



Ontologies and Reasoning

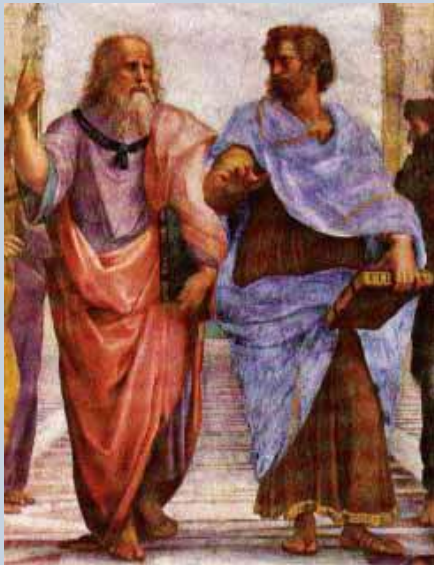
Ian Horrocks
Information Systems Group

What is an Ontology?

What is an Ontology?

A fundamental branch of **metaphysics**

- Studies “being” or “existence” and their **basic categories**
- Aims to find out what **entities** and **types of entities** exist



Supreme genus:

Differentiae:

Subordinate genera:

Differentiae:

Subordinate genera:

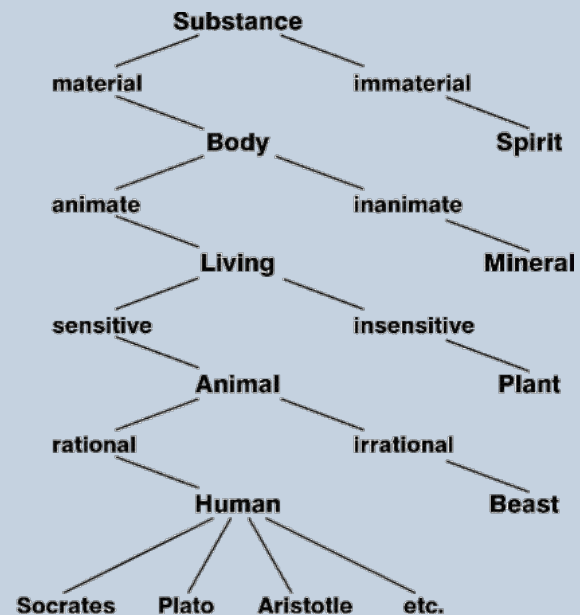
Differentiae:

Proximate genera:

Differentiae:

Species:

Individuals:



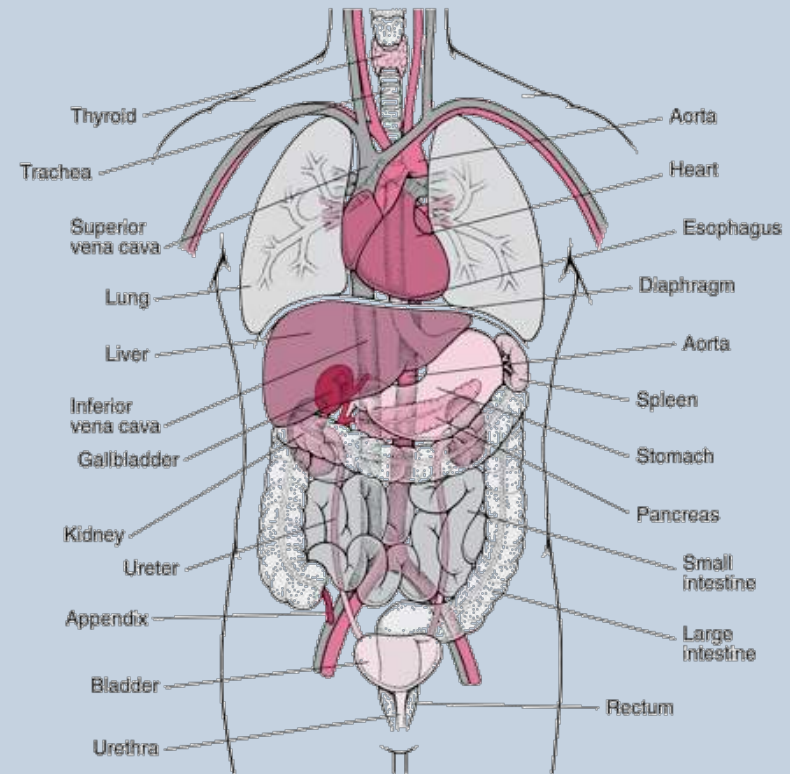
What is an Ontology?

