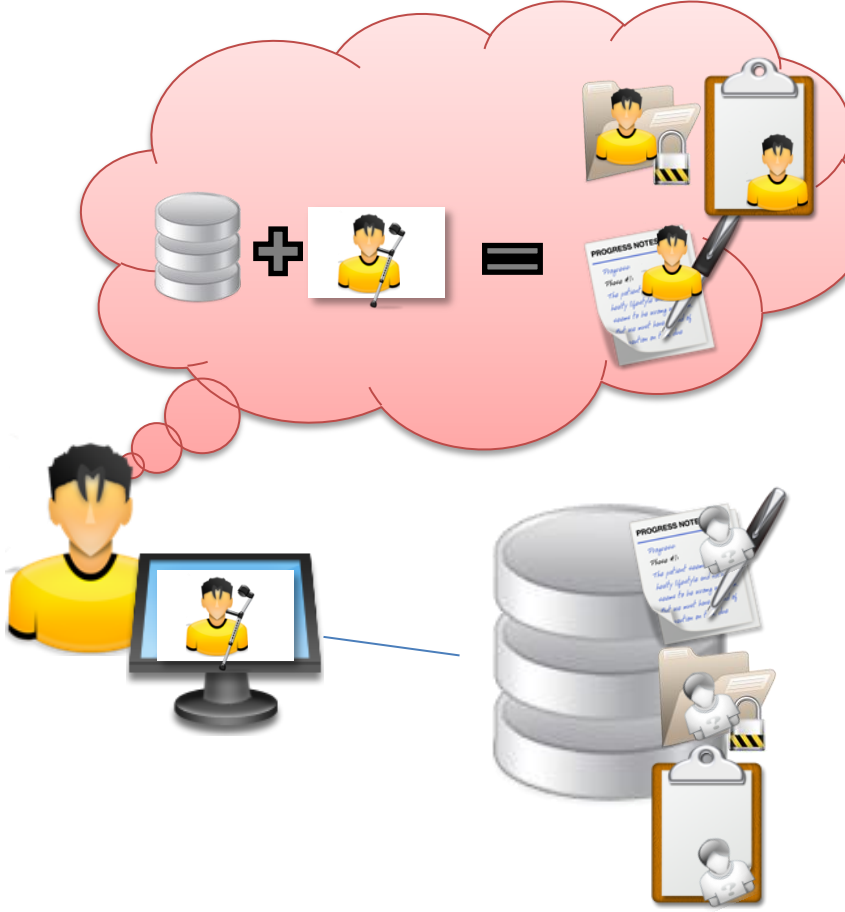
- *Undetectability: The attacker cannot sufficiently distinguish whether it exists or not (Pfitzmann)*
- *Unobservability: undetectability + anonymity of subjects involved in the IOI even against the other subjects involved in that IOI (Pfitzmann)*
- Hiding user activity
- Examples:
  - Impossible to see whether someone is accessing a web page
  - Impossible to know whether an entry in a database corresponds to a real person
  - Impossible to distinguish whether someone or no one is in a given location

Privacy

Hard privacy

- Unobservability
- Plausible deniability
- Anonymity & Pseudonymity
- Unlinkability
- Undetectability

→ requires



- User awareness

User awareness

Users are aware of the consequences of sharing information

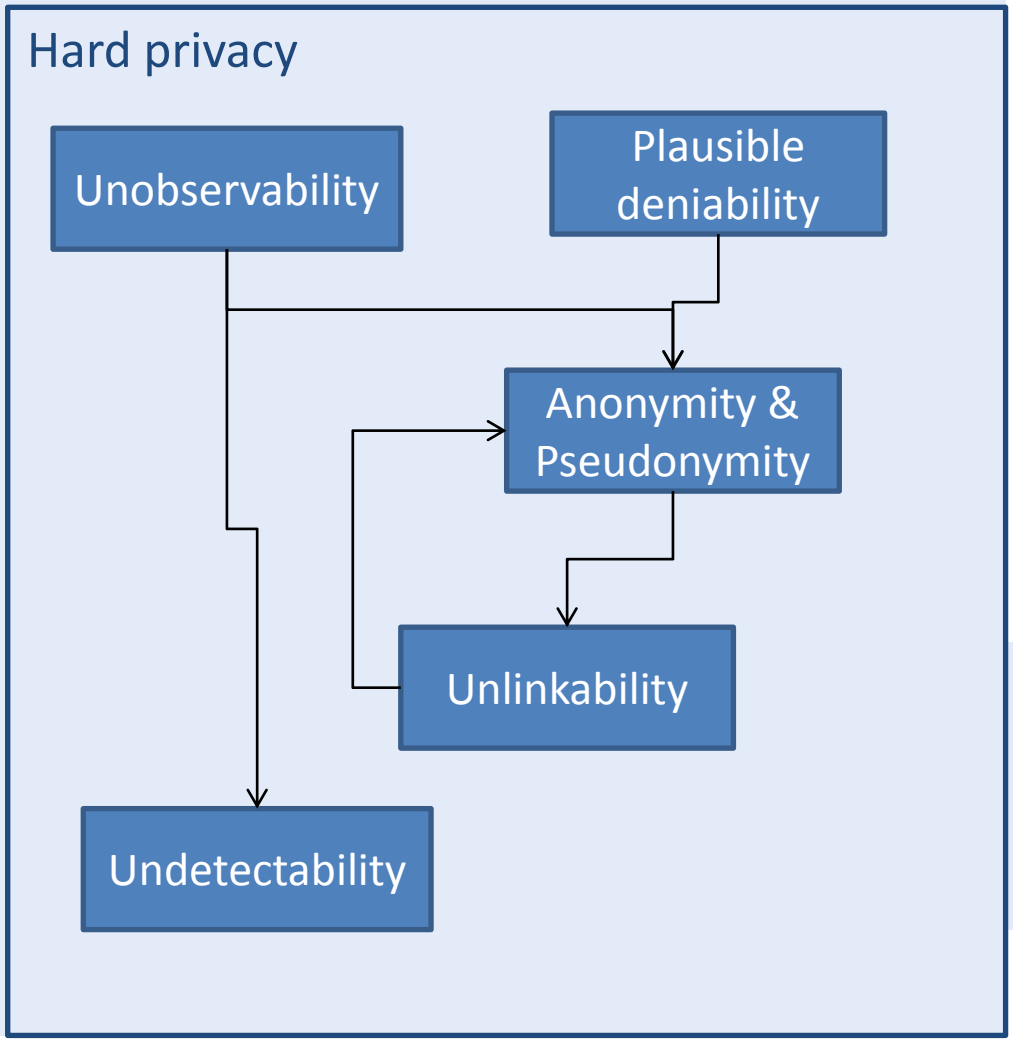
Content Awareness

- Users should be made aware of the consequences of sharing information

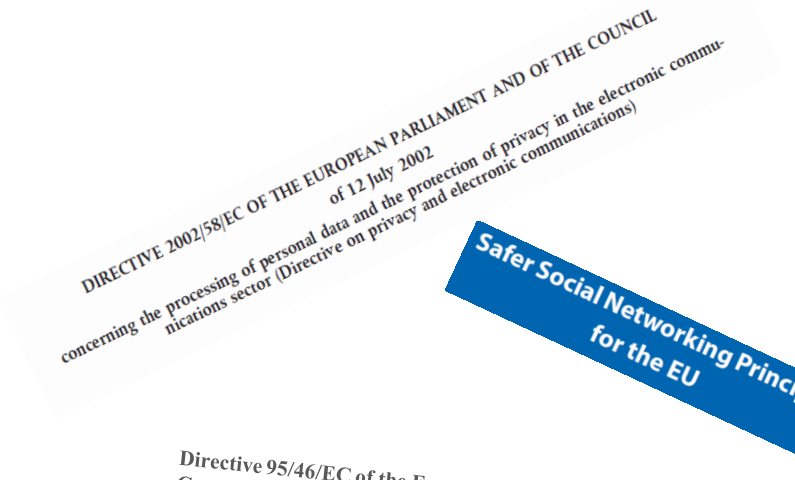
- Suggested solution: Feedback & awareness tools

Privacy

Hard privacy



⟶ requires



User awareness

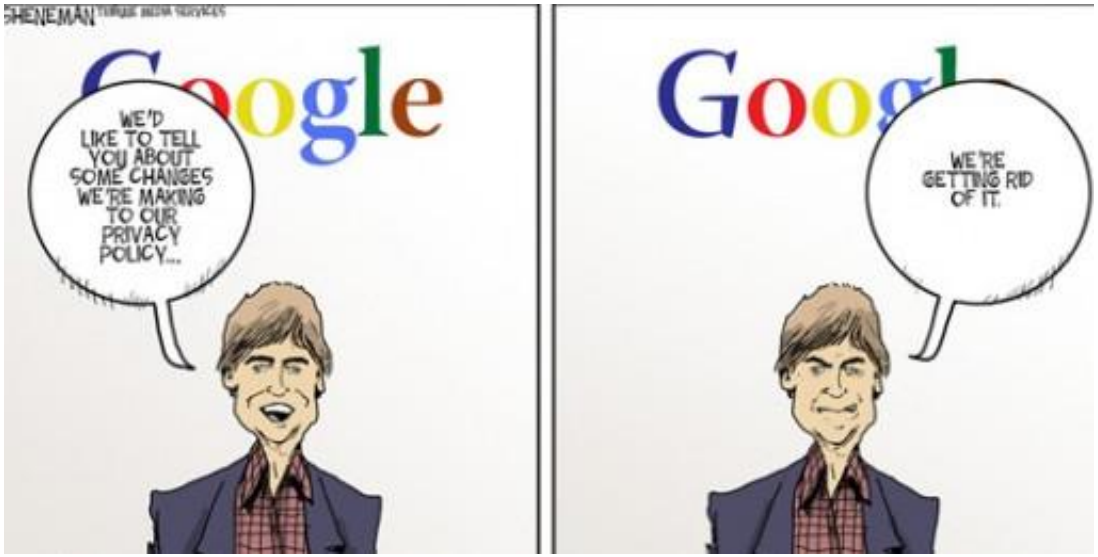
Compliance

Compliance

Legal compliance is obligated.
e.g. consents

Policy & Consent compliance

- Policies
  - Coorporate
  - Privacy



- Openness to users + control

Policy & Consent compliance

- Legal compliance is obligated
  - E.g. European Data Protection Directive
    - fair and lawful processing
    - Consent
    - purpose specification
    - minimality
    - minimal disclosure
    - information quality
    - data subject control
    - sensitivity
    - information security

Compliance example: User consent

- ***personal data*** = any information relating to an identified or identifiable natural person ("data subject“)
- **Sensitive data** = personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership, and the processing of data concerning health or sex life
- Processing of **sensitive data** *prohibited* unless
  - the processing is necessary for the protection of the vital interests of the data subject,
  - the processing is necessary for purposes of preventive medicine, medical diagnosis, provision of care or treatment or
  - the data subject has given his **explicit, written consent** to the processing of the data
  - …(art. 7)

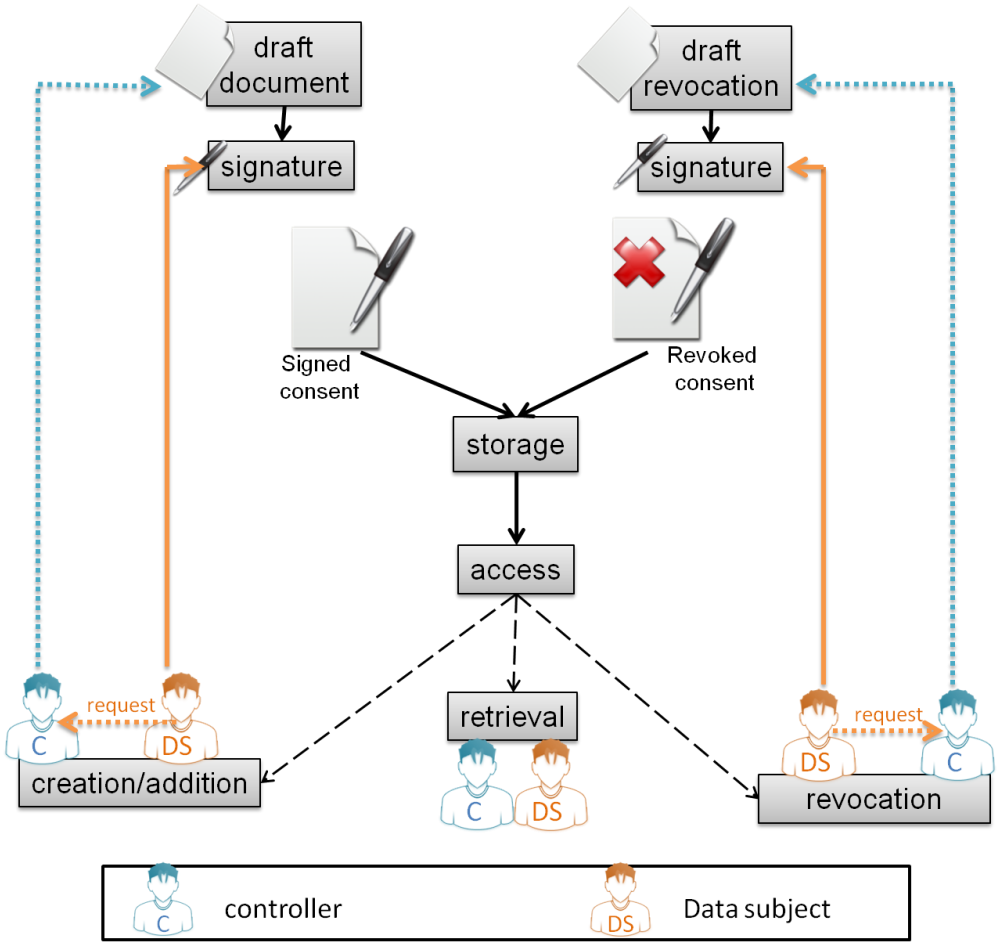
User consent

Legal requirements

- Informed
- Freely given
- Specific

Consent structure

- Data subject
- Controller
- Receiver
- Types of data
- Action (Upload or share)
- Purpose of sharing
- Type of consent (opt-in/opt-out)
- Revoked
- (Context (e.g., “emergency”))
- (Location)



Privacy

Hard privacy

- Unobservability
- Plausible deniability
- Anonymity & Pseudonymity
- Unlinkability
- Undetectability

⟶ requires

- Confidentiality
- User awareness
- Compliance

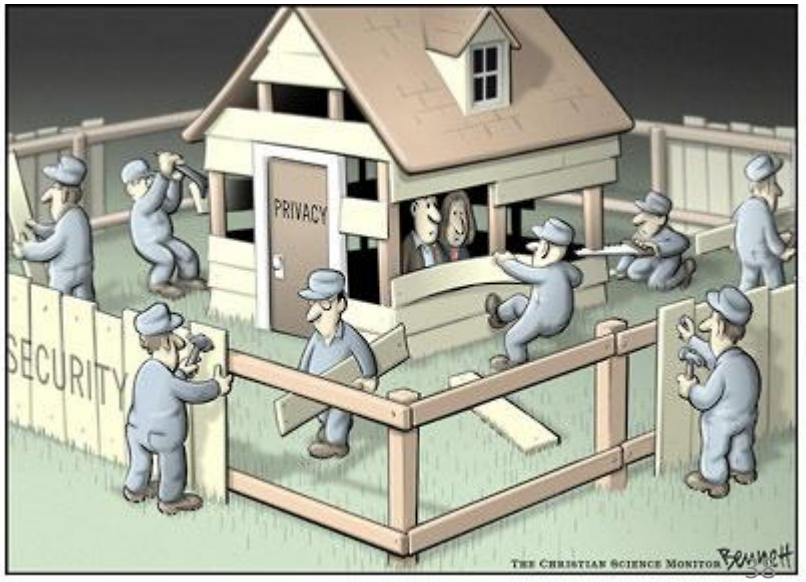
Confidentiality

authorized restrictions on information access and disclosure

Confidentiality

- Preserving authorized restrictions on information access and disclosure, including means for protecting personal privacy and proprietary information *(NIST)*

