Computer-aided Privacy Analysis of Functional Requirements with ProPAn

Rene Meis

paluno - The Ruhr Institute for Software Technology
University of Duisburg-Essen, Germany

February 18, 2016

Introduction

FoMSESS '16

Rene Meis

Introduction

Running Example

Background

Problem Frames

ProPAn

Overview

Identification of Personal Data

Personal Data Flow Analysis

Generate Privacy Requirements

Conclusion

Research field: Privacy-aware requirements engineering

Research questions:

- How to derive privacy requirements from regulations and standards?
- What kinds of privacy requirements exist?
- How to identify privacy-relevant parts of the software?
- Which additional knowledge is needed for a privacy analysis?
- How to operationalize privacy requirements?
- How to test or verify that a software satisfies its privacy requirements?

Introduction

FoMSESS '16

Rene Meis

Introduction

Running Example

Background

Problem Frames

ProPAn

Overview

Identification of Personal Data

Personal Data Flow Analysis

Generate Privacy Requirements

Conclusion

Research field: Privacy-aware requirements engineering

Research questions:

- How to derive privacy requirements from regulations and standards?
- What kinds of privacy requirements exist?
- How to identify privacy-relevant parts of the software?
- Which additional knowledge is needed for a privacy analysis?
- How to operationalize privacy requirements?
- How to test or verify that a software satisfies its privacy requirements?

Problem-based **P**rivacy **An**alysis (ProPAn)

FoMSESS '16
Rene Meis
Introduction
Running Example
Background
Problem Frames
ProPAn
Overview
Identification of Personal Data
Personal Data Flow Analysis
Generate Privacy Requirements
Conclusion

The ProPAn method provides a tool supported privacy analysis of a set of functional requirements that are represented as problem diagrams [Jackson, 2001].

ProPAn aims at the identification of:

- Privacy-relevant domain knowledge
  [Meis, 2014, Beckers et al., 2014a]
- Potential privacy concerns
  [Beckers et al., 2014b]
- Personal data processed by the system-to-be
  [Meis and Heisel, 2015]
- Flow of personal data in the system
  [Meis and Heisel, 2015]
- Privacy requirements
  [Meis et al., 2015]

eHealth Scenario

Provided by industrial partners of the EU project NESSoS

FoMSESS '16
Rene Meis
Introduction
Running Example
Background
Problem Frames
ProPAn
Overview
Identification of Personal Data
Personal Data Flow Analysis
Generate Privacy Requirements
Conclusion

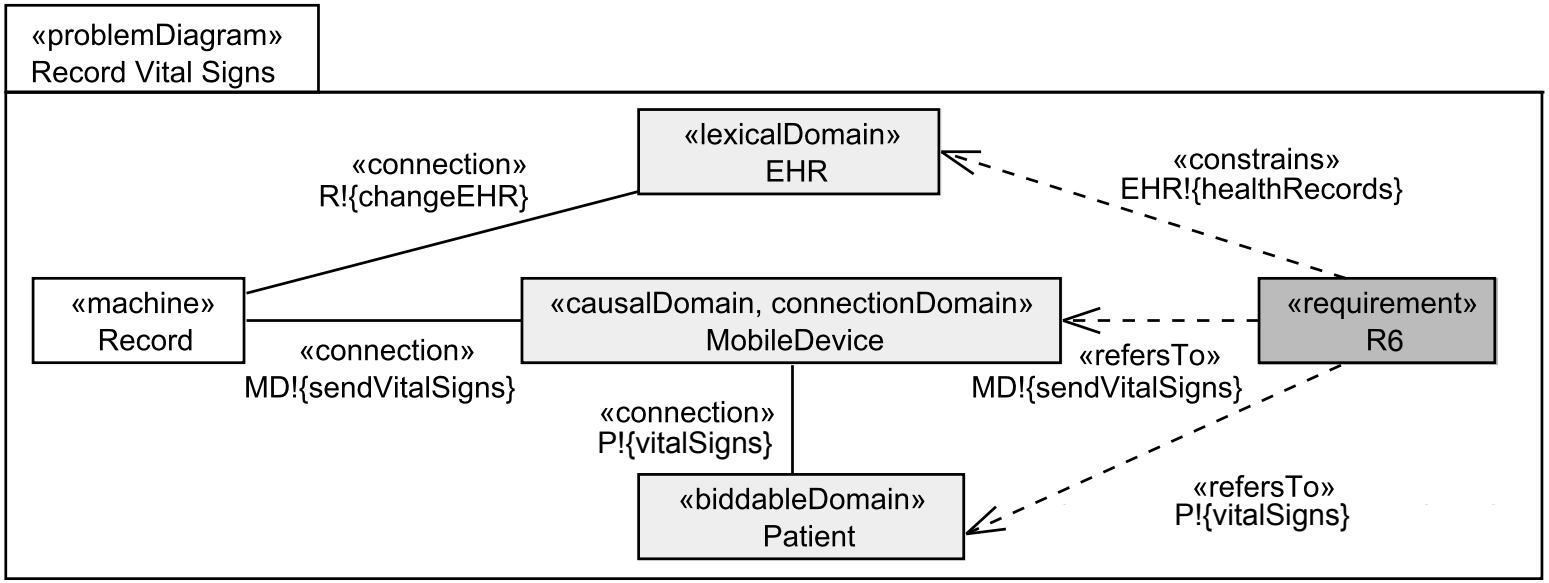
Concerns the management and usage of Electronic Health Records (EHRs).

- R1 Doctors shall be able to create and modify EHRs.
- R2 Doctors shall be able to browse EHRs.
- R3 The accounting of patients shall be performed using an insurance application based on the EHRs. If necessary invoices shall be created.
- R4 The billing of patients shall be performed using a financial application based on the invoices.
- R5 Appointments, instructions and alarms shall be sent to the mobile devices of patients based on the EHRs.
- R6 Vital signs of patients sent via their mobile devices shall be recorded in the EHRs.

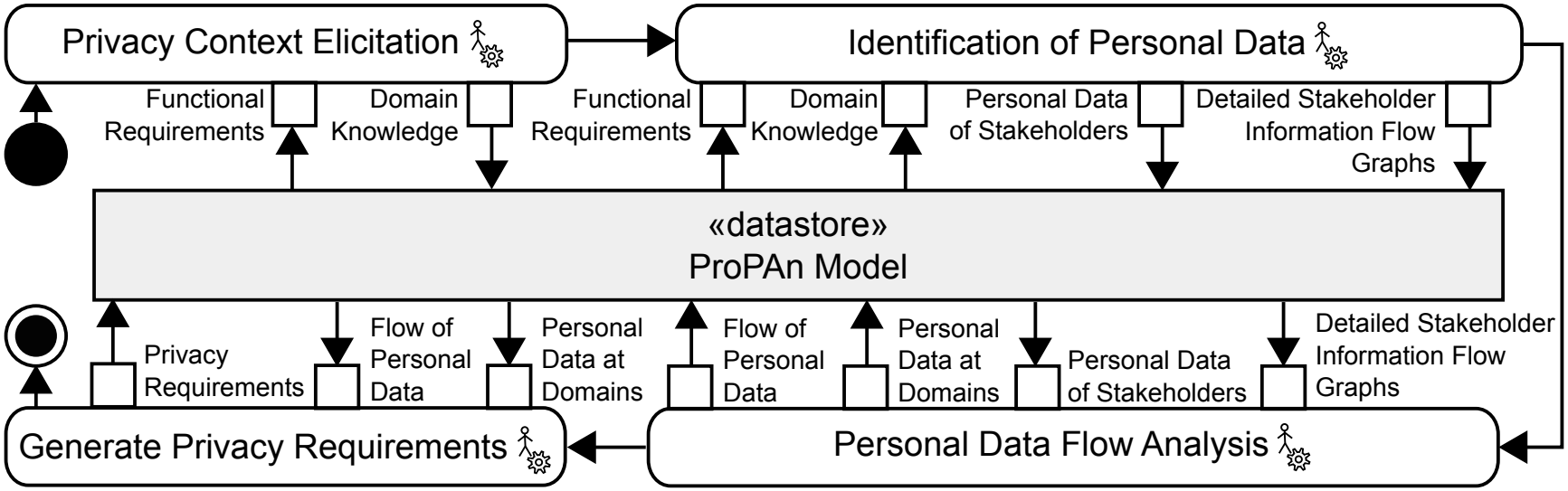
Problem Frames Approach
[Jackson, 2001]

FoMSESS '16
Rene Meis
Introduction
Running Example
Background
Problem Frames
ProPAn
Overview
Identification of Personal Data
Personal Data Flow Analysis
Generate Privacy Requirements
Conclusion

- System consists of the machine and its environment. The environment is structured into domains.
- Requirements are represented in problem diagrams.
- Requirements refer to and constrain phenomena of domains.
- Phenomena can be causal or symbolic.



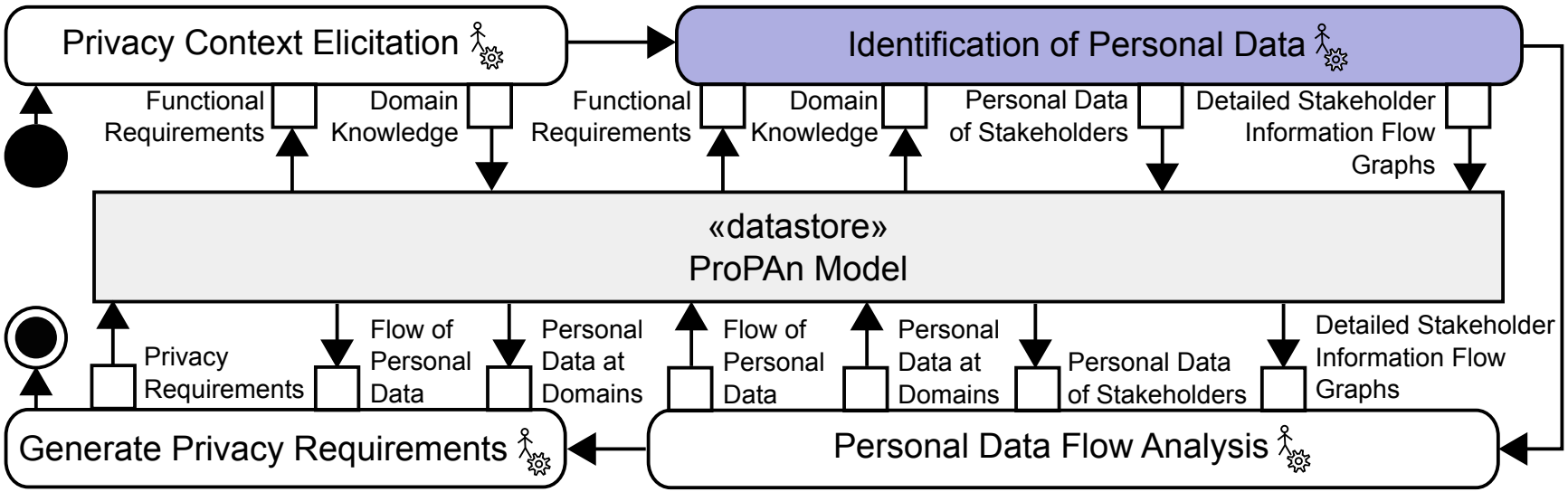
Overview of the ProPAn Method



- All steps make use of one central UML model.
- All steps are partly automated by the ProPAn tool[1]

