NC STATE UNIVERSITY

CSC 495.002 – Lecture 7
AI for Privacy: Privacy Requirements

Dr. Özgür Kafalı

North Carolina State University
Department of Computer Science

Fall 2017

NC STATE UNIVERSITY

PREVIOUSLY ON SOCIAL NETWORKS

Web/Social Networks Privacy

- Inference
- Sharing and disclosure
- Violations and regret
- Targeted advertising
- K-anonymity

What You Will Learn

- Privacy requirements engineering
- Autonomous agents and reasoning
  - Argumentation
  - Negotiation
- Privacy norms
- Reasoning about privacy breaches
  - Ontologies
  - Semantic similarity

Requirements

- Software requirements: Software has to provide solutions to establish the needs of its stakeholders
  - Satisfy a capability needed by a user to achieve an objective
  - Functionality to comply with a contract, regulation, or standard

- Example requirements from an electronic health records (EHR) software:
- The physician shall alter the current prescriptions of a patient or add new prescriptions after a routine visit
- The system shall respond to a patient scheduling request within 30 seconds

Security and Privacy Requirements

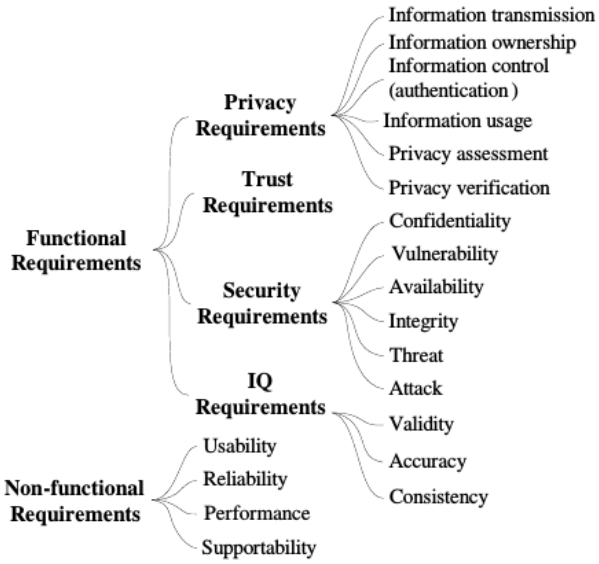
- Typically non-functional requirements, though might change depending on the domain
- Can be implied from functional requirements
- Requirement: The physician shall alter the current prescriptions of a patient or add new prescriptions after a routine visit

- What are the security and privacy implications of this requirement?
- Patients' prescription list should be encrypted
- Patients' prescription list should not be taken out of the hospital without being anonymized
- Physicians should only access those patients that they are currently treating

Access Control Requirements

- Describe who can access what using a role-based access control mechanism
- Can be implemented as part of the EHR software
- In an emergency, relax the access control mechanism
- Instead, a norm prohibits physicians from accessing EHR of other patients
- You can also log each access for auditing