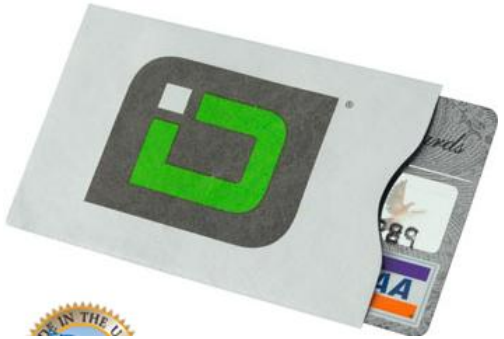
- Security property



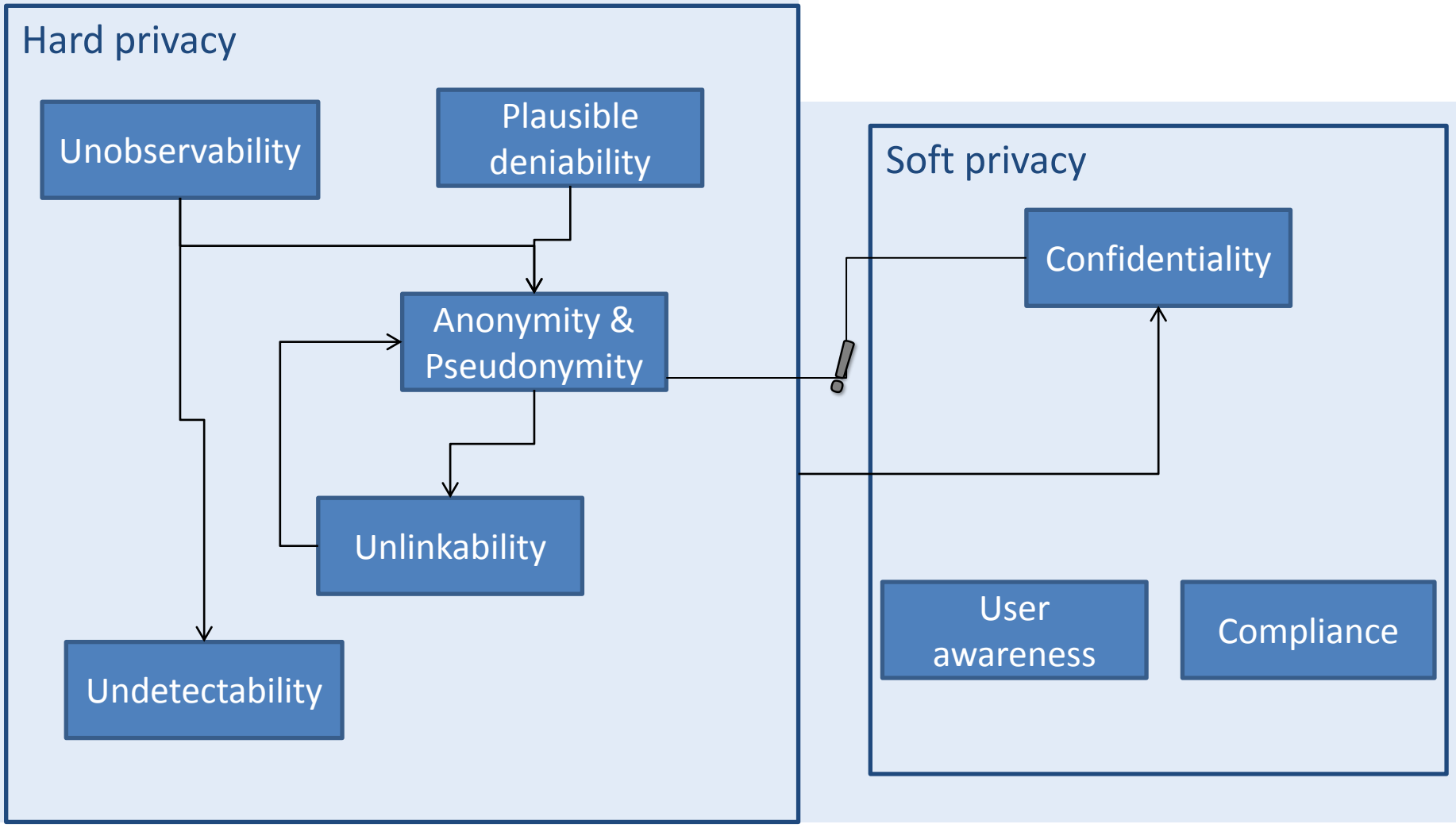
Confidentiality: example

Problem:
Electronic pickpocketing

Solution:
Confidentiality

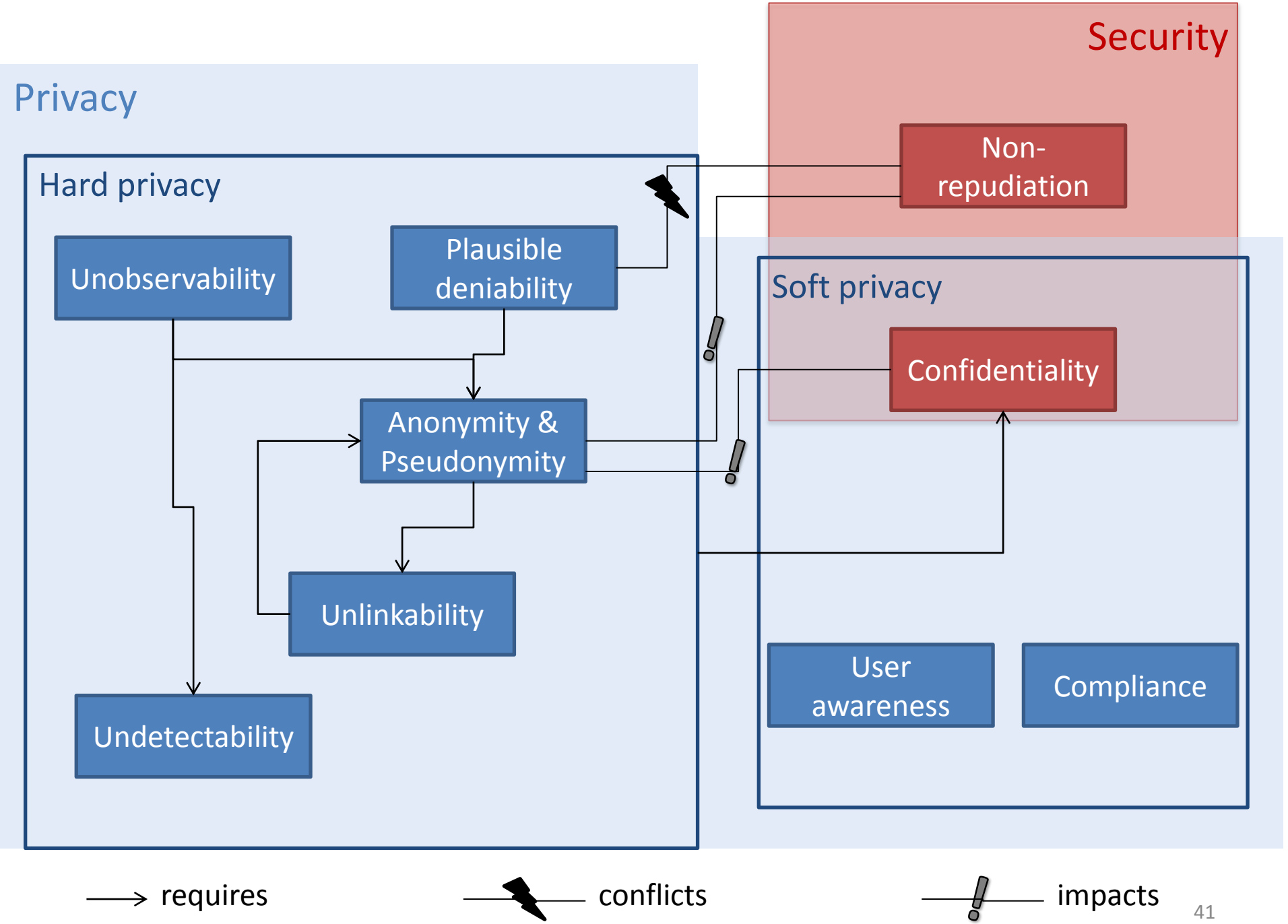


Privacy



→ requires

! impacts


Security
Privacy
Hard privacy
Non-repudiation
Unobservability
Plausible deniability
Soft privacy
Confidentiality
Anonymity & Pseudonymity
Unlinkability
User awareness
Compliance
Undetectability
requires
conflicts
impacts

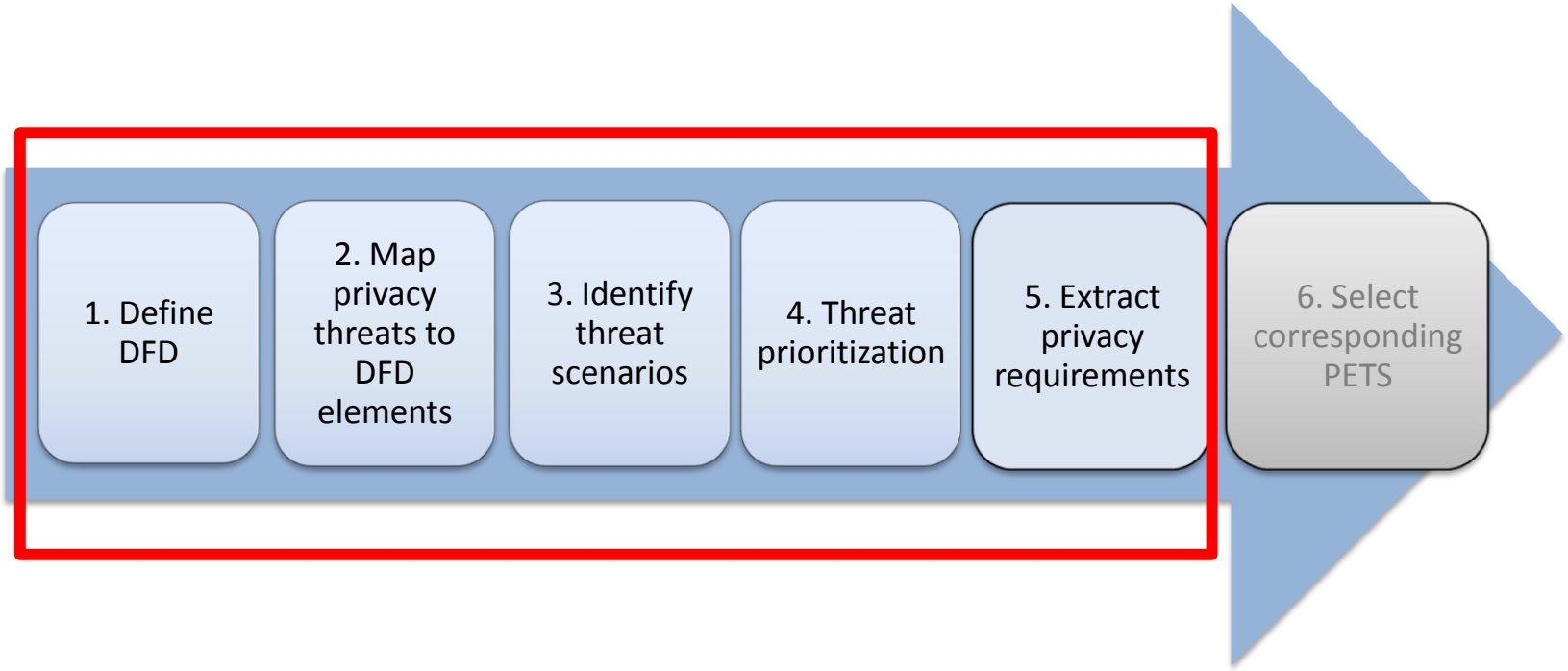
LINDDUN - Privacy threat analysis

PRIVACY METHODOLOGY

Integrating privacy in the system

- Not straight-forward
- Should be part of Software development lifecycle
- Methodology based on STRIDE

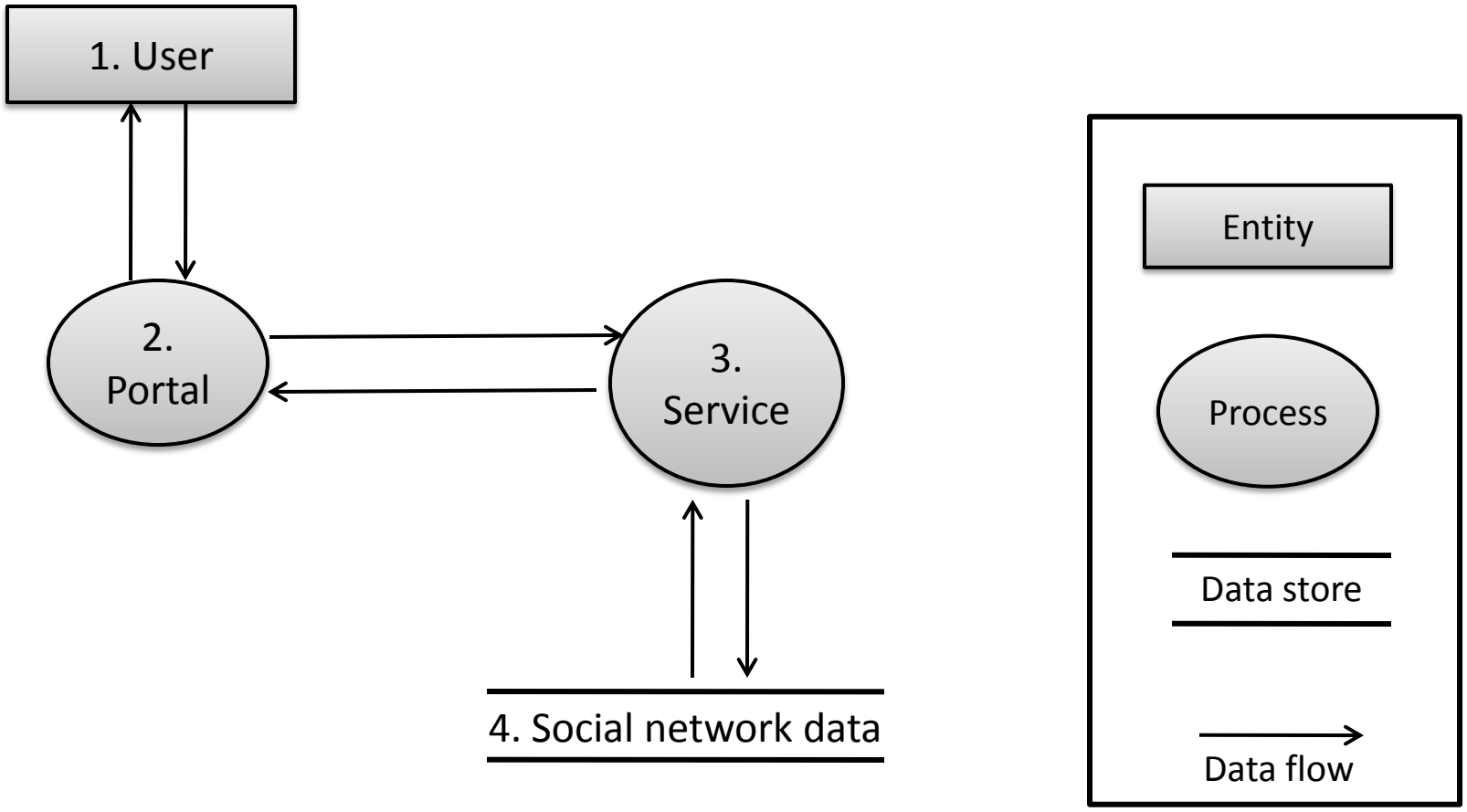
LINDDUN Methodology



LINDDUN Methodology

- **Step 1**
  - **Create the DFD diagram (assets)**
- Step 2
  - Map LINDDUN to DFD element types
- Step 3
  - Refine threats via threat tree patterns
  - Document assumptions
  - Document the threats with template
- Step 4
  - Assign priorities
- Step 5
  - Extract privacy requirements

DFD: social network scenario



LINDDUN Methodology

- Step 1
  - Create the DFD diagram (assets)
- **Step 2**
  - **Map LINDDUN to DFD element types**
- Step 3
  - Refine threats via threat tree patterns
  - Document assumptions
  - Document the threats with template
- Step 4
  - Assign priorities
- Step 5
  - Extract privacy requirements

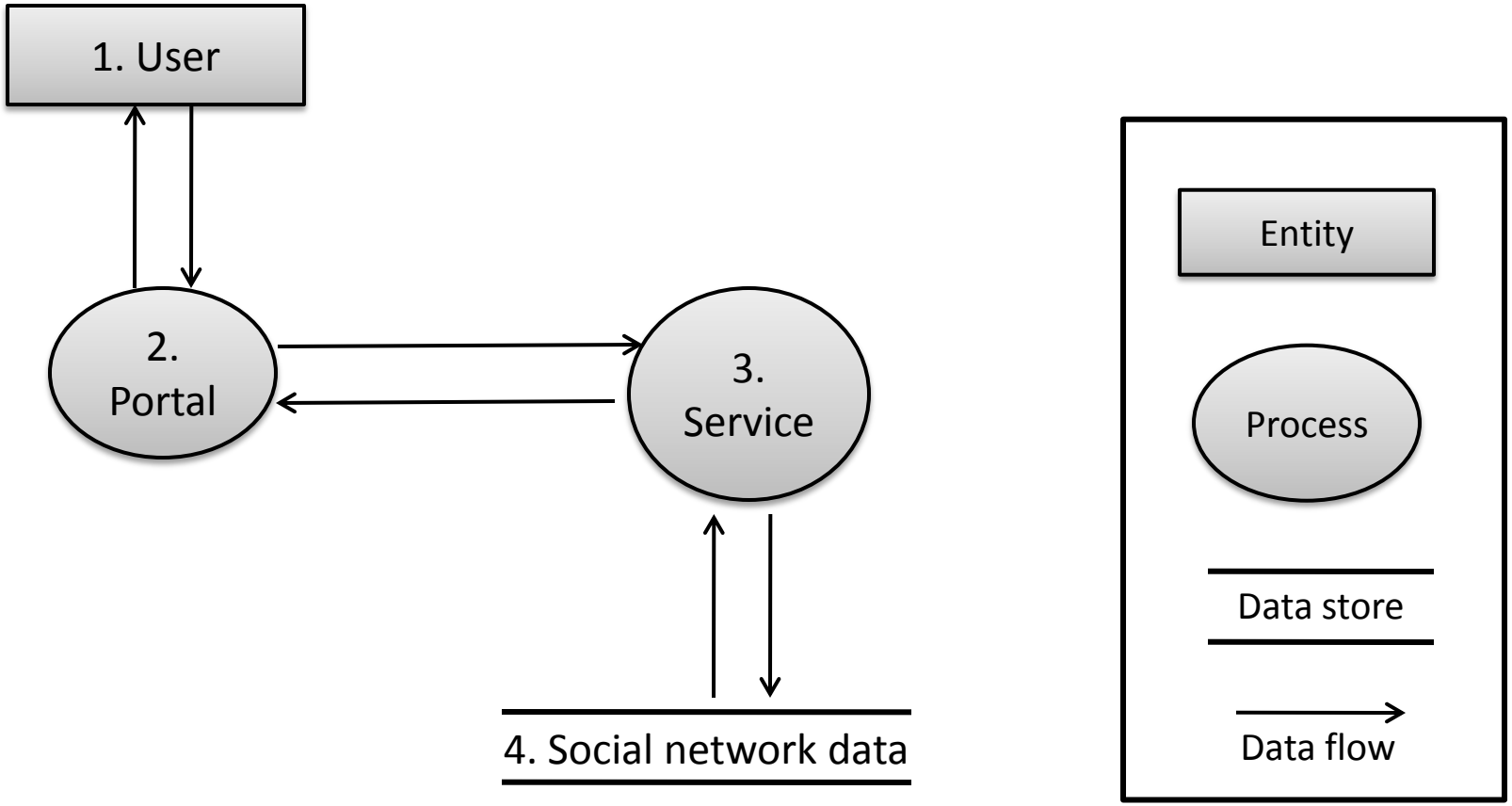
LINDDUN privacy threats

- **Linkability**
  - Sufficiently distinguish whether 2 IOI are linked or not
- **Identifiability**
  - Possible to identify the subject within a set of subjects
- **Non-repudiation**
  - Possible to gather evidence to counter the claims of the repudiating party
- **Detectability**
  - sufficiently distinguish whether IOI exists or not
- **Disclosure of Information**
  - Exposal of information to individuals who are not suppose to have access to it
- **Unawareness of the content**
  - user is unaware of the information he is supplying to the system
- **Noncompliance of policy/consent**
  - System is not compliant with its advertised policies/consents

Mapping threats to DFD

	Linkability	Identifiability	Non-repudiation	Detectability	Information Disclosure	Content Unawareness	Policy & Consent Non-compliance
Data store	X	X	X	X	X		X
Data flow	X	X	X	X	X		X
Process	X	X	X	X	X		X
Entity	X	X				X	

DFD: social network scenario



Mapping Example scenario

	Threat target	**L**	**I**	**N**	**D**	**D**	**U**	**N**
Data store	Social network db	X	X	X	X	X		X
Data flow	User data stream (user-portal)	X	X	X	X	X		X
	Service data stream (portal-service)	X	X	X	X	X		X
	DB data stream (service – DB)	X	X	X	X	X		X
Process	Portal	X	X	X	X	X		X
	Social network service	X	X	X	X	X		X
Entity	User	X	X				X	

LINDDUN Methodology

- Step 1
  - Create the DFD diagram (assets)
- Step 2
  - Map LINDDUN to DFD element types
- **Step 3**
  - Refine threats via threat tree patterns
  - **Document assumptions**
  - Document the threats with template
- Step 4
  - Assign priorities
- Step 5
  - Extract privacy requirements

Assumptions

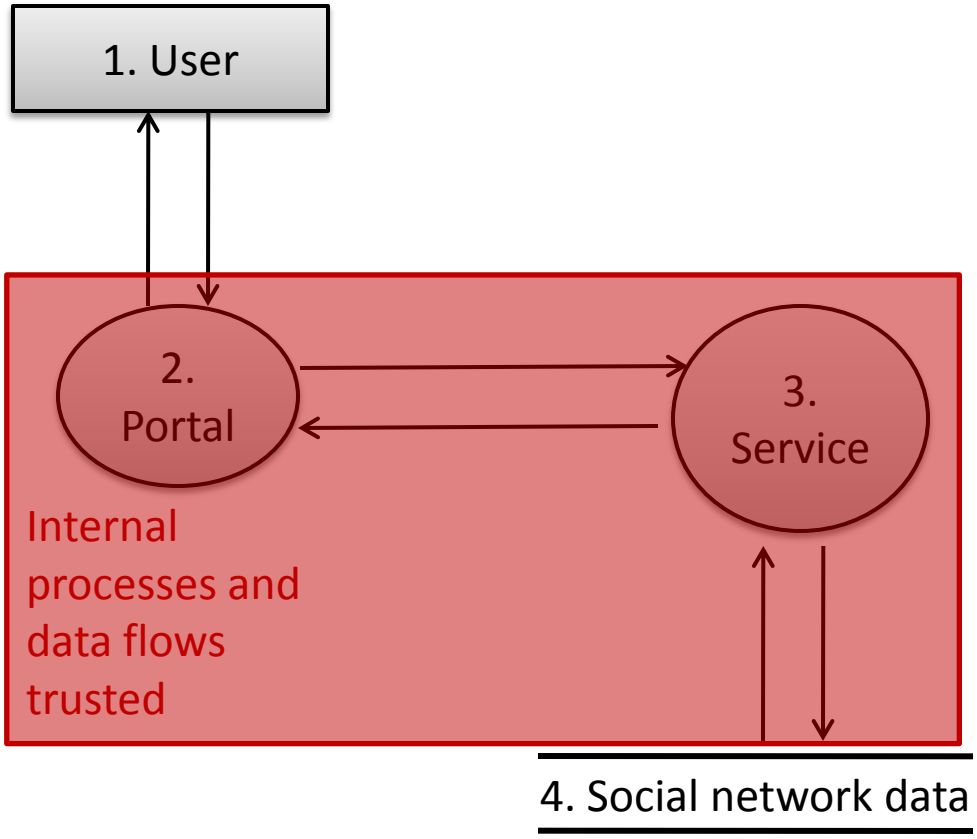
- Assumptions are explicit or implicit choices to trust an element of the system (e.g., human, piece of software) to behave as expected
- The privacy analyst trusts the assumption to be true
- These assumed properties or assertions act as domain restrictions, i.e., they restrict the domain in some way

