How Good is a Security Policy against Breaches?

Özgür Kafalı†, Jasmine Jones*, Megan Petruso*,
Laurie Williams, and Munindar P. Singh

North Carolina State University
Department of Computer Science
†rkafali@ncsu.edu
*Supported by REU grant at NCSU

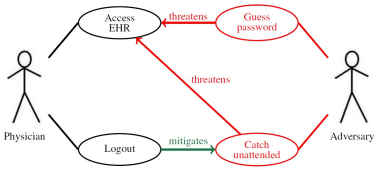
February 2017

Policies vs Breaches

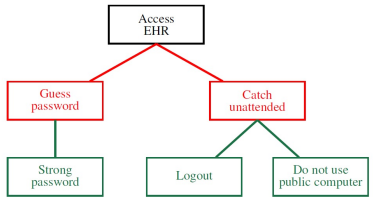
Design time Artifacts

Security Policies

Threat Models



Misuse Cases



A/D Trees

Policies vs Breaches

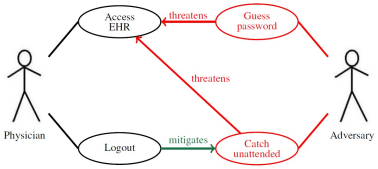
Design time Artifacts

Run time Artifacts

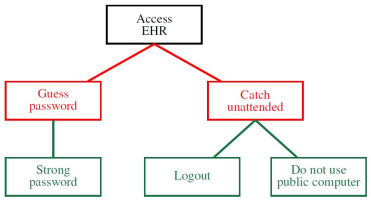
Security Policies

Breaches

Threat Models



Misuse Cases



A/D Trees

Policies vs Breaches

Design time Artifacts

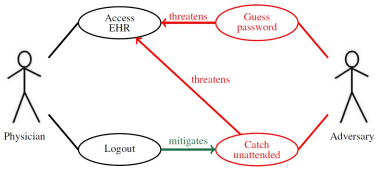
Run time Artifacts

Security Policies

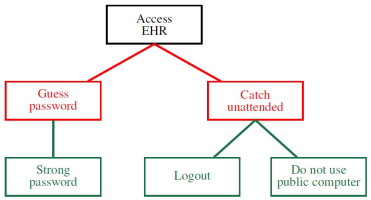
⟵ Connection ⟶

Breaches

Threat Models



Misuse Cases



A/D Trees

Policies vs Breaches

Design time Artifacts

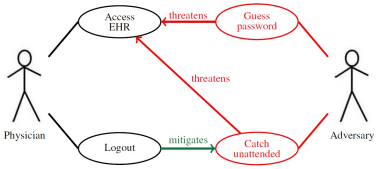
Run time Artifacts

Security Policies

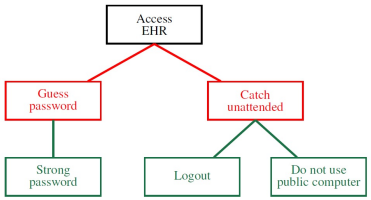
No such breach →

Breaches

Threat Models



Misuse Cases



A/D Trees

Policies vs Breaches

Design time Artifacts

Run time Artifacts

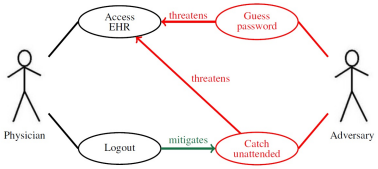
Security Policies

No such breach

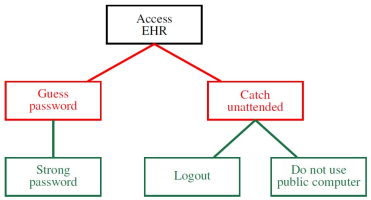
Severe sanction

Breaches

Threat Models



Misuse Cases



A/D Trees

Policies vs Breaches

Design time Artifacts

Run time Artifacts

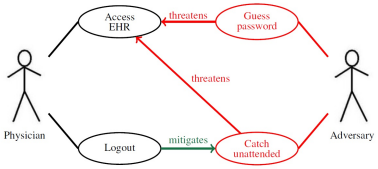
Security Policies

No such breach

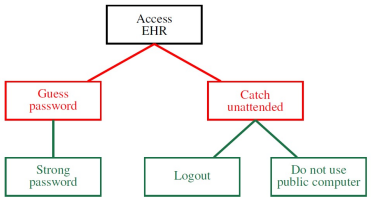
Nothing worth protecting

Breaches

Threat Models



Misuse Cases



A/D Trees

Policies vs Breaches

Design time Artifacts

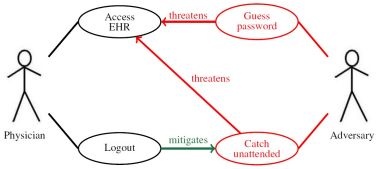
Run time Artifacts

Security Policies

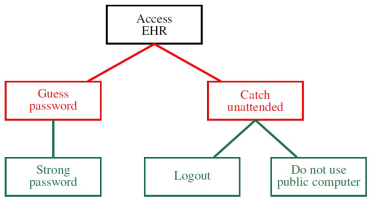
No such policy clause

Breaches

Threat Models



Misuse Cases



A/D Trees

Example

- HHS breach incident: In 2010, a failure to erase data contained on disposed photocopiers' hard drives led to the disclosure of patient records.

- HIPAA clause 45 CFR 164.310–(d)(2)(i): "Implement policies and procedures to address the final disposition of electronic protected health information, and/or the hardware or electronic media on which it is stored."

Example

- HHS breach incident: In 2010, a failure to erase data contained on disposed photocopiers' hard drives led to the disclosure of patient records.

- HIPAA clause 45 CFR 164.310–(d)(2)(i): "Implement policies and procedures to address the final disposition of electronic protected health information, and/or the hardware or electronic media on which it is stored."

Research Questions

- Representation: How can we formalize security policies and breaches to bring out their mutual correspondence?
- Similarity: What are the commonalities and differences between concepts in security policies and breach descriptions? How do those correspond to gaps in between?
- Analysis: How prevalent are accidental misuses among reported breaches, and do security policies account for them?