A conceptual model of (some aspect of) the world

What is an Ontology?

A conceptual model of (some aspect of) the world

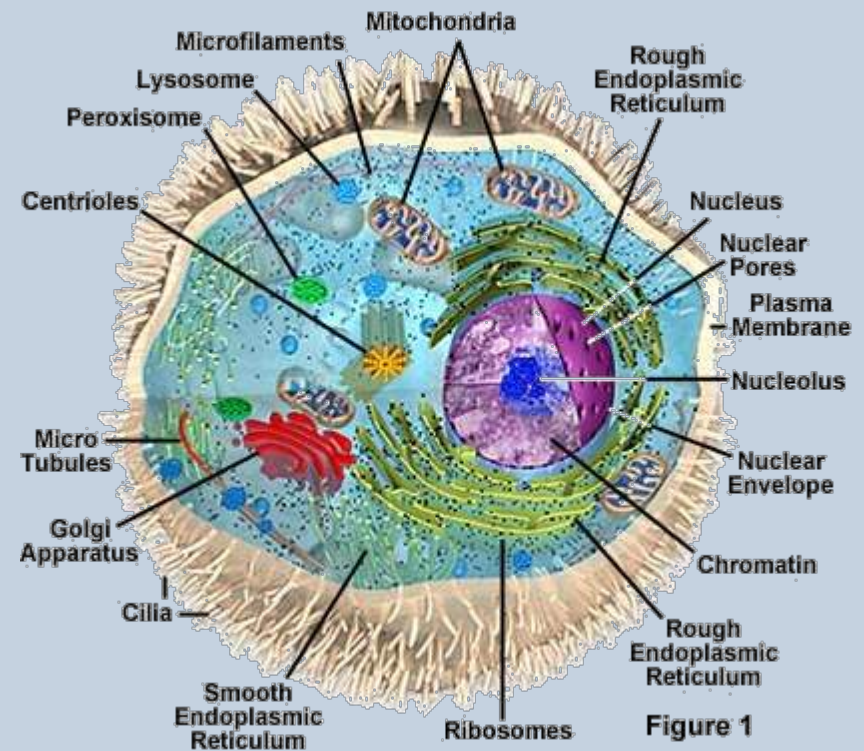
- Introduces **vocabulary** relevant to domain, e.g.:
 - Anatomy



What is an Ontology?