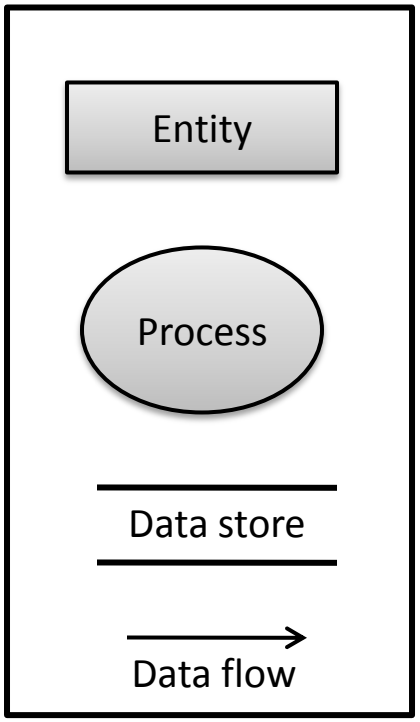
Assumptions

- When adding DFD elements, the number of threats grows exponentially
  - Limit by making assumptions

- Example: assumptions for Social network 2.0:
  1. Internal DFD elements are trustworthy.
     A. trust the processes and data flows in the back-end system.
     B. do not trust the user and its communication with the portal or the data store containing all the user's information.
  2. non-repudiation and detectability threats are considered irrelevant for social networks. (based on threat trees)
  3. non-compliance threats are not specific to a specific DFD element, but are applicable to the entire system

DFD: social network scenario





Impact assumptions on example scenario

	Threat target	L	I	N	D	D	U	N
Data store	Social network db	1	4	X	X	7		10*
Data flow	User data stream (user-portal)	2	5	X	X	8		10*
	Service data stream (portal-service)	X	X	X	X	X		10*
	DB data stream (service – DB)	X	X	X	X	X		10*
Process	Portal	X	X	X	X	X		10*
	Social network service	X	X	X	X	X		10*
Entity	User	3	6				9	

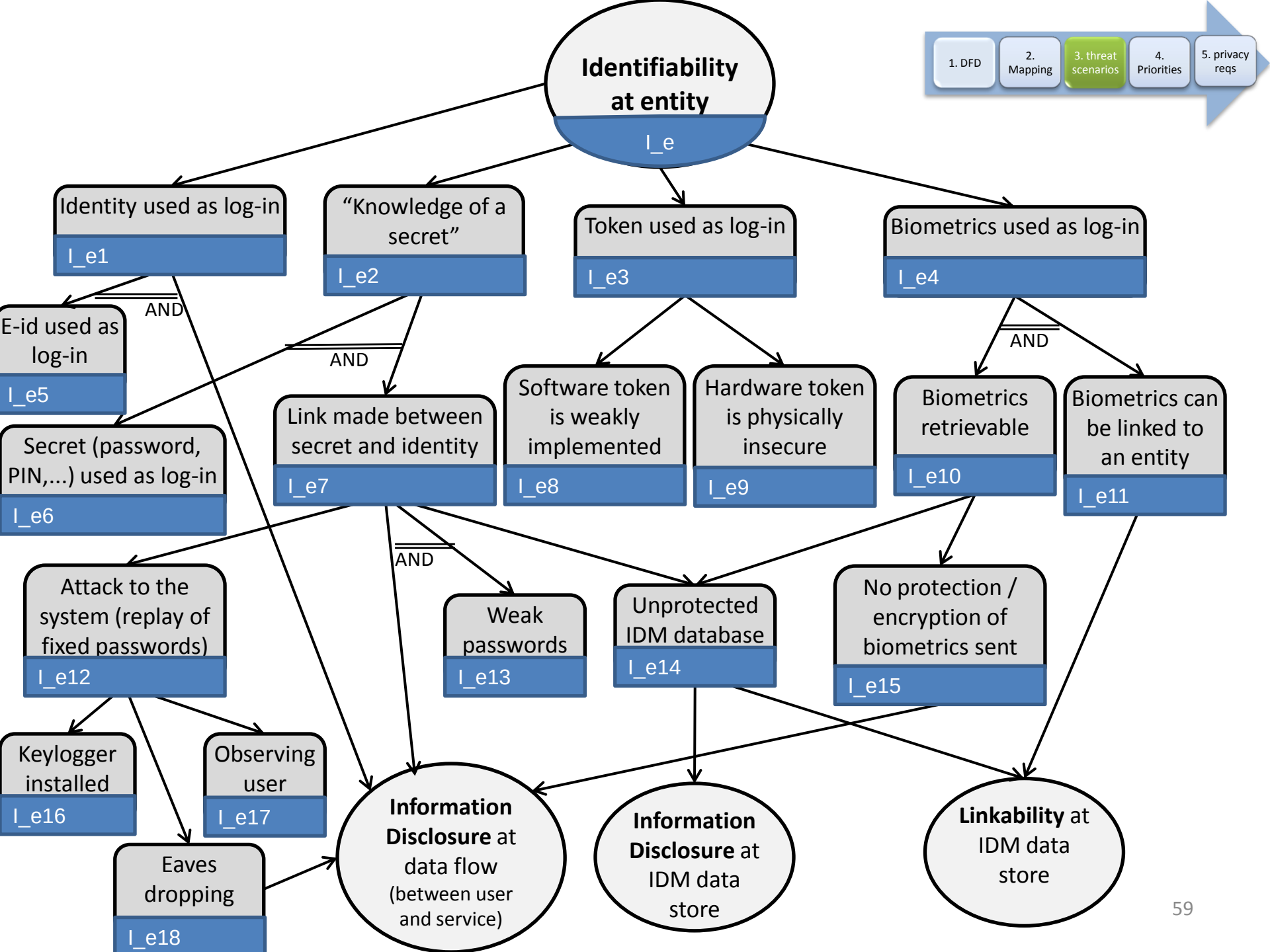
LINDDUN Methodology

- Step 1
  - Create the DFD diagram (assets)
- Step 2
  - Map LINDDUN to DFD element types
- **Step 3**
  - **Refine threats via threat tree patterns**
  - Document assumptions
  - Document the threats with template
- Step 4
  - Assign priorities
- Step 5
  - Extract privacy requirements

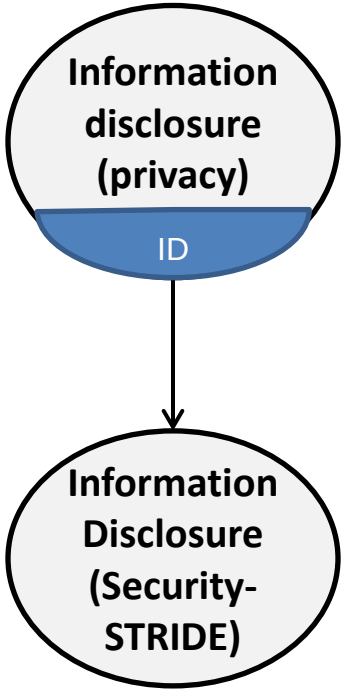
Privacy threat tree patterns

- Illustrate the most common attack patterns
- Used to determine threat applies to system

- Note:
  - Do not limit your analysis to these trees

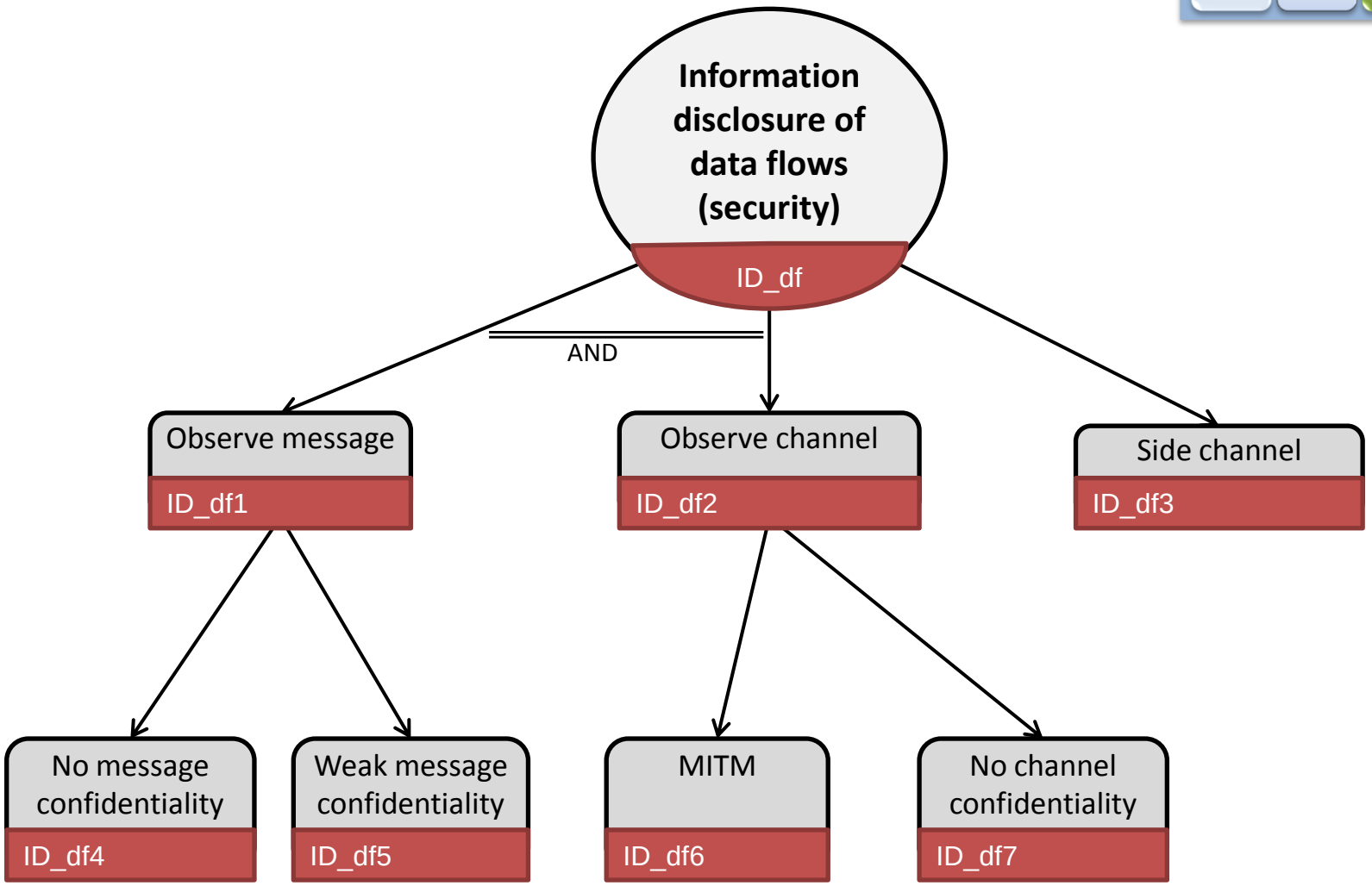

1. DFD
2. Mapping
3. threat scenarios
4. Priorities
5. privacy reqs
Identifiability at entity
I_e
Identity used as log-in
I_e1
"Knowledge of a secret"
I_e2
Token used as log-in
I_e3
Biometrics used as log-in
I_e4
AND
E-id used as log-in
I_e5
AND
AND
Secret (password, PIN,...) used as log-in
I_e6
Link made between secret and identity
I_e7
Software token is weakly implemented
I_e8
Hardware token is physically insecure
I_e9
Biometrics retrievable
I_e10
Biometrics can be linked to an entity
I_e11
AND
Attack to the system (replay of fixed passwords)
I_e12
Weak passwords
I_e13
Unprotected IDM database
I_e14
No protection / encryption of biometrics sent
I_e15
Keylogger installed
I_e16
Observing user
I_e17
Eaves dropping
I_e18
Information Disclosure at data flow (between user and service)
Information Disclosure at IDM data store
Linkability at IDM data store


1. DFD
2. Mapping
3. threat scenarios
4. Priorities
5. privacy reqs


Information disclosure (privacy)
ID
Information Disclosure (Security-STRIDE)

STRIDE revisited

- Systematic approach for security threat identification
- Spoofing, Tampering, Repudiation, Information disclosure, Denial of service, Elevation of privilege


Information disclosure of data flows (security)
ID_df
AND
Observe message
ID_df1
Observe channel
ID_df2
Side channel
ID_df3
No message confidentiality
ID_df4
Weak message confidentiality
ID_df5
MITM
ID_df6
No channel confidentiality
ID_df7

LINDDUN Methodology

- Step 1
  - Create the DFD diagram (assets)
- Step 2
  - Map LINDDUN to DFD element types
- **Step 3**
  - Refine threats via threat tree patterns
  - Document assumptions
  - **Document the threats with template**
- Step 4
  - Assign priorities
- Step 5
  - Extract privacy requirements

Threat description

Inspired by Misuse Cases template

- ID & Title
- Summary
- Misactor profile
- Basic path
- Consequence
- Leaf node(s)
- Root node(s)
- DFD element(s)
- Remarks

In your report

✓ Mention the threats in the order you found them

Threat description Example (naive) 1/3

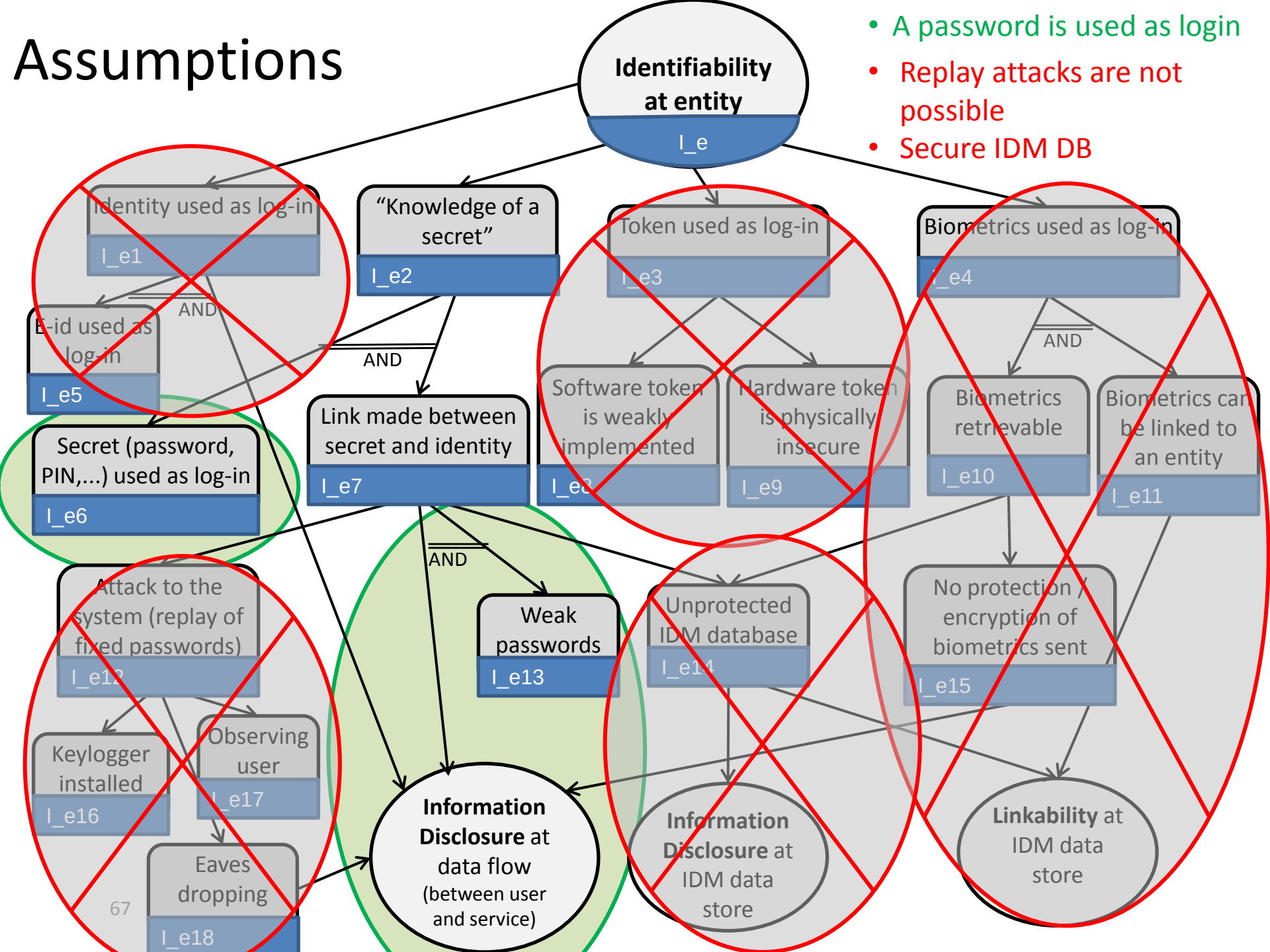
- **ID & Title**
  - T01 . Identify users of the social network system
- **Summary**
  - A misactor gains access to the “secret” sent by the user to log-in and deduces the user’s identity from it
- **Misactor profile**
  - skilled outsider

Threat description Example (naive) 2/3

- **Basic path**
  1. The misactor gains access to the data flow between the user and the portal
  2. The data contains the user's password
  3. The misactor can directly link the password to the user due to weak password use (e.g. initials + birthdate)
- **Consequence**
  - The user's identity is compromised

Assumptions

- A password is used as login
- Replay attacks are not possible
- Secure IDM DB



Threat description Example (naive) 3/3

- **Reference to leaf node(s):** I_e6, I_e13
- **Reference to root node:** I_e
- **DFD element:** User
- **Remarks:**
  - the data flow between the user and the portal is susceptible to information disclosure threats (assumption 1B). This threat is described in T06.
  - A password is used as log-in (Assumption 4)
  - Replay attacks are not considered a threat (Assumption 5)
  - The IDM database is considered secure (Assumption 6)

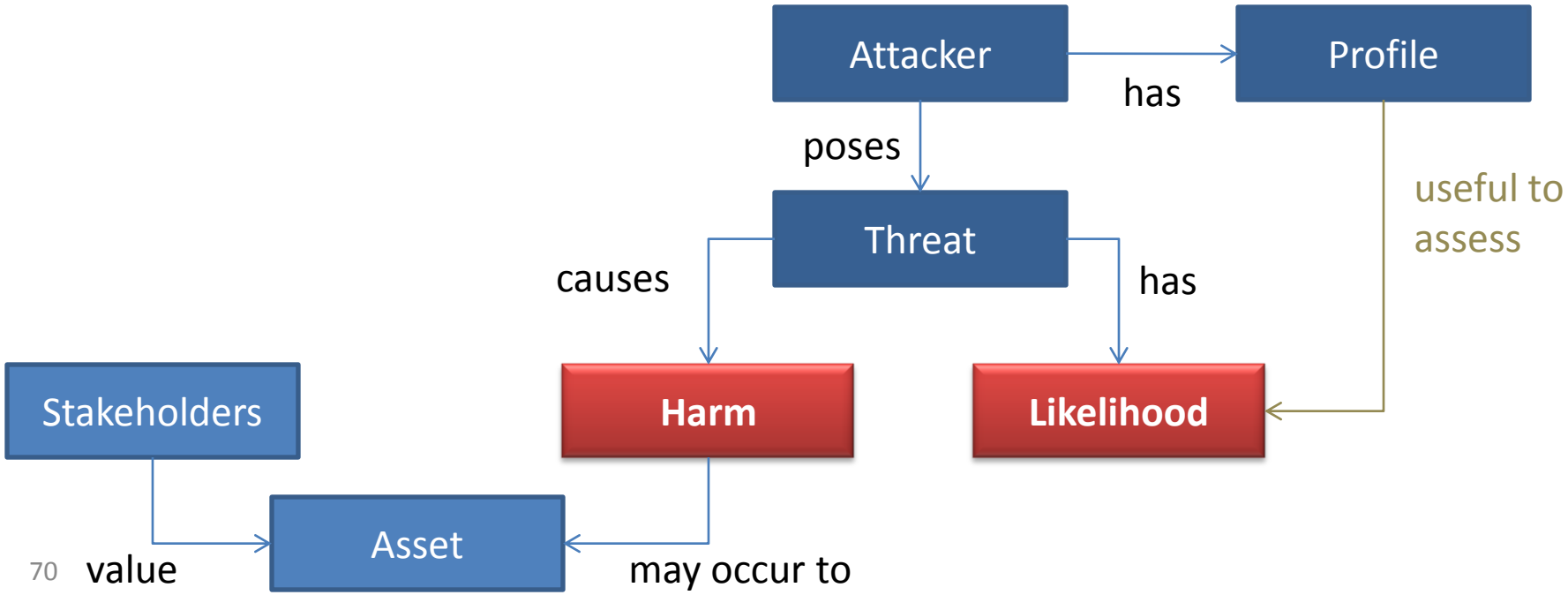
If these assumptions do not hold, the threat tree leaf nodes will result in additional threats

