Reasoning in Description Logic Ontologies for Privacy Management

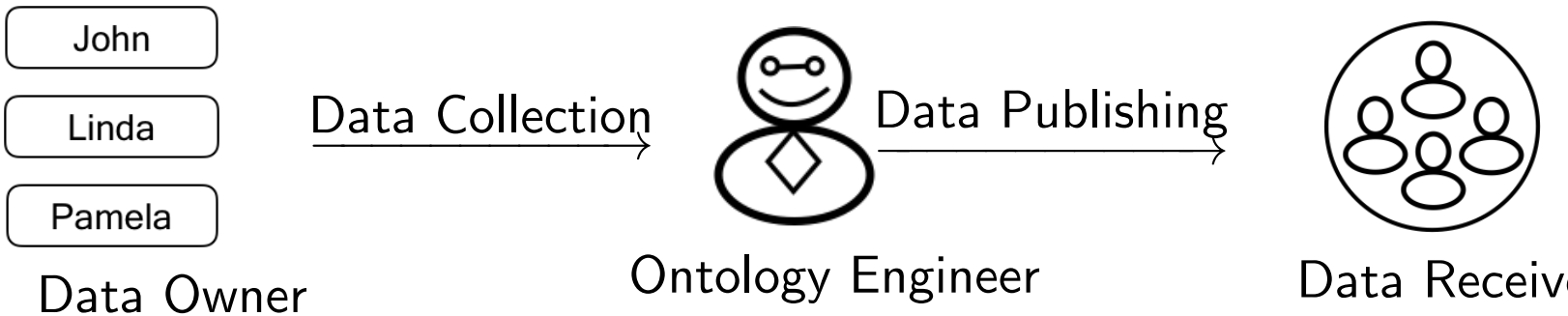
Adrian Nuradiansyah

Technische Universität Dresden

October 6, 2020

Data Collection and Data Publishing for Ontologies

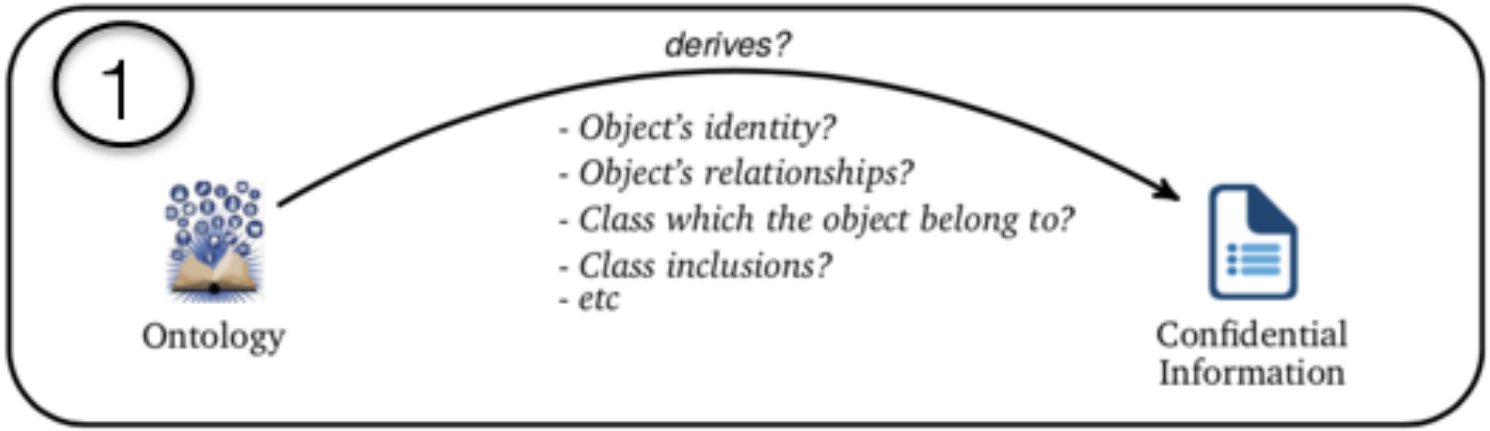
What **[Fung et. al, 2010]** illustrate ...

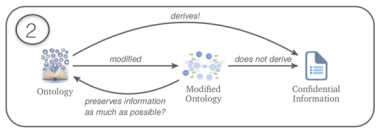


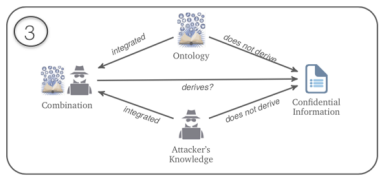
In the context of Description Logic Ontologies, **[Grau, 2010]** concerns ...

- A rise in the number of ontologies integrated in mainstream applications, e.g., medical systems
- Possible **unauthorized disclosures** of medical information may occur
- Designing privacy-preserving systems is being a critical requirement

What Should the Engineer Do Before Publishing?

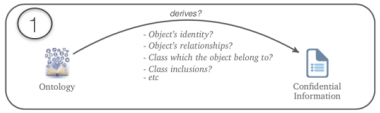


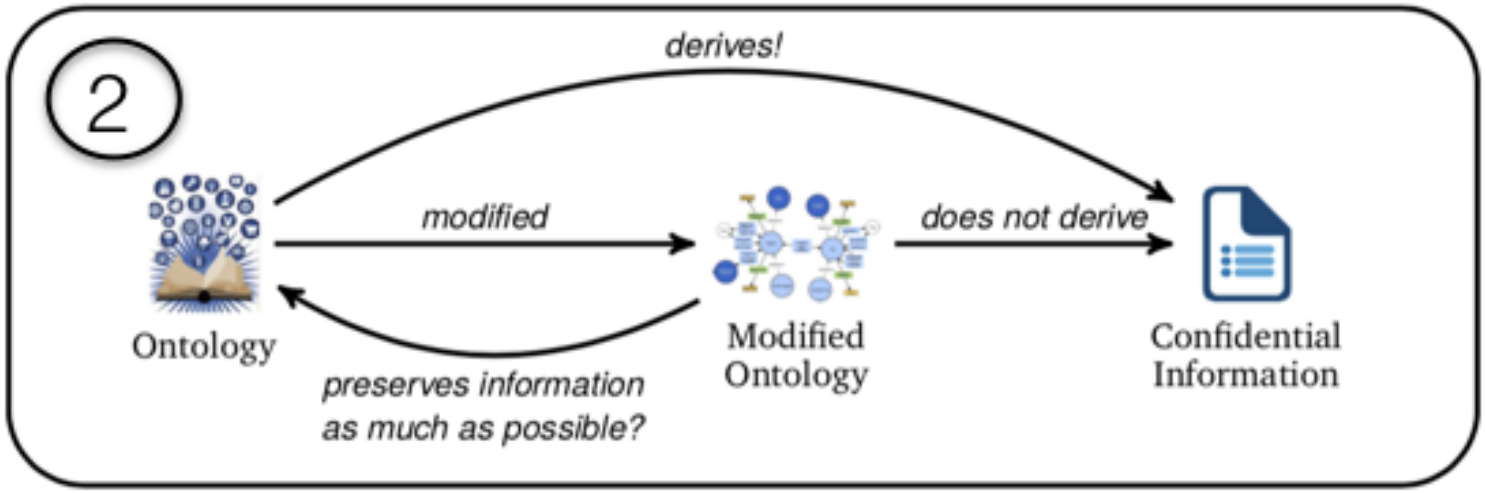


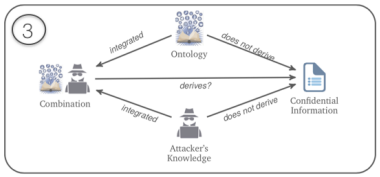


Detect Privacy Breach

What Should the Engineer Do Before Publishing?

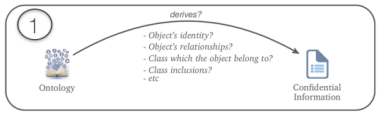


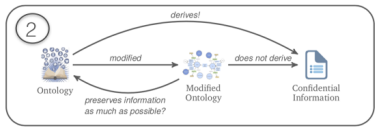


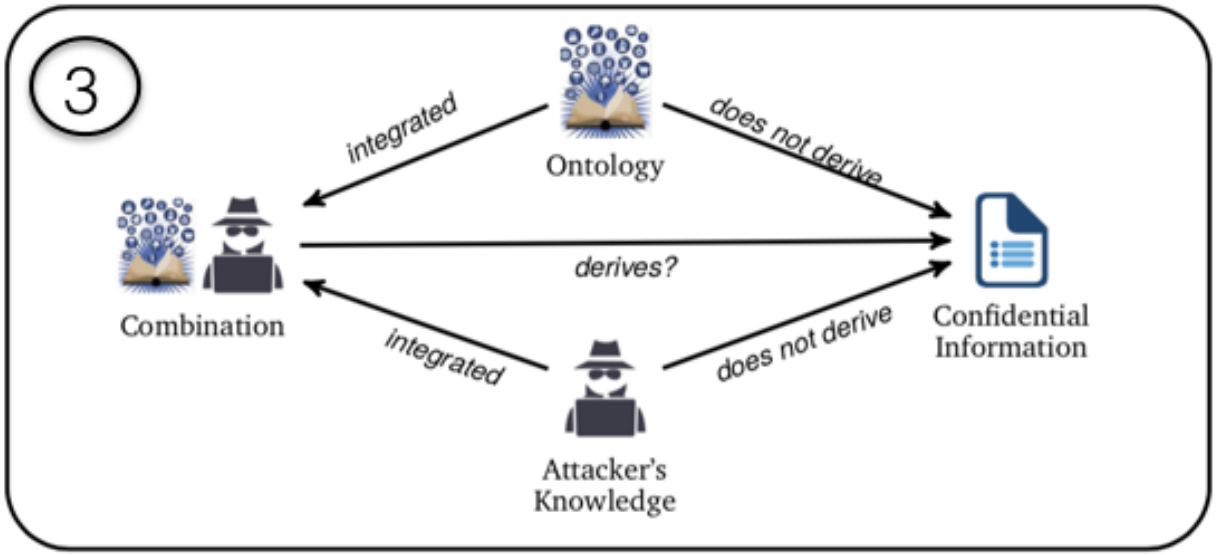


Detect Privacy Breach → Ontology Repair

What Should the Engineer Do Before Publishing?







Detect Privacy Breach → Ontology Repair → Avoid ILnkage Attacks

What People Have Done

- Confidential information $\Rightarrow$ *property of individuals*
- Membership of individuals (tuple of individuals) in the answers to certain queries
  (e.g., [Calvanesse et. al., 2008], [Stouppa & Studer, 2009], [Tao et.al., 2010])

What People Have Done

- Confidential information $\Rightarrow$ *property of individuals*
- Membership of individuals (tuple of individuals) in the answers to certain queries
  (e.g., [Calvanesse et. al., 2008], [Stouppa & Studer, 2009], [Tao et.al., 2010])

Focus on Identity? What is "identity"?