Sample Requirements Taxonomy

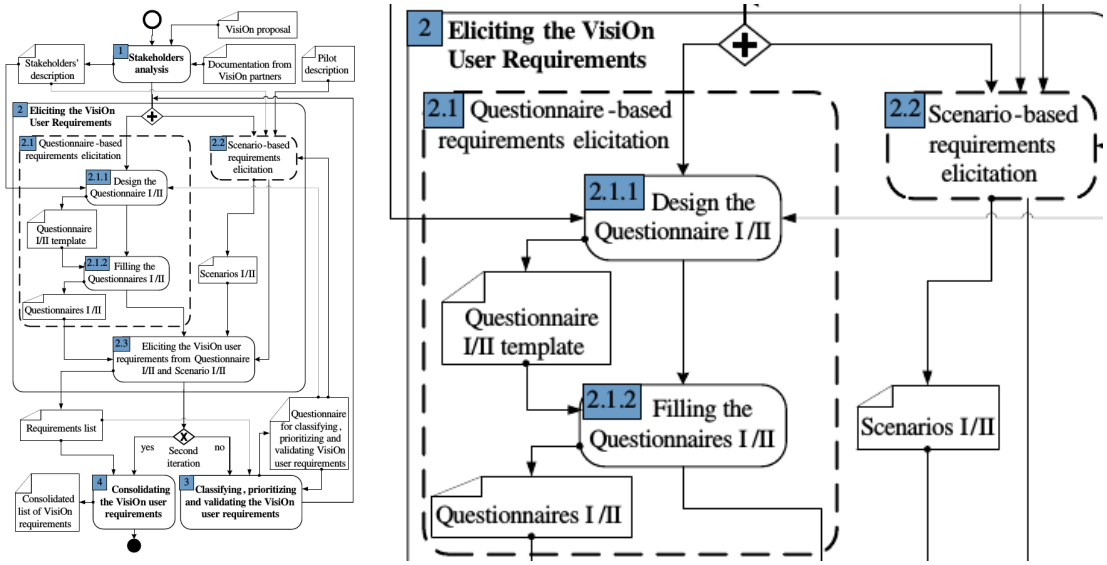


Gharib et al. Privacy Requirements: Findings and Lessons Learned in Developing a Privacy Platform. Requirements Engineering Conference (RE), pages 256–265, 2016

Phases of Requirements Engineering

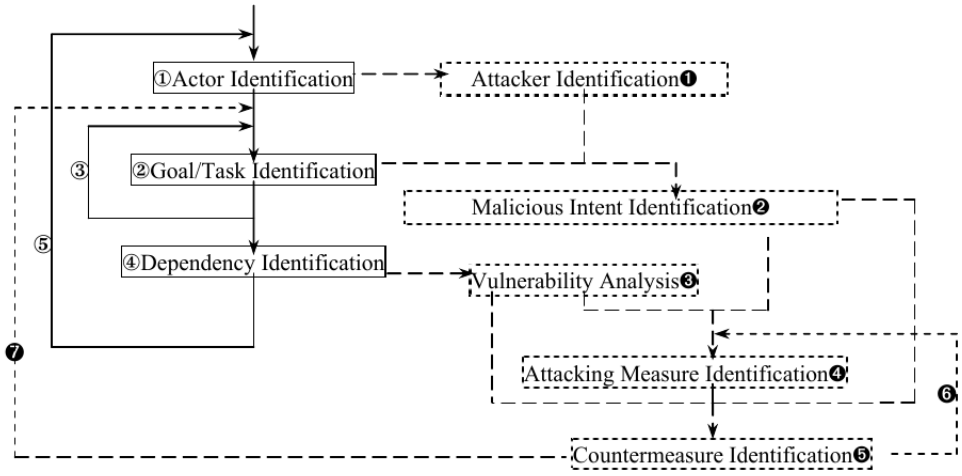
- Requirements elicitation
- Requirements analysis
  - Classification
  - Prioritization
  - Negotiation
- Requirements specification
- Requirements validation

Sample Elicitation Process: VisiOn



Gharib et al. Privacy Requirements: Findings and Lessons Learned in Developing a Privacy Platform. Requirements Engineering Conference (RE), pages 256–265, 2016

Sample Elicitation Process: i*



Liu et al. Security and privacy requirements analysis within a social setting. Requirements Engineering Conference (RE), pages 151–161, 2003