LINDDUN Methodology

- Step 1
  - Create the DFD diagram (assets)
- Step 2
  - Map LINDDUN to DFD element types
- Step 3
  - Refine threats via threat tree patterns
  - Document assumptions
  - Document the threats with template
- **Step 4**
  - **Assign priorities**
- Step 5
  - Extract privacy requirements

The role of risk

- Risk is a function of the likelihood of a threat and the severity of its impact on the organization
  - R = f(likelihood , impact)



LINDDUN Methodology

- Step 1
  - Create the DFD diagram (assets)
- Step 2
  - Map LINDDUN to DFD element types
- Step 3
  - Refine threats via threat tree patterns
  - Document assumptions
  - Document the threats with template
- Step 4
  - Assign priorities
- **Step 5**
  - **Extract privacy requirements**

Privacy requirements

- Possible mitigation techniques [1]
  - Do nothing
  - Remove feature
  - Turn off feature
  - Warn user
  - Counter threats
    - with preventive or reactive technology

1. M. Howard and S. Lipner, ***The Security Development Lifecycle****.*
Microsoft Press, 2006.

From threats to requirements

Threat	Requirement
Linkability	Unlinkability
Identifiability	Anonymity(/pseudonymity)
Non-repudiation	Plausible deniability
Detectability	Undetectability
Disclosure of information	Confidentiality
Unawareness of content	Content awareness
Noncompliance of policy/consent	Consent/policy compliance of the system

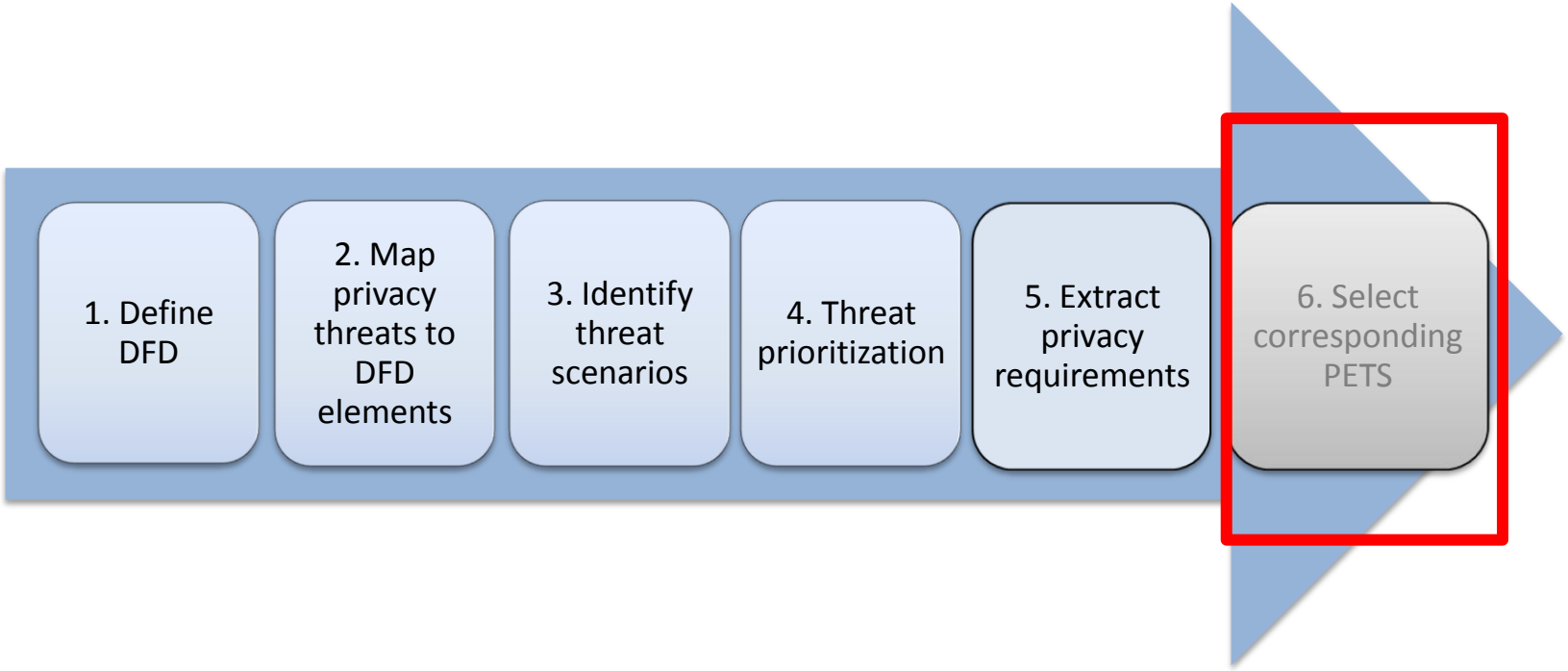
Straight-forward mapping… BUT

From threats to requirements

- Requirements should **not be limited to** straight-forward **mapping**
  - Look at each leaf node that causes threat
  - Determine for each node the proper mitigation

Threats (misuse cases)	**Caused by** (leaf nodes)	**Mitigated by** (requirements)
Deducing identity from password	Information disclosure of data flow	Ensure confidential communication channel (encryption)
	Weak passwords	System should reject weak passwords (at registration) OR
		Users should be made aware of consequences of weak passwords (e.g. feedback given at registration)

LINDDUN Methodology



	Mitigation techniques: PETs	U	A	P	D	C	W	O
Anonymous communication	Mix-networks (1981) [29], DC-networks (1985) [30,31], ISDN-mixes [32], Onion Routing (1996) [33], Crowds (1998) [34], Single proxy (90s) (Penet pseudonymous remailer (1993-1996), Anonymizer, SafeWeb), anonymous Remailer (Cipherpunk Type 0, Type 1 [35], Mixmaster Type 2 (1994) [36], Mixminion Type 3 (2003) [37]), and Low-latency communication (Freedom Network (1999-2001) [38], Java Anon Proxy (JAP) (2000) [39], Tor (2004) [40])	×	×			×		
	DC-net & MIX-net + dummy traffic, ISDN-mixes [32]	×	×		×	×		
	Broadcast systems [41,42] + dummy traffic	×	×		×			
Privacy preserving authentication	Private authentication [43,44]	×	×					
	Anonymous credentials (single show [45], multishow [46])	×	×					
	Deniable authentication [47]	×	×	×				
	Off-the-record messaging [48]	×	×	×		×		
Privacy preserving cryptographic protocols	Multi-party computation (Secure function evaluation) [49, 50]	×				×		
	Anonymous buyer-seller watermarking protocol [51]	×	×			×		
Information retrieval	Private information retrieval [52] + dummy traffic	×	×		×			
	Oblivious transfer [53,54])	×	×			×		
	Privacy preserving data mining [55,56]	×	×			×		
	Searchable encryption [57] / Private search [58]		×			×		
Data anonymization	K-anonymity model [23,59], l-Diversity [60]	×	×					
Information hiding	Steganography [61]	×	×		×			
	Covert communication [62]	×	×		×			
	Spread spectrum [63]	×	×		×			
Pseudonymity systems	Privacy-enhancing identity management system [64]	×	×					
	User-controlled identity management system [65]	×	×					
	Privacy-preserving biometrics [66]	×	×			×		
Encryption techniques	Symmetric key & public key encryption [67]					×		
	Deniable encryption			×		×		
	Homomorphic encryption [68]					×		
	Verifiable encryption [69]					×		
Access control techniques	Context-based access control [70]					×		
	Privacy-aware access control [71,72]					×		
Policy and feedback tools	Policy communication (P3P [19])							×
	Policy enforcement (XACML [73], EPAL [74])							×
	Feedback tools for user privacy awareness [12,13,75]						×	
	Data removal tools (spyware detection and removal, browser cleaning tools, activity traces eraser, harddisk data eraser)						×	

U: unlinkability

A: anonymity

P: plausible deniability

D: undetectability

C: confidentiality

W: content awareness

O: consent/policy compliance

Social network example

Nr	MUC	Requirements	Solutions
1	Linkability data store	Unlinkability of data entries within the social network database	Data anonymization techniques: K-anonymity
		Protection of data store	Access control: relationship-based
2	Linkability data flow	Unlinkability of messages	Anonymous communication system: TOR
		Channel confidentiality	
…			
9	Content unawareness of user	Make users aware that they only need to provide minimal set of information	Use feedback tools to raise awareness
10	Policy and consent non-compliance	Design system in compliance with legal guidelines	Assign policy compliance responsibility to employee
		Ensure user aware of legitimate actions to perform	User can sue organization
		Employee contracts specify internal rules	Employees disclosing information are penalized

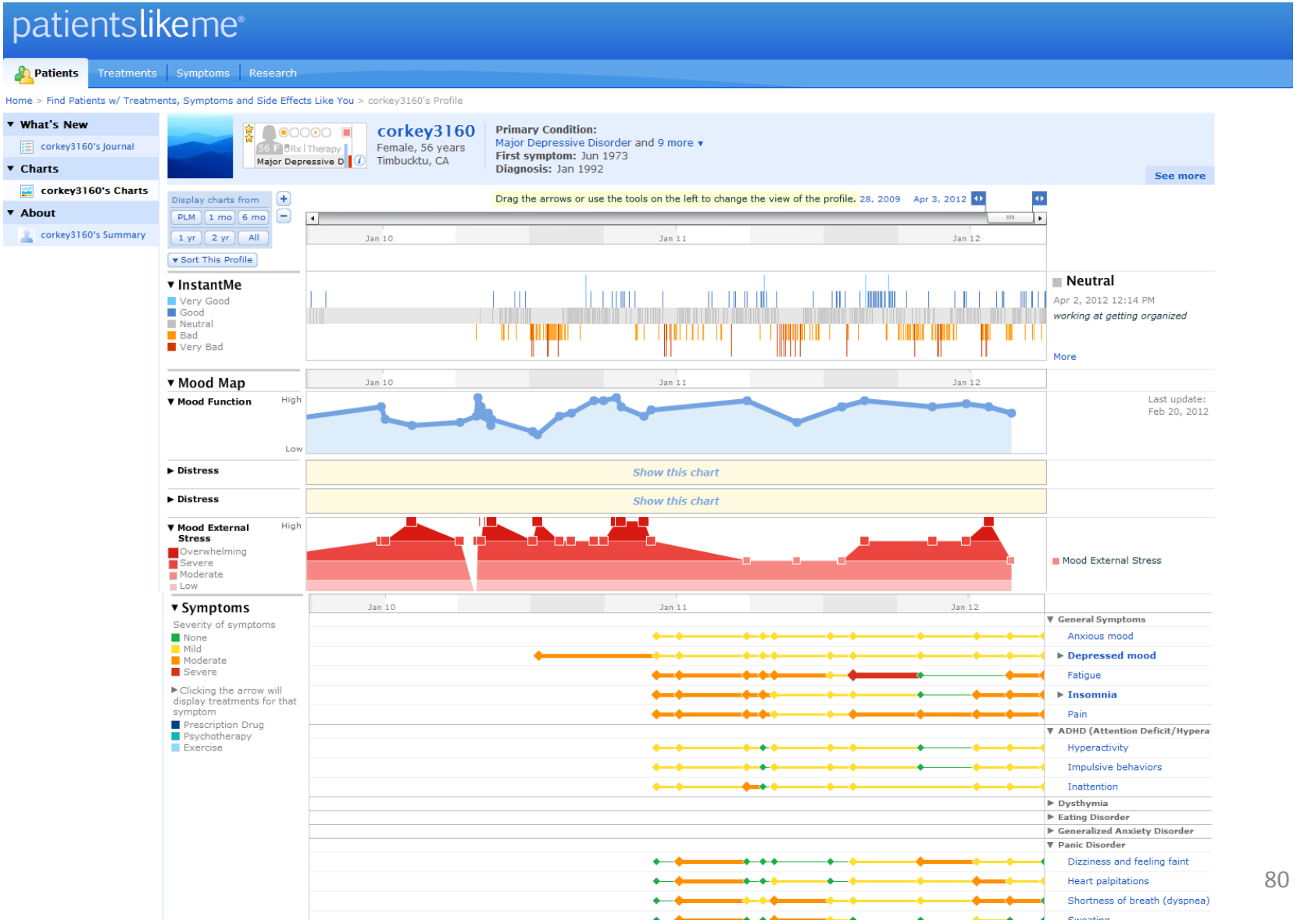
Suggested reading

- Privacy
  - **Pfitzmann & Hansen** (2010): A terminology for talking about privacy by data minimization: Anonymity, Unlinkability, Undetectability, Unobservability, Pseudonymity, and Identity Management
    - defines *anonymity*, *pseudonymity*, *undetectability*, *unobservability*, *unlinkability*
  - **Guerses** (2010): Multilateral Privacy Requirements Analysis in Online Social Network Services (PhD thesis)
    - section 2.2 (pg. 22-32) provide an interesting overview of privacy. Especially interesting for the following concepts: *confidentiality*, *feedback and awareness*
  - **Guarda and Zannone** (2008): Towards the development of privacy-aware systems
    - for those interested in the legal aspects of privacy
    - summarize the privacy principles from a legislation perspective, as it is clearly also important that a system is compliant with law, policies, and user consent (*policy and consent compliance*)
- Methodology
  - LINDDUN: a privacy threat analysis framework
    - http://people.cs.kuleuven.be/~kim.wuyts/ERISE/LINDDUN.pdf

LINDDUN - Privacy threat analysis

EXAMPLE – PATIENT COMMUNITY SYSTEM

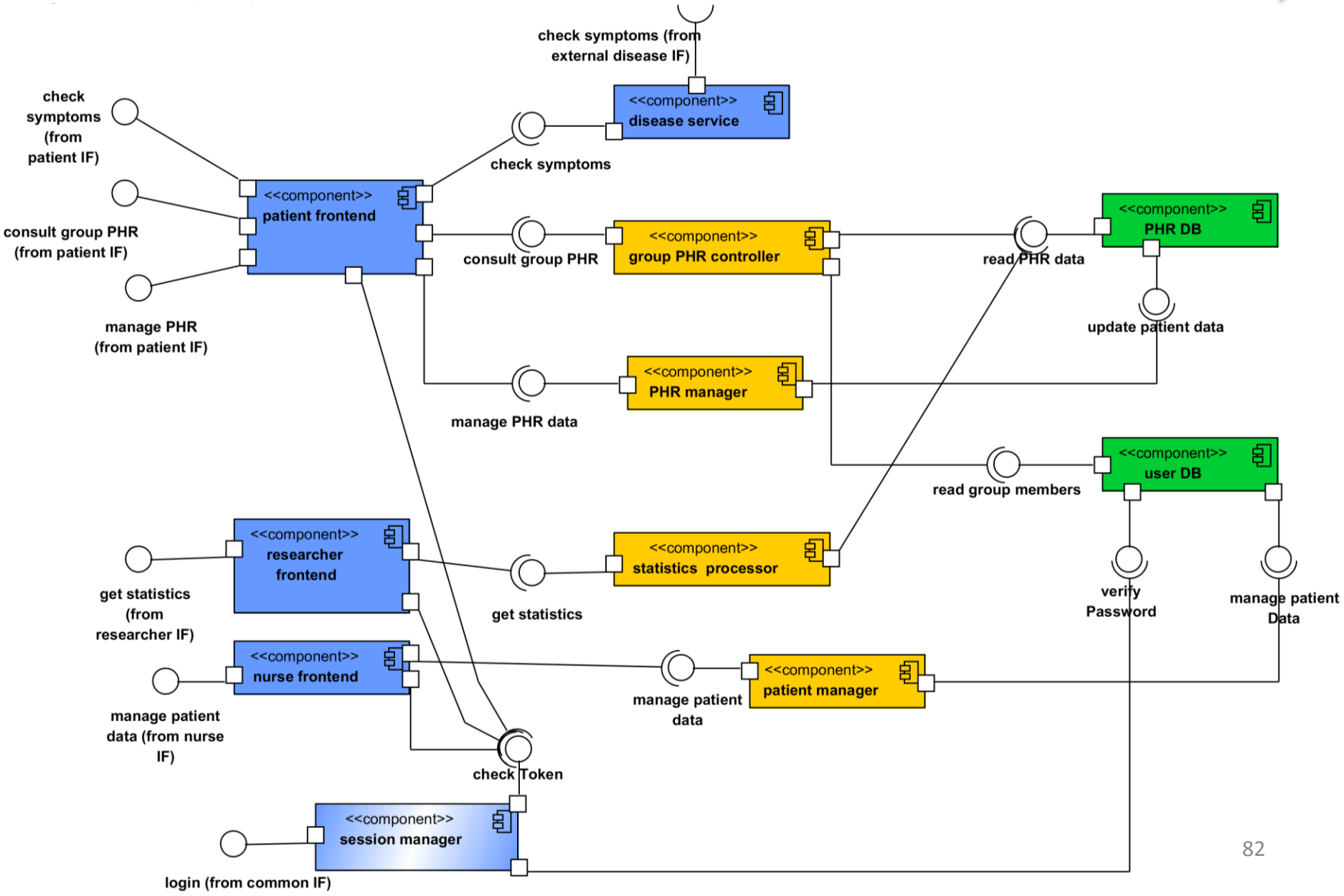
Existing patient communities



Patient communities case study

- Patient
  - Store personal health data (PHR)
  - Retrieve (pseudonymized) PHR from other patients (group members) with same condition
  - Retrieve trustworthy information on diseases and treatments (from external service)
- Nurse
  - Add users and manage groups
- Researcher
  - Retrieve (anonymized) PHR data to use in analysis