[1] http://www.uml4pf.org/ext-propan/index.html

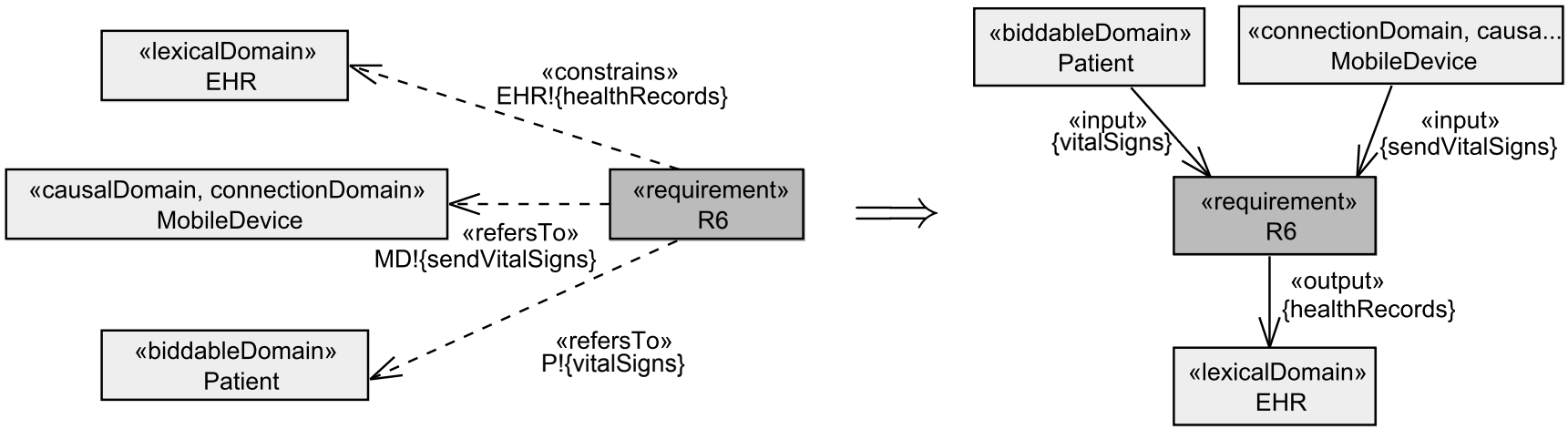
Overview of our Method
Identification of Personal Data



Graph Generation
Principle idea

FoMSESS '16
Rene Meis
Introduction
Running Example
Background
Problem Frames
ProPAn
Overview
Identification of Personal Data
Personal Data Flow Analysis
Generate Privacy Requirements
Conclusion

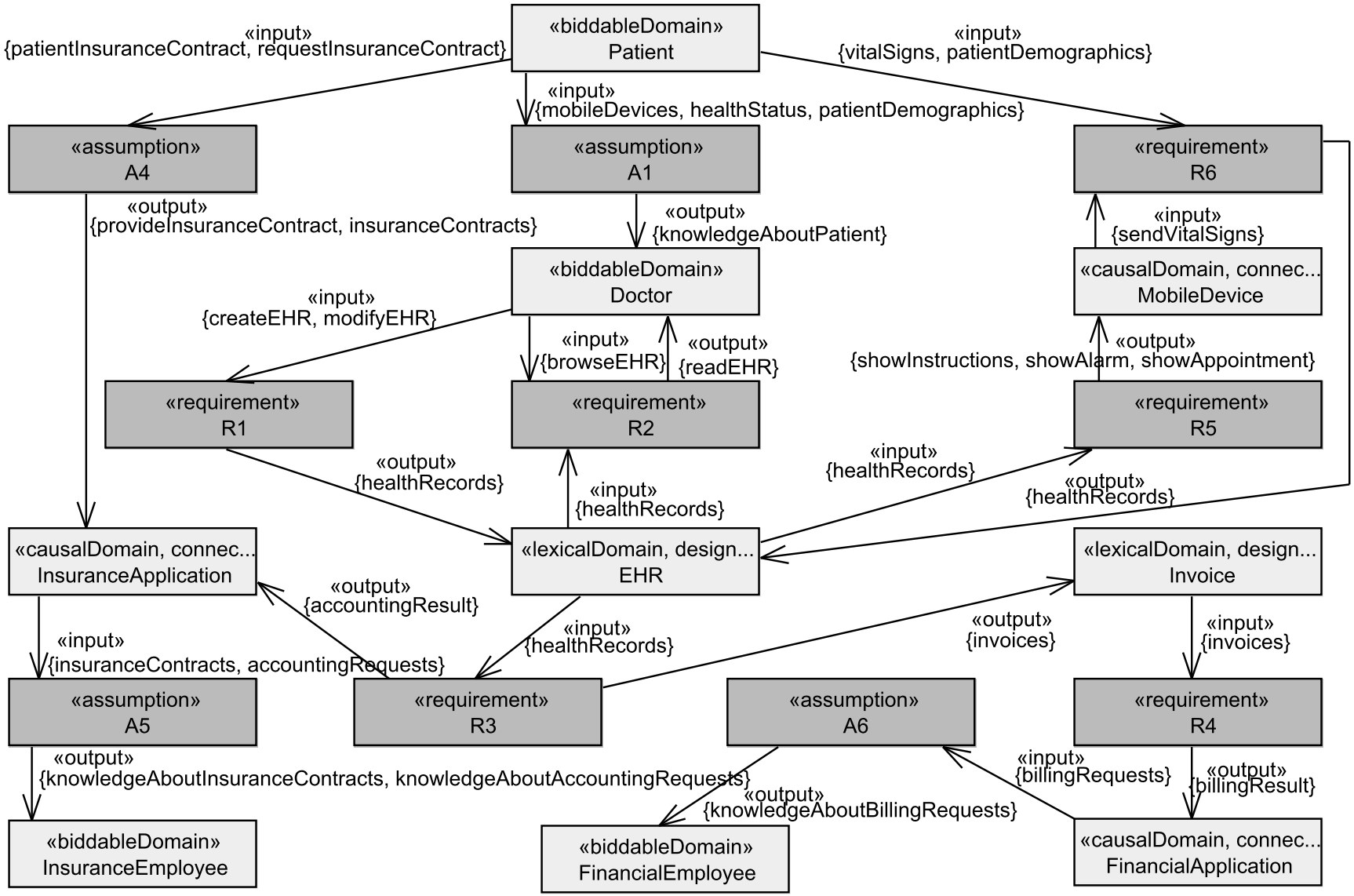
Statements (requirements, facts, and assumptions) imply possible information flows from referred to to constrained domains.



Graph Generation

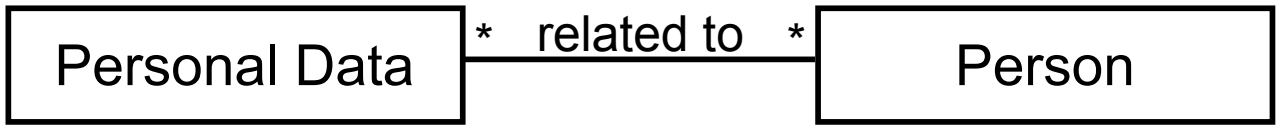
Aggregation of all Information Flows starting at Patient

FoMSESS '16
Rene Meis
Introduction
Running Example
Background
Problem Frames
ProPAn
Overview
Identification of Personal Data
Personal Data Flow Analysis
Generate Privacy Requirements
Conclusion



Relations that are elicited and documented

FoMSESS '16
Rene Meis
Introduction
Running Example
Background
Problem Frames
ProPAn
Overview
Identification of Personal Data
Personal Data Flow Analysis
Generate Privacy Requirements
Conclusion

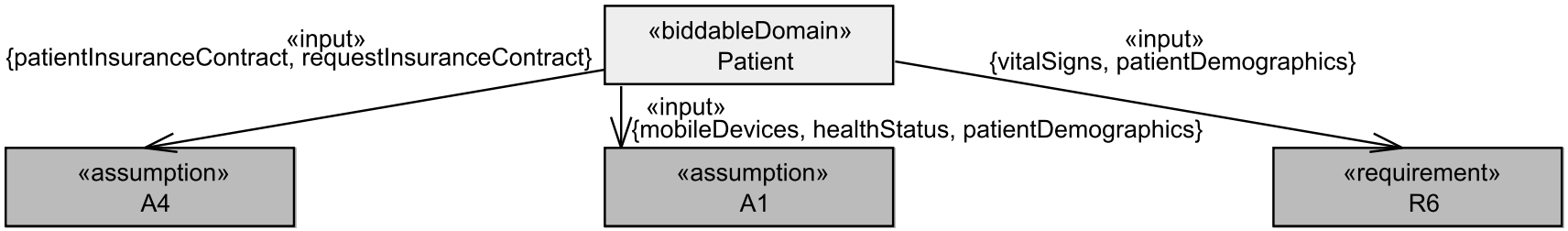


- Symbolic phenomena are used to represent personal data
- Biddable domains are persons

Identification of Personal Data

FoMSESS '16
Rene Meis
Introduction
Running Example
Background
Problem Frames
ProPAn
Overview
Identification of Personal Data
Personal Data Flow Analysis
Generate Privacy Requirements
Conclusion

Candidates for personal data can be derived from the Detailed Stakeholder Informaion Flow Graph (DSIFG).