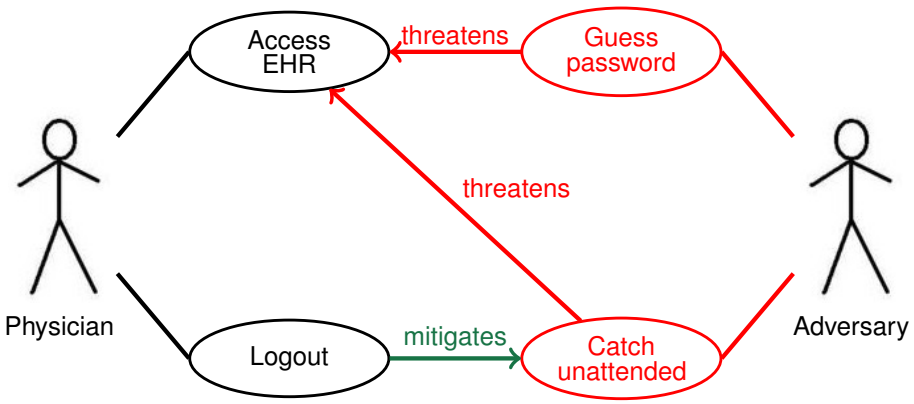
Attacker Analysis

- Assumption: "All actors are guilty until proven innocent"
- Any actor (roles, positions, agents) can be a potential attacker
  - To the system
  - To other actors
- For example, in what ways a physician can misuse the system?
- What benefit will the physician gain from an information disclosure?

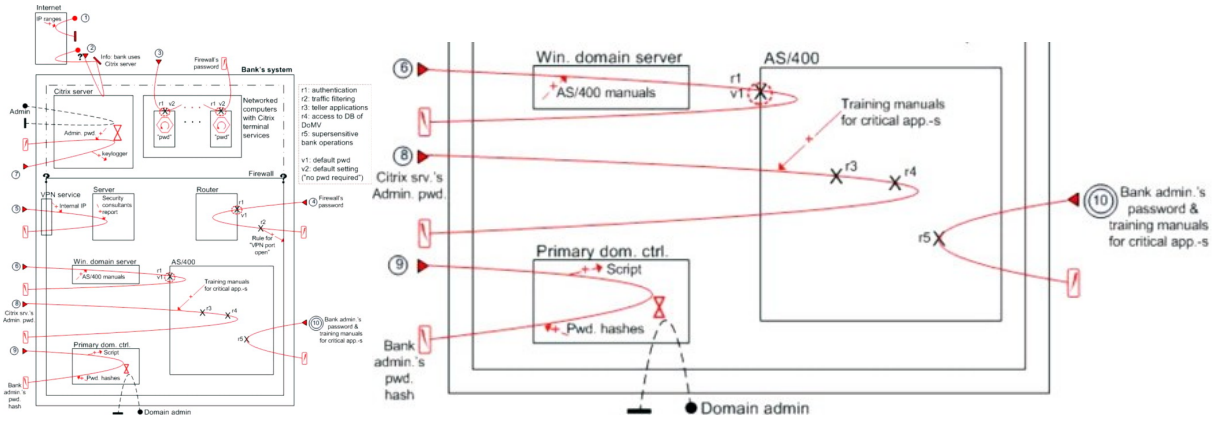
Threat Modeling

- Enumerate potential ways that your system might be attacked
- Typically include only attack nodes
- But, defense nodes can also be included that mitigate such attacks