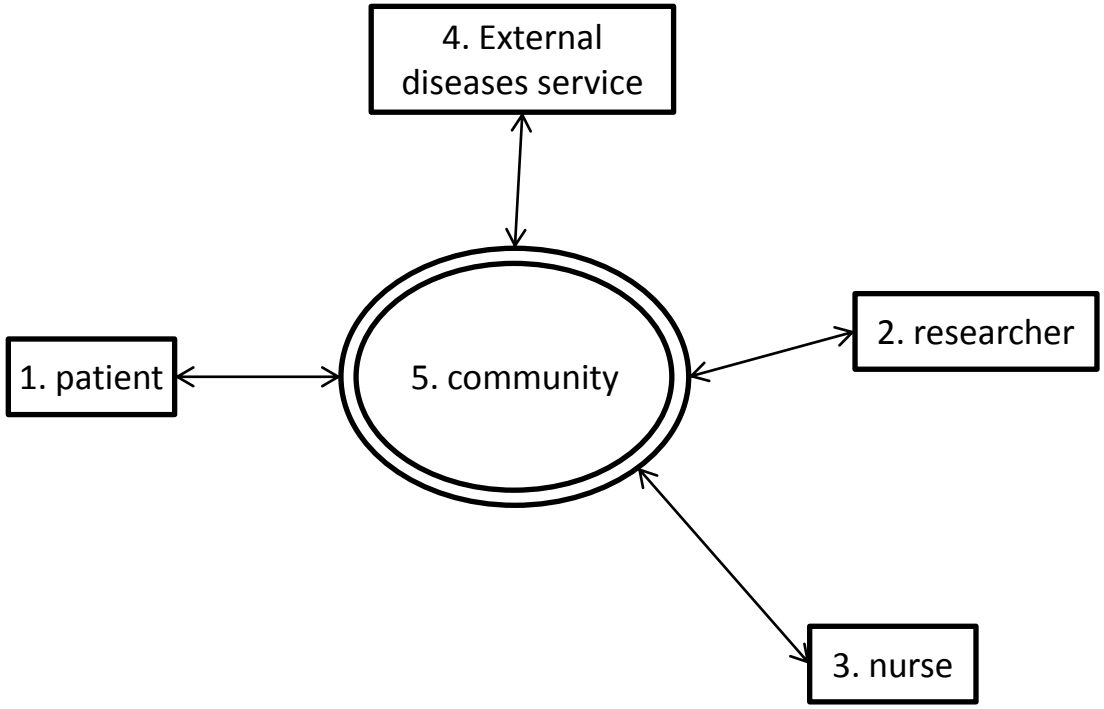
Client-server view



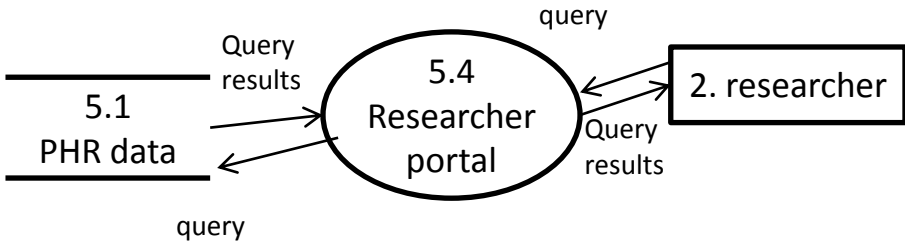
LINDDUN Methodology

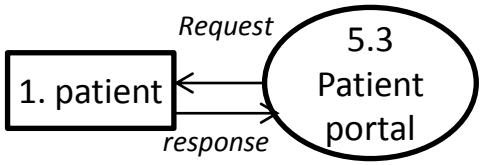
- **Step 1**
  - **Create the DFD diagram (assets)**
- Step 2
  - Map LINDDUN to DFD element types
- Step 3
  - Refine threats via threat tree patterns
  - Document assumptions
  - Document the threats with template
- Step 4
  - Assign priorities
- Step 5
  - Extract privacy requirements

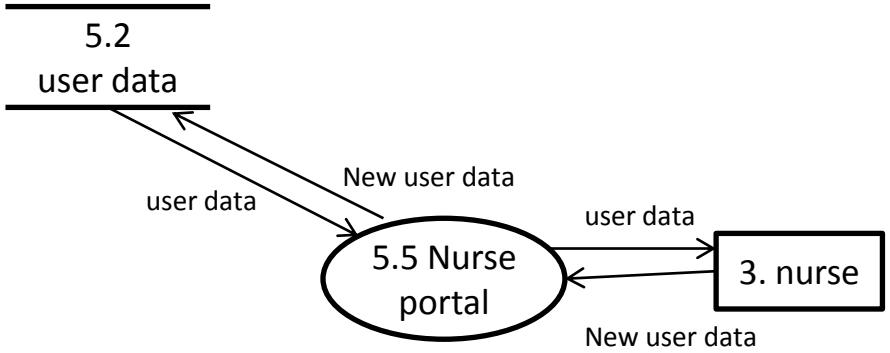
DFD level 0

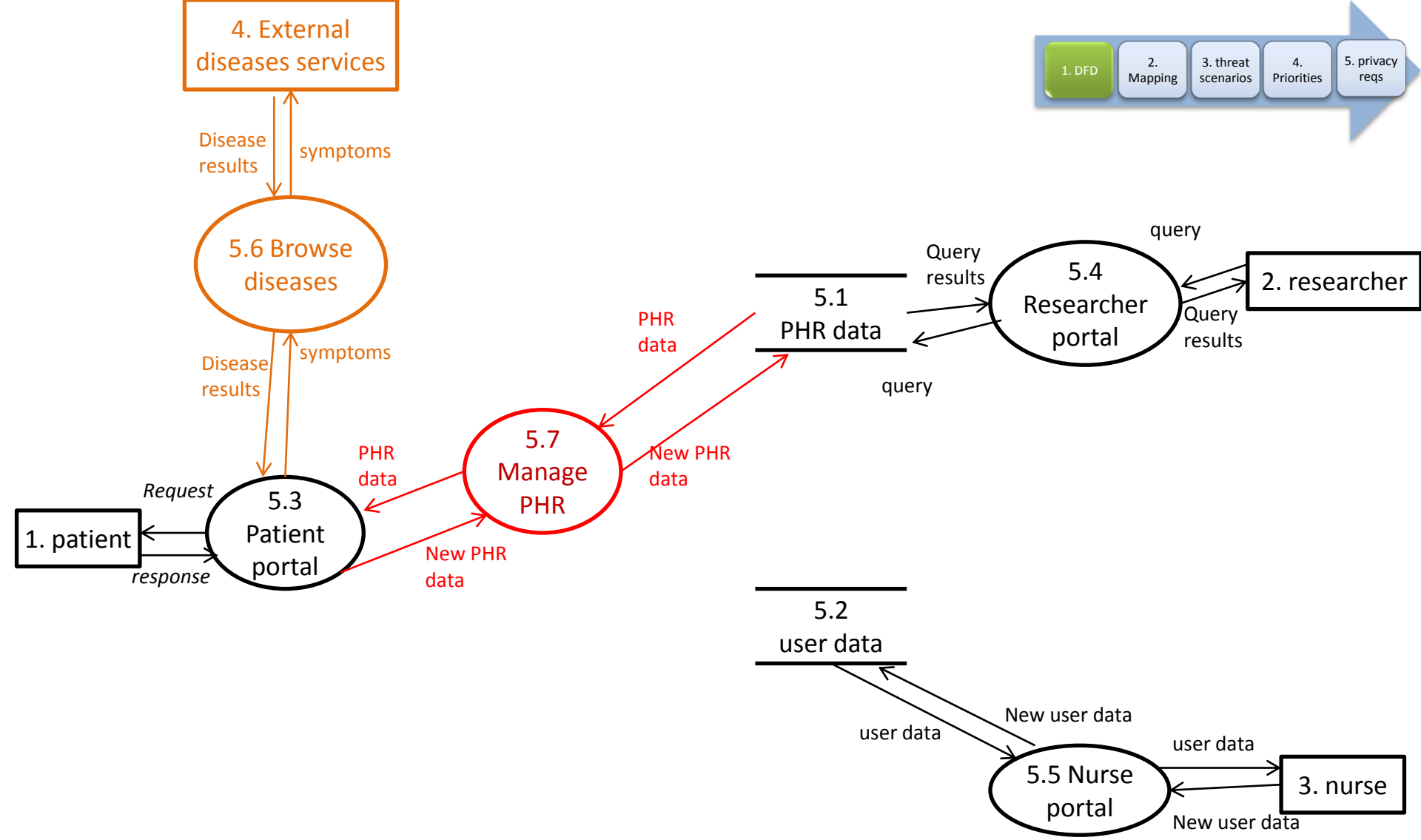

4. External diseases service
1. patient
5. community
2. researcher
3. nurse

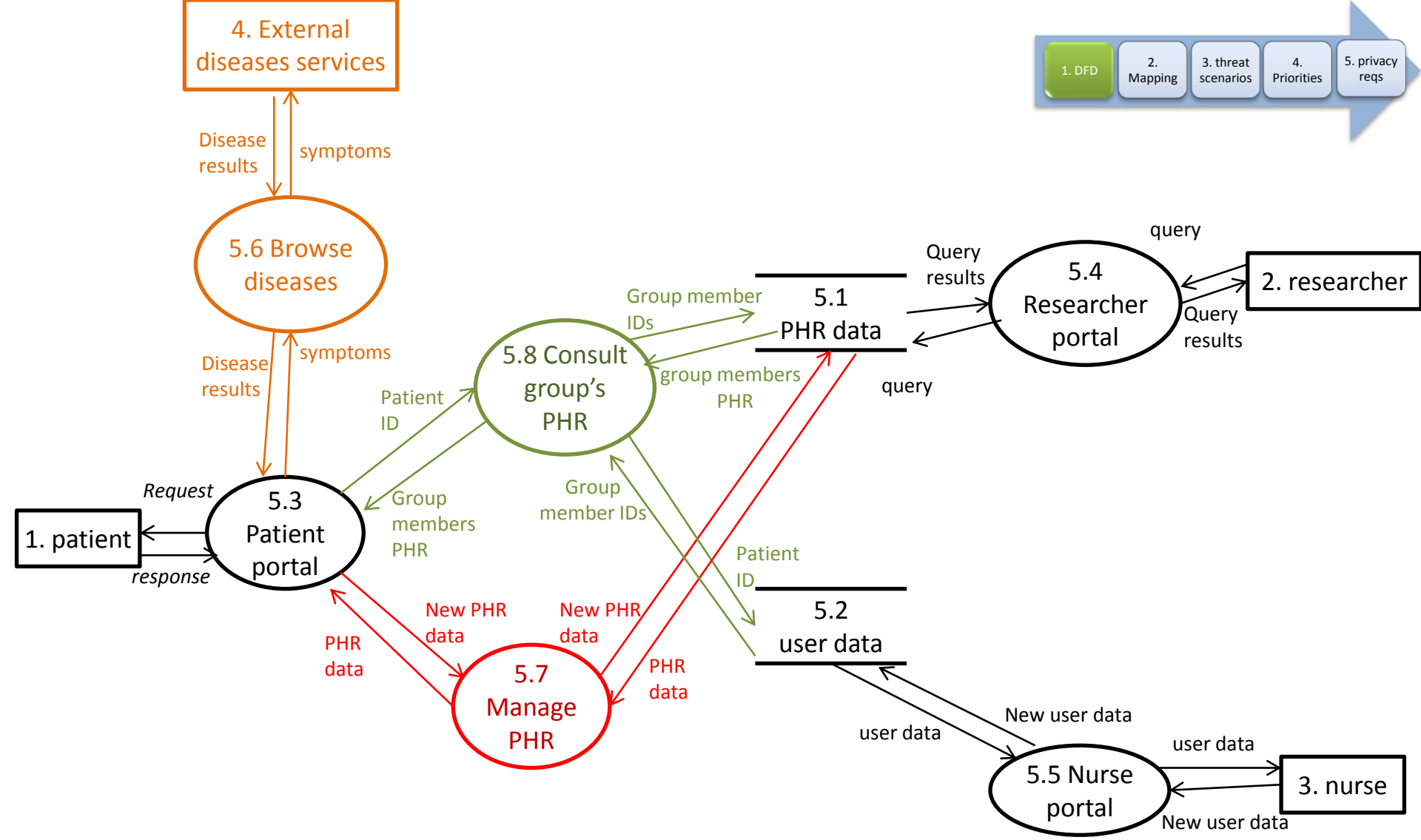
4. External diseases services

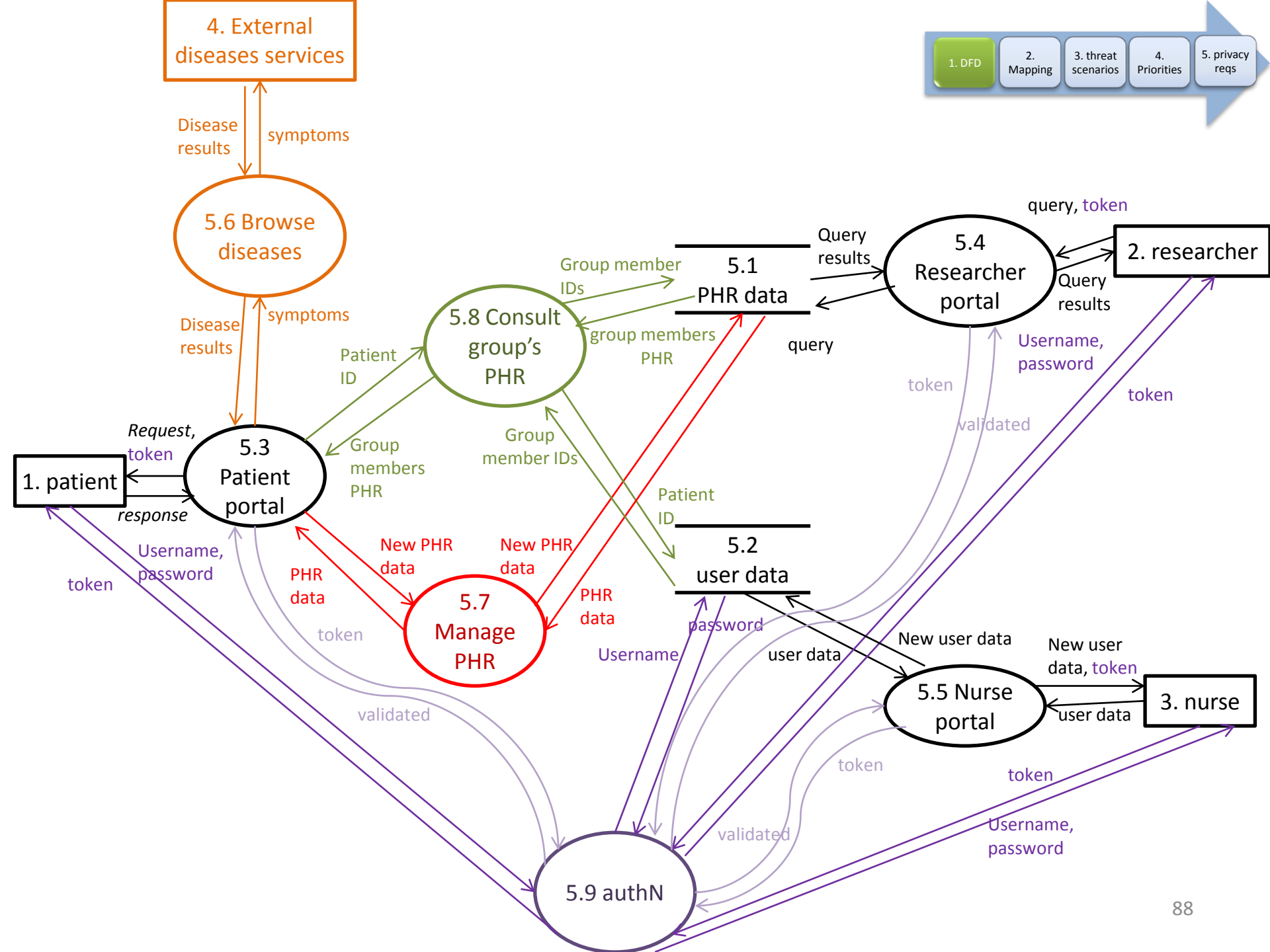







4. External diseases services
Disease results
symptoms
5.6 Browse diseases
Disease results
symptoms
Request
1. patient
response
5.3 Patient portal
PHR data
5.7 Manage PHR
New PHR data
New PHR data
PHR data
5.1 PHR data
Query results
5.4 Researcher portal
query
2. researcher
Query results
query
5.2 user data
user data
New user data
5.5 Nurse portal
user data
3. nurse
New user data
1. DFD
2. Mapping
3. threat scenarios
4. Priorities
5. privacy reqs


4. External diseases services
Disease results
symptoms
5.6 Browse diseases
Disease results
symptoms
Request
1. patient
response
5.3 Patient portal
Patient ID
Group members PHR
5.8 Consult group's PHR
Group member IDs
group members PHR
5.1 PHR data
Query results
5.4 Researcher portal
query
2. researcher
Query results
query
Group member IDs
Patient ID
5.2 user data
New PHR data
PHR data
New PHR data
PHR data
5.7 Manage PHR
New user data
user data
user data
5.5 Nurse portal
3. nurse
New user data
1. DFD
2. Mapping
3. threat scenarios
4. Priorities
5. privacy reqs

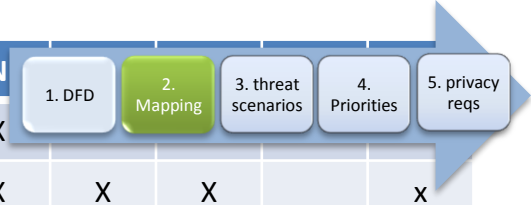

1. DFD
2. Mapping
3. threat scenarios
4. Priorities
5. privacy reqs
4. External diseases services
Disease results
symptoms
5.6 Browse diseases
Disease results
symptoms
Request, token
1. patient
response
5.3 Patient portal
Patient ID
Group members PHR
5.8 Consult group's PHR
Group member IDs
group members PHR
5.1 PHR data
Query results
query
5.4 Researcher portal
query, token
2. researcher
Query results
Username, password
token
validated
token
Group member IDs
Patient ID
5.2 user data
New PHR data
PHR data
5.7 Manage PHR
New PHR data
PHR data
Username, password
token
token
validated
password
Username
New user data
user data
5.5 Nurse portal
New user data, token
user data
3. nurse
token
validated
token
Username, password
5.9 authN

LINDDUN Methodology

- Step 1
  - Create the DFD diagram (assets)
- **Step 2**
  - **Map LINDDUN to DFD element types**
- Step 3
  - Refine threats via threat tree patterns
  - Document assumptions
  - Document the threats with template
- Step 4
  - Assign priorities
- Step 5
  - Extract privacy requirements

Mapping threats to DFD

	Linkability	Identifiability	Non-repudiation	Detectability	Information Disclosure	Content Unawareness	Policy & Consent Non-compliance
Data store	X	X	X	X	X		X
Data flow	X	X	X	X	X		X
Process	X	X	X	X	X		X
Entity	X	X				X	



		L	I	N				
Data store	PHR data (5.1)	X	X	X				
	User data (5.2)	X	X	X	X	X		x
flow	Patient – portal flow (1 -5.3)	X	X	X	X	X		x
	Portal – patient flow (5.3-1)	X	X	X	X	X		x
	Researcher – portal flow (2-5.4)	X	X	X	X	X		x
	Portal – researcher flow (5.4-2)	X	X	X	X	X		x
	Nurse – portal flow (3-5.5)	X	X	X	X	X		x
	Portal – nurse flow (5.5-3)	X	X	X	X	X		x
	Disease service – browse diseases flow (4-5.6)	X	X	X	X	X		x
	Browse disease – disease service flow (5.6-4)	X	X	X	X	X		x
	Patient portal - browse diseases flow (5.3-5.6)	X	X	X	X	X		x
	Browse diseases – patient portal flow (5.6-5.3)	X	X	X	X	X		x
	Patient portal – manage PHR flow (5.3-5.7)	X	X	X	X	X		x
	Manage PHR flow (5.7-5.3)	X	X	X	X	X		x
	Patient portal – consult group PHR (5.3-5.8)	X	X	X	X	X		x
	Consult group PHR – patient portal flow (5.8-5.3)	X	X	X	X	X		x
	Researcher portal – PHR data flow (5.4-5.1)	X	X	X	X	X		x
	PHR data – researcher portal flow (5.1-5.4)	X	X	X	X	X		x
	Nurse portal – user data flow (5.5-5.2)	X	X	X	X	X		x
	User data – nurse portal flow(5.2-5.5)	X	X	X	X	X		x
	Manage PHR – PHR data (5.7-5.1)	X	X	X	X	X		x
	PHR data – manage PHR (5.1-5.7)	X	X	X	X	X		x