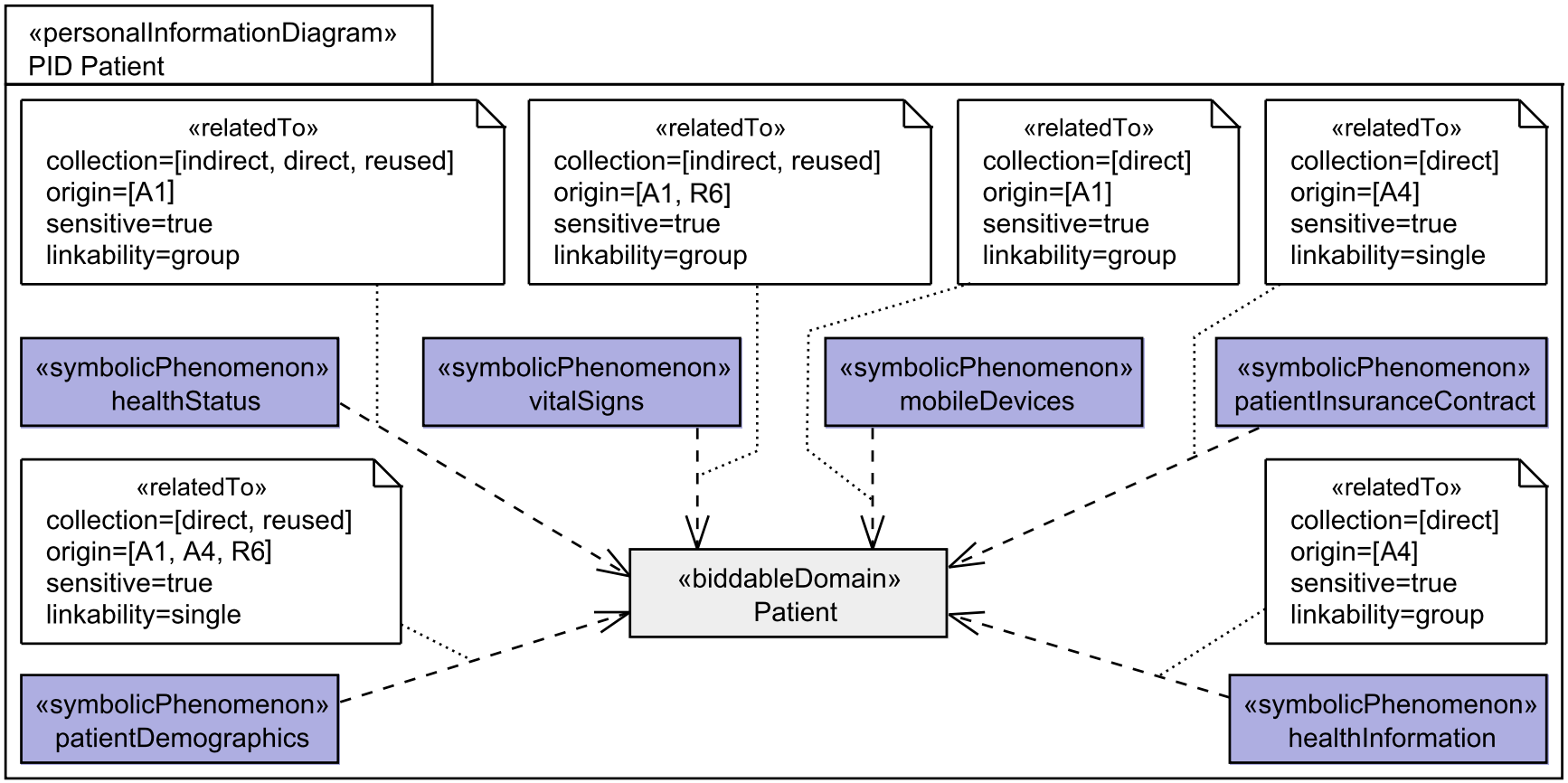
We distinguish two cases:

Symbolic phenomena can directly represent personal data.

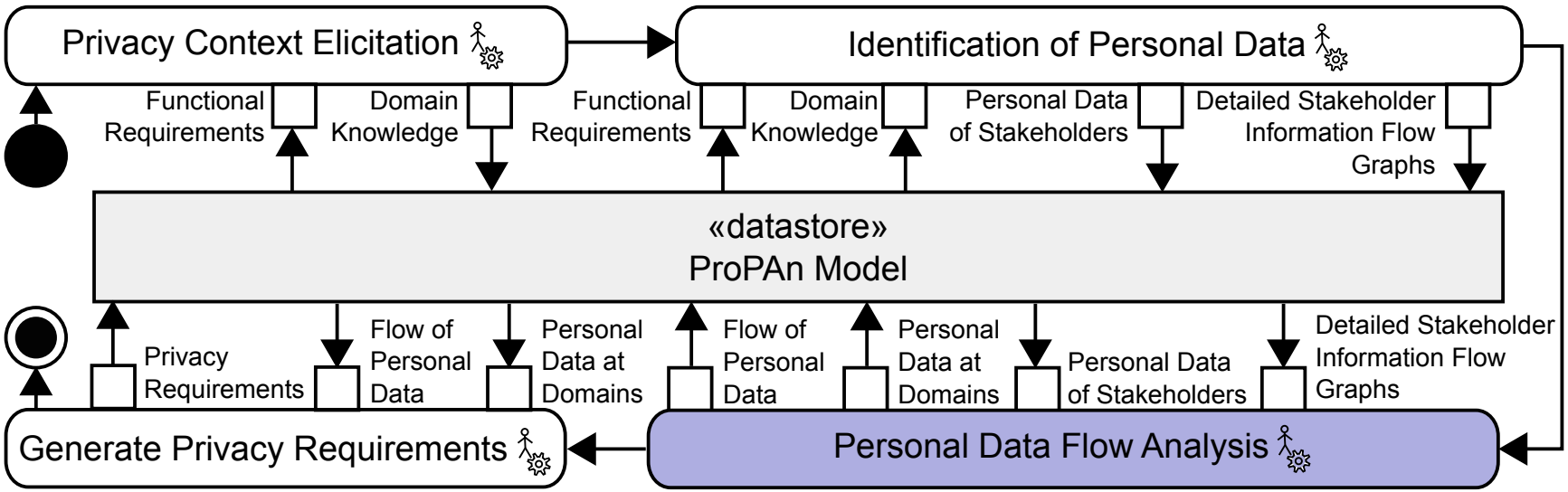
Causal phenomena may transmit/contain personal data. This transmitted/contained personal data is elicited.

Identification of Personal Data

Identified personal data is modelled in personal information diagrams.