Misuse Cases

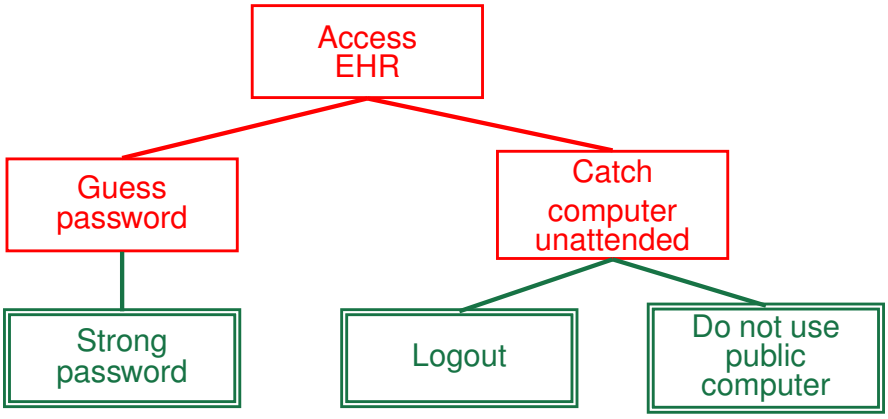


Misuse Case Maps

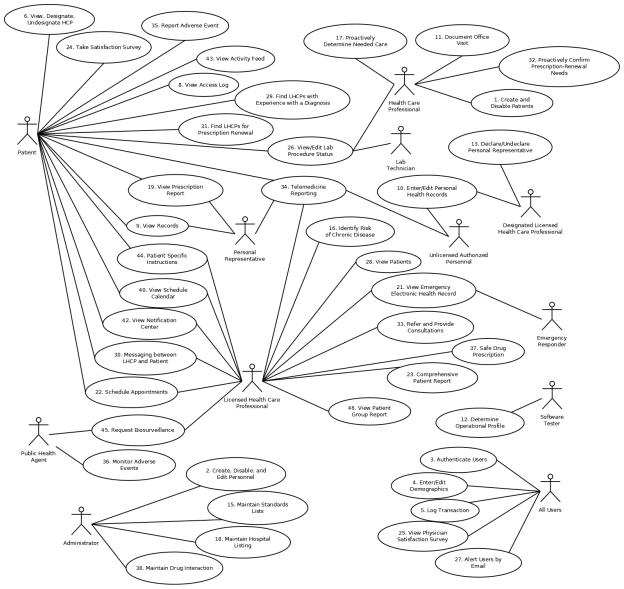


Karpati et al. Investigating security threats in architectural context: Experimental evaluations of misuse case maps. Journal of Systems and Software, 104(C):90–111, 2015

Attack/Defense Trees



Exercise: Healthcare Threat Model



http://agile.csc.ncsu.edu/iTrust/wiki/doku.php?id=requirements

Exercise: Internet of Things Threat Model

http://www.devolo.com/en/Products/devolo-Home-Control-Key-Fob-Switch/

Eddy: A Formal Language for Privacy Requirements

Eddy, a formal language for specifying and analyzing data flow specifications for conflicting privacy requirements

Travis D. Breaux · Hanan Hibshi · Ashwini Rao

Received: 25 August 2013 / Accepted: 26 October 2013 / Published online: 18 December 2013

Abstract Increasingly, companies use multi-source data to operate new information systems, such as social networking, e-commerce, and location-based services. These systems leverage complex, multi-stakeholder data supply chains in which each stakeholder (e.g., users, developers, companies, and government) must manage privacy and security requirements that cover their practices. US regulator and European regulator expect companies to ensure consistency between their privacy policies and their data practices, including restrictions on what data may be collected, how it may be used, to whom it may be transferred, and for what purposes. To help developers check consistency, we identified a strict subset of commonly found privacy requirements and we developed a methodology to map these requirements from natural language text to a formal language in description logic, called Eddy. Using this language, developers can detect conflicting privacy requirements within a policy and enable the tracing of data flows within these policies. We derived our methodology from an exploratory case study of the Facebook platform policy and an extended case study using privacy policies from Zynga and AOL Advertising. In this paper, we report results from multiple analysts in a literal replication study, which includes a refined methodology and set of heuristics that we used to extract privacy requirements from policy texts. In addition to providing the method, we report results from performing automated conflict detection within the Facebook, Zynga, and AOL privacy specifications, and results from a computer simulation that demonstrates the scalability of our formal language toolset to specifications of reasonable size.

Keywords Privacy · Requirements · Standardization · Description logic · Formal analysis