What People Have Done

- Finding justifications why the (unwanted) consequences can be derived (e.g., [Schlobach, 2003], [Parsia et. al., 2007], [Baader et. al., 2008])
- Remove axioms that are responsible for the entailment (e.g., [Kalyanpur et. al., 2006])

What People Have Done

- Finding justifications why the (unwanted) consequences can be derived (e.g., [Schlobach, 2003], [Parsia et. al., 2007], [Baader et. al., 2008])
- Remove axioms that are responsible for the entailment (e.g., [Kalyanpur et. al., 2006])

Do these approaches also remove useful consequences?
Can we do it more "gentle"?

What People Have Done

- Learning type of attackers' background knowledge
- Investigating *attribute linkage*, *table linkage*, etc thoroughly in e.g., [Fung et. al., 2010]
- Introducing the notion of *policy-compliance and policy-safety* in the context of RDF graphs/Linked Data in e.g., [Grau & Kostylev, 2016]

What People Have Done

- Learning type of attackers' background knowledge
- Investigating *attribute linkage*, *table linkage*, etc thoroughly in e.g., [Fung et. al., 2010]
- Introducing the notion of *policy-compliance and policy-safety* in the context of RDF graphs/Linked Data in e.g., [Grau & Kostylev, 2016]

Is such setting already considered in DL ontologies?

Problem Descriptions

Detecting Privacy Breach

The Identity Problem and its Variants in Description Logic Ontologies

Ontology Repair

Repairing Description Logic Ontologies via Axiom Weakening

Avoiding Linkage Attacks

Privacy-Preserving Ontology Publishing

Description Logics (DLs)

The **logical underpinning** of **Web Ontology Language (OWL)**

Decidable fragments of First Order Logics

Representing the conceptual knowledge of an application domain in a well-understood way.

Description Logics (DLs)

The **logical underpinning** of **Web Ontology Language (OWL)**

Decidable fragments of First Order Logics

Representing the conceptual knowledge of an application domain in a well-understood way.

Non-German people who work at an IT Department whose all locations are either in Germany or in Austria

$$\Downarrow$$

$$\neg German \sqcap \exists worksAt.(ITDept \sqcap \forall located.(Germany \sqcup Austria))$$

DL Concepts

Name	**Syntax**	**Example**
Top	$\top$	*tautology*
Concept Name	A	*Germany*
Conjunction	$C \sqcap D$	$German \sqcap Female$
Disjunction	$C \sqcup D$	$Germany \sqcup Austria$
Existential Restriction	$\exists r.C$	$German \sqcap \exists worksAt.ITDept$
Universal Restriction	$\forall r.C$	$ITDept \sqcap \forall located.Germany$
Negation	$\neg C$	$\neg German$
(One of) Nominal	$\{a_1, \ldots, a_n\}$	$\{LINDA, JOHN, JIM\}$

DL Concepts

Name	Syntax	Example
Top	$\top$	*tautology*
Concept Name	A	*Germany*
Conjunction	$C \sqcap D$	$German \sqcap Female$
Disjunction	$C \sqcup D$	$Germany \sqcup Austria$
Existential Restriction	$\exists r.C$	$German \sqcap \exists worksAt.ITDept$
Universal Restriction	$\forall r.C$	$ITDept \sqcap \forall located.Germany$
Negation	$\neg C$	$\neg German$
(One of) Nominal	$\{a_1, \ldots, a_n\}$	$\{LINDA, JOHN, JIM\}$

$\mathcal{ALC}$

- Closed under Boolean operators
- Intractable

DL Concepts

Name	Syntax	Example
Top	$\top$	*tautology*
Concept Name	A	*Germany*
Conjunction	$C \sqcap D$	$German \sqcap Female$
Disjunction	$C \sqcup D$	$Germany \sqcup Austria$
Existential Restriction	$\exists r.C$	$German \sqcap \exists worksAt.ITDept$
Universal Restriction	$\forall r.C$	$ITDept \sqcap \forall located.Germany$
Negation	$\neg C$	$\neg German$
(One of) Nominal	$\{a_1, \ldots, a_n\}$	$\{LINDA, JOHN, JIM\}$

$\mathcal{EL}$

inexpressive, but reasoning is in PTime

DL Concepts

Name	Syntax	Example
Top	$\top$	*tautology*
Concept Name	A	*Germany*
Conjunction	$C \sqcap D$	$German \sqcap Female$
Disjunction	$C \sqcup D$	$Germany \sqcup Austria$
Existential Restriction	$\exists r.C$	$German \sqcap \exists worksAt.ITDept$
Universal Restriction	$\forall r.C$	$ITDept \sqcap \forall located.Germany$
Negation	$\neg C$	$\neg German$
(One of) Nominal	$\{a_1, \ldots, a_n\}$	$\{LINDA, JOHN, JIM\}$

$\mathcal{FL}_0$

The dual of $\mathcal{EL}$

DL Concepts

Name	Syntax	Example
Top	$\top$	*tautology*
Concept Name	A	*Germany*
Conjunction	$C \sqcap D$	$German \sqcap Patient$
Disjunction	$C \sqcup D$	$Germany \sqcup Austria$
Existential Restriction	$\exists r.C$	$German \sqcap \exists worksAt.ITDept$
Universal Restriction	$\forall r.C$	$ITDept \sqcap \forall located.Germany$
Negation	$\neg C$	$\neg German$
(One of) Nominal	$\{a_1, \ldots, a_n\}$	$\{LINDA, JOHN, JIM\}$

$\mathcal{FLE}$

Combination of $\mathcal{EL}$ and $\mathcal{FL}_0$

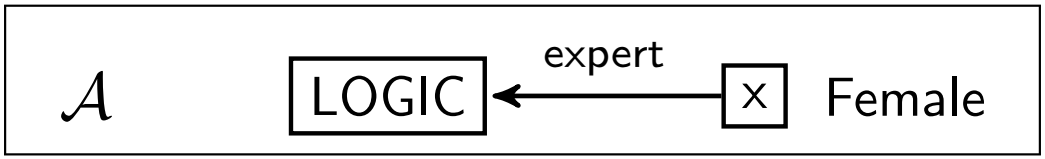
DL Ontologies

A **DL ontology** $\mathfrak{O}$ consists of an **ABox** $\mathcal{A}$ and a **TBox** $\mathcal{T} \Longleftrightarrow \mathfrak{O} = (\mathcal{A}, \mathcal{T})$

DL Ontologies

A **DL ontology** $\mathfrak{O}$ consists of an **ABox** $\mathcal{A}$ and a **TBox** $\mathcal{T} \iff \mathfrak{O} = (\mathcal{A}, \mathcal{T})$

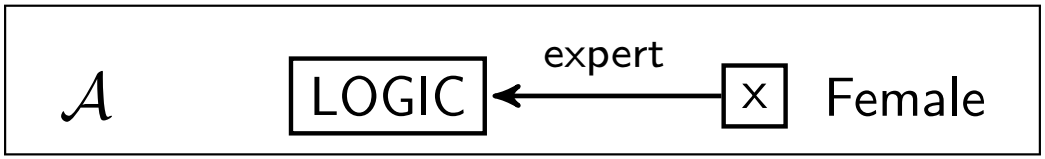
An ABox $\mathcal{A}$: **knowledge about individuals** (instance relationships $C(a)$ and individual relationships $r(a, b)$)



DL Ontologies

A **DL ontology** $\mathfrak{O}$ consists of an **ABox** $\mathcal{A}$ and a **TBox** $\mathcal{T} \iff \mathfrak{O} = (\mathcal{A}, \mathcal{T})$

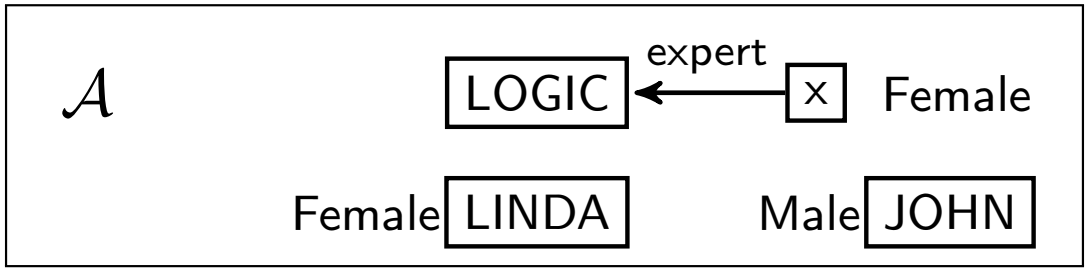
An ABox $\mathcal{A}$: **knowledge about individuals** (instance relationships $C(a)$ and individual relationships $r(a, b)$)



A TBox $\mathcal{T}$: inclusion relationships/constraints between concepts $C \sqsubseteq D$ (**General Concept Inclusions (GCIs)**)

$\mathcal{T}$ $\quad \exists expert.\{LOGIC\} \sqsubseteq VerTeam \quad Female \sqsubseteq \neg Male$

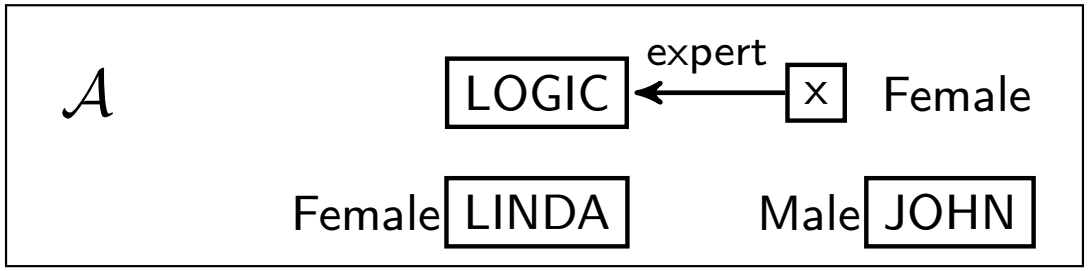
What can I Infer from an Ontology?



$\mathcal{T}$ $\quad \exists expert.\{LOGIC\} \sqsubseteq VerTeam \qquad Female \sqsubseteq \neg Male$

$$VerTeam \equiv \{LINDA, JOHN\}$$

What can I Infer from an Ontology?