Fundamental Elements

- Norms: Commitments, Authorizations, Prohibitions
  - Represent policy clauses
  - Represent breach incidents
- Breach ontology
- Coverage metric

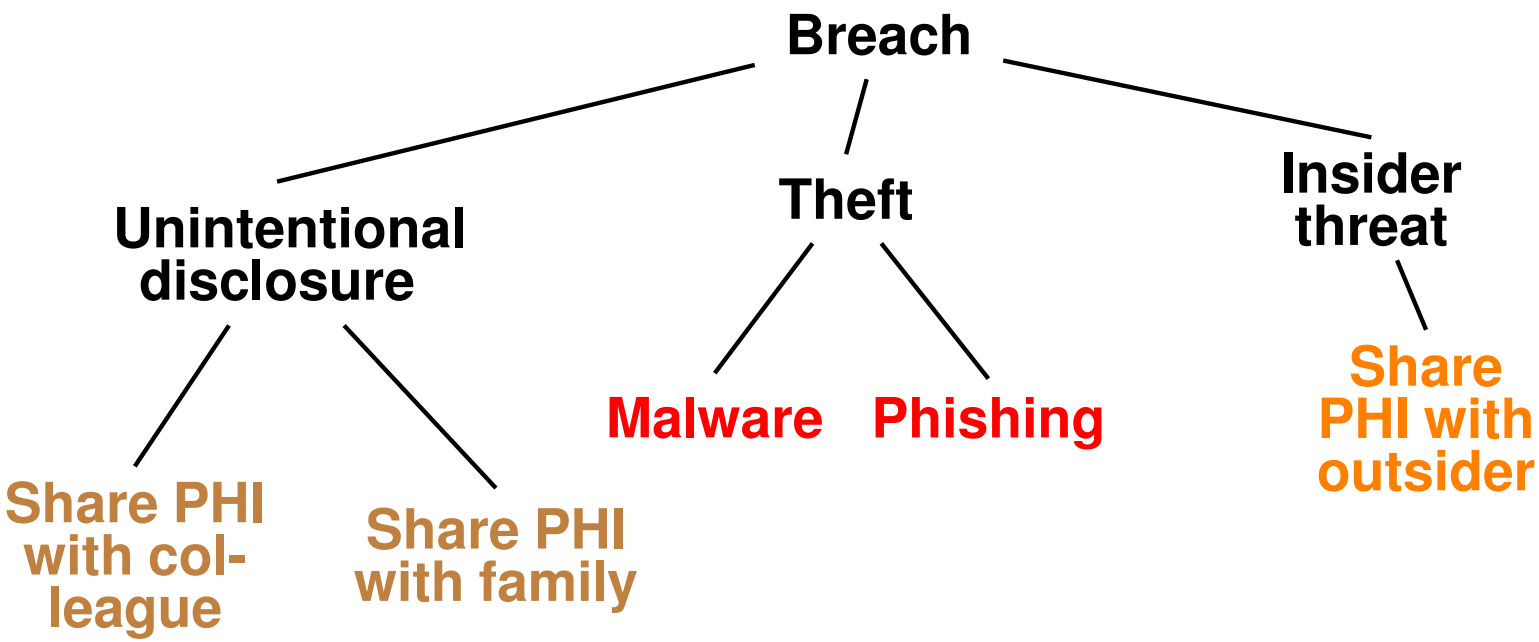
Norms

- Generic form: *N*(SUBJECT, OBJECT, antecedent, consequent)
- *N* = {Commitment, Authorization, Prohibition}

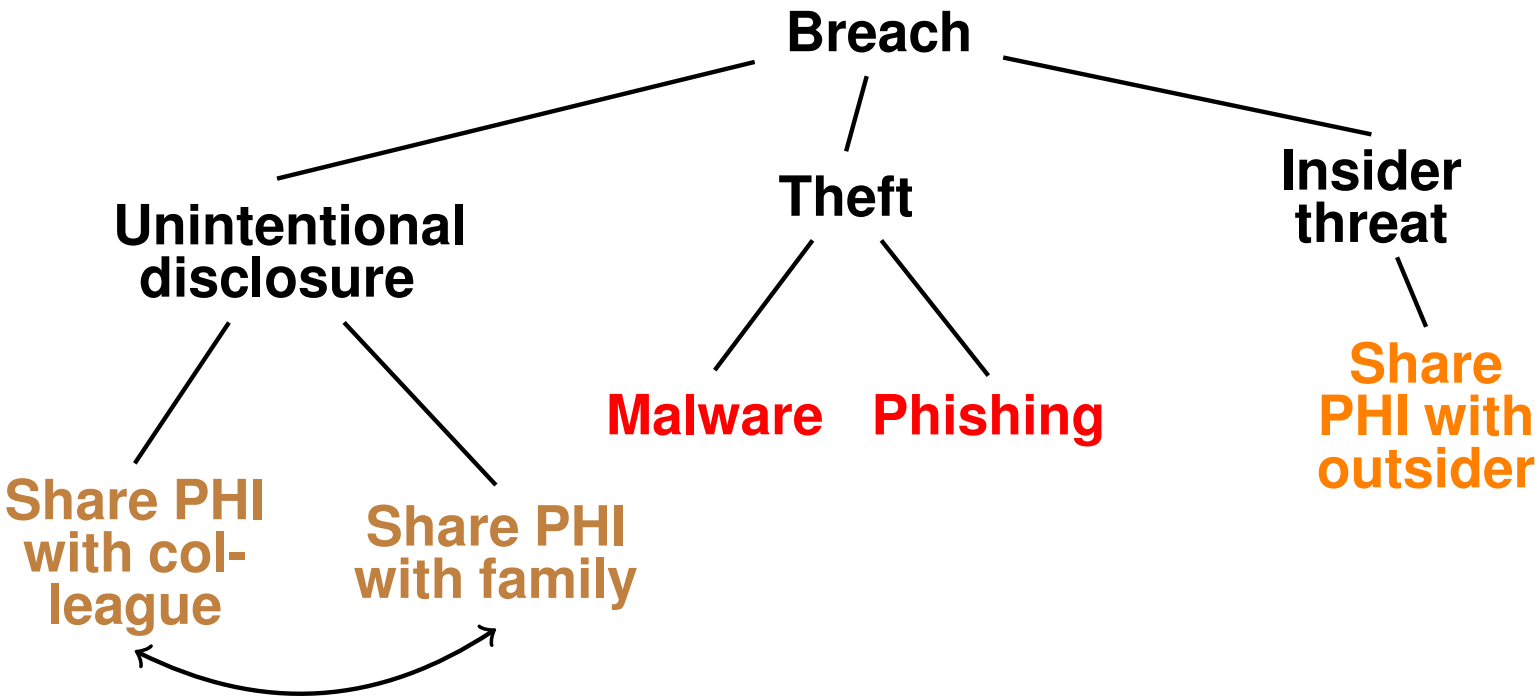
- HIPAA clause 45 CFR 164.310–(d)(2)(i): "Healthcare workers must erase patients' PHI stored on disposed electronic media."