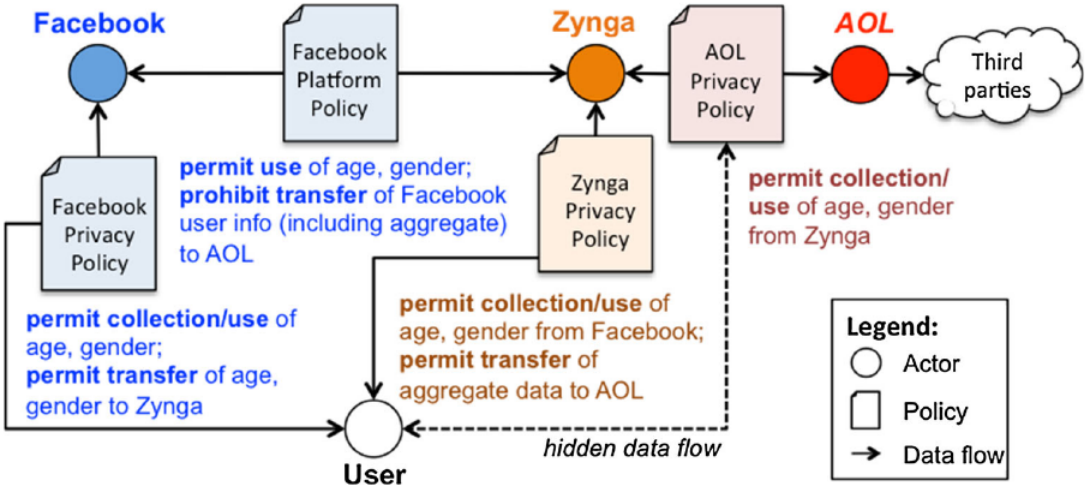
1 Introduction

Emerging Web and mobile information systems leverage user data that are collected from multiple sources without a clear understanding of data provenance or the privacy requirements that should follow these data. These systems are increasingly based on multi-tier platforms in which each "tier" may be owned and operated by a different party, such as cellular and wireless network providers, mobile and desktop operating system manufacturers, and mobile or Web application developers. In addition, user services developed on these tiers are abstracted into platforms to be extensible by other developers, such as Google Maps and the Facebook and LinkedIn social networking platforms. Application marketplaces, such as Amazon Appstore, Google Play, and iTunes, have also emerged to provide small

Breaux et al. Eddy, a Formal Language for Specifying and Analyzing Data Flow Specifications for Conflicting Privacy Requirements. Requirements Engineering, 19(3):281–307, 2014