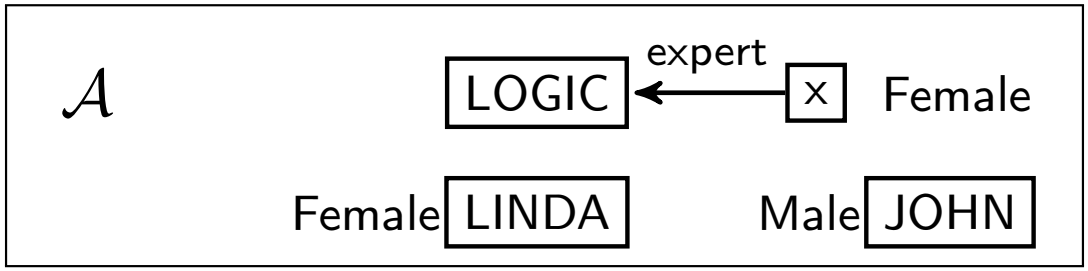
$\mathcal{T}$ $\quad \exists expert.\{LOGIC\} \sqsubseteq VerTeam \qquad Female \sqsubseteq \neg Male$

$VerTeam \equiv \{LINDA, JOHN\}$

Does $\exists expert.\{LOGIC\} \sqsubseteq \{LINDA, JOHN\}$ hold w.r.t. $(\mathcal{A},\mathcal{T})$? ✓ (**Subsumption Problem**)

Is x an instance of $VerTeam$ w.r.t. $(\mathcal{A}, \mathcal{T})$? ✓ (**Instance Problem**)

What can I Infer from an Ontology?



$\mathcal{T}$ $\exists expert.\{LOGIC\} \sqsubseteq VerTeam$ $Female \sqsubseteq \neg Male$
$VerTeam \equiv \{LINDA, JOHN\}$

Does $\exists expert.\{LOGIC\} \sqsubseteq \{LINDA, JOHN\}$ hold w.r.t. $(\mathcal{A},\mathcal{T})$? ✓ (**Subsumption Problem**)

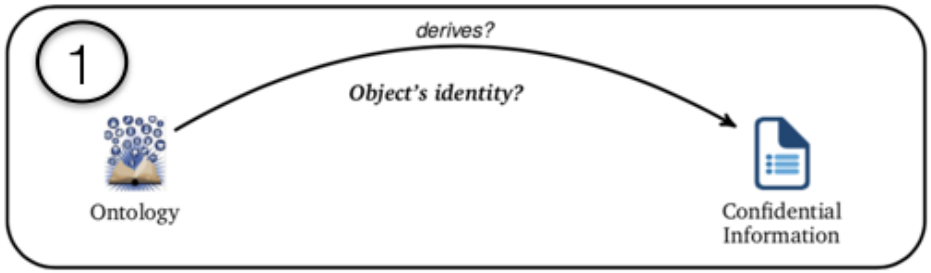
Is x an instance of $VerTeam$ w.r.t. $(\mathcal{A}, \mathcal{T})$? ✓ (**Instance Problem**)

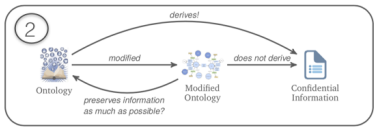
Wait, how do you know it?

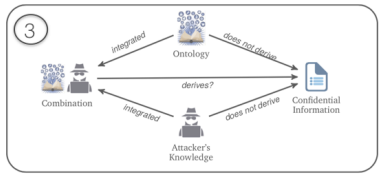
Is LINDA the **identity** of anonymous x? ✓

how do we call this problem?

Problem 1: Is My Identity Safe?

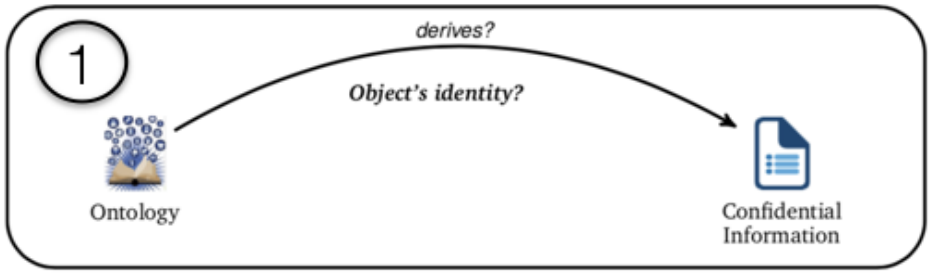


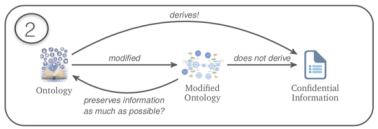


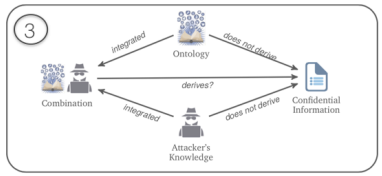


Identity Problem ($\mathfrak{O} \models x \doteq a$) [DL 2017], [JIST 2017]

Problem 1: Is My Identity Safe?



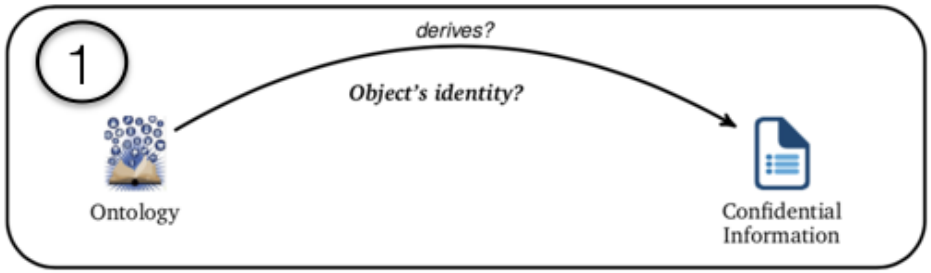


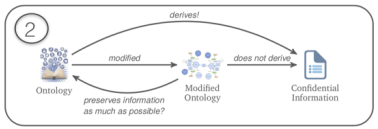


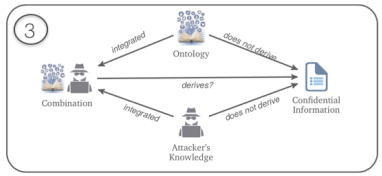
Identity Problem ($\mathfrak{O} \models x \doteq a$) [DL 2017], [JIST 2017]

- Not all DLs are able to derive equalities between individuals, e.g. $\mathcal{ALC}$.
- **DLs with equality power**: nominals, number restrictions, and functional dependencies.

Problem 1: Is My Identity Safe?





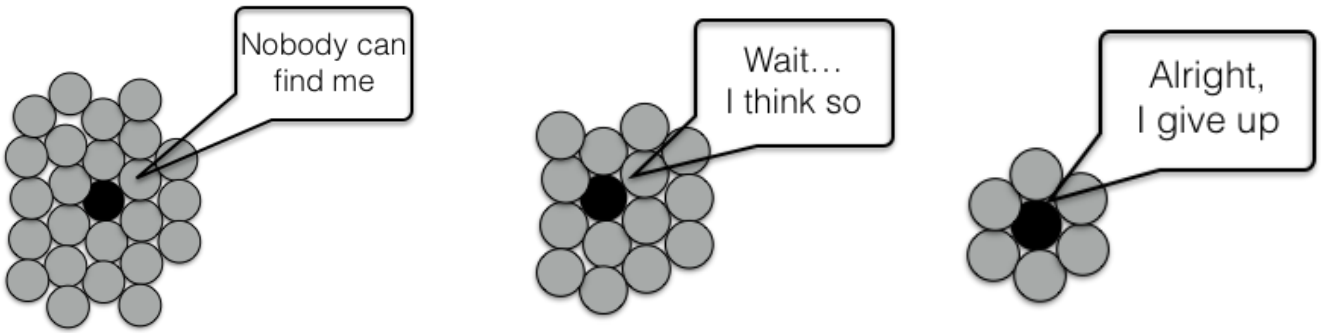


Identity Problem ($\mathfrak{O} \models x \doteq a$) [DL 2017], [JIST 2017]

- Not all DLs are able to derive equalities between individuals, e.g. $\mathcal{ALC}$.
- **DLs with equality power**: nominals, number restrictions, and functional dependencies.
- **Identity to Instance**: Given two individuals x, a, and an ontology $\mathfrak{O}$ formulated in a DL with equality power, it holds

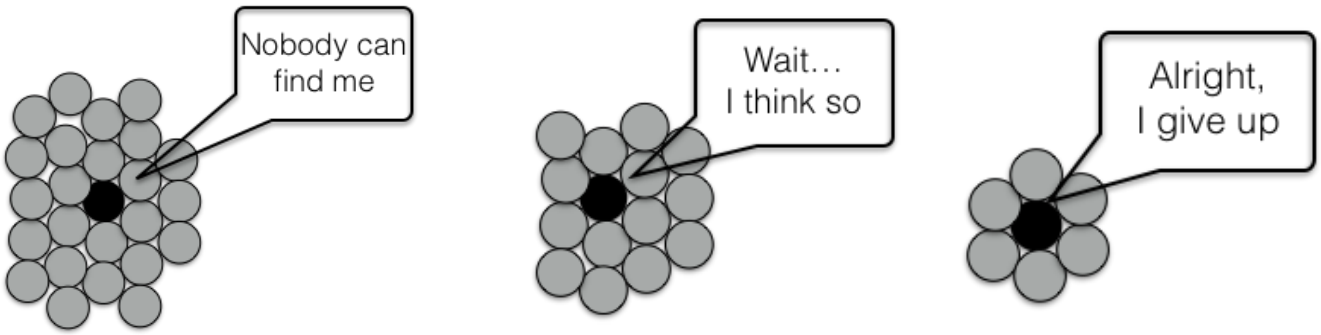
$$\mathfrak{O} \models x \doteq a \text{ iff } (\mathfrak{O} \cup \{Q(x)\}) \models Q(a), \text{ where } Q \text{ is a fresh concept name}$$

The Identity is one of k Known Individuals



"Hiding in the crowd" . . .