  Commitment(HEALTHCARE_WORKER, COVERED_ENTITY, media_disposal, erase_PHI)

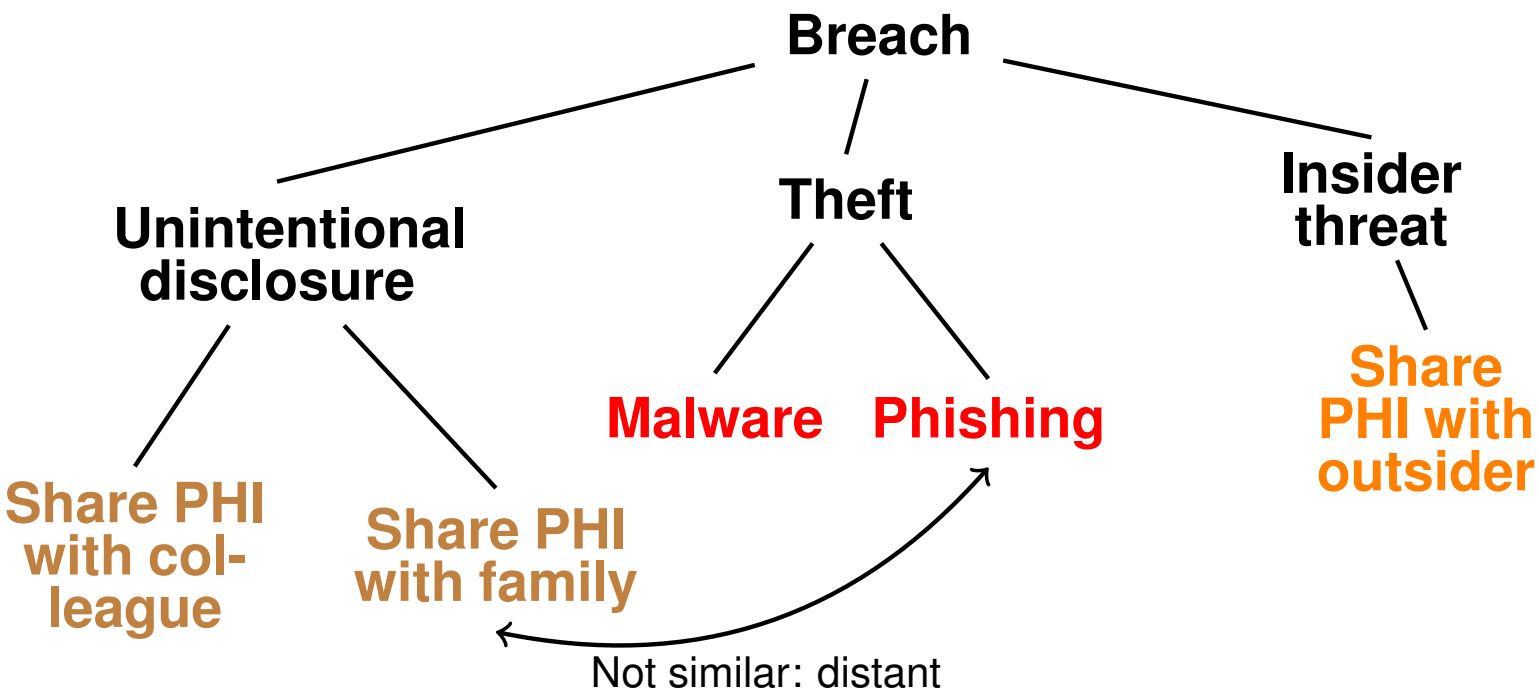
Ontologies: Breach Concepts



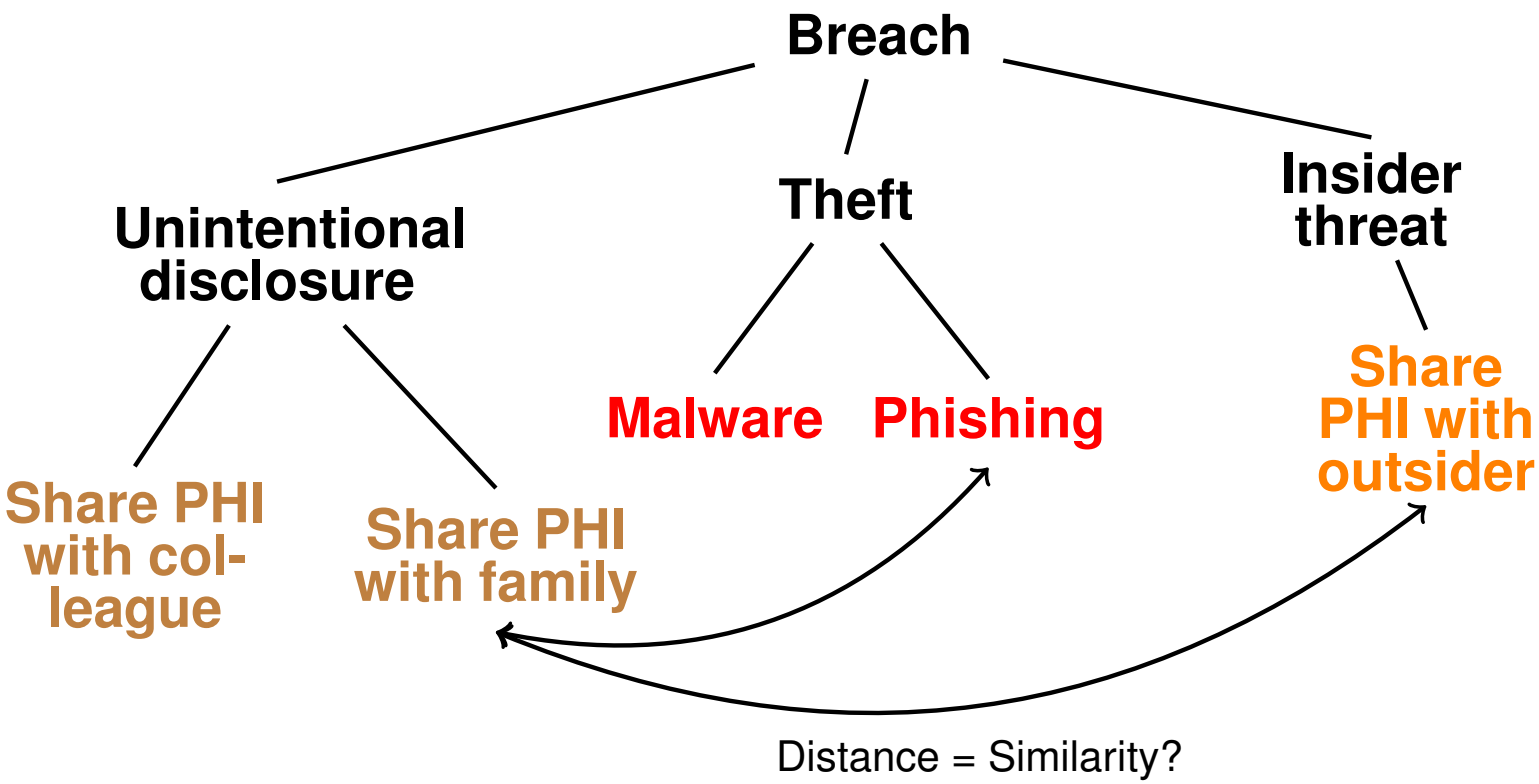
Ontologies: Breach Concepts



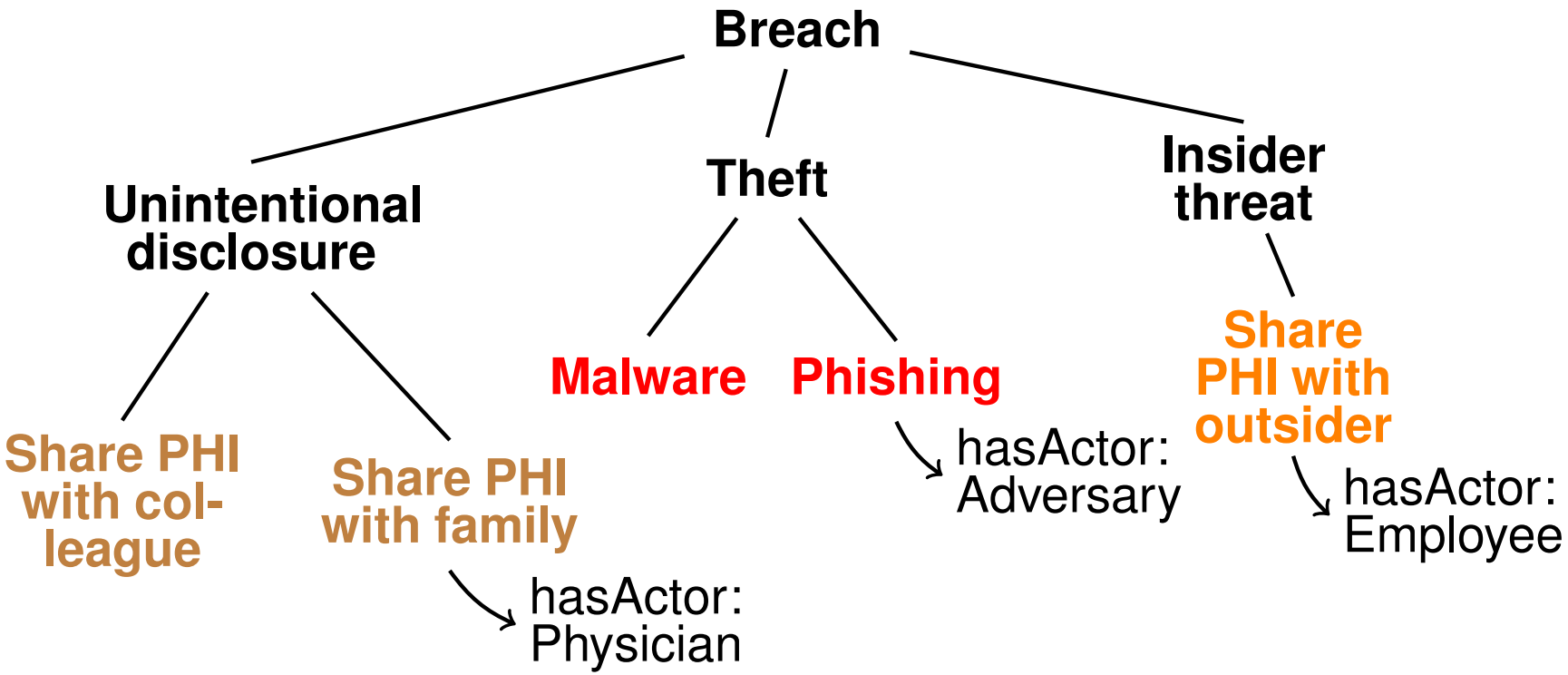
Ontologies: Breach Concepts



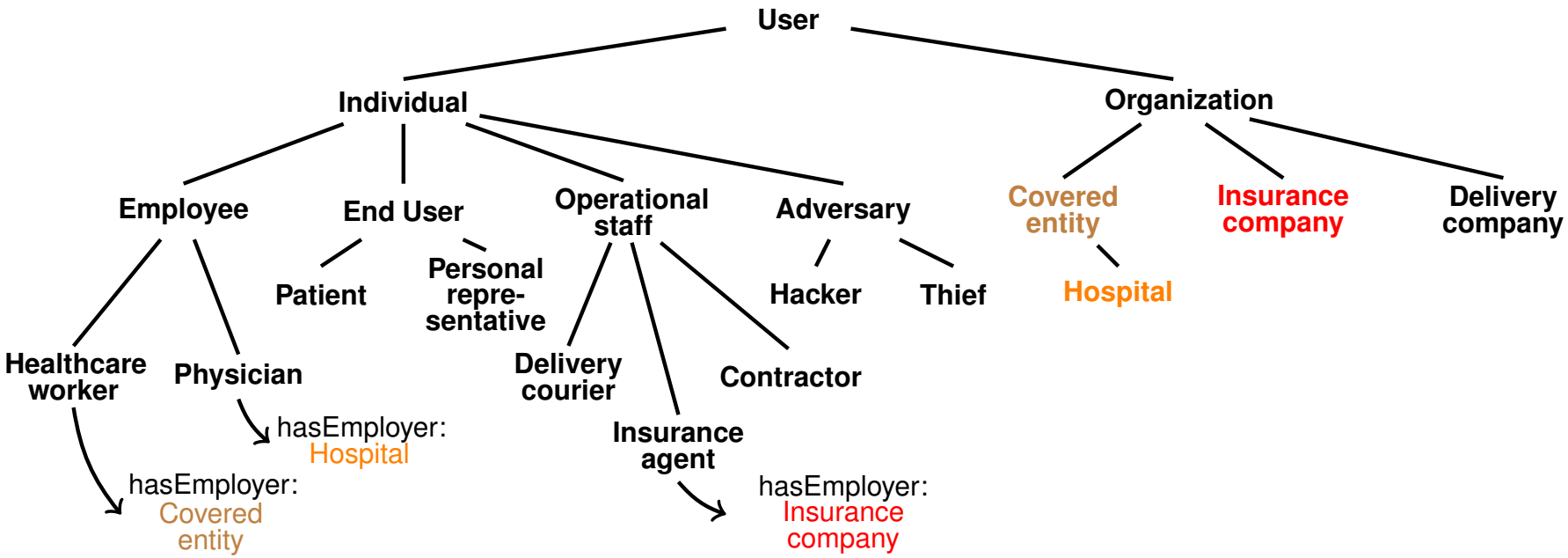
Ontologies: Breach Concepts



Ontologies: Breach Concepts



Ontologies: Healthcare Users



Semantic Reasoning