		L	I	N	D	D	U	N
	Consult group PHR – PHR data flow (5.8-5.1)	X	X	X	X	X		x
	PHR data – consult group PHR flow (5.1-5.8)	X	X	X	X	X		x
	Consult group PHR – user data (5.8-5.2)	X	X	X	X	X		x
	User data – consult group PHR (5.2-5.8)	X	X	X	X	X		x
	Patient – authN flow (1-5.9)	X	X	X	X	X		x
	authN - patient flow (5.9-1)	X	X	X	X	X		x
	Research – authN flow (2-5.9)	X	X	X	X	X		x
	authN flow – researcher (5.9-2)	X	X	X	X	X		x
	Nurse – authN flow (3-5.9)	X	X	X	X	X		x
	authN – nurse (3-5.9)	X	X	X	X	X		x
	User data – authN flow (5.2-5.9)	X	X	X	X	X		x
	authN – user data flow (5.9-5.2)	X	X	X	X	X		x
	Patient portal – authN (5.3-5.9)	X	X	X	X	X		x
	authN – patient portal (5.9-5.3)	X	X	X	X	X		x
	Researcher portal – authN (5.4-5.9)	X	X	X	X	X		x
	authN – researcher portal (5.9-5.4)	X	X	X	X	X		x
	Nurse portal – authN (5.5-5.9)	X	X	X	X	X		x
	authN – nurse portal (5.9-5.5)	X	X	X	X	X		x

		L	I	N	D	D	U	N
process	Patient portal (5.3)	X	X	X	X	X		x
	Researcher portal (5.4)	X	X	X	X	X		x
	Nurse portal (5.5)	X	X	X	X	X		x
	Browse disease (5.6)	X	X	X	X	X		x
	Manage PHR (5.7)	X	X	X	X	X		x
	Consult group PHR (5.8)	X	X	X	X	X		x
	authN (5.9)	X	X	X	X	X		x
Entity	Patient (1)	X	X				X	
	Researcher (2)	X	X				X	
	Nurse (3)	X	X				X	
	External disease service (4)	X	X				X	

LINDDUN Methodology

- Step 1
  - Create the DFD diagram (assets)
- Step 2
  - Map LINDDUN to DFD element types
- **Step 3**
  - **Refine threats via threat tree patterns**
  - **Document assumptions**
  - Document the threats with template
- Step 4
  - Assign priorities
- Step 5
  - Extract privacy requirements

General assumptions

1. all internal processes are only susceptible to insider threats, as we consider the back-end sufficiently protected against outsider threats. We will therefore combine the process threats and examine only one, as the threats apply to all of them
2. all data flows between internal processes and between internal processes and internal data stores are only susceptible to insider threats, as we consider the back-end sufficiently protected against outsider threats. We will therefore combine the data flow threats and examine only one, as the threats apply to all of them
3. data flows between an entity and a process are not considered trusted (as it involves transactions of an external entity to and from a trusted process over an insecure communication line)
4. data stores are not considered confidential, as no access control system is present

Positive assumptions can help understand reasoning

		L	I	N				
Data store	PHR data (5.1)	X	X	X				
	User data (5.2)	X	X	X	X	X		x
flow	Patient – portal flow (1 -5.3)	X	X	X	X	X		x
	Portal – patient flow (5.3-1)	X	X	X	X	X		x
	Researcher – portal flow (2-5.4)	X	X	X	X	X		x
	Portal – researcher flow (5.4-2)	X	X	X	X	X		x
	Nurse – portal flow (3-5.5)	X	X	X	X	X		x
	Portal – nurse flow (5.5-3)	X	X	X	X	X		x
	Disease service – browse diseases flow (4-5.6)	X	X	X	X	X		x
	Browse disease – disease service flow (5.6-4)	X	X	X	X	X		x
	Patient portal - browse diseases flow (5.3-5.6)	X	X	X	X	X		x
	Browse diseases – patient portal flow (5.6-5.3)	X	X	X	X	X		x
	Patient portal – manage PHR flow (5.3-5.7)	X	X	X	X	X		x
	Manage PHR flow (5.7-5.3)	X	X	X	X	X		x
	Patient portal – consult group PHR (5.3-5.8)	X	X	X	X	X		x
	Consult group PHR – patient portal flow (5.8-5.3)	X	X	X	X	X		x
	Researcher portal – PHR data flow (5.4-5.1)	X	X	X	X	X		x
	PHR data – researcher portal flow (5.1-5.4)	X	X	X	X	X		x
	Nurse portal – user data flow (5.5-5.2)	X	X	X	X	X		x
	User data – nurse portal flow(5.2-5.5)	X	X	X	X	X		x
	Manage PHR – PHR data (5.7-5.1)	X	X	X	X	X		x
	PHR data – manage PHR (5.1-5.7)	X	X	X	X	X		x

Internal data flows combined

Internal data flows combined

Internal data flows combined

		L	I	N	D	D	U	N
	Consult group PHR – PHR data flow (5.8-5.1)	X	X	X	X	X		x
	PHR data – consult group PHR flow (5.1-5.8)	X	X	X	X	X		x
	Consult group PHR – user data (5.8-5.2)	X	X	X	X	X		x
	User data – consult group PHR (5.2-5.8)	X	X	X	X	X		x
	Patient – authN flow (1-5.9)	X	X	X	X	X		x
	authN - patient flow (5.9-1)	X	X	X	X	X		x
	Research – authN flow (2-5.9)	X	X	X	X	X		x
	authN flow – researcher (5.9-2)	X	X	X	X	X		x
	Nurse – authN flow (3-5.9)	X	X	X	X	X		x
	authN – nurse (3-5.9)	X	X	X	X	X		x
	User data – authN flow (5.2-5.9)	X	X	X	X	X		x
	authN – user data flow (5.9-5.2)	X	X	X	X	X		x
	Patient portal – authN (5.3-5.9)	X	X	X	X	X		x
	authN – patient portal (5.9-5.3)	X	X	X	X	X		x
	Researcher portal – authN (5.4-5.9)	X	X	X	X	X		x
	authN – researcher portal (5.9-5.4)	X	X	X	X	X		x
	Nurse portal – authN (5.5-5.9)	X	X	X	X	X		x
	authN – nurse portal (5.9-5.5)	X	X	X	X	X		x

		L	I	N	D	D	U	N
process	Patient portal (5.3)	X	X	X	X	X		x
	Researcher portal (5.4)	X	X	X	X	X		x
	Nurse portal (5.5)	X	X	X	X	X		x
	Browse disease (5.6)	X	X	X	X	X		x
	Manage PHR (5.7)	X	X	X	X	X		x
	Consult group PHR (5.8)	X	X	X	X	X		x
	authN (5.9)	X	X	X	X	X		x
Entity	Patient (1)	X	X				X	
	Researcher (2)	X	X				X	
	Nurse (3)	X	X				X	
	External disease service (4)	X	X				X	

Internal processes combined

		L	I	N				
Data store	PHR data (5.1)	X	X	X	X	X		x
	User data (5.2)	X	X	X	X	X		x
flow	Patient – portal flow (1 -5.3)	X	X	X	X	X		x
	Portal – patient flow (5.3-1)	X	X	X	X	X		x
	Researcher – portal flow (2-5.4)	X	X	X	X	X		x
	Portal – researcher flow (5.4-2)	X	X	X	X	X		x
	Nurse – portal flow (3-5.5)	X	X	X	X	X		x
	Portal – nurse flow (5.5-3)	X	X	X	X	X		x
	Disease service – browse diseases flow (4-5.6)	X	X	X	X	X		x
	Browse disease – disease service flow (5.6-4)	X	X	X	X	X		x
	Patient portal - browse diseases flow (5.3-5.6)	X	X	X	X	X		x
	Patient – authN flow (1-5.9)	X	X	X	X	X		x
	authN - patient flow (5.9-1)	X	X	X	X	X		x
	Research – authN flow (2-5.9)	X	X	X	X	X		x
	authN flow – researcher (5.9-2)	X	X	X	X	X		x
	Nurse – authN flow (3-5.9)	X	X	X	X	X		x
	authN – nurse (3-5.9)	X	X	X	X	X		x
process	Patient portal (5.3)	X	X	X	X	X		x
Entity	Patient (1)	X	X				X	
	Researcher (2)	X	X				X	
	Nurse (3)	X	X				X	
	External disease service (4)	X	X				X	

General internal DF

General internal P

General assumptions

5. No non-repudiation threats exist in the system, as the data flows, processes and data stores do not require plausible deniability
6. detectability is not considered a threat for this specific system. The privacy concerns of this system are all focused on the data itself, not on the detectability of it
7. non-compliance is an important threat, however, it is not specific to one part of the system, but poses to the system as a whole. We will therefore not make a distinction between the different DFD elements for this threat.

		L	I	N	D	D	U	N
Data store	PHR data (5.1)	X	X	X	X	X		x
	User data (5.2)	X	X	X	X	X		x
flow	Patient – portal flow (1 -5.3)	X	X	X	X	X		x
	Portal – patient flow (5.3-1)	X	X	X	X	X		x
	Researcher – portal flow (2-5.4)	X	X	X	X	X		x
	Portal – researcher flow (5.4-2)	X	X	X	X	X		x
	Nurse – portal flow (3-5.5)	X	X	X	X	X		x
	Portal – nurse flow (5.5-3)	X	X	X	X	X		x
	Disease service – browse diseases flow (4-5.6)	X	X	X	X	X		x
	Browse disease – disease service flow (5.6-4)	X	X	X	X	X		x
	Patient portal - browse diseases flow (5.3-5.6)	X	X	X	X	X		x
	Patient – authN flow (1-5.9)	X	X	X	X	X		x
	authN - patient flow (5.9-1)	X	X	X	X	X		x
	Research – authN flow (2-5.9)	X	X	X	X	X		x
	authN flow – researcher (5.9-2)	X	X	X	X	X		x
	Nurse – authN flow (3-5.9)	X	X	X	X	X		x
	authN – nurse (3-5.9)	X	X	X	X	X		x
process	Patient portal (5.3)	X	X	X	X	X		x
Entity	Patient (1)	X	X				X	
	Researcher (2)	X	X				X	
	Nurse (3)	X	X				X	
	External disease service (4)	X	X				X	

No non-repudiation or detectability threat

Assumptions

8. Identifiability of entities (researchers, nurses, patients or the external service) is not considered a threat, as all entities should have their own unique (long-term) identifier and there is no need to hide the entity's identity. Knowing that an entity is using the community service is not considered an issue.
11. Linkability of entities (sensors, cardiologists, nurses, or patients) is not considered a threat, as all entities should have their own unique (long-term) identifier and there is no need to hide the entity's identity. Knowing that an entity is using the community service is not considered an issue.

		L	I	N	D	D	U	N
Data store	PHR data (5.1)	X	X	X	X	X		x
	User data (5.2)	X	X	X	X	X		x
flow	Patient – portal flow (1 -5.3)	X	X	X	X	X		x
	Portal – patient flow (5.3-1)	X	X	X	X	X		x
	Researcher – portal flow (2-5.4)	X	X	X	X	X		x
	Portal – researcher flow (5.4-2)	X	X	X	X	X		x
	Nurse – portal flow (3-5.5)	X	X	X	X	X		x
	Portal – nurse flow (5.5-3)	X	X	X	X	X		x
	Disease service – browse diseases flow (4-5.6)	X	X	X	X	X		x
	Browse disease – disease service flow (5.6-4)	X	X	X	X	X		x
	Patient portal - browse diseases flow (5.3-5.6)	X	X	X	X	X		x
	Patient – authN flow (1-5.9)	X	X	X	X	X		x
	authN - patient flow (5.9-1)	X	X	X	X	X		x
	Research – authN flow (2-5.9)	X	X	X	X	X		x
	authN flow – researcher (5.9-2)	X	X	X	X	X		x
	Nurse – authN flow (3-5.9)	X	X	X	X	X		x
	authN – nurse (3-5.9)	X	X	X	X	X		x
process	Patient portal (5.3)	X	X	X	X	X		x
Entity	Patient (1)	X	X				X	
	Researcher (2)	X	X				X	
	Nurse (3)	X	X				X	
	External disease service (4)	X	X				X	

Assumptions

14. Linkability and identifiability do not pose a threat to the data flows between entities (patient, nurse, and researcher) and (portal) processes because of assumptions 8 and 11

9. Identifiability of the data flow only poses a threat to one specific data flow: 5.6 ->4 (browse diseases to external disease services), as the external service should not be able to identify the patient that is using this disease browsing service.

10. Linkability of the data flow to the external disease service (5.6 -> 4) is the only linkability threat to data flows in the patient community system. Although less likely, when the patient identifiers are replaces by pseudonyms, linking the different symptoms (of different searches) together can still result in an identifiability threat

15. Linkability and identifiability do not apply to internal data flows as knowing that 2 requests belong to the same user, or knowing who made a request does not violate the patient's privacy. The patient's privacy is only violated when the content of the communication is revealed (information disclosure threat)

16. Linkability and identifiability do not apply to internal processes as knowing that 2 actions belong to the same user does not violate the patient's privacy. The patient's privacy is only violated when the content of the action is revealed (information disclosure threat)

		L	I	N	D	D	U	N
Data store	PHR data (5.1)	X	X	X	X	X		x
	User data (5.2)	X	X	X	X	X		x
flow	Patient – portal flow (1 -5.3)	X	X	X	X	X		x
	Portal – patient flow (5.3-1)	X	X	X	X	X		x
	Researcher – portal flow (2-5.4)	X	X	X	X	X		x
	Portal – researcher flow (5.4-2)	X	X	X	X	X		x
	Nurse – portal flow (3-5.5)	X	X	X	X	X		x
	Portal – nurse flow (5.5-3)	X	X	X	X	X		x
	Disease service – browse diseases flow (4-5.6)	X	X	X	X	X		x
	Browse disease – disease service flow (5.6-4)	X	X	X	X	X		x
	Patient portal - browse diseases flow (5.3-5.6)	X	X	X	X	X		x
	Patient – authN flow (1-5.9)	X	X	X	X	X		x
	authN - patient flow (5.9-1)	X	X	X	X	X		x
	Research – authN flow (2-5.9)	X	X	X	X	X		x
	authN flow – researcher (5.9-2)	X	X	X	X	X		x
	Nurse – authN flow (3-5.9)	X	X	X	X	X		x
	authN – nurse (3-5.9)	X	X	X	X	X		x
process	Patient portal (5.3)	X	X	X	X	X		x
Entity	Patient (1)	X	X				X	
	Researcher (2)	X	X				X	
	Nurse (3)	X	X				X	
	External disease service (4)	X	X				X	

Assumptions

12. Information disclosure between the external disease service and the browse disease process does not pose a threat, as it does not contain any sensitive or personal information

19. Content unawareness only applies to the patient, as the researcher does not add any information, a nurse only registers patients, and the external disease service does not directly input any data

17. Identifiability and linkability are applicable to both data stores, and will therefore be examined in a combined fashion

		L	I	N	D	D	U	N
Data store	PHR data (5.1)	X	X	X	X	X		x
	User data (5.2)	X	X	X	X	X		x
flow	Patient – portal flow (1 -5.3)	X	X	X	X	X		x
	Portal – patient flow (5.3-1)	X	X	X	X	X		x
	Researcher – portal flow (2-5.4)	X	X	X	X	X		x
	Portal – researcher flow (5.4-2)	X	X	X	X	X		x
	Nurse – portal flow (3-5.5)	X	X	X	X	X		x
	Portal – nurse flow (5.5-3)	X	X	X	X	X		x
	Disease service – browse diseases flow (4-5.6)	X	X	X	X	X		x
	Browse disease – disease service flow (5.6-4)	X	X	X	X	X		x
	Patient portal - browse diseases flow (5.3-5.6)	X	X	X	X	X		x
	Patient – authN flow (1-5.9)	X	X	X	X	X		x
	authN - patient flow (5.9-1)	X	X	X	X	X		x
	Research – authN flow (2-5.9)	X	X	X	X	X		x
	authN flow – researcher (5.9-2)	X	X	X	X	X		x
	Nurse – authN flow (3-5.9)	X	X	X	X	X		x
	authN – nurse (3-5.9)	X	X	X	X	X		x
process	Patient portal (5.3)	X	X	X	X	X		x
Entity	Patient (1)	X	X				X	
	Researcher (2)	X	X				X	
	Nurse (3)	X	X				X	
	External disease service (4)	X	X				X	

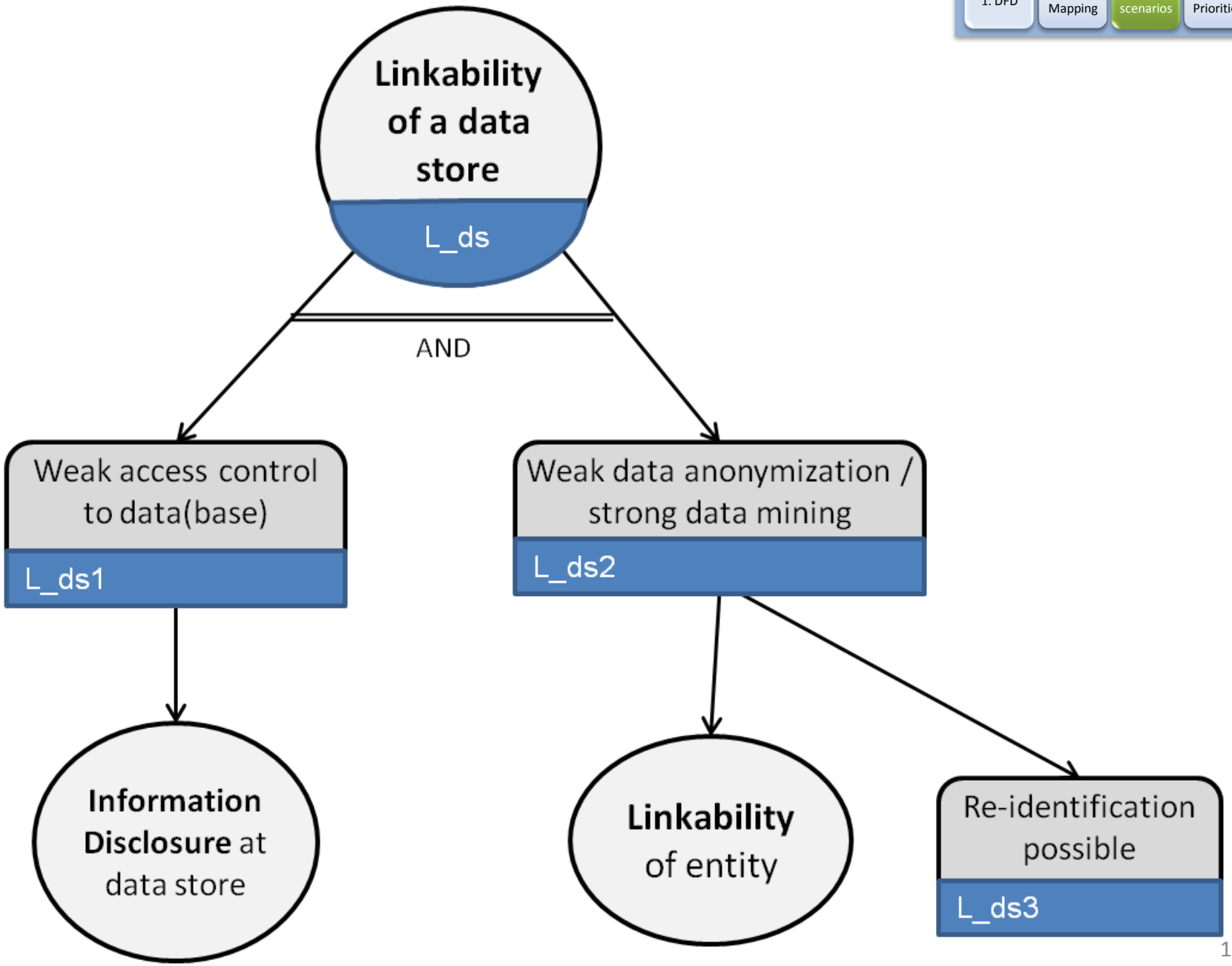
		L	I	N	D	D	U	N
Data store	PHR data (5.1)	X	X			X		x
	User data (5.2)	X	X			X		x
flow	Patient – portal flow (1 -5.3)					X		x
	Portal – patient flow (5.3-1)					X		x
	Researcher – portal flow (2-5.4)					X		x
	Portal – researcher flow (5.4-2)					X		x
	Nurse – portal flow (3-5.5)					X		x
	Portal – nurse flow (5.5-3)					X		x
	Disease service – browse diseases flow (4-5.6)					X		x
	Browse disease – disease service flow (5.6-4)	X	X					x
	Patient portal - browse diseases flow (5.3-5.6)					X		x
	Patient – authN flow (1-5.9)					X		x
	authN - patient flow (5.9-1)					X		x
	Research – authN flow (2-5.9)					X		x
	authN flow – researcher (5.9-2)					X		x
	Nurse – authN flow (3-5.9)					X		x
	authN – nurse (3-5.9)					X		x
process	Patient portal (5.3)					X		x
Entity	Patient (1)						X	
	Researcher (2)							
	Nurse (3)							
	External disease service (4)							