The Identity is one of k Known Individuals



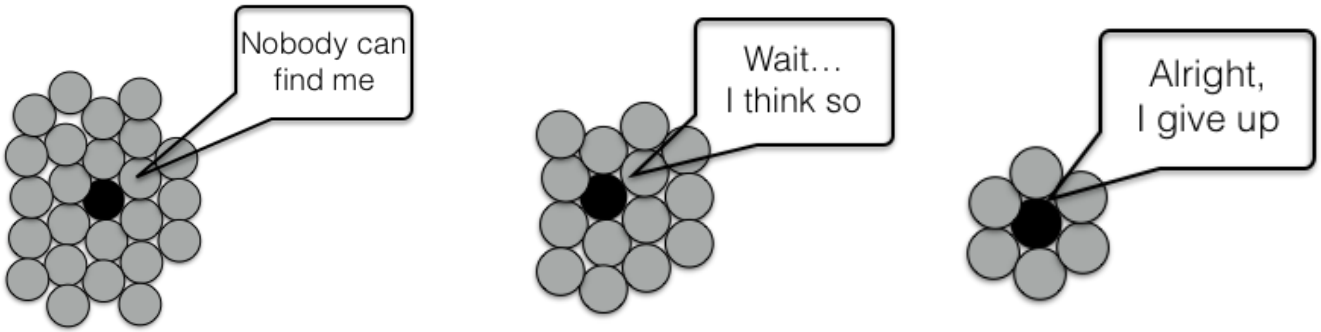
"Hiding in the crowd" . . .

k-Hiding

The anonymous individual x is **not k-hidden** w.r.t. $\mathfrak{O}$ iff there are known individuals $a_1, \ldots, a_{k-1}$ such that

$$x \text{ belongs to } \{a_1, \ldots, a_{k-1}\} \text{ w.r.t. } \mathfrak{O}$$

The Identity is one of k Known Individuals



"Hiding in the crowd" . . .

k-Hiding

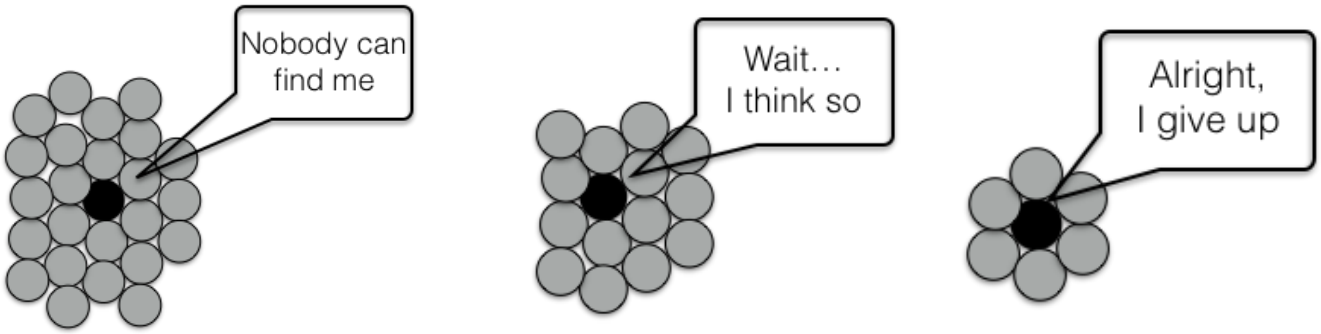
The anonymous individual x is **not k-hidden** w.r.t. $\mathfrak{O}$ iff there are known individuals $a_1, \ldots, a_{k-1}$ such that

$$x \text{ belongs to } \{a_1, \ldots, a_{k-1}\} \text{ w.r.t. } \mathfrak{O}$$

How to solve it

- Reduce it to the instance problem for *all* DLs with equality power
- Reduce it to the identity problem for *some* convex DLs with equality power

The Identity is one of k Known Individuals



"Hiding in the crowd" . . .

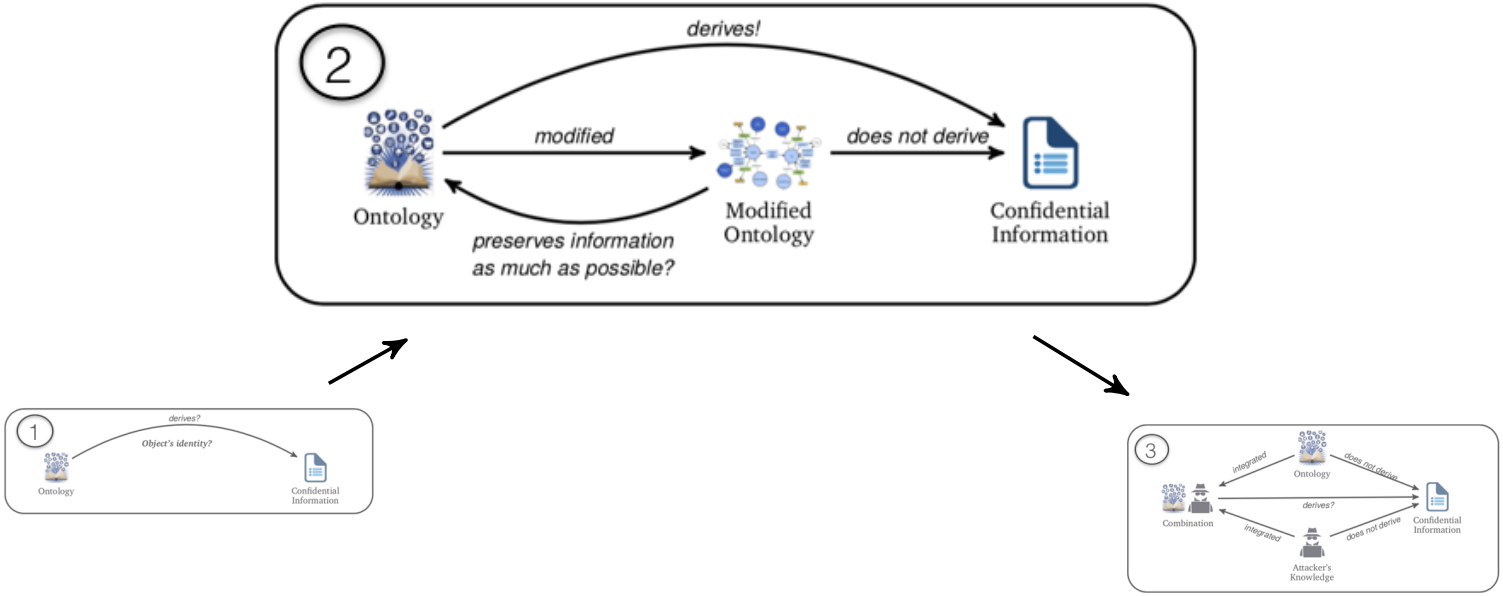
k-Hiding

The anonymous individual x is **not k-hidden** w.r.t. $\mathfrak{O}$ iff there are known individuals $a_1, \ldots, a_{k-1}$ such that

$$x \text{ belongs to } \{a_1, \ldots, a_{k-1}\} \text{ w.r.t. } \mathfrak{O}$$

If (variants) of the identity problem can be reduced to classical reasoning problems in DLs, then now let's consider ***more general types of confidential axioms*** *(e.g., instance relationships, subsumptions, CQs, etc).*

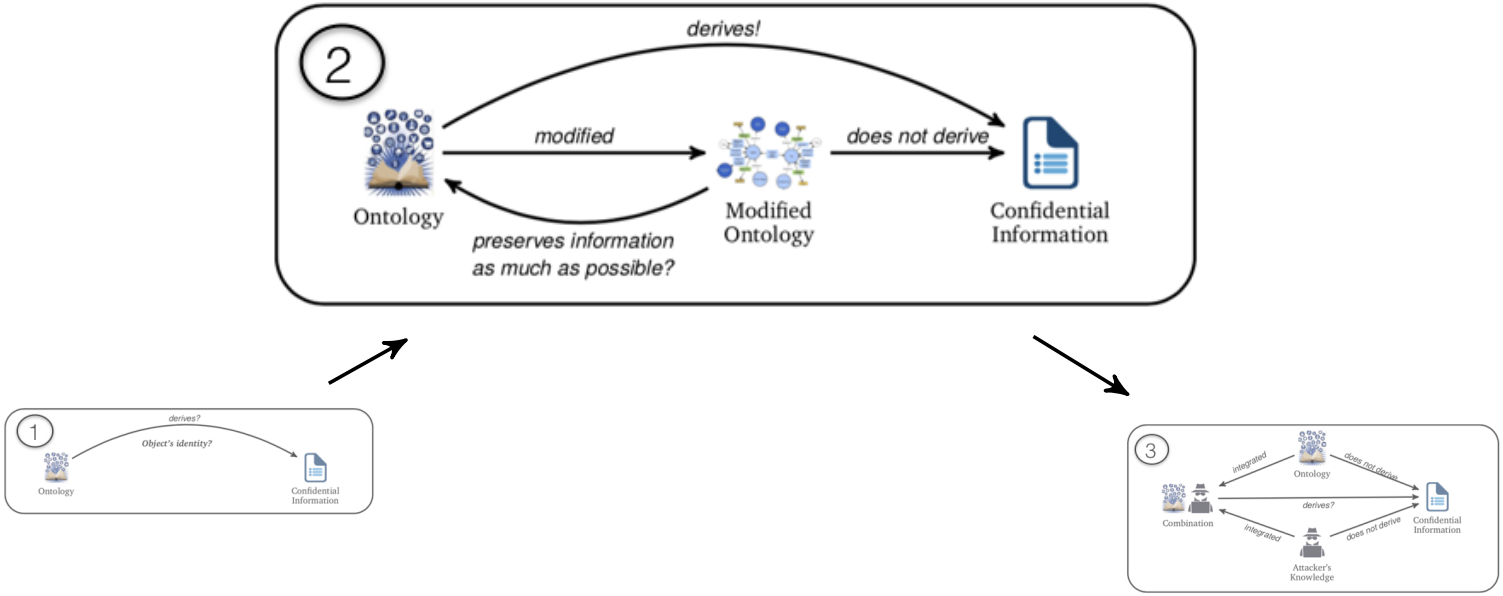
Problem 2: How to Protect the Confidential Information?



Ontology Repair ([KR 2018], [DL 2018])

- $\mathfrak{O} = \mathfrak{O}_s \cup \mathfrak{O}_r$, where $\mathfrak{O}_s$ is a **static ontology** and $\mathfrak{O}_r$ is a **refutable ontology**.
- Let $Con(\mathfrak{O}) := \{\alpha \mid \mathfrak{O} \models \alpha\}$ be the set of all **consequences** of $\mathfrak{O}$.

Problem 2: How to Protect the Confidential Information?