Overview of our Method
Identification of Personal Data

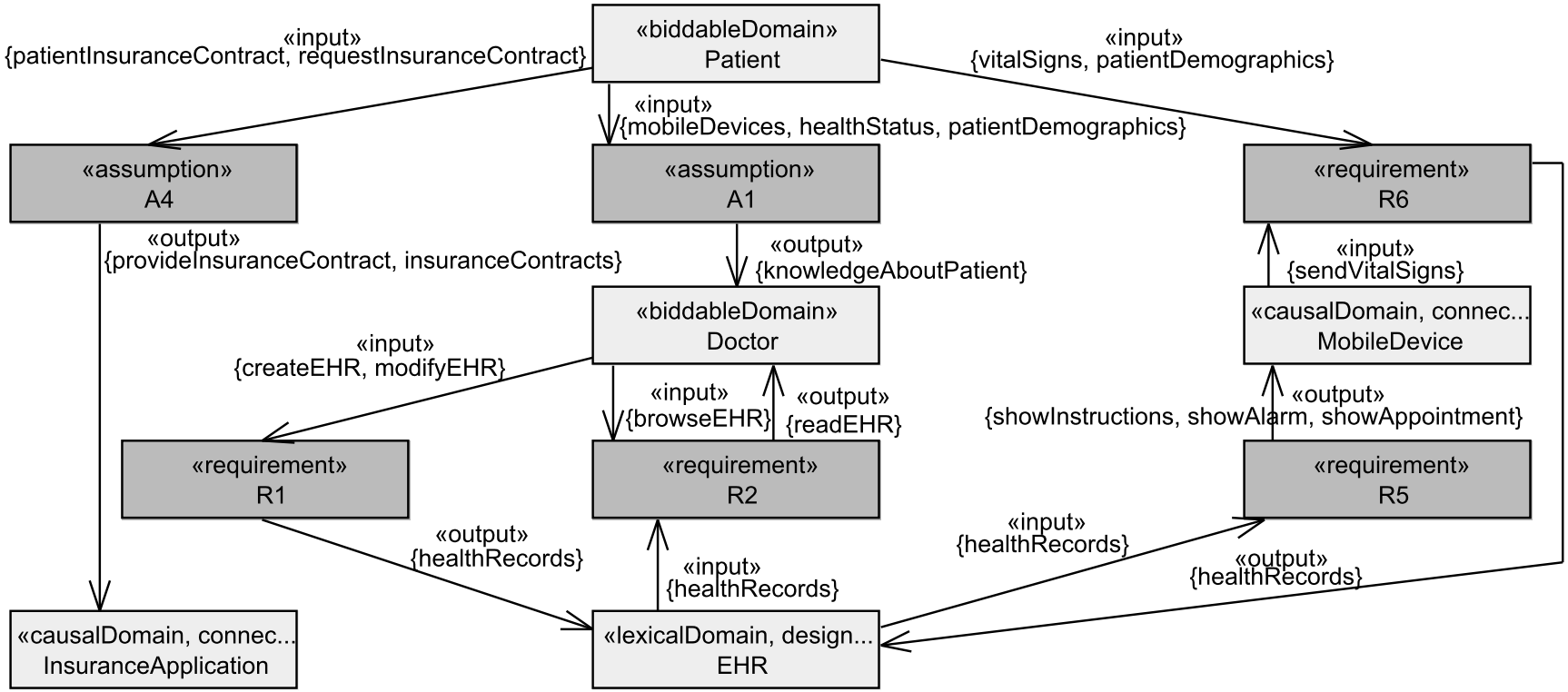


Personal Data Flow Analysis

Principle idea of a step in the analysis

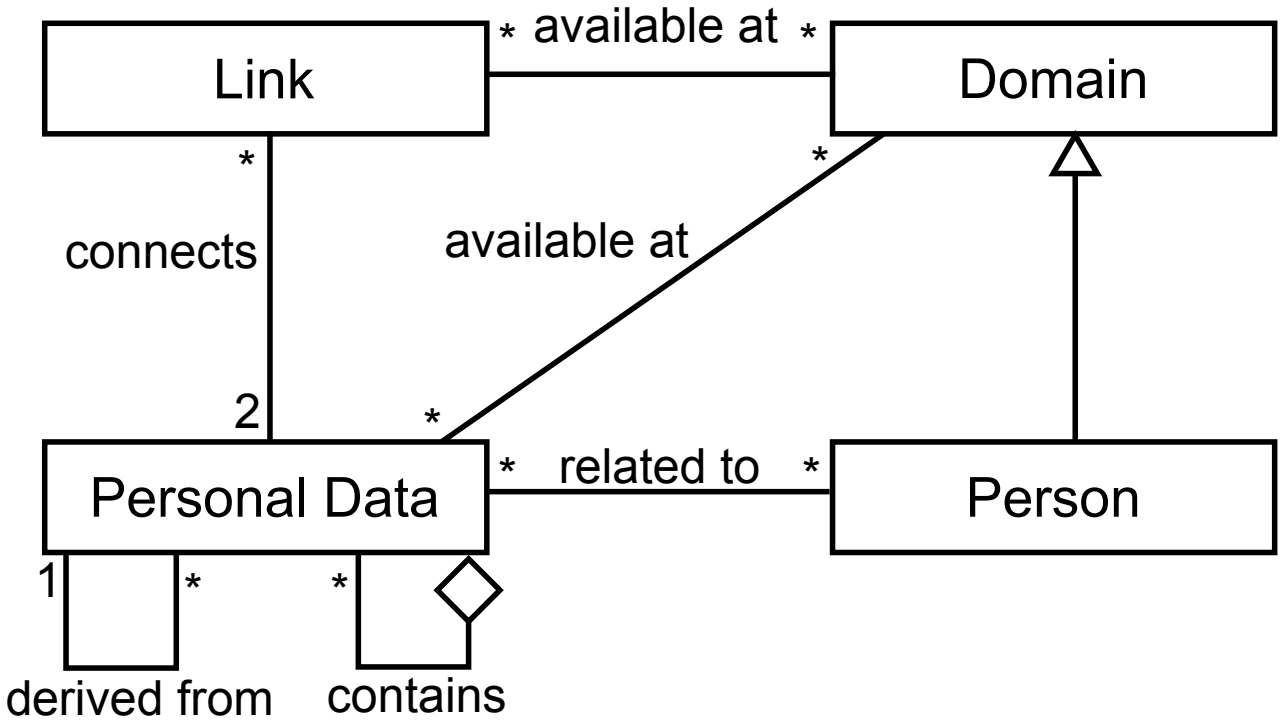
FoMSESS '16
Rene Meis
Introduction
Running Example
Background
Problem Frames
ProPAn
Overview
Identification of Personal Data
Personal Data Flow Analysis
Generate Privacy Requirements
Conclusion

1. Select a statement in the DSIFG that has an input domain at which personal data is available.
2. Decide which of the available personal data flows to the output domains of the statement.
3. Repeat until all statements have been considered



Relations that are elicited and documented

FoMSESS '16
Rene Meis
Introduction
Running Example
Background
Problem Frames
ProPAn
Overview
Identification of Personal Data
Personal Data Flow Analysis
Generate Privacy Requirements
Conclusion



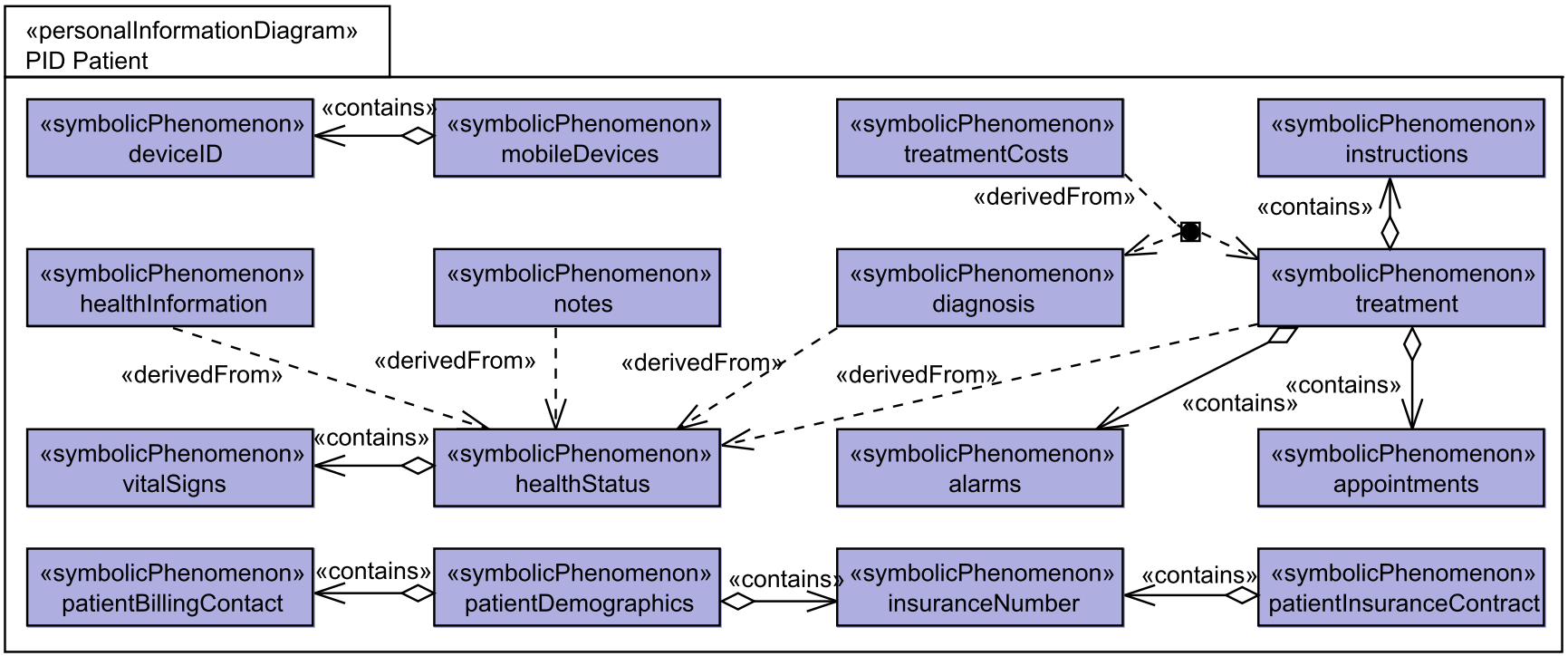
- Symbolic phenomena are used to represent personal data
- Biddable domains are persons

Personal Data Flow Analysis I

Result of the Analysis

FoMSESS '16
Rene Meis
Introduction
Running Example
Background
Problem Frames
ProPAn
Overview
Identification of Personal Data
Personal Data Flow Analysis
Generate Privacy Requirements
Conclusion

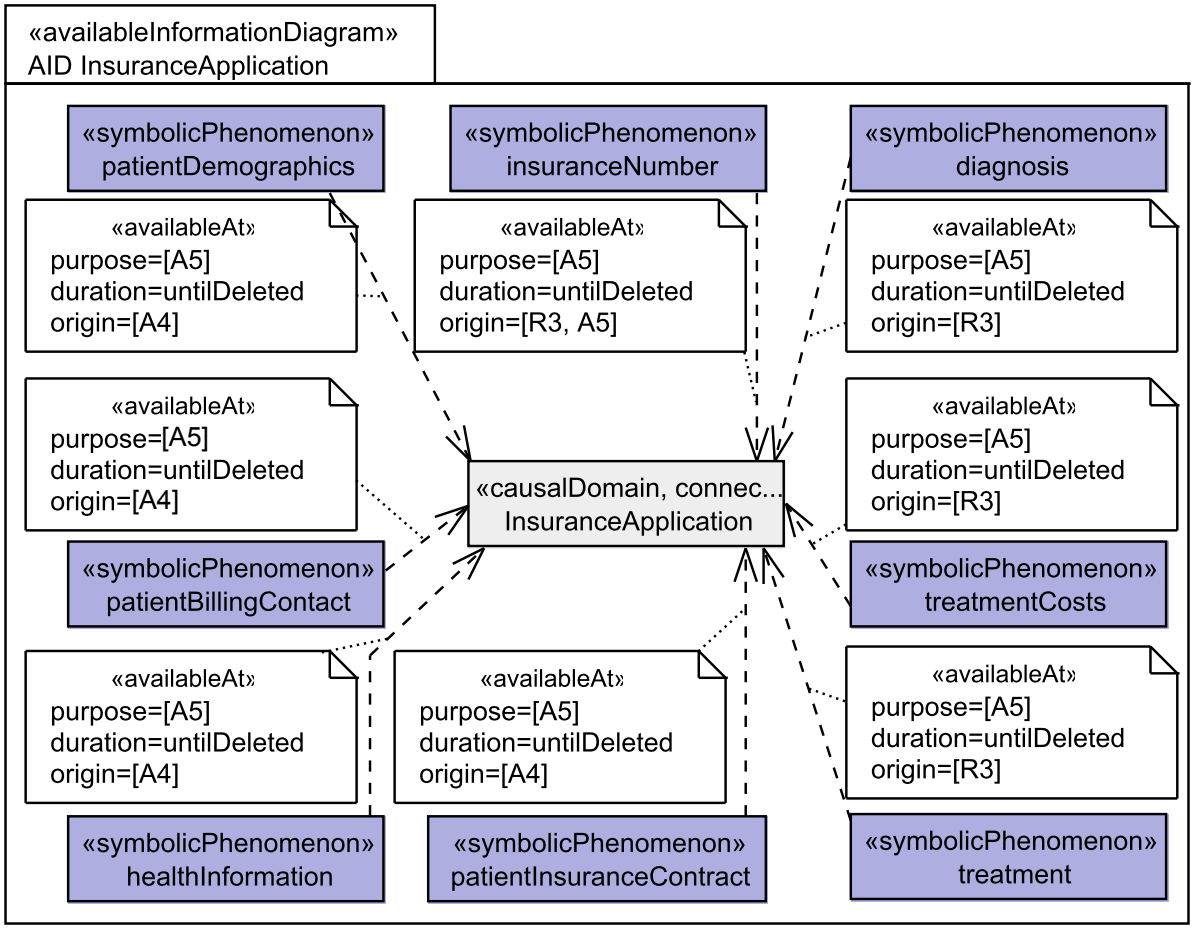
View on the resulting personal information diagram for the Patient.