- Norm similarity:

$$sim_{n_1,n_2} = (sim_{\text{SBJ}_1,\text{SBJ}_2} + sim_{\text{OBJ}_1,\text{OBJ}_2} + sim_{\text{ant}_1,\text{ant}_2} + sim_{\text{con}_1,\text{con}_2}) \,/\, 4$$

Semantic Reasoning

- Norm similarity:
  $sim_{n_1,n_2} = (sim_{\text{SBJ}_1,\text{SBJ}_2} + sim_{\text{OBJ}_1,\text{OBJ}_2} + sim_{\text{ant}_1,\text{ant}_2} + sim_{\text{con}_1,\text{con}_2}) \,/\, 4$

- Distance between concepts: $\Delta_{c_1,c_2} = \text{edge\_count}(c_1, c_2)$

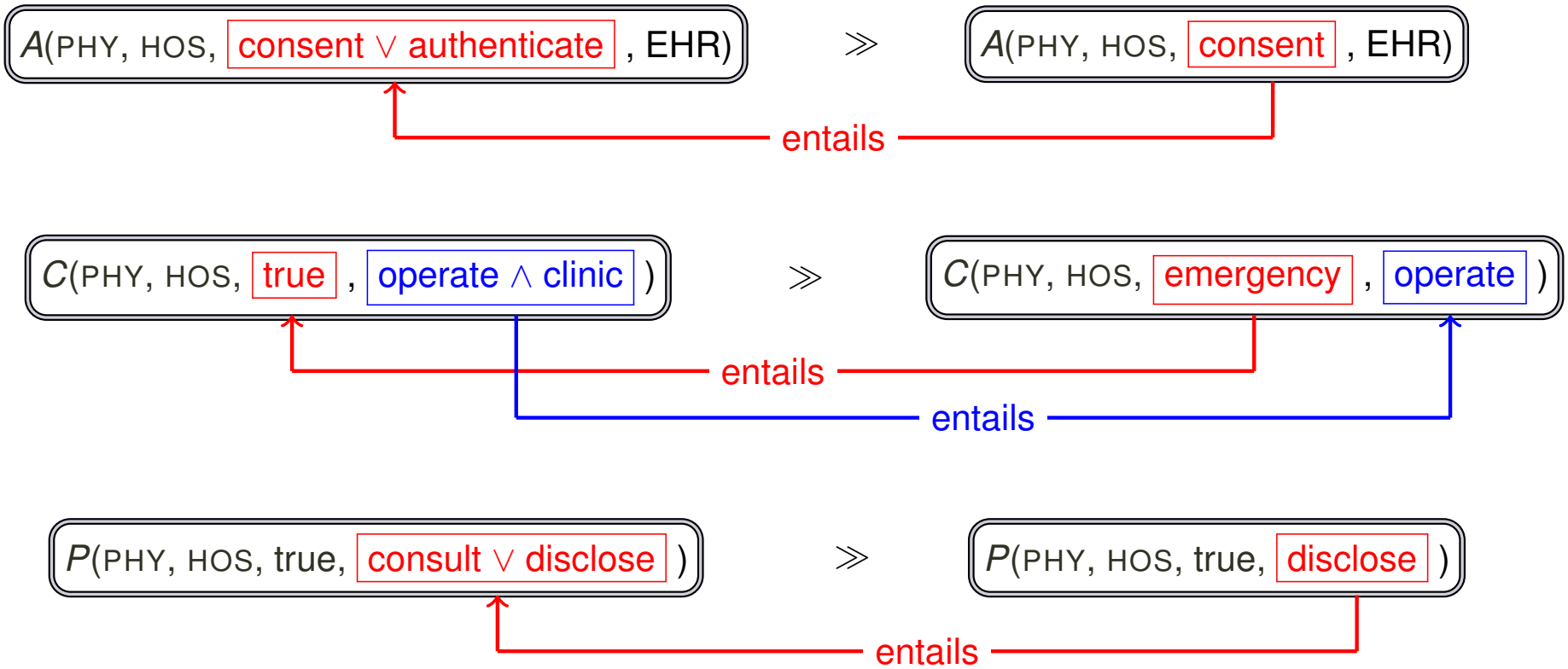
Semantic Reasoning

- Norm similarity:
  $sim_{n_1,n_2} = (sim_{\text{SBJ}_1,\text{SBJ}_2} + sim_{\text{OBJ}_1,\text{OBJ}_2} + sim_{\text{ant}_1,\text{ant}_2} + sim_{\text{con}_1,\text{con}_2}) \,/\, 4$
- Distance between concepts: $\Delta_{c_1,c_2} = \text{edge\_count}(c_1, c_2)$
- Similarity between concepts: $sim_{c_1,c_2} = \frac{1}{1+\Delta_{c_1,c_2}} \times sim^{prop}_{c_1,c_2}$