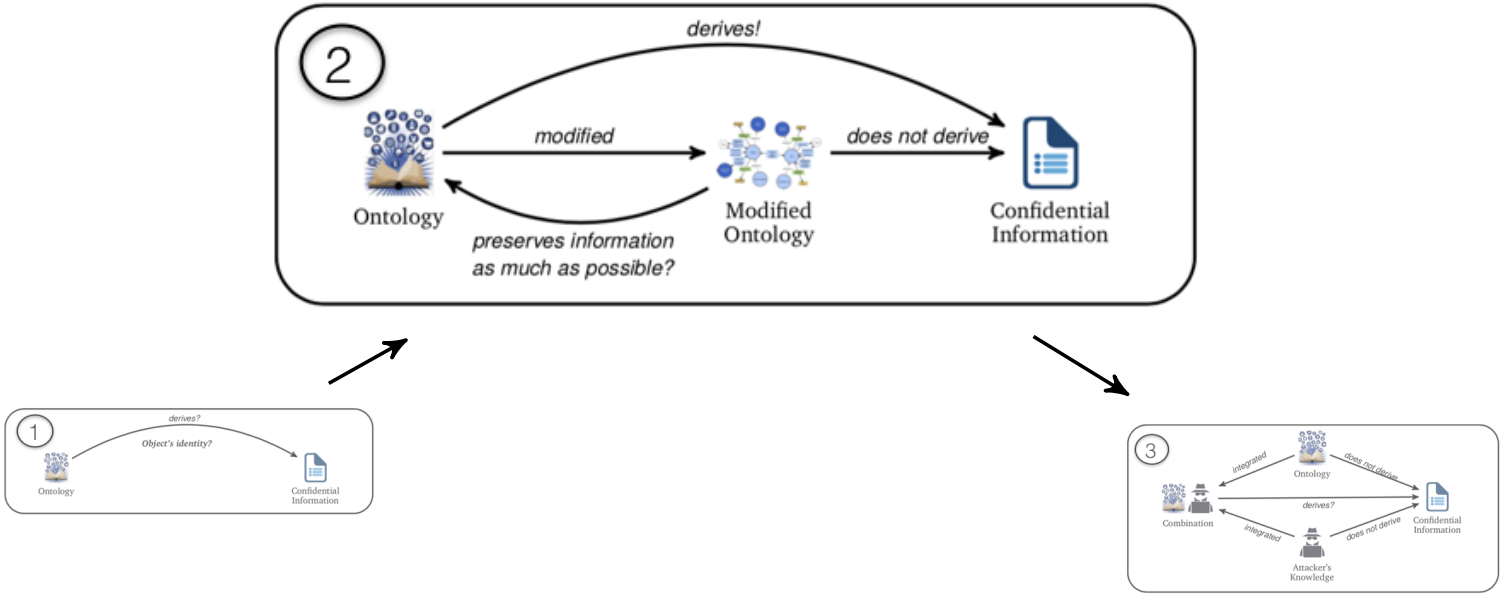


Ontology Repair ([KR 2018], [DL 2018])

- $\mathfrak{O} = \mathfrak{O}_s \cup \mathfrak{O}_r$, where $\mathfrak{O}_s$ is a **static ontology** and $\mathfrak{O}_r$ is a **refutable ontology**.
- Let $\mathit{Con}(\mathfrak{O}) := \{\alpha \mid \mathfrak{O} \models \alpha\}$ be the set of all **consequences** of $\mathfrak{O}$.
- Let $\mathfrak{O} \models \alpha$ and $\mathfrak{O}_s \not\models \alpha$. The ontology $\mathfrak{O}'$ is a **repair** of $\mathfrak{O}$ w.r.t. α if

$$\mathit{Con}(\mathfrak{O}_s \cup \mathfrak{O}') \subseteq \mathit{Con}(\mathfrak{O}) \setminus \{\alpha\}$$

Problem 2: How to Protect the Confidential Information?



Ontology Repair ([KR 2018], [DL 2018])

- $\mathfrak{O} = \mathfrak{O}_s \cup \mathfrak{O}_r$, where $\mathfrak{O}_s$ is a **static ontology** and $\mathfrak{O}_r$ is a **refutable ontology**.
- Let $\mathit{Con}(\mathfrak{O}) := \{\alpha \mid \mathfrak{O} \models \alpha\}$ be the set of all **consequences** of $\mathfrak{O}$.
- Let $\mathfrak{O} \models \alpha$ and $\mathfrak{O}_s \not\models \alpha$. The ontology $\mathfrak{O}'$ is a **repair** of $\mathfrak{O}$ w.r.t. α if

$$\mathit{Con}(\mathfrak{O}_s \cup \mathfrak{O}') \subseteq \mathit{Con}(\mathfrak{O}) \setminus \{\alpha\}$$

- **Optimal repair** $\mathfrak{O}'$ of $\mathfrak{O}$ w.r.t. α:
  No Repair $\mathfrak{O}''$ of $\mathfrak{O}$ w.r.t. α such that $\mathit{Con}(\mathfrak{O}' \cup \mathfrak{O}_s) \subset \mathit{Con}(\mathfrak{O}'' \cup \mathfrak{O}_s)$.

Optimal Classical Repairs

Optimal Repairs need not exist in general!

Optimal Classical Repair

A maximum subset $\mathfrak{O}'$ of $\mathfrak{O}_r$ such that $\mathfrak{O}_s \cup \mathfrak{O}' \not\models \alpha$

Optimal Classical Repairs

Optimal Repairs need not exist in general!

Optimal Classical Repair

A maximum subset $\mathfrak{O}'$ of $\mathfrak{O}_r$ such that $\mathfrak{O}_s \cup \mathfrak{O}' \not\models \alpha$

- Optimal classical repairs always exist $\rightarrow$ **Justification** and **Hitting Set** **(Reiter, 1987)**
- Let $\mathfrak{O} \models \alpha$. A **justification** J of $\mathfrak{O}$ w.r.t. α is a minimal subset of $\mathfrak{O}_r$ s.t. $\mathfrak{O}_s \cup J \models \alpha$.
- Let $J_1, \ldots, J_k$ be the justifications of $\mathfrak{O}$ w.r.t. α.
  A **hitting set** $\mathcal{H}$ of these justifications is a set of axioms such that $\mathcal{H} \cap J_i \neq \emptyset$
- A hitting set $\mathcal{H}_{min}$ is **minimal** if there is no $\mathcal{H}'$ of $J_1, \ldots, J_k$ such that $\mathcal{H}' \subset \mathcal{H}_{min}$.
- $\mathfrak{O}' := \mathfrak{O}_r \setminus \mathcal{H}_{min}$ is an **optimal classical repair** of $\mathfrak{O}$ w.r.t. α such that

$$\mathfrak{O}_s \cup \mathfrak{O}' \not\models \alpha$$

Gentle Repair

Obtaining Classical Repairs $\rightarrow$ **removing axioms** from $\mathfrak{O}$.

Instead, we want to **weaken axioms** in $\mathcal{H} \Rightarrow$ **Gentle Repair**!

Given axioms β, γ, an axiom γ is **weaker than** β if $Con(\{\gamma\}) \subset Con(\{\beta\})$

Illustration

Gentle Repair

Illustration

$$\mathfrak{O}_s := \{\exists receives.(Gift \sqcap Deluxe) \sqsubseteq \exists gets.Bribe\}$$

$$\mathfrak{O}_r := \{IndonesianPolitician \sqsubseteq \exists receives.(Gift \sqcap Deluxe)\}$$

- Every Indonesian politician is bribed w.r.t. $\mathfrak{O}_s \cup \mathfrak{O}_r$.

Gentle Repair

Illustration

$$\mathfrak{O}_s := \{\exists receives.(Gift \sqcap Deluxe) \sqsubseteq \exists gets.Bribe\}$$

$$\mathfrak{O}_r := \{IndonesianPolitician \sqsubseteq \exists receives.(Gift \sqcap Deluxe)\}$$

- Every Indonesian politician is bribed w.r.t. $\mathfrak{O}_s \cup \mathfrak{O}_r$.
- **Classical**: Removes a "common knowledge":
  $IndonesianPolitician \sqsubseteq \exists receives.(Gift \sqcap Deluxe)$

Gentle Repair

Illustration

$$\mathfrak{O}_s := \{\exists receives.(Gift \sqcap Deluxe) \sqsubseteq \exists gets.Bribe\}$$

$$\mathfrak{O}_r := \{IndonesianPolitician \sqsubseteq \exists receives.(Gift \sqcap Deluxe)\}$$

- Every Indonesian politician is bribed w.r.t. $\mathfrak{O}_s \cup \mathfrak{O}_r$.
- **Classical**: Removes a "common knowledge":
  $IndonesianPolitician \sqsubseteq \exists receives.(Gift \sqcap Deluxe)$
- **Gentle**: Weaken $\beta \in \mathfrak{O}_r$ to $IndonesianPolitician \sqsubseteq \exists receives.Gift$
  But, this consequence $IndonesianPolitician \sqsubseteq \exists receives.Deluxe$ is also gone.

Gentle Repair

Illustration

$$\mathfrak{O}_s := \{\exists receives.(Gift \sqcap Deluxe) \sqsubseteq \exists gets.Bribe\}$$
$$\mathfrak{O}_r := \{IndonesianPolitician \sqsubseteq \exists receives.(Gift \sqcap Deluxe)\}$$

- Every Indonesian politician is bribed w.r.t. $\mathfrak{O}_s \cup \mathfrak{O}_r$.
- **Classical**: Removes a "common knowledge":
  $IndonesianPolitician \sqsubseteq \exists receives.(Gift \sqcap Deluxe)$
- **Gentle**: Weaken $\beta \in \mathfrak{O}_r$ to $IndonesianPolitician \sqsubseteq \exists receives.Gift$
  But, this consequence $IndonesianPolitician \sqsubseteq \exists receives.Deluxe$ is also gone.
- **More gentle**: Weaken β to
  $IndonesianPolitician \sqsubseteq \exists receives.Gift \sqcap \exists receives.Deluxe$

How to Make it Gentle?

Gentle Repair Algorithm: [BaKrNuPe, KR 2018]

- Take all justifications and one minimal hitting set $\mathcal{H}_{min}$
- For each $\beta \in \mathcal{H}_{min}$ and all $J_1, \ldots, J_k$ containing β,
  replace β with **exactly one** γ, where γ is weaker than β such that

$$\mathfrak{O}_s \cup (J_i \setminus \{\beta\}) \cup \{\gamma\} \not\models \alpha \text{ for } i = 1, \ldots, k. \tag{1}$$

  γ always exists.
- **Construct** $\mathfrak{O}'$ **obtained** from $\mathfrak{O}_r$ by **replacing** each $\beta \in \mathcal{H}_{min}$ with an appropriate weaker γ satisfying (1).
- **Check** if α is a consequence of $\mathfrak{O}_s \cup \mathfrak{O}'$.

How to Make it Gentle?

Gentle Repair Algorithm: [BaKrNuPe, KR 2018]

- Take all justifications and one minimal hitting set $\mathcal{H}_{min}$
- For each $\beta \in \mathcal{H}_{min}$ and all $J_1, \ldots, J_k$ containing β, **replace** β with **exactly one** γ, where γ is weaker than β such that

$$\mathfrak{O}_s \cup (J_i \setminus \{\beta\}) \cup \{\gamma\} \not\models \alpha \text{ for } i = 1, \ldots, k. \tag{1}$$

  γ always exists.
- **Construct** $\mathfrak{O}'$ **obtained** from $\mathfrak{O}_r$ by **replacing** each $\beta \in \mathcal{H}_{min}$ with an appropriate weaker γ satisfying (1).
- **Check** if α is a consequence of $\mathfrak{O}_s \cup \mathfrak{O}'$.

Obtaining Gentle Repairs needs Iterations

- Using the algorithm above, α still can be a consequence of $\mathfrak{O}_s \cup \mathfrak{O}'$.
- Solution: Just **iterate** Gentle Repair Algorithm until $\mathfrak{O}_s \cup \mathfrak{O}' \not\models \alpha$.
- The iterative algorithm yields **an exponential upper bound** on the number of iterations.

Weakening Relations

To obtain better bounds on the number of iterations, introduce weakening relations on axioms.

Weakening Relation

Weakening Relations

To obtain better bounds on the number of iterations, introduce weakening relations on axioms.