Personal Data Flow Analysis II

Result of the Analysis

View on the resulting available information diagram for the Insurance Application.

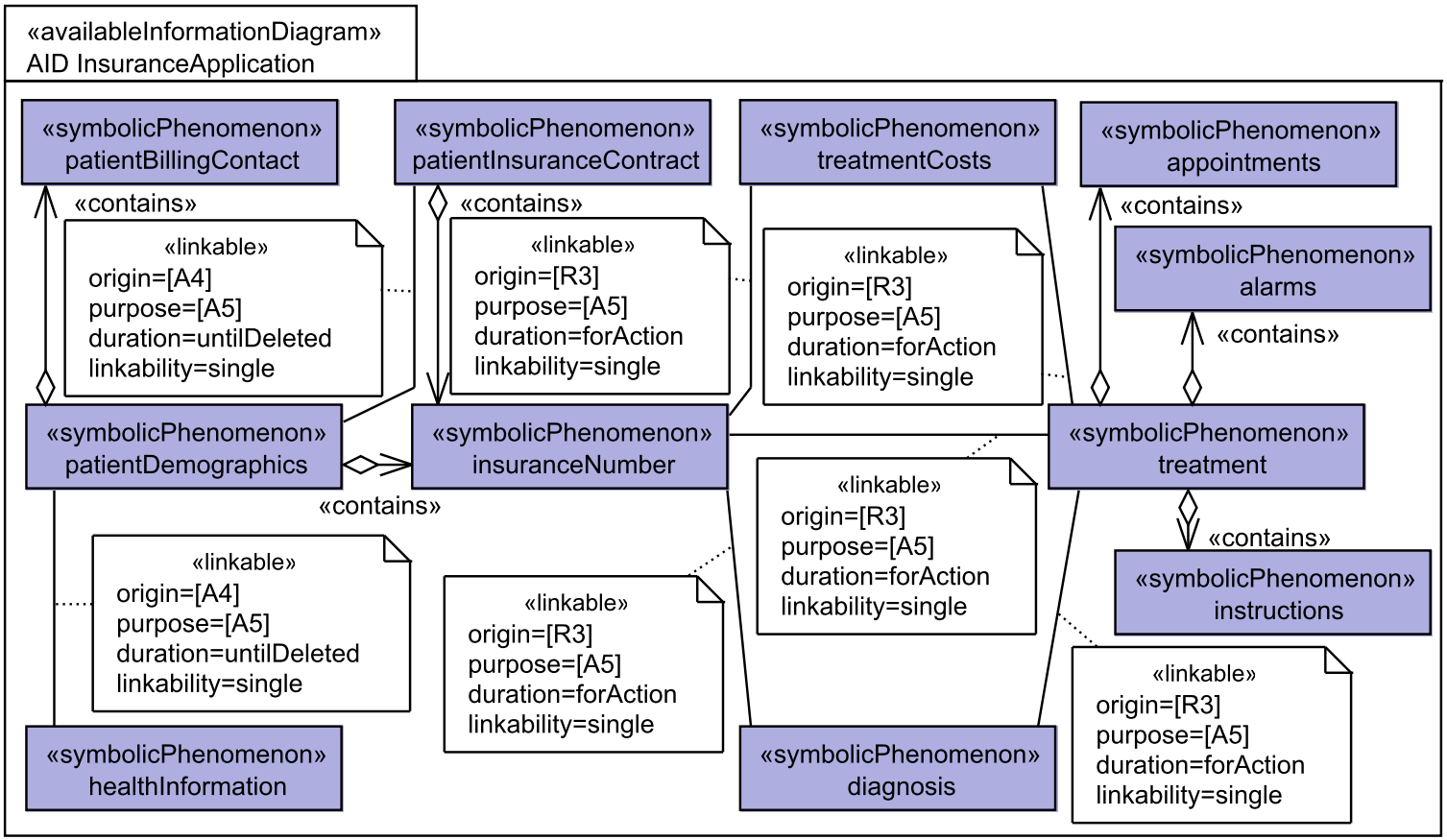


Personal Data Flow Analysis III

Result of the Analysis

FoMSESS '16
Rene Meis
Introduction
Running Example
Background
Problem Frames
ProPAn
Overview
Identification of Personal Data
Personal Data Flow Analysis
Generate Privacy Requirements
Conclusion

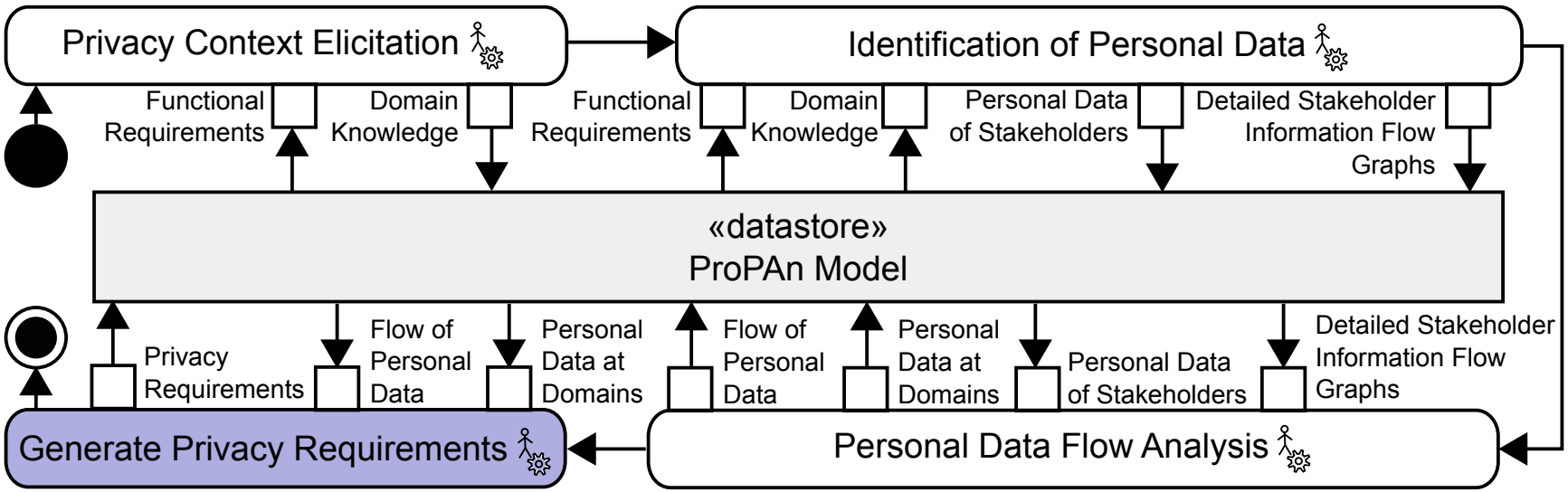
View on the resulting available information diagram for the Insurance Application.



Overview of our Method

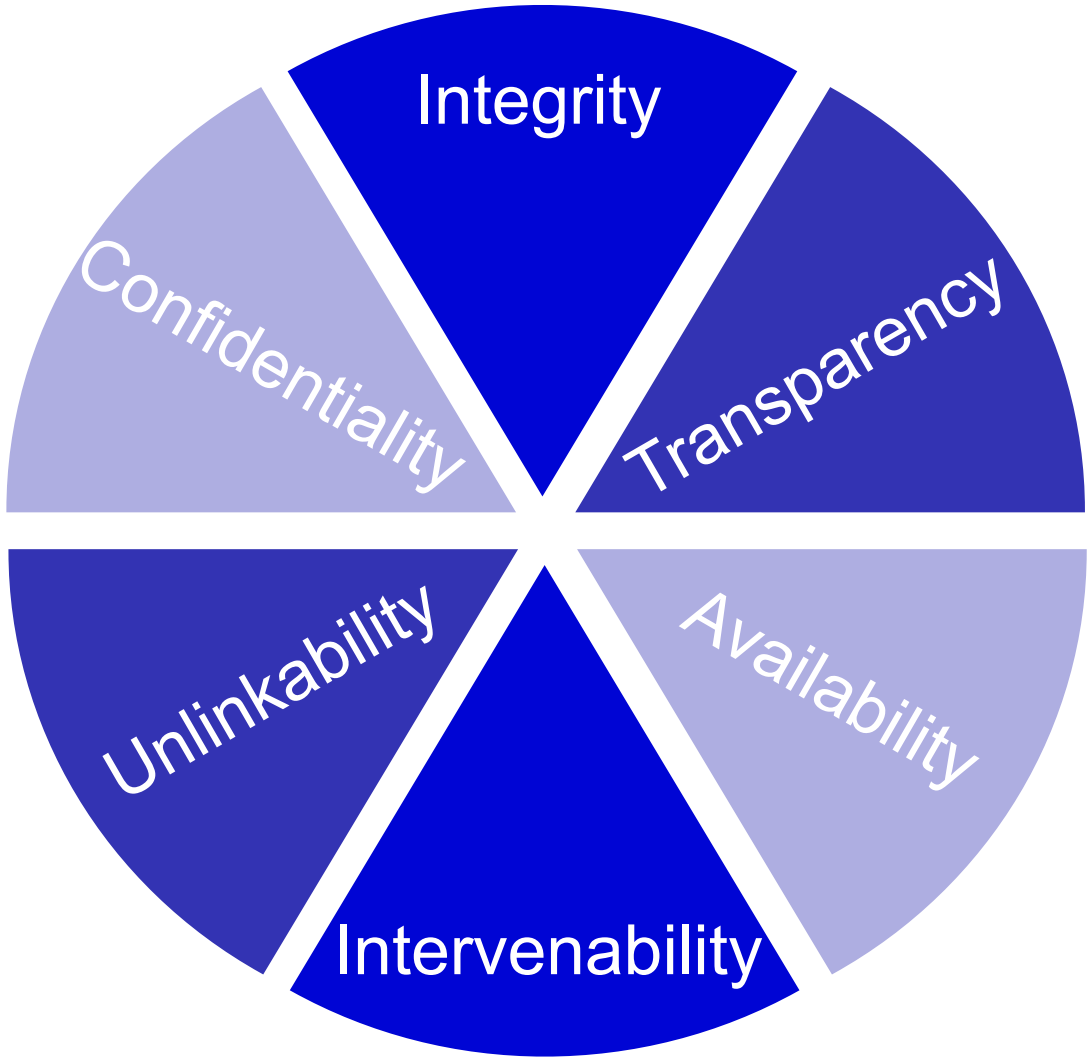
Generate Privacy Requirements

FoMSESS '16
Rene Meis
Introduction
Running Example
Background
Problem Frames
ProPAn
Overview
Identification of Personal Data
Personal Data Flow Analysis
Generate Privacy Requirements
Conclusion