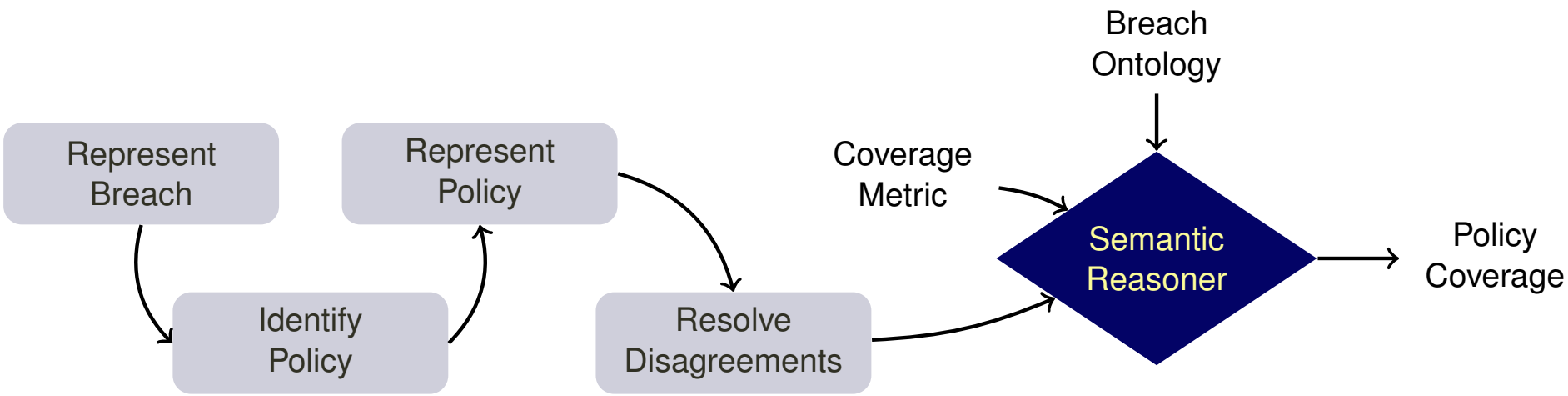
Semantic Reasoning

- Norm similarity:
  $sim_{n_1,n_2} = (sim_{\text{SBJ}_1,\text{SBJ}_2} + sim_{\text{OBJ}_1,\text{OBJ}_2} + sim_{\text{ant}_1,\text{ant}_2} + sim_{\text{con}_1,\text{con}_2}) \,/\, 4$
- Distance between concepts: $\Delta_{c_1,c_2} = \text{edge\_count}(c_1, c_2)$
- Similarity between concepts: $sim_{c_1,c_2} = \frac{1}{1+\Delta_{c_1,c_2}} \times sim^{prop}_{c_1,c_2}$
- Policy coverage: $coverage = \dfrac{\sum_{b_i \in \text{B}} \begin{cases} 1 & \text{if } n_{\text{policy}} \text{ covers } n_{b_i} \\ sim_{n_{\text{policy}},n_{b_i}} & \text{otherwise} \end{cases}}{|\text{B}|}$