Weakening Relation

The binary relation $\succ$ on axioms is

- a **weakening relation** if $\beta \succ \gamma$ implies that γ is weaker than β;
- **well-founded** if there is no infinite $\succ$-chain $\beta_1 \succ \beta_2 \succ \beta_3 \succ \ldots$;
- **complete** if for any β that is not a tautology, there is a tautology γ s.t. $\beta \succ \gamma$.
- **linear (polynomial)** if for every axiom β, the length of the longest chain $\succ$-generated from β is **linearly (polynomially)** bounded by the size of β;

Weakening Relations

Weakening Relation

The binary relation $\succ$ on axioms is

- a **weakening relation** if $\beta \succ \gamma$ implies that γ is weaker than β;
- **well-founded** if there is no infinite $\succ$-chain $\beta_1 \succ \beta_2 \succ \beta_3 \succ \ldots$;
- **complete** if for any β that is not a tautology, there is a tautology γ s.t. $\beta \succ \gamma$.
- **linear (polynomial)** if for every axiom β, the length of the longest chain $\succ$-generated from β is **linearly (polynomially)** bounded by the size of β;

Weakening relations making **larger steps** may **decrease** the number of iterations

Weakening relations making **smaller steps** may make the repair more gentle

Maximally Strong Weakening Axioms

Replace β with exactly one weaker γ s.t.

$$\mathfrak{O}_s \cup (J_i \setminus \{\beta\}) \cup \{\gamma\} \not\models \alpha \text{ for } i = 1, \ldots, k$$

If γ is a tautology, then it is the same as classical repair.

To make this repair as gentle as possible, γ should be **maximally strong**

$$\mathfrak{O}_s \cup (J_i \setminus \{\beta\}) \cup \{\gamma\} \not\models \alpha$$
but for all δ such that $\beta \succ \delta \succ \gamma$, we have
$$\mathfrak{O}_s \cup (J_i \setminus \{\beta\}) \cup \{\delta\} \models \alpha$$

Maximally Strong Weakening Axioms

Replace β with exactly one weaker γ s.t.

$$\mathfrak{O}_s \cup (J_i \setminus \{\beta\}) \cup \{\gamma\} \not\models \alpha \text{ for } i = 1, \ldots, k$$

If γ is a tautology, then it is the same as classical repair.

To make this repair as gentle as possible, γ should be **maximally strong**

$$\mathfrak{O}_s \cup (J_i \setminus \{\beta\}) \cup \{\gamma\} \not\models \alpha$$
but for all δ such that $\beta \succ \delta \succ \gamma$, we have
$$\mathfrak{O}_s \cup (J_i \setminus \{\beta\}) \cup \{\delta\} \models \alpha$$

Do they always exists?

How to compute them?

Weakening Relations in $\mathcal{EL}$

Focus on GCIs and generalize the right-hand side of GCIs.

A Weakening Relation $\succ^{sub}$

$C \sqsubseteq D \succ^{sub} C' \sqsubseteq D'$ if $C' = C$, $D \sqsubset D'$, and
$\{C' \sqsubseteq D'\} \not\models C \sqsubseteq D$.

$D \sqsubset^{syn} D' \Rightarrow$ removing occurrences of subconcepts of D.

A Weakening Relation $\succ^{syn}$

$C \sqsubseteq D \succ^{syn} C' \sqsubseteq D'$ if $C' = C$ and $D \sqsubset^{syn} D'$, and
$\{C' \sqsubseteq D'\} \not\models C \sqsubseteq D$.

Weakening Relations in $\mathcal{EL}$

Focus on GCIs and generalize the right-hand side of GCIs.

A Weakening Relation $\succ^{sub}$

$C \sqsubseteq D \succ^{sub} C' \sqsubseteq D'$ if $C' = C$, $D \sqsubset D'$, and
$\{C' \sqsubseteq D'\} \not\models C \sqsubseteq D$.

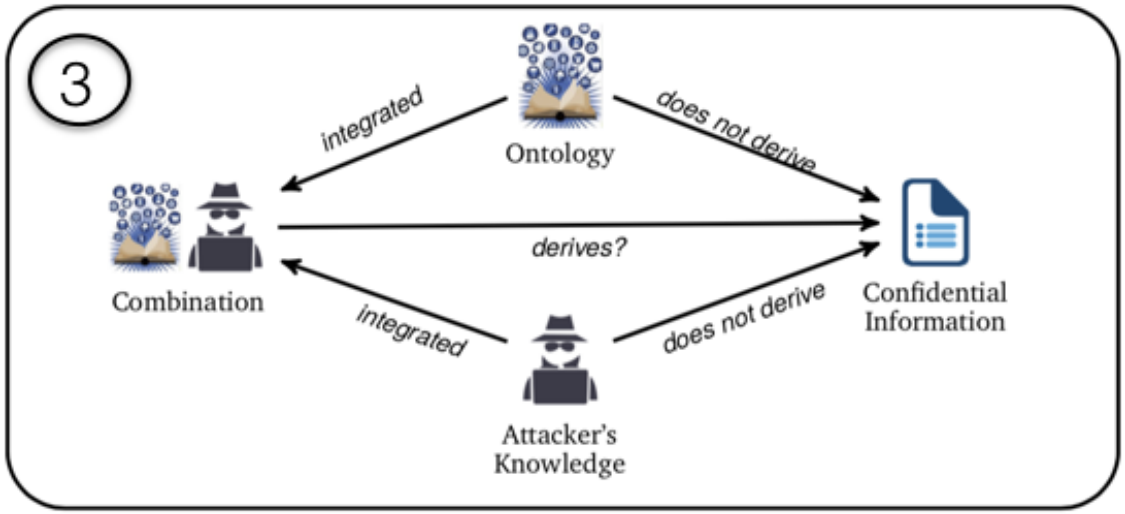
Employing both, maximally strong weakenings can be effectively computed

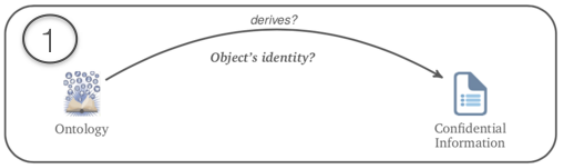
$D \sqsubset^{syn} D' \Rightarrow$ removing occurrences of subconcepts of D.

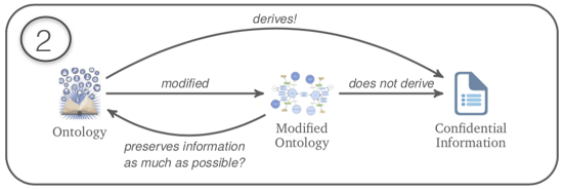
A Weakening Relation $\succ^{syn}$

$C \sqsubseteq D \succ^{syn} C' \sqsubseteq D'$ if $C' = C$ and $D \sqsubset^{syn} D'$, and
$\{C' \sqsubseteq D'\} \not\models C \sqsubseteq D$.

Problem 3: Privacy-Preserving Ontology Publishing (PPOP)







PPOP for $\mathcal{EL}$ Ontologies ([DL 2018], [JELIA 2019], [KI 2019])

Restricting the ontology:

- $\mathcal{EL}$ Instance Stores & $\mathcal{EL}$ ABoxes (**No TBoxes**)
- **Instance Stores**: Ontologies without individual relationships

PPOP for $\mathcal{EL}$ Instance Stores

$\mathcal{EL}$ Instance Stores without TBox

$C_1(a), C_2(a)$ implies $(C_1 \sqcap C_2)(a)$

only one concept assertion speaking about one individual

Published Information

(an $\mathcal{EL}$ Concept C)

Attacker's Knowledge

(an $\mathcal{EL}$ / $\mathcal{FL}_0$ / $\mathcal{FLE}$ Concept E)

Confidential Information

(a finite set of $\mathcal{EL}$ concepts)

$\{D_1, \ldots, D_p\}$

Privacy Attacks in $\mathcal{EL}$ Instance Stores

Confidential Information $\mathcal{P} = \{D\}$ about *LINDA*

$$D = Patient \sqcap \exists seenBy.(Doctor \sqcap \exists worksIn.Oncology)$$

Original **Published Information** C about *LINDA*

$$C = Patient \sqcap Female \sqcap \exists seenBy.(Doctor \sqcap Male \sqcap \exists worksIn.Oncology)$$

Note C is not **compliant with** D

Privacy Attacks in $\mathcal{EL}$ Instance Stores

Confidential Information $\mathcal{P} = \{D\}$ about *LINDA*

$$D = Patient \sqcap \exists seenBy.(Doctor \sqcap \exists worksIn.Oncology)$$

Original **Published Information** C about *LINDA*

$$C = Patient \sqcap Female \sqcap \exists seenBy.(Doctor \sqcap Male \sqcap \exists worksIn.Oncology)$$

Note C is not **compliant with** D

Modification

$$C_1 = Female \sqcap \exists seenBy.(Doctor \sqcap Male \sqcap \exists worksIn.Oncology)$$

Note $C \sqsubseteq C_1$ and C_1 **complies with** D

Privacy Attacks in $\mathcal{EL}$ Instance Stores

Confidential Information $\mathcal{P} = \{D\}$ about *LINDA*

$$D = Patient \sqcap \exists seenBy.(Doctor \sqcap \exists worksIn.Oncology)$$

Original **Published Information** C about *LINDA*

$$C = Patient \sqcap Female \sqcap \exists seenBy.(Doctor \sqcap Male \sqcap \exists worksIn.Oncology)$$

Note C is not **compliant with** D

$\mathcal{EL}$-Attacker is Coming!

$$C_1 = Female \sqcap \exists seenBy.(Doctor \sqcap Male \sqcap \exists worksIn.Oncology)$$

He knows $Patient(LINDA)$

Privacy Attacks in $\mathcal{EL}$ Instance Stores

Confidential Information $\mathcal{P} = \{D\}$ about *LINDA*

$$D = Patient \sqcap \exists seenBy.(Doctor \sqcap \exists worksIn.Oncology)$$

Original **Published Information** C about *LINDA*

$$C = Patient \sqcap Female \sqcap \exists seenBy.(Doctor \sqcap Male \sqcap \exists worksIn.Oncology)$$

Note C is not **compliant with** D

Linked and Revealed!

$$C_1' = Female \sqcap \exists seenBy.(Doctor \sqcap Male \sqcap \exists worksIn.Oncology) \sqcap \textbf{\textit{Patient}}$$

Note $D(LINDA)$ is **revealed** and C_1 is not $\mathcal{EL}$-safe for D

Privacy Attacks in $\mathcal{EL}$ Instance Stores

Confidential Information $\mathcal{P} = \{D\}$ about *LINDA*

$$D = Patient \sqcap \exists seenBy.(Doctor \sqcap \exists worksIn.Oncology)$$

Original **Published Information** C about *LINDA*

$$C = Patient \sqcap Female \sqcap \exists seenBy.(Doctor \sqcap Male \sqcap \exists worksIn.Oncology)$$