Privacy Protection Goals [Hansen et al., 2015]

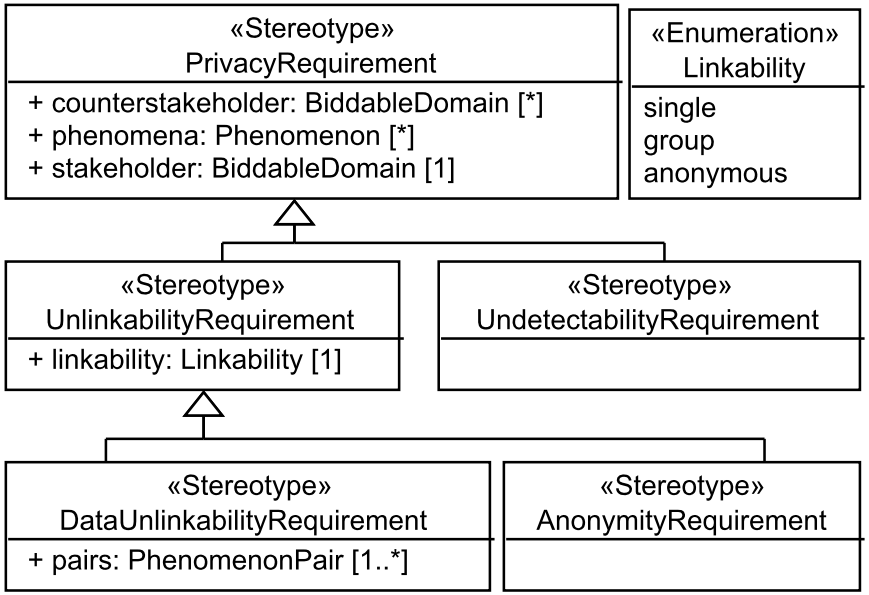
FoMSESS '16
Rene Meis
Introduction
Running Example
Background
Problem Frames
ProPAn
Overview
Identification of Personal Data
Personal Data Flow Analysis
Generate Privacy Requirements
Conclusion

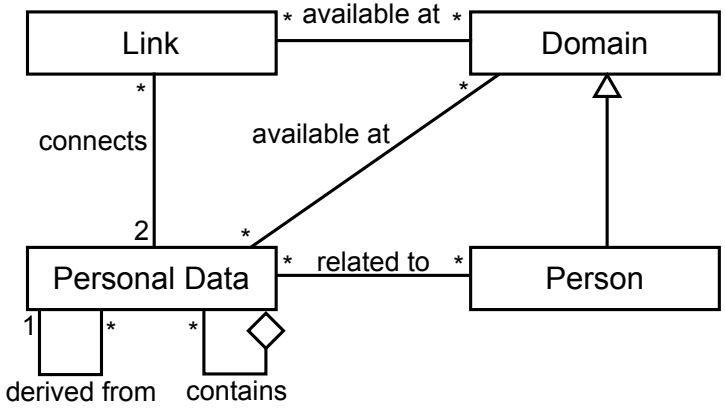


Generation of Unlinkability Requirements

Undetectability

FoMSESS '16
Rene Meis
Introduction
Running Example
Background
Problem Frames
ProPAn
Overview
Identification of Personal Data
Personal Data Flow Analysis
Generate Privacy Requirements
Conclusion





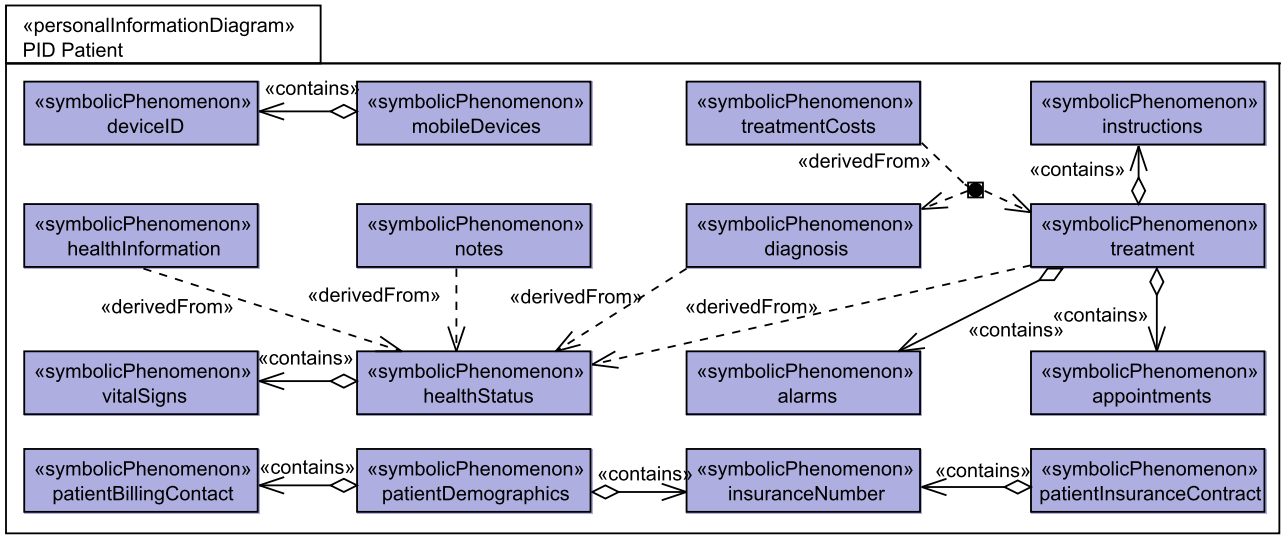
Undetectability:

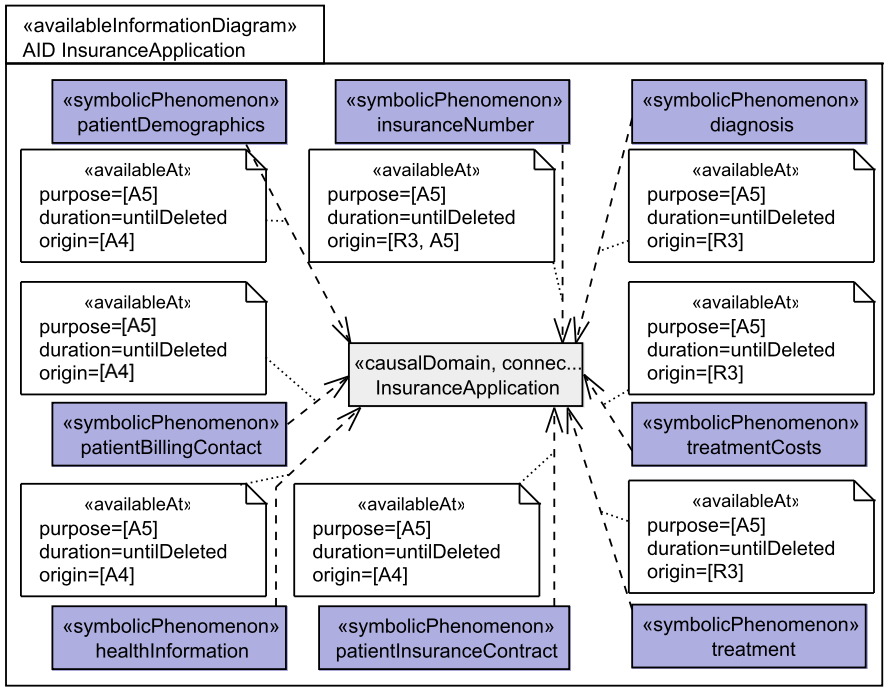
The counterstakeholders shall not be able to sufficiently distinguish whether the personal information phenomena of the stakeholder exists or not.

Generation of Unlinkability Requirements

Example of an Undetectability Requirement

FoMSESS '16
Rene Meis
Introduction
Running Example
Background
Problem Frames
ProPAn
Overview
Identification of Personal Data
Personal Data Flow Analysis
Generate Privacy Requirements
Conclusion





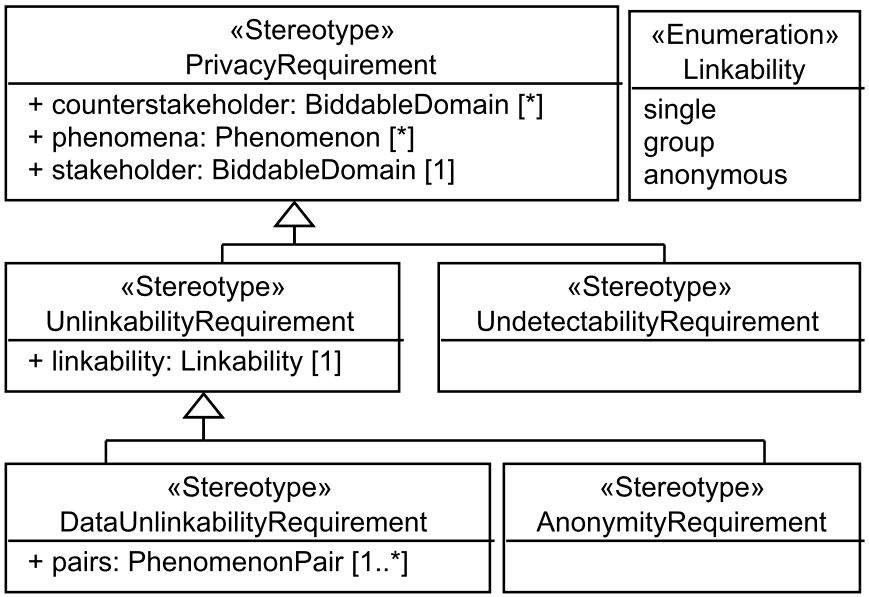
«UndetectabilityRequirement»
Undetectabilty_Patient_InsuranceEmployee