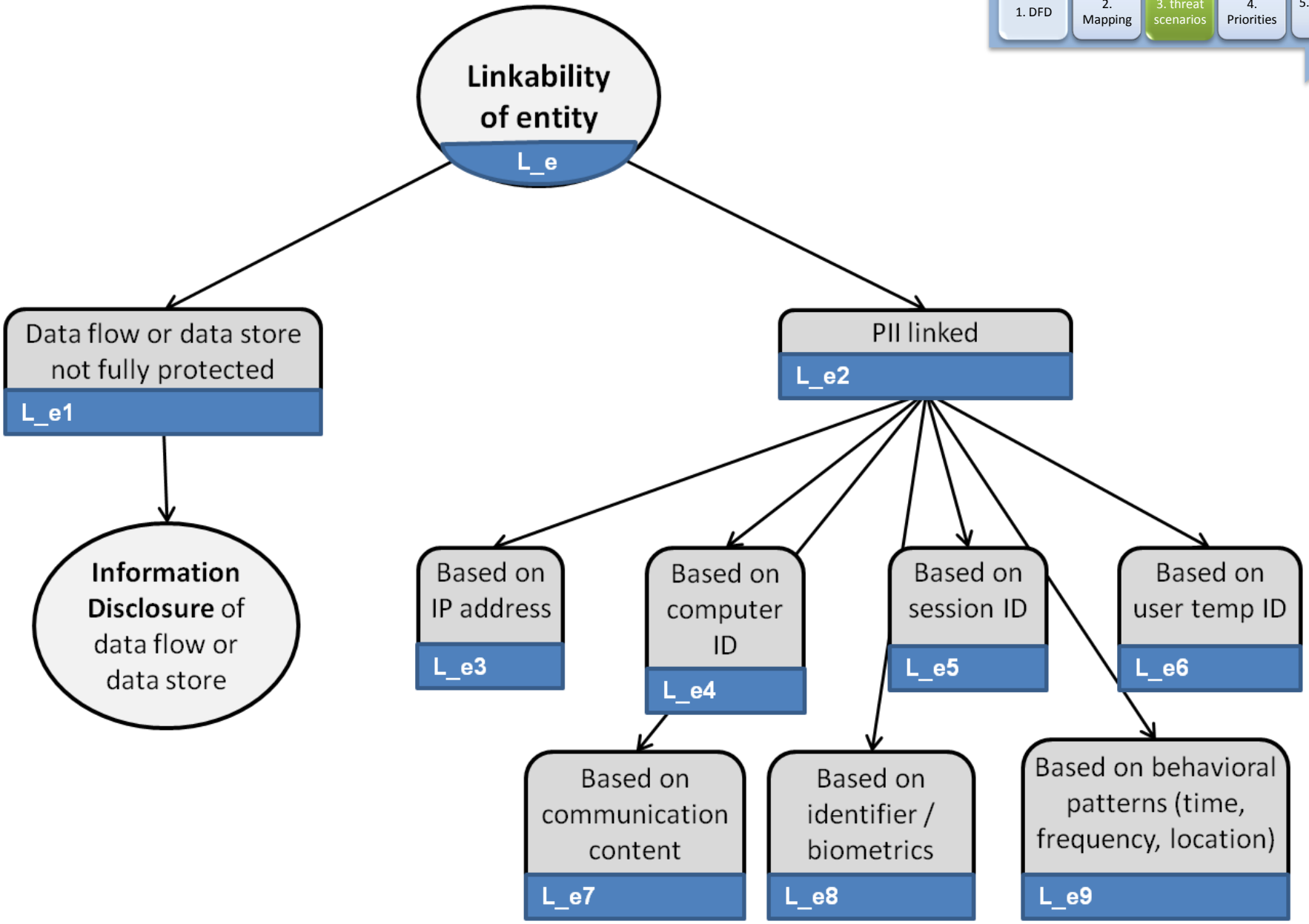
LINDDUN Methodology

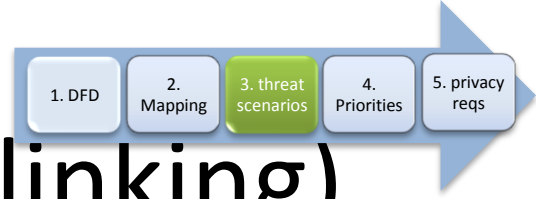
- Step 1
  - Create the DFD diagram (assets)
- Step 2
  - Map LINDDUN to DFD element types
- **Step 3**
  - **Refine threats via threat tree patterns**
  - Document assumptions
  - **Document the threats with template**
- Step 4
  - Assign priorities
- Step 5
  - Extract privacy requirements

Linking community data

- Assumption: *Identifiability and linkability are applicable to both data stores, and will therefore be examined in a combined fashion*
- linking PHR data
  - Applies also to user data


Linkability of a data store
L_ds
AND
Weak access control to data(base)
L_ds1
Weak data anonymization / strong data mining
L_ds2
Information Disclosure at data store
Linkability of entity
Re-identification possible
L_ds3


Linkability of entity
L_e
Data flow or data store not fully protected
L_e1
Information Disclosure of data flow or data store
PII linked
L_e2
Based on IP address
L_e3
Based on computer ID
L_e4
Based on session ID
L_e5
Based on user temp ID
L_e6
Based on communication content
L_e7
Based on identifier / biometrics
L_e8
Based on behavioral patterns (time, frequency, location)
L_e9



T01 - Profiling PHR data (linking)

Summary: A researcher or other insider with malicious intent links PHR data (or user data)

Primary mis-actor: unskilled insider (authenticated user, e.g. researcher)

Basic path:

bf1. The misactor performs a set of targeted queries on the PHR data or user data store and retrieves very detailed results

bf2. The misactor links the results of the queries together (e.g. based on medication which is usually combined, medical conditions which occur together, or pseudo-identifiers like street and age)

Consequence: By combining the query results, the misactor has access to more information about the patient than anticipated

Reference to threat tree node(s): L_ds2, L_e2

Parent threat tree(s): L_ds, I_ds

DFD element(s): 5.1 PHR data, 5.2 user data

Remarks:

r1. This threat can be used as precondition for the identifiability threat at the data store (T03 - Identifying a patient from his PHR data)

r2. This threat was inspired by L_ds2 and L_e2, however none of L_e2's leaf nodes matched

r3. The (weak) access requirement (L ds1) is fulfilled because the misactor is an insider who has access to the database

r4. Although this threat mainly describes the PHR data case, it also applies to the user data store (assumption 4)

Linking community data

- Assumption: *Identifiability and linkability are applicable to both data stores, and will therefore be examined in a combined fashion*
- linking PHR data
  - Applies also to user data
- Linking PHR data to user data

T02 - Linking PHR data to user data

Summary: The administrator or other insider with access to both the PHR data store and user data store is able to link the data from both databases (and sell this information to advertisers, insurance companies, etc.)

Primary mis-actor: unskilled insider with access to both data stores

Basic path:

bf1. The misactor retrieves information from both the PHR data store and the user data store

bf2. The misactor links both sets of data (e.g. based on a shared foreign key)

Consequence: The combined set of data contains (possibly sensitive) personal identifiable information and especially poses a privacy threat when the misactor sells the information (e.g. to a company selling medication, to the patient's insurance company, etc.)

Reference to threat tree node(s): L_ds2, L_e6

Parent threat tree(s): L_ds, I_ds

DFD element(s): 5.1 PHR data, 5.2 user data

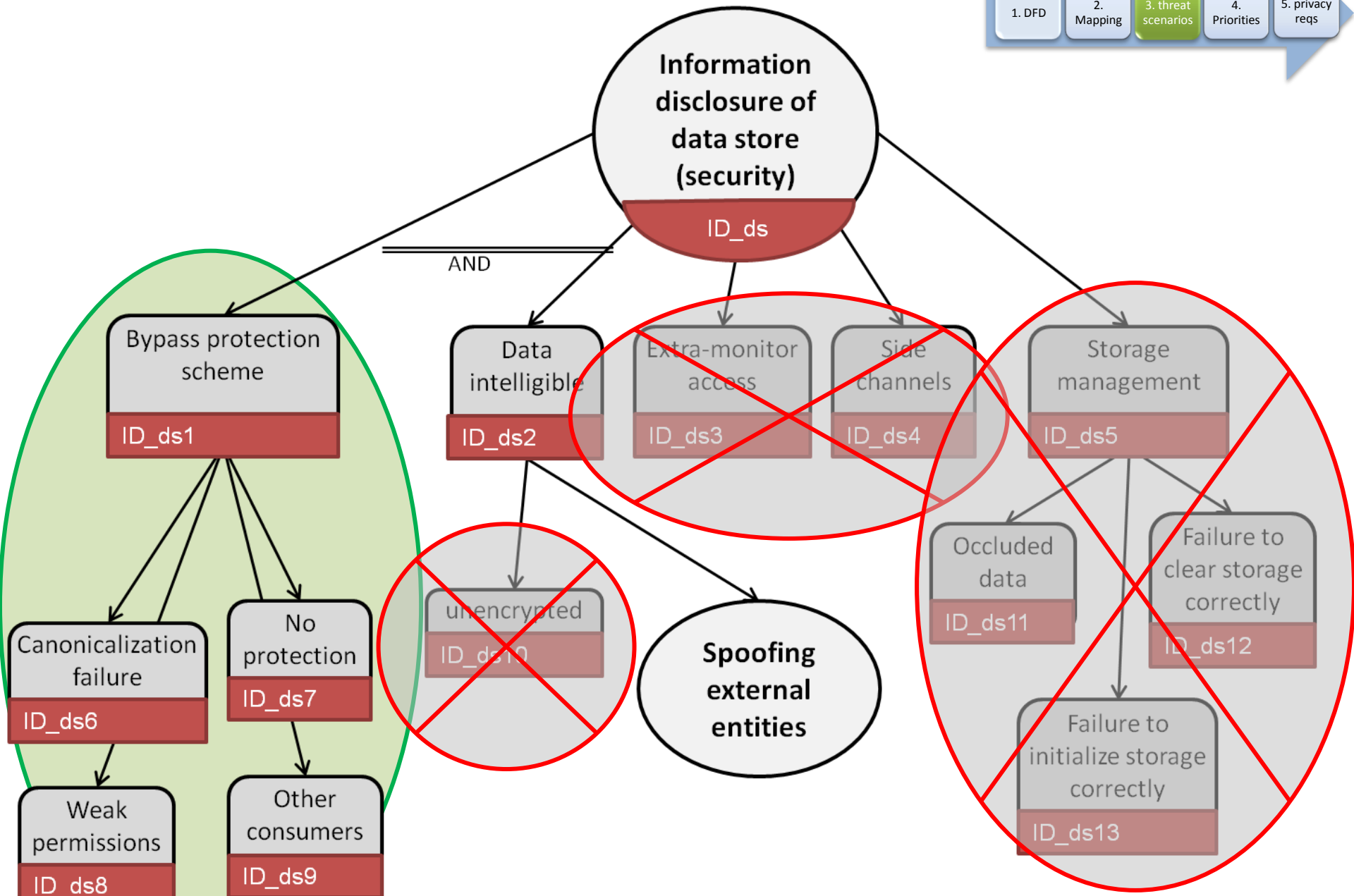
Remarks:

r1. The L_ds1 requirement of (weak) access is fulfilled, as this threat only involves insiders who have access to the data stores

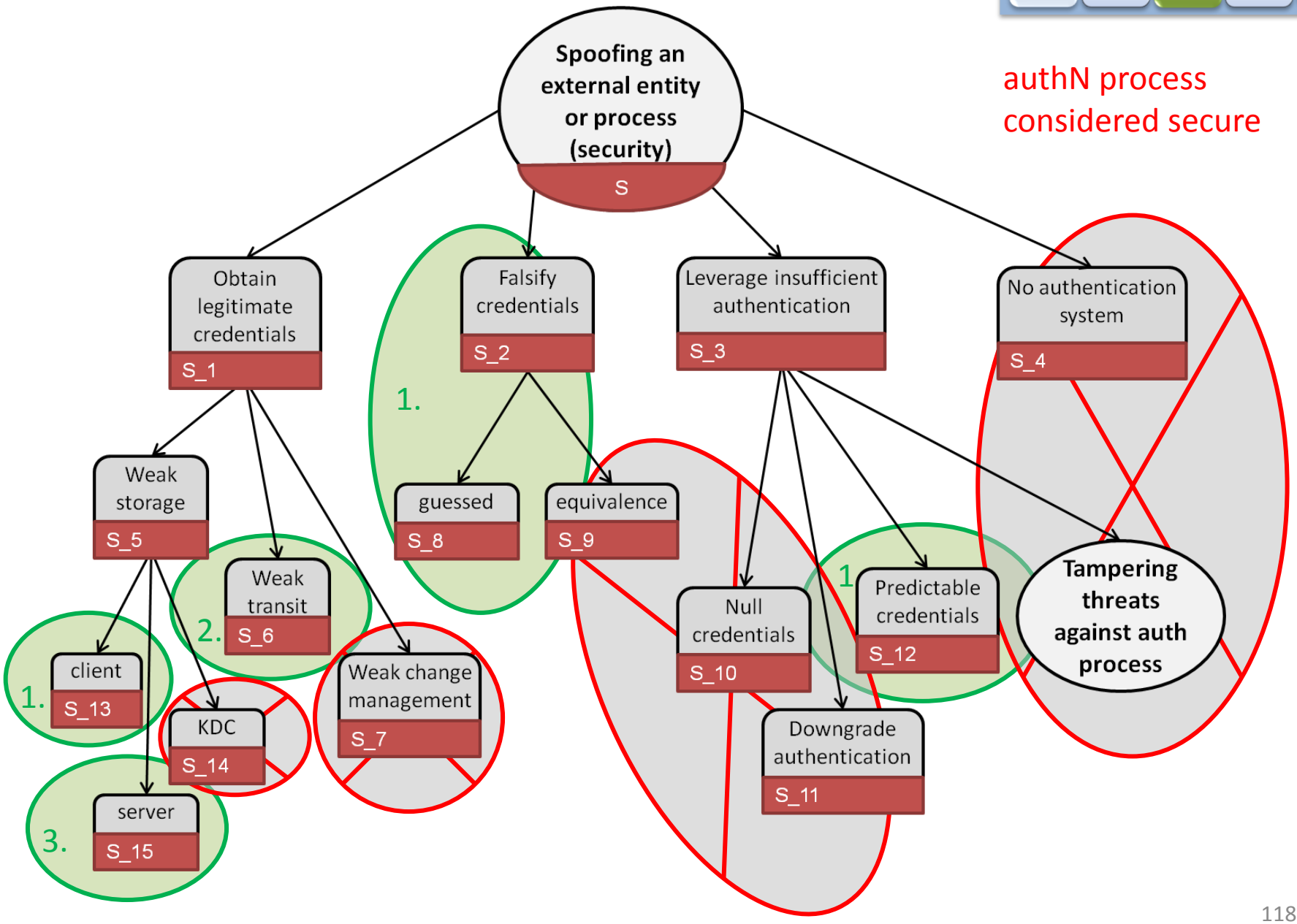
r2. The linkability of entity leaf node L_e6, indicating linkability based on the user's temporary ID inspired to this data store linkability threat

Information disclosure of community data

- no access control system is present (assumption 4)
- We assume that the data stores are sufficiently protected and that side-channel attacks, extra-monitor and bad storage management are not possible (assumption 20)

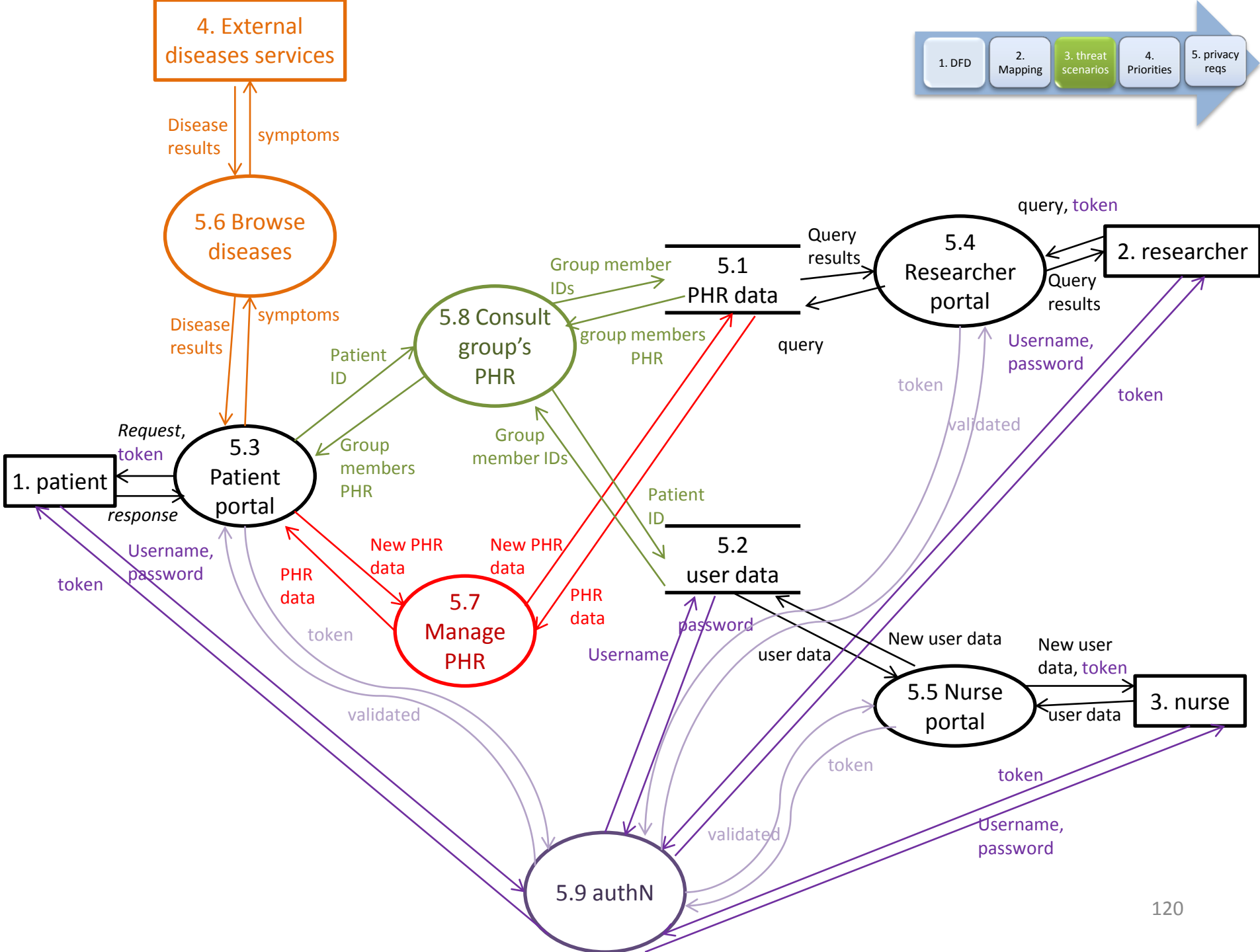

Information disclosure of data store (security)
ID_ds
AND
Bypass protection scheme
ID_ds1
Data intelligible
ID_ds2
Extra-monitor access
ID_ds3
Side channels
ID_ds4
Storage management
ID_ds5
Canonicalization failure
ID_ds6
No protection
ID_ds7
Weak permissions
ID_ds8
Other consumers
ID_ds9
unencrypted
ID_ds10
Spoofing external entities
Occluded data
ID_ds11
Failure to clear storage correctly
ID_ds12
Failure to initialize storage correctly
ID_ds13

authN process considered secure



Spoofing users (patients, researchers, nurses)