Note C is not **compliant with** D

Modification

$$C_2 = Female \sqcap \exists seenBy.(Doctor \sqcap Male \sqcap \exists worksIn.\top) \sqcap \exists seenBy.(Male \sqcap worksIn.Oncology)$$

C_2 is $\mathcal{EL}$-**safe** for D

Privacy Attacks in $\mathcal{EL}$ Instance Stores

Confidential Information $\mathcal{P} = \{D\}$ about *LINDA*

$$D = Patient \sqcap \exists seenBy.(Doctor \sqcap \exists worksIn.Oncology)$$

Original **Published Information** C about *LINDA*

$$C = Patient \sqcap Female \sqcap \exists seenBy.(Doctor \sqcap Male \sqcap \exists worksIn.Oncology)$$

Note C is not **compliant with** D

$\mathcal{FL}_0$**-Attacker is Coming!**

$$C_2 = Female \sqcap \exists seenBy.(Doctor \sqcap Male \sqcap \exists worksIn.\top) \sqcap \exists seenBy.(Male \sqcap worksIn.Oncology)$$

He knows $(Patient \sqcap \forall seenBy.\forall worksIn.Oncology)(LINDA)$

Privacy Attacks in $\mathcal{EL}$ Instance Stores

Confidential Information $\mathcal{P} = \{D\}$ about *LINDA*

$$D = Patient \sqcap \exists seenBy.(Doctor \sqcap \exists worksIn.Oncology)$$

Original **Published Information** C about *LINDA*

$$C = Patient \sqcap Female \sqcap \exists seenBy.(Doctor \sqcap Male \sqcap \exists worksIn.Oncology)$$

Note C is not **compliant with** D

Linked and Revealed!

$$C_2' = Female \sqcap \exists seenBy.(Doctor \sqcap Male \sqcap \exists worksIn.\top) \sqcap \exists seenBy.(Male \sqcap worksIn.Oncology) \sqcap \mathbf{Patient \sqcap \forall seenBy.\forall worksIn.Oncology}$$

$D(LINDA)$ is revealed again and C_2 is not $\mathcal{FL}_0$-safe for D

Privacy Attacks in $\mathcal{EL}$ Instance Stores

Confidential Information $\mathcal{P} = \{D\}$ about *LINDA*

$$D = Patient \sqcap \exists seenBy.(Doctor \sqcap \exists worksIn.Oncology)$$

Original **Published Information** C about *LINDA*

$$C = Patient \sqcap Female \sqcap \exists seenBy.(Doctor \sqcap Male \sqcap \exists worksIn.Oncology)$$

Note C is not **compliant with** D

Modification

$$C_3 = Female \sqcap Patient \sqcap \exists seenBy.(Doctor \sqcap Male)$$

C_3 is $\mathcal{FL}_0$-**safe** for D

Decision & Computational Problems for Instance Stores

Given $\mathcal{L} \in \{\mathcal{EL}, \mathcal{FL}_0, \mathcal{FLE}\}$, a published information ($\mathcal{EL}$ concept) C, an $\mathcal{EL}$ confidential information $\mathcal{P}$.

Decision Problems

- **Compliance**:
  Is an $\mathcal{EL}$ concept C compliant with $\mathcal{P}$?
- **$\mathcal{L}$-Safety**:
  Is an $\mathcal{EL}$ concept C $\mathcal{L}$-safe for $\mathcal{P}$?
- **OptCom**:
  Is an $\mathcal{EL}$ concept C_1 an optimal compliant generalization of C w.r.t. $\mathcal{P}$?
- **$\mathcal{L}$-Optimality**:
  Is an $\mathcal{EL}$ concept C_1 an optimal $\mathcal{L}$-safe generalization of C for $\mathcal{P}$?

Note

Optimal: For all C_2, if $C_2 \sqsubset C_1$, then C_2 is not (compliant) $\mathcal{L}$-safe w.r.t. $\mathcal{P}$.

Decision & Computational Problems for Instance Stores

Given $\mathcal{L} \in \{\mathcal{EL}, \mathcal{FL}_0, \mathcal{FLE}\}$, a published information ($\mathcal{EL}$ concept) C, an $\mathcal{EL}$ confidential information $\mathcal{P}$.

Decision Problems

- **Compliance**:
  Is an $\mathcal{EL}$ concept C compliant with $\mathcal{P}$?
- **$\mathcal{L}$-Safety**:
  Is an $\mathcal{EL}$ concept C $\mathcal{L}$-safe for $\mathcal{P}$?
- **OptCom**:
  Is an $\mathcal{EL}$ concept C_1 an optimal compliant generalization of C w.r.t. $\mathcal{P}$?
- **$\mathcal{L}$-Optimality**:
  Is an $\mathcal{EL}$ concept C_1 an optimal $\mathcal{L}$-safe generalization of C for $\mathcal{P}$?

Computational Problem

Find an $\mathcal{EL}$ concept C_1 s.t C_1 is an optimal (compliant) $\mathcal{L}$-safe generalization of C for $\mathcal{P}$!

Complexity Results on PPOP for $\mathcal{EL}$ Instance Stores

Compliance is in PTime, whereas OptCom is in coNP, but Dual-hard.

Decision Problems	$\mathcal{L} = \mathcal{EL}$	$\mathcal{L} = \mathcal{FL}_0$	$\mathcal{L} = \mathcal{FLE}$
$\mathcal{L}$-safety	PTime	PTime	PTime
$\mathcal{L}$-optimality	coNP and Dual-hard	coNP and Dual-hard	PTime

Table: Complexity results of $\mathcal{L}$-safety and $\mathcal{L}$-optimality on PPOP for $\mathcal{EL}$ instance stores

Optimal Compliance Generalization(s) can be computed in ExpTime.

Computational Problems	$\mathcal{L} = \mathcal{EL}$	$\mathcal{L} = \mathcal{FL}_0$	$\mathcal{L} = \mathcal{FLE}$
Optimal $\mathcal{L}$-safe Generalization(s)	ExpTime	ExpTime	PTime

Table: Complexity of computing one/all optimal Q-safe generalizations for $\mathcal{P}$

PPOP for $\mathcal{EL}$ ABoxes

Including relationships between individuals in $\mathcal{EL}$ ABoxes.

Published Information

(an $\mathcal{EL}$ ABox)

Attacker's Knowledge

(an $\mathcal{EL}$ ABox)

Confidential Information

(an $\mathcal{EL}$ concept or a conjunctive query)

PPOP for $\mathcal{EL}$ ABoxes

Including relationships between individuals in $\mathcal{EL}$ ABoxes.

Published Information
(an $\mathcal{EL}$ ABox)

Attacker's Knowledge
(an $\mathcal{EL}$ ABox)

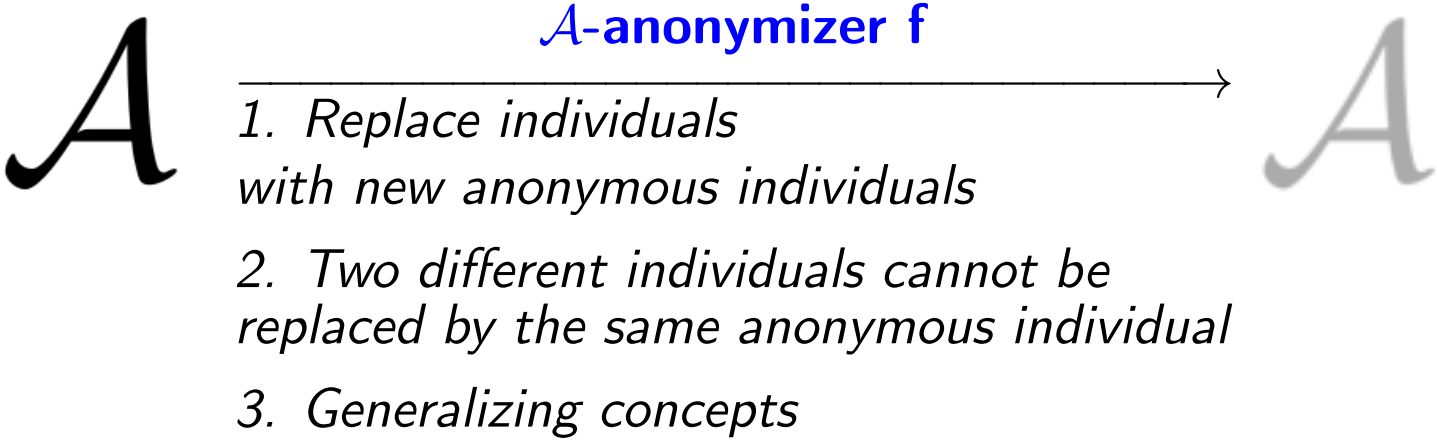
Confidential Information
(an $\mathcal{EL}$ concept or a conjunctive query)

Given an $\mathcal{EL}$ ABox $\mathcal{A}$, and a confidential information $\mathcal{P}$ that is either an **instance query** ($\mathcal{EL}$ concept) D or a **conjunctive query** q.

- $\mathcal{A}$ is **compliant** with D iff $\mathcal{A} \not\models D(a)$ for all individuals a.
- $\mathcal{A}$ is **compliant** with q iff $\mathcal{A} \not\models q(\vec{a})$ for all tuples $\vec{a}$ of individuals.
- $\mathcal{A}$ is **safe** for $\mathcal{P}$ iff for all (attackers' knowledge) $\mathcal{A}'$ complying with $\mathcal{P}$, $\mathcal{A} \cup \mathcal{A}'$ complies with $\mathcal{P}$

Anonymizing EL ABoxes

How to modify $\mathcal{EL}$ ABoxes?