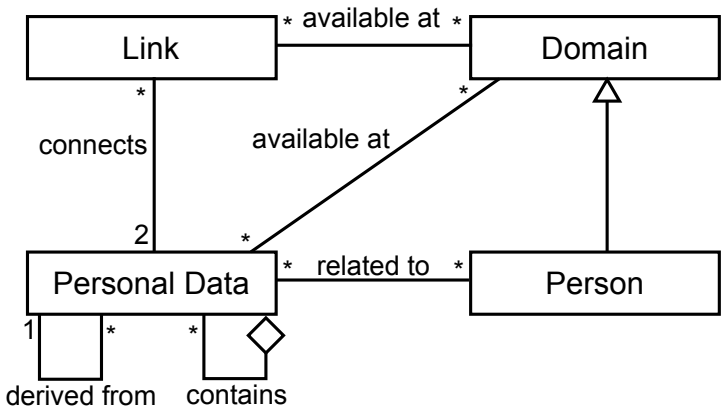
stakeholder = Patient
counterstakeholder = [Insurance Employee]
phenomena = [healthStatus, mobileDevices, deviceID, vitalSigns, notes]

Generation of Unlinkability Requirements

Data Unlinkability

FoMSESS '16
Rene Meis
Introduction
Running Example
Background
Problem Frames
ProPAn
Overview
Identification of Personal Data
Personal Data Flow Analysis
Generate Privacy Requirements
Conclusion





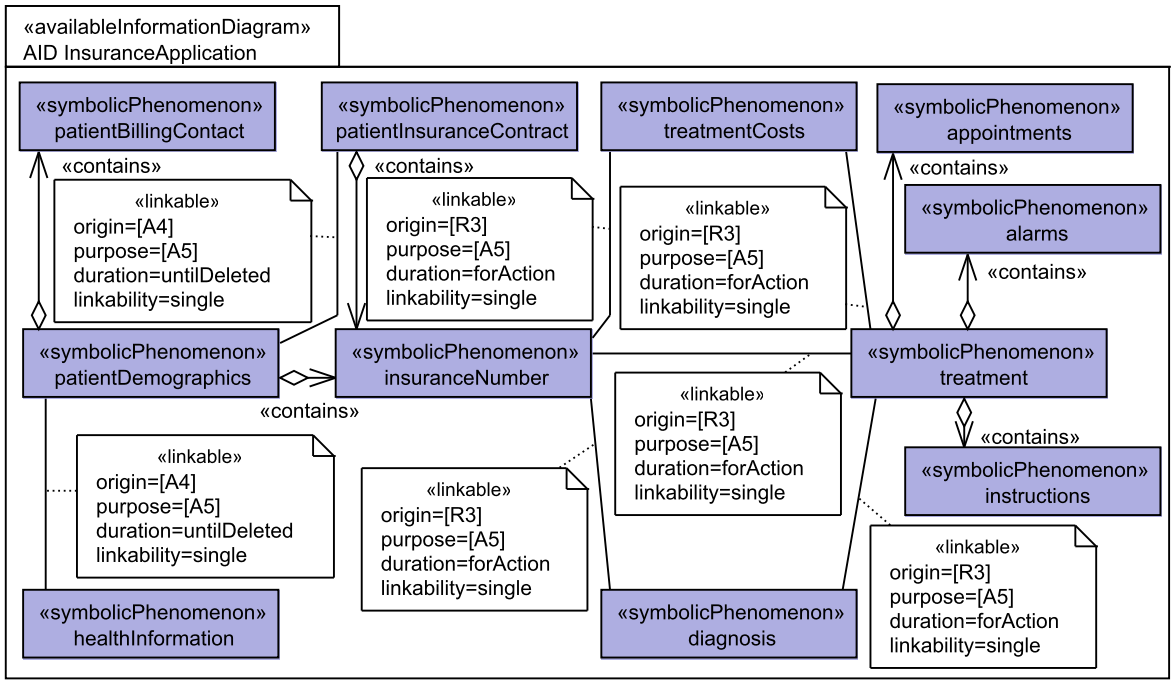
Data Unlinkability:

For each pair of personal information pairs of the stakeholder, the counterstakeholders shall at most be able to link instances of the two elements of the pair to each other with linkability linkability.

Generation of Unlinkability Requirements

Example of a Data Unlinkability Requirement

FoMSESS '16
Rene Meis
Introduction
Running Example
Background
Problem Frames
ProPAn
Overview
Identification of Personal Data
Personal Data Flow Analysis
Generate Privacy Requirements
Conclusion



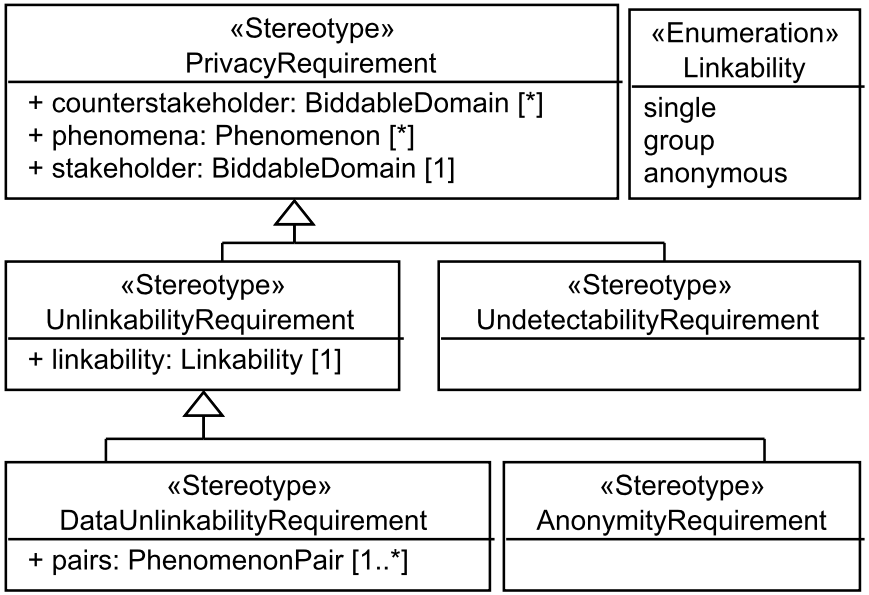
«DataUnlinkabilityRequirement»
DataUnlinkability_Patient_InsuranceEmployee

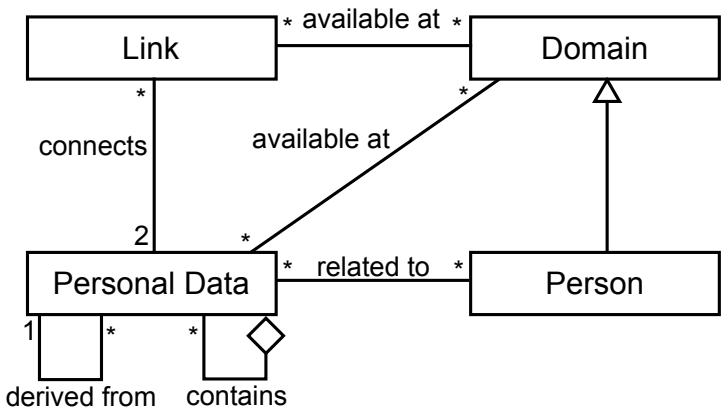
stakeholder = Patient
counterstakeholder = [Insurance Employee]
pairs = [(treatment, diagnosis), ...,
(patientBillingContact, healthInformation)]
linkability = single

Generation of Unlinkability Requirements

Anonymity

FoMSESS '16
Rene Meis
Introduction
Running Example
Background
Problem Frames
ProPAn
Overview
Identification of Personal Data
Personal Data Flow Analysis
Generate Privacy Requirements
Conclusion





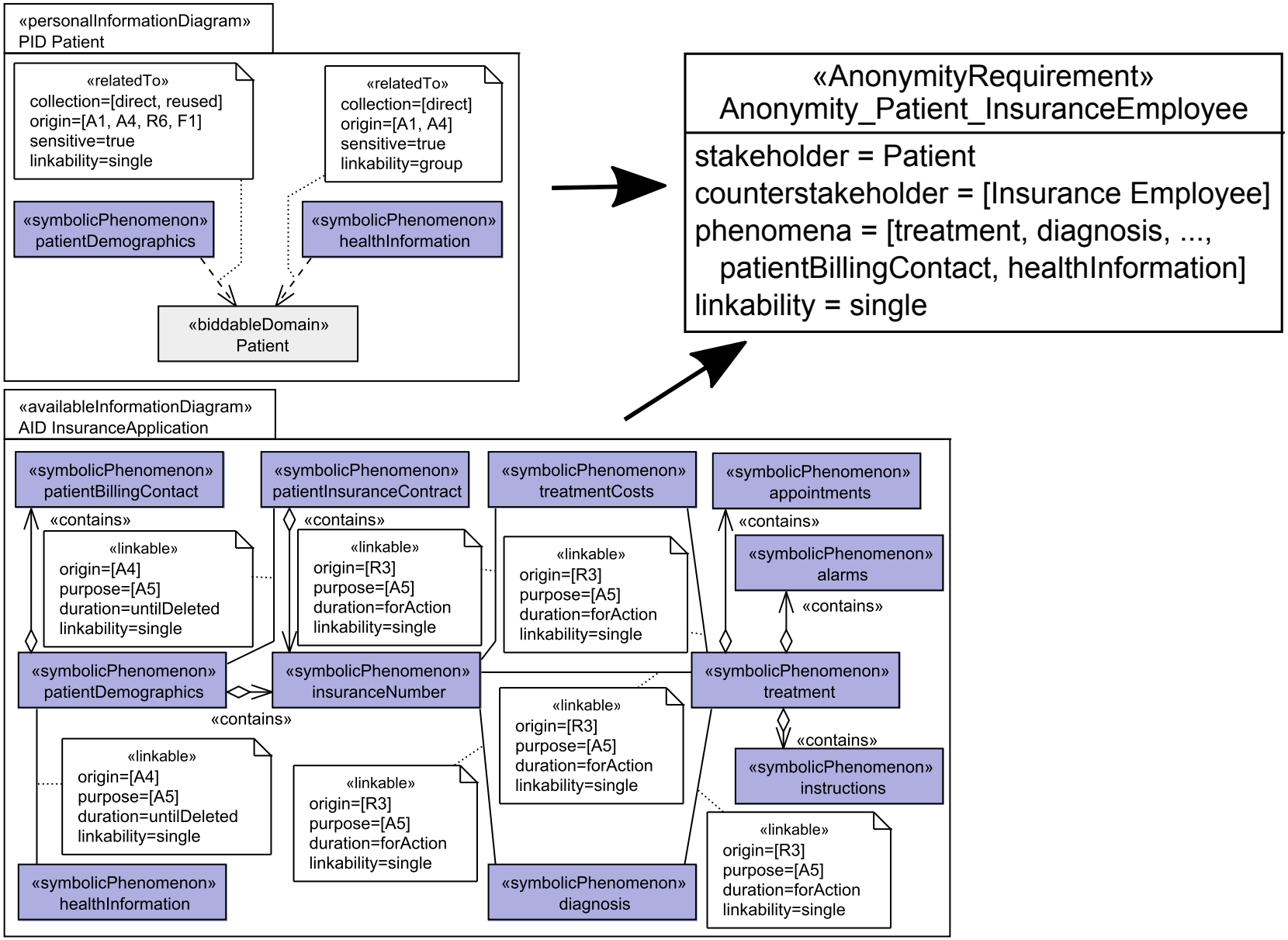
Anonymity:

The counterstakeholders shall at most be able to link the personal information phenomena to the stakeholder with linkability linkability.