A conceptual model of (some aspect of) the world

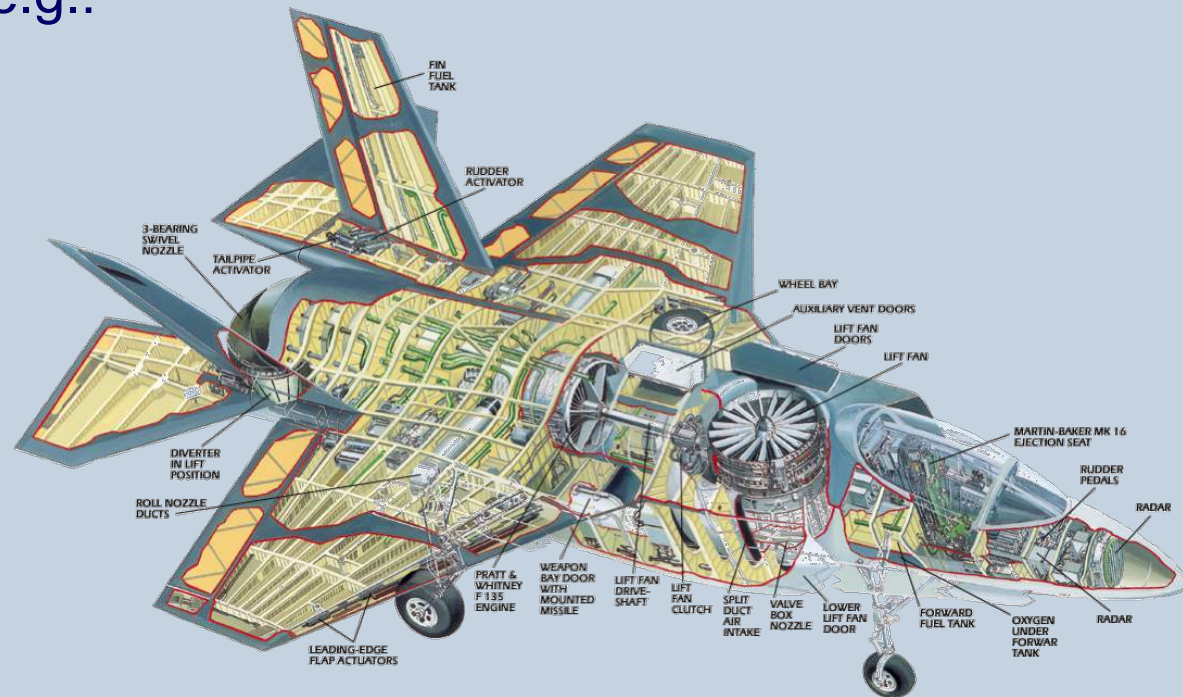
- Introduces **vocabulary** relevant to domain, e.g.:
 - Anatomy
 - Cellular biology



What is an Ontology?

A conceptual model of (some aspect of) the world

- Introduces **vocabulary** relevant to domain, e.g.:
 - Anatomy
 - Cellular biology
 - Aerospace

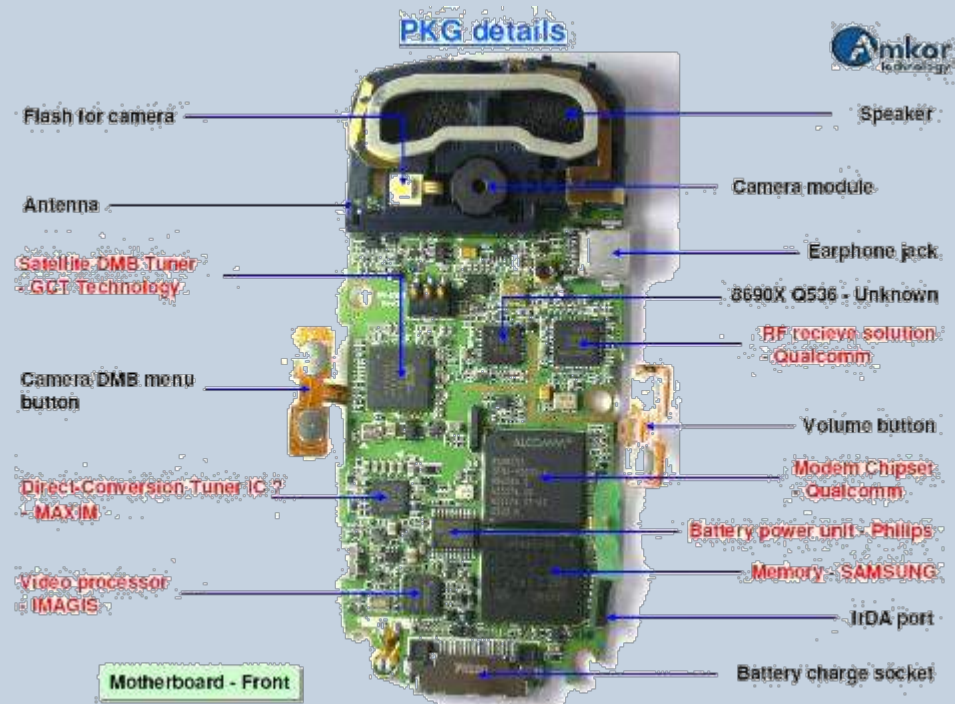


What is an Ontology?