- Spoofing by falsifying credentials
- Spoofing by eavesdropping communication
  - Information disclosure of transmitted credentials
  - Information disclosure of transmitted session token
- Spoofing because of weak credential storage
  - Information disclosure of community data


4. External diseases services
Disease results
symptoms
5.6 Browse diseases
Disease results
symptoms
Request, token
1. patient
response
5.3 Patient portal
Patient ID
Group members PHR
5.8 Consult group's PHR
Group member IDs
group members PHR
5.1 PHR data
Query results
5.4 Researcher portal
query, token
2. researcher
Query results
query
Username, password
token
validated
token
Group member IDs
Patient ID
5.2 user data
New PHR data
New PHR data
PHR data
PHR data
5.7 Manage PHR
Username, password
token
token
validated
password
Username
New user data
user data
New user data, token
5.5 Nurse portal
user data
3. nurse
token
token
validated
Username, password
5.9 authN

spoofing external disease service

Spoofing external disease service

Summary: The external disease service is spoofed (e.g. by a competitor or a advertising company for medication)

Primary mis-actor: Skilled outsider

Basic path:

bf1. The misactor pretends to be the disease service

bf2. The community browse service contacts the spoofed disease service with symptoms

bf3. The misactor returns false information

Consequence: The patient community system returns false disease information to the patient which has an impact on the system's reputation (as one of the benefits of the provided service is the trustworthiness of the information)

Reference to threat tree node(s): S 4

Parent threat tree(s): S

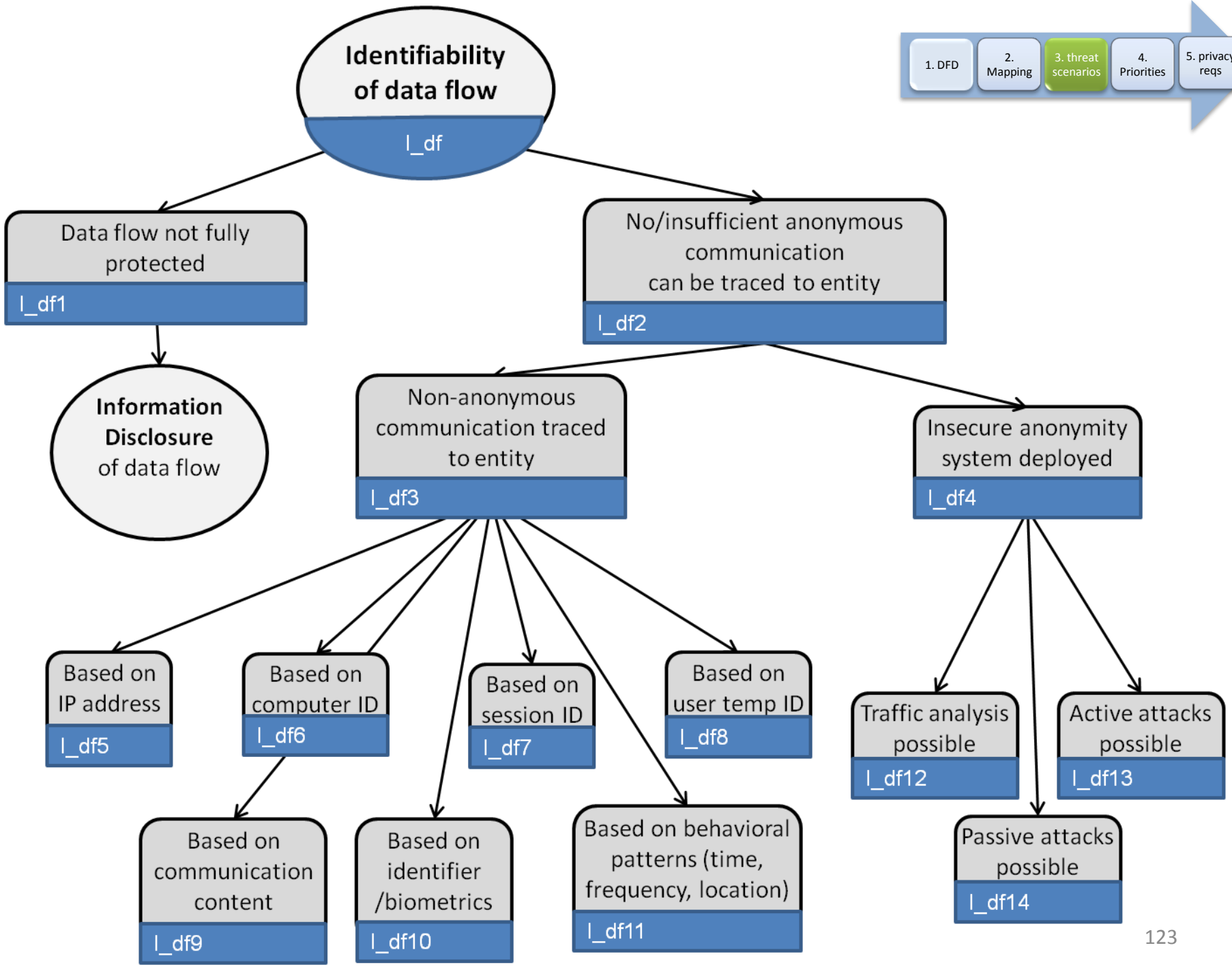
DFD element(s): 4. external disease service

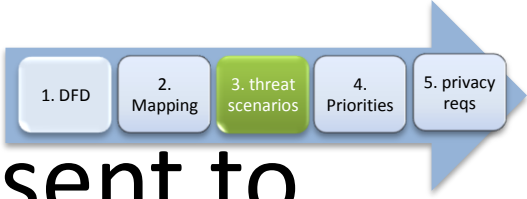
This is **NOT** a **privacy** threat.

It is a security threat (against integrity)
And should not be included

External disease service

- Linkability & Identifiability of data flow
  - NOT during transit
  - When arrived at external disease service
    - Always information disclosure


1. DFD
2. Mapping
3. threat scenarios
4. Priorities
5. privacy reqs
Identifiability of data flow
l_df
Data flow not fully protected
l_df1
No/insufficient anonymous communication can be traced to entity
l_df2
Information Disclosure of data flow
Non-anonymous communication traced to entity
l_df3
Insecure anonymity system deployed
l_df4
Based on IP address
l_df5
Based on computer ID
l_df6
Based on session ID
l_df7
Based on user temp ID
l_df8
Traffic analysis possible
l_df12
Active attacks possible
l_df13
Based on communication content
l_df9
Based on identifier /biometrics
l_df10
Based on behavioral patterns (time, frequency, location)
l_df11
Passive attacks possible
l_df14



T12 - Identifiability of data sent to external disease service

Summary: The misactor extracts the patient's identity from the request and links it to the symptoms

Primary mis-actor: unskilled insider/skilled outsider

Basic path:

bf1. The patient searches diseases by providing his symptoms to the patient portal, which forwards the request (include the patient's identifiable information (e.g. SSN, address, etc.) to the external disease service

bf2. The misactor intercepts the data flow (threat T10 – Information disclosure of transmitted medical or personal data) or is (or has access to) the external disease service

Consequence: The misactor knows which patient has which symptoms

Reference to threat tree node(s): I_df1, I_df8

Parent threat tree(s): I_df

DFD element(s): data flow from browse service to external disease service (5.6 -> 4)

Remarks:

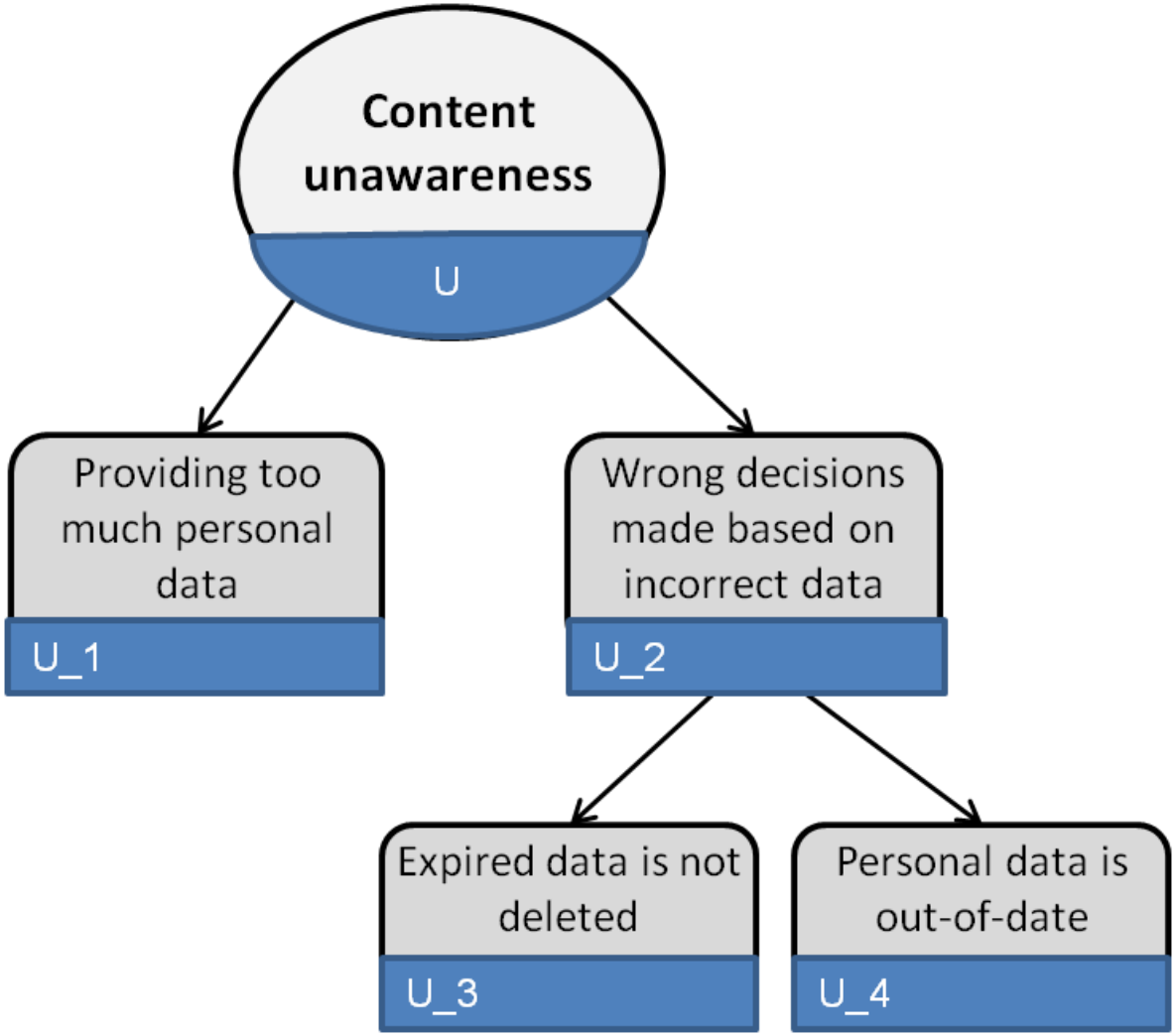
r1. I_df1 requires an unprotected data flow, which is currently present (assumption 3) and misactor is receiver, thus assumption always applies

r2. The different requests are traced back based on the transmitted (temporary/internal) user ID (I_df8)

r3. The right branch of the tree (insecure anonymity system (I_df4)) and the other leaf nodes of the non-anonymous communication branch (I_df3) are not considered, as it is not the sender (browse service) whose identity should be protected, but the patient, who is not directly part of the data flow

Soft privacy

- Non-compliance of employees
- Non-compliance of management
- Missing consent system
- Patient unawareness
- Content inaccuracy


Content unawareness
U
Providing too much personal data
U_1
Wrong decisions made based on incorrect data
U_2
Expired data is not deleted
U_3
Personal data is out-of-date
U_4

T19 - User unwareness

Summary: The user is unaware of the consequences of sharing information (e.g. by sharing too much information even anonymized data can reveal the user's identity)

Primary mis-actor: Management

Basic path:

bf1. The management fails to add as requirement the need of notifications and warnings when the patients intends to upload sensitive and/or identifiable content (e.g. picture of his broken arm which also shows his face)

bf2. The user adds information to the system which can easily identify him (e.g. a picture of himself) as he is unaware of the consequences

bf3. Group members retrieve information and can still identify the pseudonymized user

Consequence: When group members retrieve information, the identifiable information. The user's privacy is violated as he assumes that his information stays confidential and his identity will not be revealed

Reference to threat tree node(s): U 1

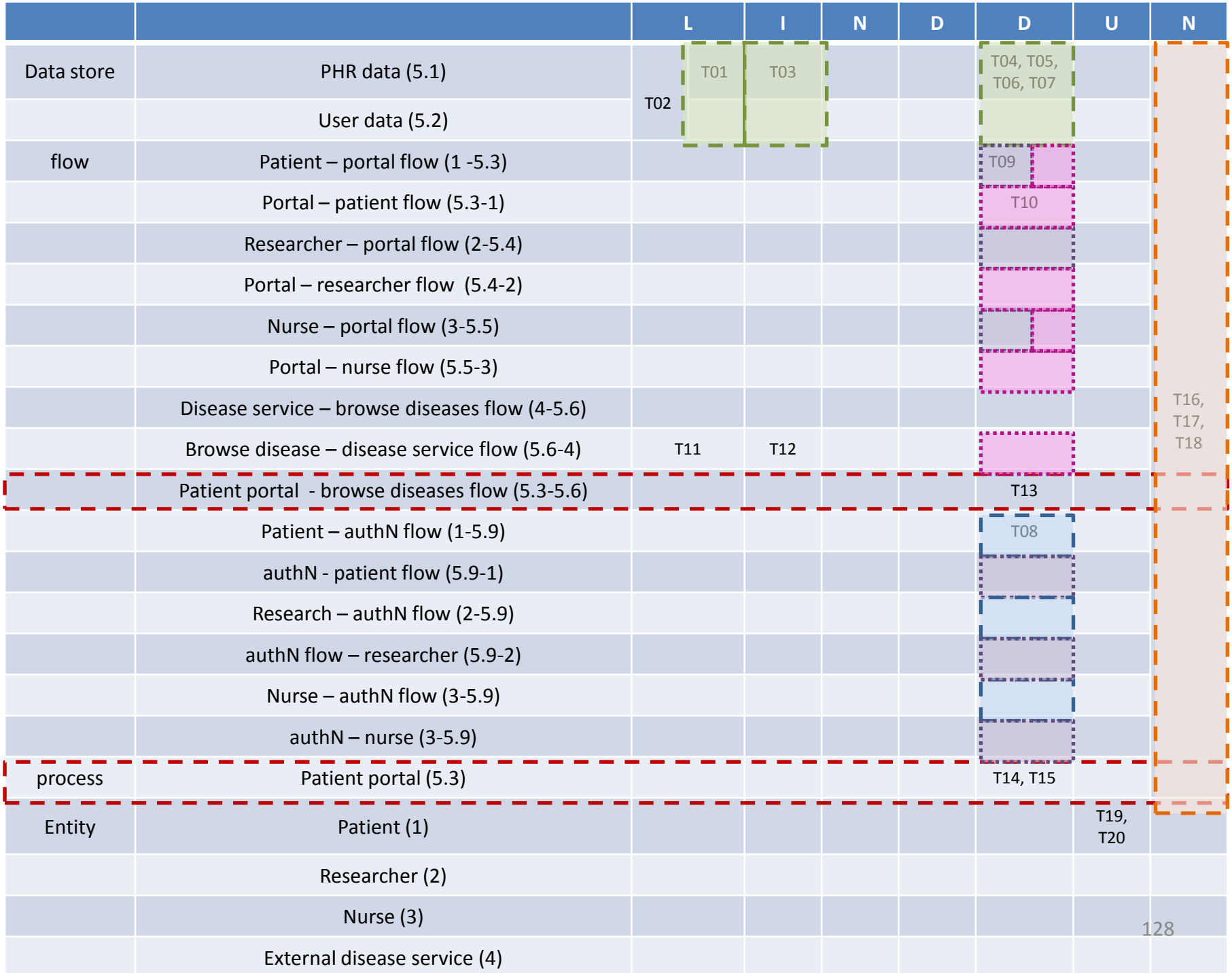
Parent threat tree(s): U

DFD element(s): 1 patient

Remarks:

r1. This threat only applies to the patient (assumption 19)

r2. The threat concerning inaccurate user information is described in T20 - content inaccuracy


L
I
N
D
D
U
N
Data store
PHR data (5.1)
T01
T03
T04, T05, T06, T07
T02
User data (5.2)
flow
Patient – portal flow (1 -5.3)
T09
Portal – patient flow (5.3-1)
T10
Researcher – portal flow (2-5.4)
Portal – researcher flow (5.4-2)
Nurse – portal flow (3-5.5)
Portal – nurse flow (5.5-3)
Disease service – browse diseases flow (4-5.6)
T16, T17, T18
Browse disease – disease service flow (5.6-4)
T11
T12
Patient portal - browse diseases flow (5.3-5.6)
T13
Patient – authN flow (1-5.9)
T08
authN - patient flow (5.9-1)
Research – authN flow (2-5.9)
authN flow – researcher (5.9-2)
Nurse – authN flow (3-5.9)
authN – nurse (3-5.9)
process
Patient portal (5.3)
T14, T15
Entity
Patient (1)
T19, T20
Researcher (2)
Nurse (3)
External disease service (4)

LINDDUN Methodology

- Step 1
  - Create the DFD diagram (assets)
- Step 2
  - Map LINDDUN to DFD element types
- Step 3
  - Refine threats via threat tree patterns
  - Document assumptions
  - Document the threats with template
- **Step 4**
  - **Assign priorities**
- Step 5
  - Extract privacy requirements

Priorities

Information disclosure and identifiability of stored data violate privacy more than disclosure of "partial" transmitted data

High

- T04 - Information disclosure of patient community data
- T03 - Identifying a patient from his PHR data
- T08 - Disclosure of the transmitted log-in credentials
- T09 - Disclosure of the transmitted session token
- T10 - Disclosure of transmitted medical/personal information
- T05 - Spoofing a user of the social network system by falsifying credentials
- T07 - Spoofing a user of the social network system because of weak credential storage
- T06 - Spoofing a user of the social network system by eavesdropping communication

Spoofing leads to information disclosure which is a high risk threat

Data in system is purely informative, and not used for important decisions, thus impact of threat is low

Low

- T16 - Non-compliance of employees
- T20 - content inaccuracy
- T14 - Information disclosure internal process
- T13 - Disclosure of internal transmitted medical/personal information
- T15 - Side channel information disclosure internal process

There is a trust relationship with the employees, thus likelihood of threats is low

Medium

- T12 - Identifiability of data sent to external disease service
- T11 - Linkability of symptoms sent to external disease service
- T01 - Profiling PHR data (linking)
- T02 - Linking PHR data to user data
- T18 - Non-compliance management
- T17 - Missing user consents
- T19 - User unwareness

Only partial data and patient deniability

Linkability can lead to identifiability

Non-compliance can still have "part" in place + reputation

LINDDUN Methodology

- Step 1
  - Create the DFD diagram (assets)
- Step 2
  - Map LINDDUN to DFD element types
- Step 3
  - Refine threats via threat tree patterns
  - Document assumptions
  - Document the threats with template
- Step 4
  - Assign priorities
- **Step 5**
  - **Extract privacy requirements**

Threats to requirements

T01 - Profiling PHR data

- A researcher or other insider with malicious intent links PHR data or user data
  - e.g. based on medication which is usually combined, medical conditions which occur together, or pseudo-identifiers like street and age

Threats (misuse cases)	**Caused by** (leaf nodes)	**Mitigated by** (requirements)
Profiling PHR data	Weak data anonymity (in the data store)	Apply strong data anonymization techniques in the database (for storage)
	PII linked (after retrieval)	Apply data anonymization techniques on query results (for information retrieval)