Generation of Unlinkability Requirements

Example of an Anonymity Requirement

FoMSESS '16
Rene Meis
Introduction
Running Example
Background
Problem Frames
ProPAn
Overview
Identification of Personal Data
Personal Data Flow Analysis
Generate Privacy Requirements
Conclusion



Conclusion

FoMSESS '16
Rene Meis
Introduction
Running Example
Background
Problem Frames
ProPAn
Overview
Identification of Personal Data
Personal Data Flow Analysis
Generate Privacy Requirements
Conclusion

Our contributions are:

- A method that derives flows of personal information and privacy requirements from a requirements model.
- Representation of this information in a UML model.
- An extensible UML profile.
- Tool support that guides the application of the method.

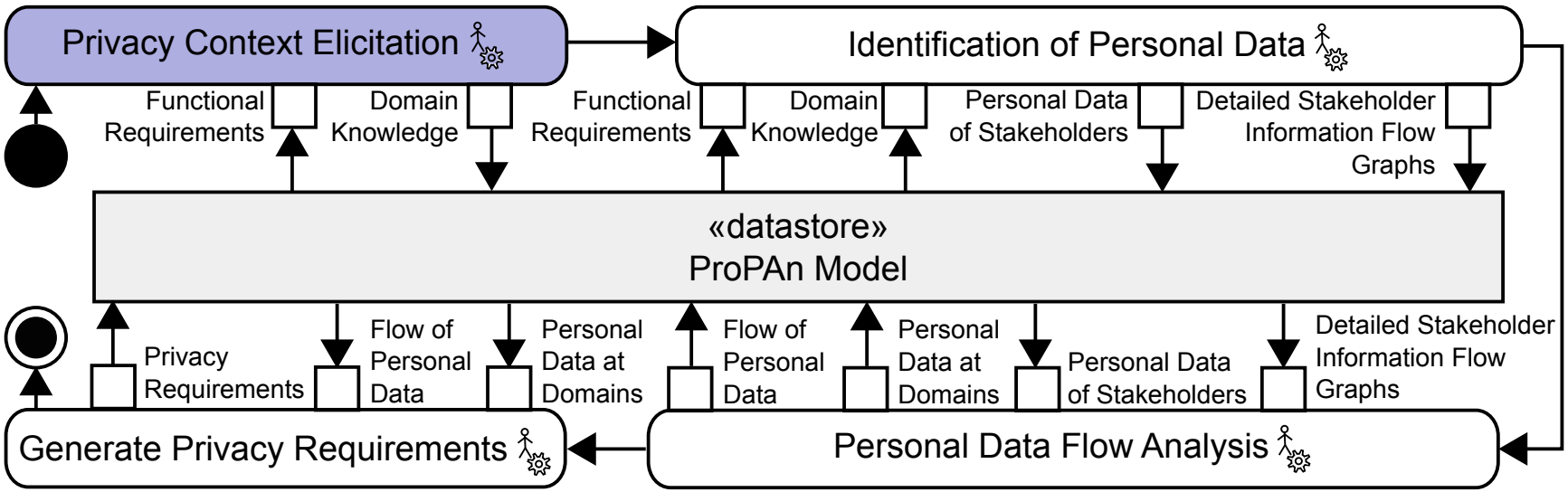
Our future directions include:

- Generation of PIA reports based on the elicited information.
- Identification of privacy threats in the model.
- Integration of PETs that mitigate the privacy threats.
- Validation of our method, the tool support, and the outputs produced by our method.

Overview of the ProPAn Method

Privacy Context Elicitation

FoMSESS '16
Rene Meis
Introduction
Running Example
Background
Problem Frames
ProPAn
Overview
Identification of Personal Data
Personal Data Flow Analysis
Generate Privacy Requirements
Conclusion



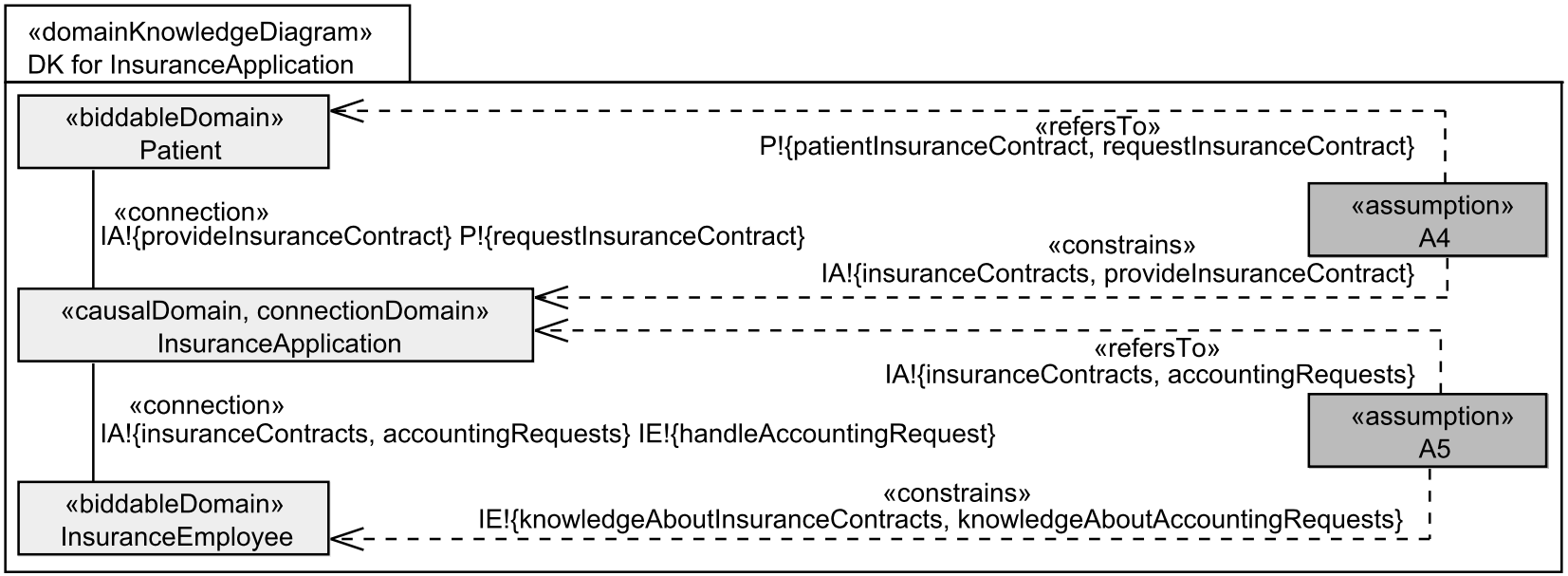
Context Elicitation

FoMSESS '16
Rene Meis
Introduction
Running Example
Background
Problem Frames
ProPAn
Overview
Identification of Personal Data
Personal Data Flow Analysis
Generate Privacy Requirements
Conclusion

The information provided by a functional requirements model does often not include information about:

- indirect stakeholders of whom data is processed
- information flows outside of the software-to-be
- indirect counterstakeholders who may have access to personal data

To systematically elicit this privacy-relevant information, we developed a method in previous work [Meis, 2014].



Bibliography I

FoMSESS '16
Rene Meis
Introduction
Running Example
Background
Problem Frames
ProPAn
Overview
Identification of Personal Data
Personal Data Flow Analysis
Generate Privacy Requirements
Conclusion


Beckers, K., Faßbender, S., Gritzalis, S., Heisel, M., Kalloniatis, C., and Meis, R. (2014a).
Privacy-aware cloud deployment scenario selection.
In *Trust, Privacy, and Security in Digital Business*, LNCS 8647, pages 94–105. Springer.


Beckers, K., Faßbender, S., Heisel, M., and Meis, R. (2014b).
A problem-based approach for computer aided privacy threat identification.
In *Annual Privacy Forum 2012*, LNCS 8319, pages 1–16. Springer.

Bibliography II

FoMSESS '16
Rene Meis
Introduction
Running Example
Background
Problem Frames
ProPAn
Overview
Identification of Personal Data
Personal Data Flow Analysis
Generate Privacy Requirements
Conclusion


Hansen, M., Jensen, M., and Rost, M. (2015).
Protection goals for privacy engineering.
In *2015 IEEE Symposium on Security and Privacy Workshops, SPW 2015, San Jose, CA, USA, May 21-22, 2015*, pages 159–166.


Jackson, M. (2001).
Problem Frames. Analyzing and structuring software development problems.
Addison-Wesley.


Meis, R. (2014).
Problem-Based Consideration of Privacy-Relevant Domain Knowledge.
In *Privacy and Identity Management for Emerging Services and Technologies*, IFIP AICT 421. Springer.

Bibliography III

FoMSESS '16
Rene Meis
Introduction
Running Example
Background
Problem Frames
ProPAn
Overview
Identification of Personal Data
Personal Data Flow Analysis
Generate Privacy Requirements
Conclusion


Meis, R. and Heisel, M. (2015).
Systematic identification of information flows from requirements to support privacy impact assessments.
In *ICSOFT-PT 2015 - Proc. of the 10th Int. Conf. on Software Paradigm Trends*, pages 43–52. SciTePress.


Meis, R., Heisel, M., and Wirtz, R. (2015).
A taxonomy of requirements for the privacy goal transparency.
In *Trust, Privacy, and Security in Digital Business*, LNCS 9264, pages 195–209. Springer.