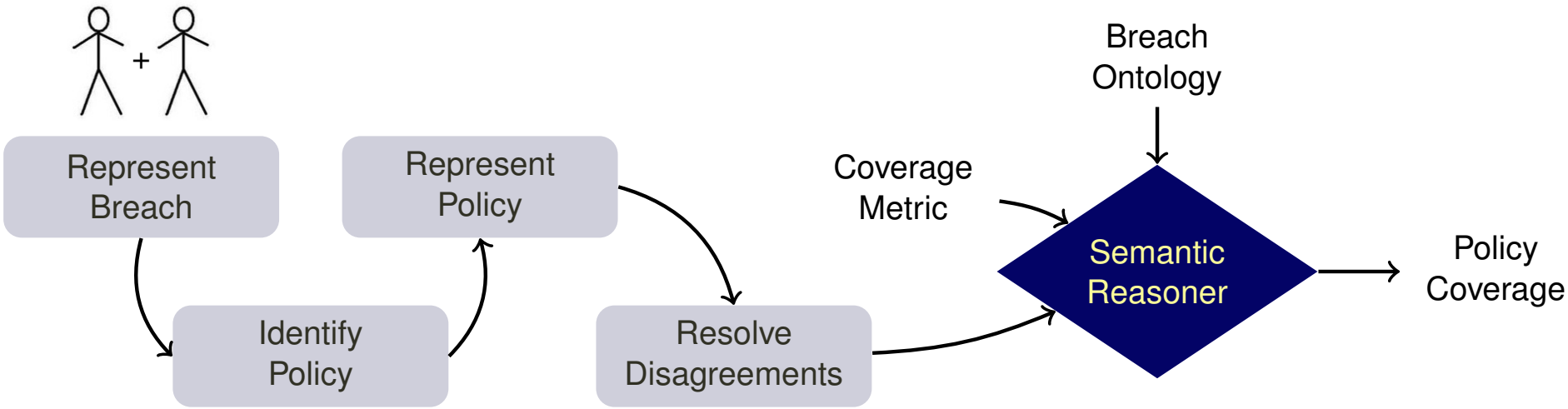
Norm Coverage



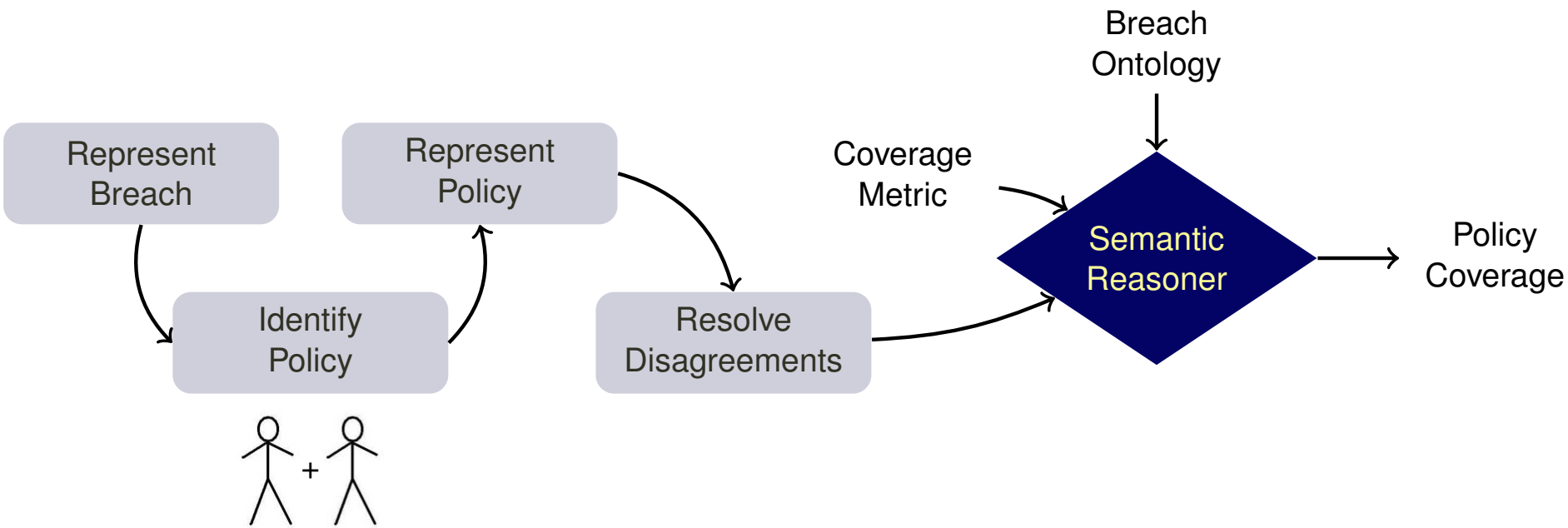
Methodology



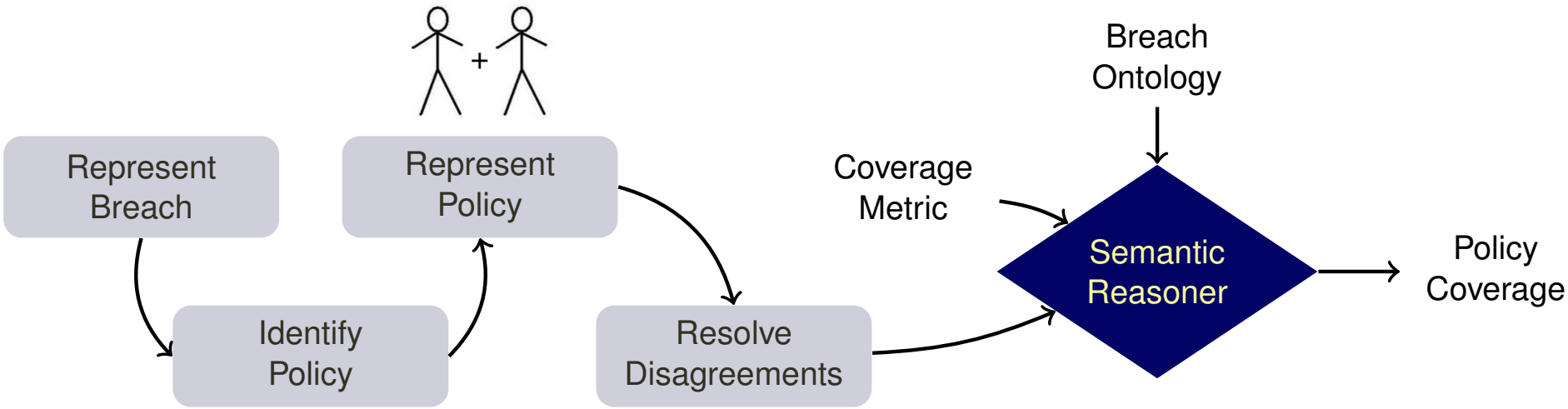
Methodology



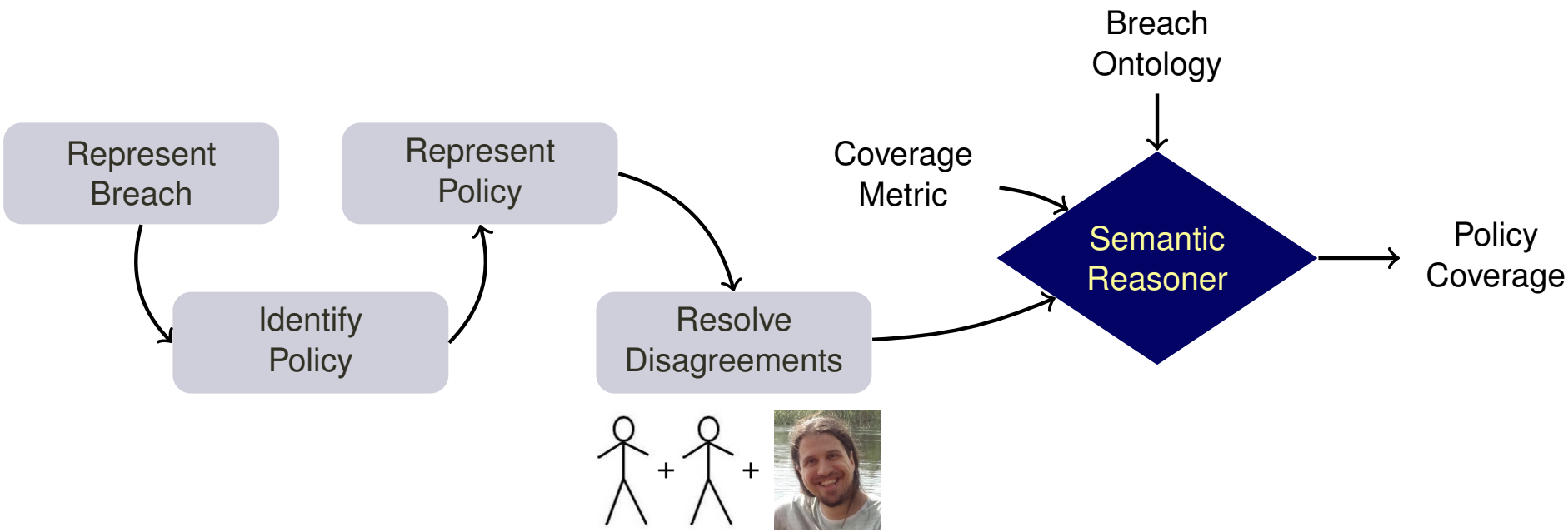
Methodology



Methodology



Methodology



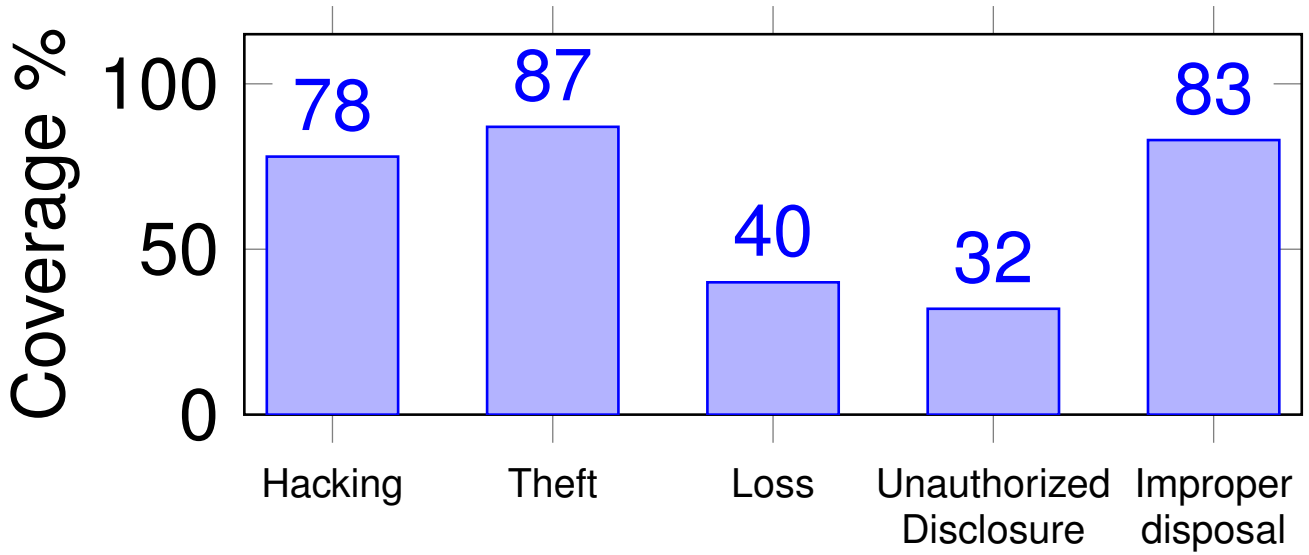
HHS Breach Report

Category	Count	Description
Hacking	191	Adversary exploits vulnerability to access EHR
Theft	642	Employee discloses PHI
Loss	129	Electronic media containing PHI are lost
Unauthorized disclosure	338	PHI is disclosed due to unauthorized access
Improper disposal	58	Employee fails to properly dispose PHI
Unclassified	219	Not classified by HHS

Classification of Breaches

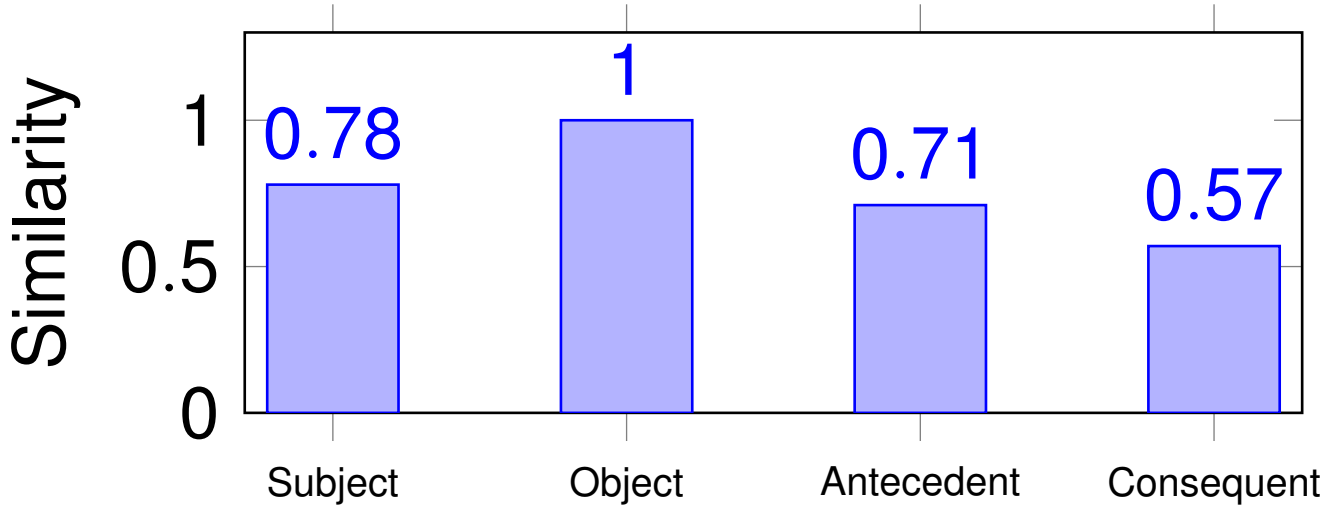
- 1,577 breaches reported by HHS
  - *Hacking* and *Theft* contain malicious misuses
  - *Loss*, *Unauthorized disclosure*, and *Improper disposal* contain accidental misuses
  - *Unclassified:* 68% accidental misuses and 13% malicious misuses
- Overall: 44% accidental misuses and 56% malicious misuses

Coverage by Breach Category



- 65% overall coverage by HIPAA
- Significantly better coverage for malicious misuses than accidental misuses

Similarity among Norm Elements



- Similarity between actors (subject/object) is higher than assets (antecedent/consequent)
- Consequent may be given a higher weight to provide a more realistic measure of coverage

Limitations

- Subjective modeling
- Assumptions on ontology, e.g., single inheritance, no instances
- Incompleteness of breaches
- Only applied to healthcare domain (though HIPAA is a dominant standard)

Future Work

- Guidelines for ontology development
- Automation and crowd for norm gathering
- Validation of coverage metric
- Narrowing the gaps with policy refinement