Example: Facebook and Zynga

Data Flow between Parties

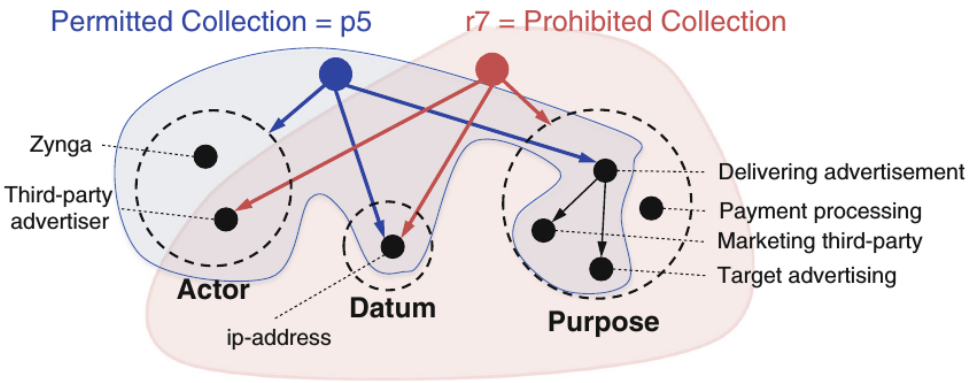


Objectives

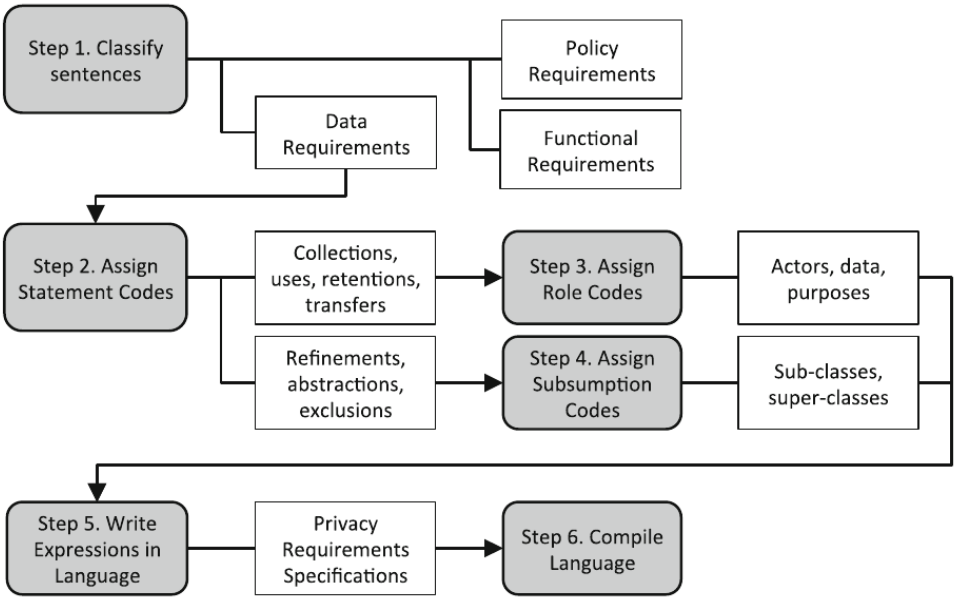
- Develop a privacy requirements specification
  - To align multi-party expectations
  - Across multi-tier applications
  - And, to formally check conflicts among requirements
- High-level design document to be used by
  - Software developers
  - Privacy law experts
  - End users

Conflicts

- permission(X) $\wedge$ prohibition(X) $\rightarrow$ conflict(X)



Methodology



Coded Policy

Step 3: Annotate policy text to identify action and role values

Modal phrase "will" indicates an assumed permission
Transfer keyword
Datum
Target
Purposes

We will provide your information to third party companies to perform services on our behalf, including payment processing, data analysis, e-mail delivery, hosting services, customer service and to assist us in our marketing efforts.

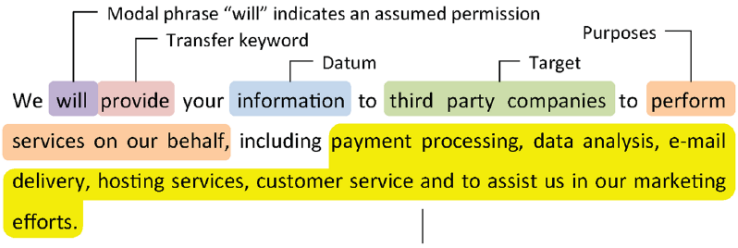
Step 4: Annotate policy text to identify other subsumption relations

Previously identified role value, in this case, a purpose
Refinement keyword

We will provide your information to third party companies to perform services on our behalf, including payment processing, data analysis, e-mail delivery, hosting services, customer service and to assist us in our marketing efforts.

List of refinements, or sub-categories of "perform services on our behalf"

Specification in Eddy Syntax



Step 5: Write expression in specification language (P = Permission)

SPEC HEADER
P performing-services > payment-processing, e-mail-delivery, hosting-services, customer-service, marketing
SPEC POLICY
P TRANSFER information TO third-party-companies FOR performing-services

Step 6: Compile language into Description Logic (OWL)

payment-processing $\sqsubseteq$ performing-services
e-mail-delivery $\sqsubseteq$ performing-services
...
Z-92 $\equiv$ TRANSFER $\sqcap$ $\exists$ hasObject.information $\sqcap$
$\exists$ hasTarget.third-party-companies $\sqcap$ $\exists$ hasPurpose.performing-services
Z-92 $\sqsubseteq$ Permission

Conflict Analysis: Between Facebook and Zynga

- PROHIBIT TRANSFER user-data
  FROM facebook TO ad-network
  FOR anything
- PERMIT TRANSFER aggregate-information, anonymous-information
  FROM anyone TO anyone

- PROHIBIT TRANSFER user-data
  FROM facebook TO third-party
  FOR merger, acquisition
- PERMIT TRANSFER information
  FOR merger, acquisition

Conflict Analysis: Within AOL

- PROHIBIT USE personally-identifiable-information
  FROM registration-environment
  FOR targeted-ads

- PERMIT COLLECT personally-identifiable-information
  FROM anyone
  FOR improving-targeted-ads