Anonymizing EL ABoxes

$\mathcal{A}$-anonymizer f

1. *Replace individuals with new anonymous individuals*
2. *Two different individuals cannot be replaced by the same anonymous individual*
3. *Generalizing concepts*

ABox Anonymization

$$\mathcal{A}_0 := \{\, Doctor \sqcap \exists worksIn.Oncology(LINDA),$$
$$seenBy(BOB, LINDA)\}$$
$$\downarrow f_1 \checkmark$$
$$\mathcal{A}_1 := \{\, Doctor \sqcap \exists worksIn.Oncology(y),$$
$$seenBy(x, LINDA)\}$$

Anonymizing EL ABoxes

$\mathcal{A}$-anonymizer f

1. *Replace individuals with new anonymous individuals*
2. *Two different individuals cannot be replaced by the same anonymous individual*
3. *Generalizing concepts*

ABox Anonymization

$$\mathcal{A}_0 := \{\, Doctor \sqcap \exists worksIn.Oncology(LINDA),\ seenBy(BOB, LINDA)\}$$

$$\downarrow f_2 \ \textbf{X}$$

$$\mathcal{A}_2 := \{\, Doctor \sqcap \exists worksIn.Oncology(z),\ seenBy(z, LINDA)\}$$

Anonymizing EL ABoxes

$\mathcal{A}$-anonymizer f

1. *Replace individuals with new anonymous individuals*
2. *Two different individuals cannot be replaced by the same anonymous individual*
3. *Generalizing concepts*

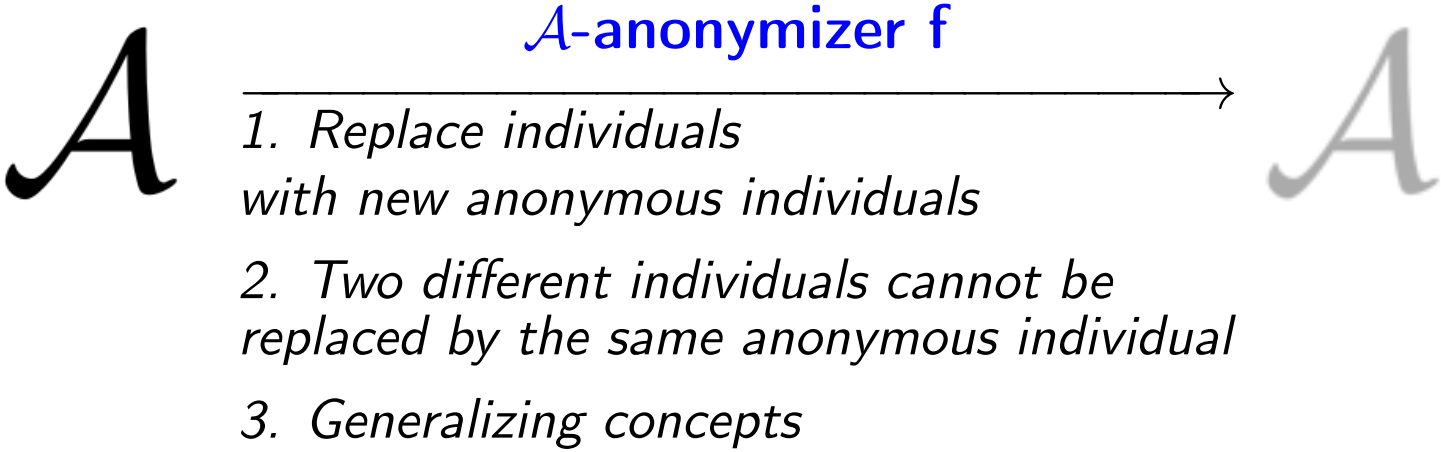
ABox Anonymization

$$\mathcal{A}_0 := \{\, \mathit{Doctor} \sqcap \exists \mathit{worksIn}.\mathit{Oncology}(\mathit{LINDA}),\ \mathit{seenBy}(\mathit{BOB}, \mathit{LINDA})\}$$

$$\downarrow f_3 \checkmark$$

$$\mathcal{A}_3 := \{\, \mathit{Doctor} \sqcap \exists \mathit{worksIn}.\top(y),\ \mathit{seenBy}(\mathit{BOB}, \mathit{LINDA})\}$$

Optimality in Anonymizations



Measuring Optimality

An $\mathcal{A}$-anonymizer f_2 is **more informative than** an $\mathcal{A}$-anonymizer f_1 ($f_2 > f_1$) if f_2 can be obtained from f_1 by:

- keeping more known individuals
- identifying more distinct anonymous individuals
- specializing more $\mathcal{EL}$ concepts

Decision Problems on PPOP for $\mathcal{EL}$ ABoxes

Given an $\mathcal{EL}$ ABox $\mathcal{A}$, an $\mathcal{EL}$ concept D, and an $\mathcal{A}$-anonymizer f,

- **Compliance**$_{IQ}$, **Safety**$_{IQ}$, and
- **Optimal-Compliance**$_{IQ}$ (**Optimal-Safety**$_{IQ}$) asks
  - if $f(\mathcal{A})$ is compliant with (safe for) D and
  - for all $\mathcal{A}$-anonymizers f', if $f' > f$, then $f'(\mathcal{A})$ is not compliant with (safe for) D

Analogous for **Compliance**$_{CQ}$, **Safety**$_{CQ}$, **Optimal-Compliance**$_{CQ}$, and **Optimal-Safety**$_{CQ}$, where the policy is a CQ

Decision Problems on PPOP for $\mathcal{EL}$ ABoxes

Given an $\mathcal{EL}$ ABox $\mathcal{A}$, an $\mathcal{EL}$ concept D, and an $\mathcal{A}$-anonymizer f,

- **Compliance**$_{IQ}$, **Safety**$_{IQ}$, and
- **Optimal-Compliance**$_{IQ}$ (**Optimal-Safety**$_{IQ}$) asks
  - if $f(\mathcal{A})$ is compliant with (safe for) D and
  - for all $\mathcal{A}$-anonymizers f', if $f' > f$, then $f'(\mathcal{A})$ is not compliant with (safe for) D

Analogous for **Compliance**$_{CQ}$, **Safety**$_{CQ}$, **Optimal-Compliance**$_{CQ}$, and **Optimal-Safety**$_{CQ}$, where the policy is a CQ

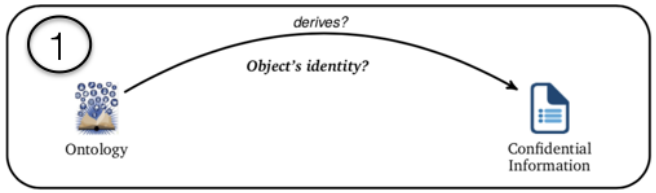
Decision Problems	**X = IQ**	**X = CQ**
Compliance$_X$	PTime	coNP-complete
Safety$_X$	PTime	Π_2^p and DP-hard
Optimal-Compliance$_X$	coNP and Dual-hard	Π_2^p and DP-hard
Optimal-Safety$_X$	coNP and Dual-hard	Π_3^p and DP-hard

Table: Complexity Results on PPOP in $\mathcal{EL}$ ABoxes

Conclusions

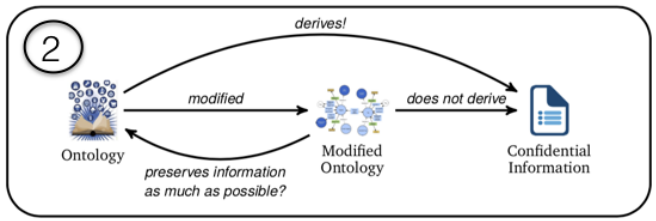
The Identity Problem:

- Non trivial for DLs with equality power
- Introducing variants of the identity problem
- Reduction to classical reasoning in DLs



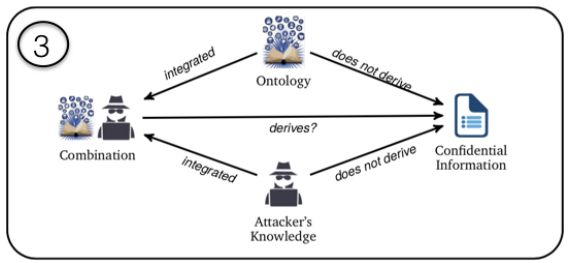
Gentle Repair:

- Introducing a framework for repair via axiom weakening
- Weakening relations
- Weakening axioms in $\mathcal{EL}$



Privacy-Preserving Ontology Publishing:

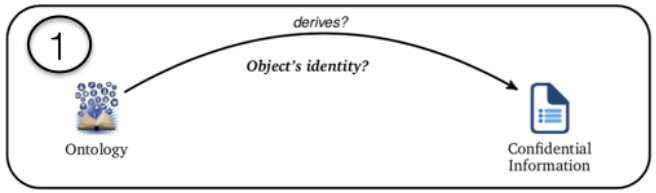
- PPOP for $\mathcal{EL}$ Instance Stores
- PPOP for $\mathcal{EL}$ ABoxes
- Applying the concepts of compliance, safety, and optimality in both settings



Future Work

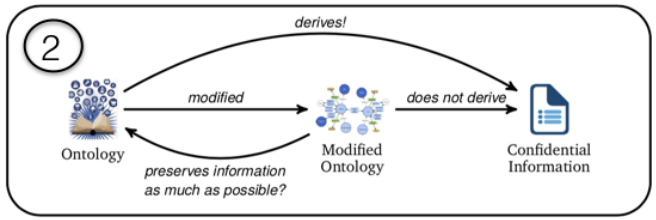
The Identity Problem:

- Formalizing the "real" definition of k-Anonymity
- Adding probability to the setting



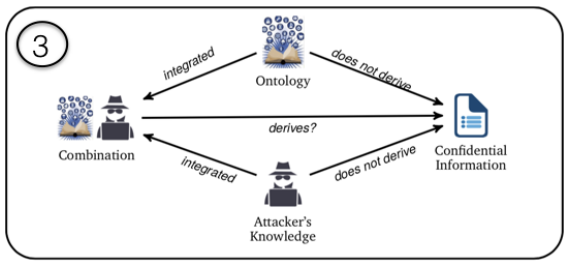
Gentle Repair:

- Choosing which axioms to be repaired
- Which maximally strong weakening is the best?
- Weakening relations for other DLs



Privacy-Preserving Ontology Publishing:

- Computing the optimal compliant (safe) anonymization
- Finding a more gentle weakening relation for ABox anonymization
- Including TBox/attackers' meta knowledge? (Bonatti et. al., 2013)



Publications

- Franz Baader, Daniel Borchmann, and **Adrian Nuradiansyah**, ***Preliminary Results on the Identity Problem in Description Logic Ontologies***, DL 2017, Montpellier, 2017.
- Franz Baader, Daniel Borchmann, and **Adrian Nuradiansyah**, ***The Identity Problem in Description Logic Ontologies and Its Applications to View-Based Information Hiding***, JIST 2017, Gold Coast, 2017.
- Franz Baader, Francesco Kriegel, **Adrian Nuradiansyah**, and Rafael Peñaloza, ***Making Repairs in Description Logics More Gentle***, KR 2018, Tempe, 2018.
- Franz Baader and **Adrian Nuradiansyah**, ***Towards Privacy-Preserving Ontology Publishing***, DL 2018, Tempe, 2018.
- Franz Baader, Francesco Kriegel, and **Adrian Nuradiansyah**, ***Privacy-Preserving Ontology Publishing for $\mathcal{EL}$ Instance Stores***, JELIA 2019, Rende, 2019.
- Franz Baader and **Adrian Nuradiansyah**, ***Mixing Description Logics in Privacy-Preserving Ontology Publishing***, KI 2019, Kassel, 2019.

Research Visits and Awards

Research Visits:

- Visiting Prof. **Rafael Peñaloza** at Free University of Bozen-Bolzano, March 1-May 16, 2018.
- Visiting Prof. **Bernardo Cuenca Grau** at the University of Oxford, UK, April 1 - June 30, 2019.

Awards:

- **The Best Student Paper Award** at the 7th Joint International Semantic Technology Conference (JIST 2017) at Gold Coast, Australia.
- Shortlisted for **The Best Paper Award** at the Künstliche Intelligenz Conference (KI 2019) at Kassel, Germany.

Thank You