A conceptual model of (some aspect of) the world

- Introduces **vocabulary** relevant to domain, e.g.:

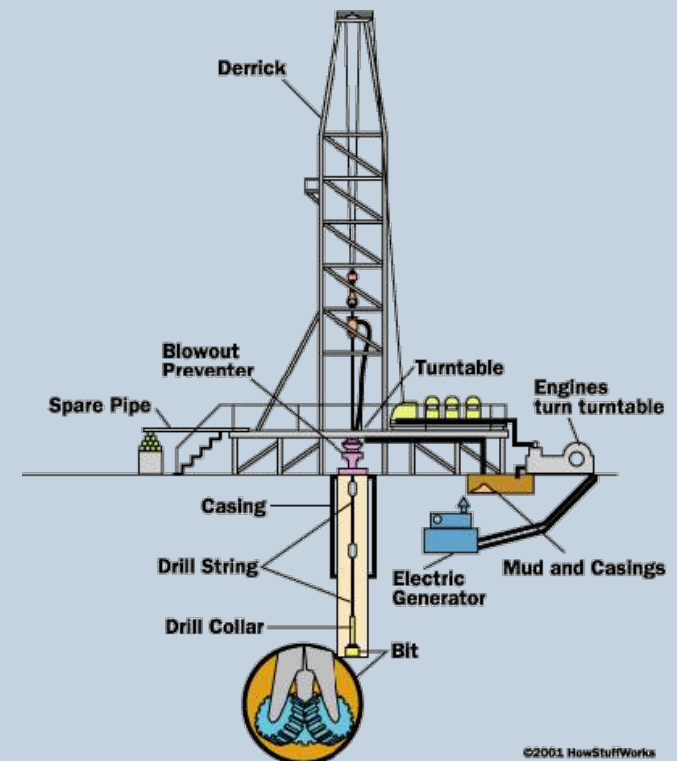
- Anatomy
- Cellular biology
- Aerospace
- Cell Phones



What is an Ontology?

A conceptual model of (some aspect of) the world

- Introduces **vocabulary** relevant to domain, e.g.:
 - Anatomy
 - Cellular biology
 - Aerospace
 - Cell Phones
 - Oil and gas

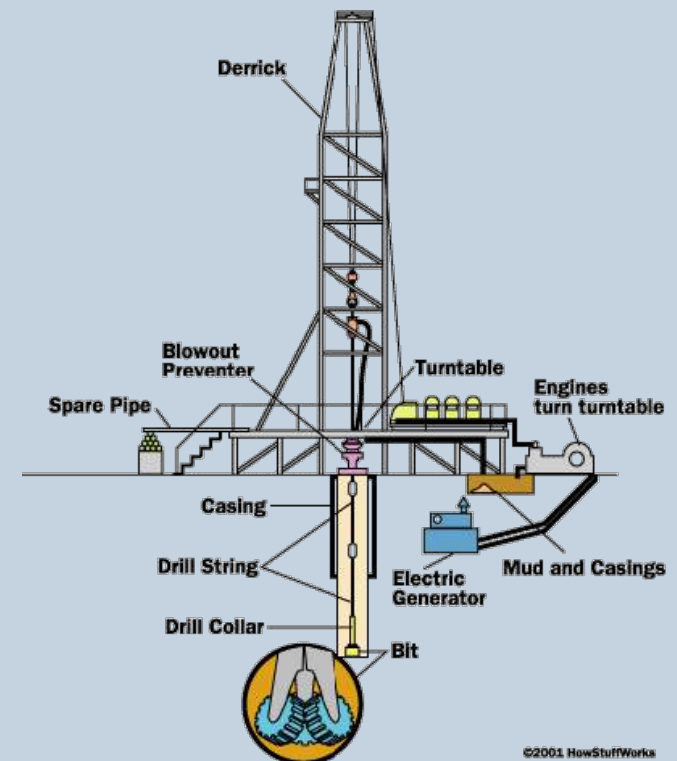


What is an Ontology?

A conceptual model of (some aspect of) the world

- Introduces **vocabulary** relevant to domain
- Specifies **meaning** (semantics) of terms

Oil pipeline is a pipeline from a facility that is an oil facility



What is an Ontology?

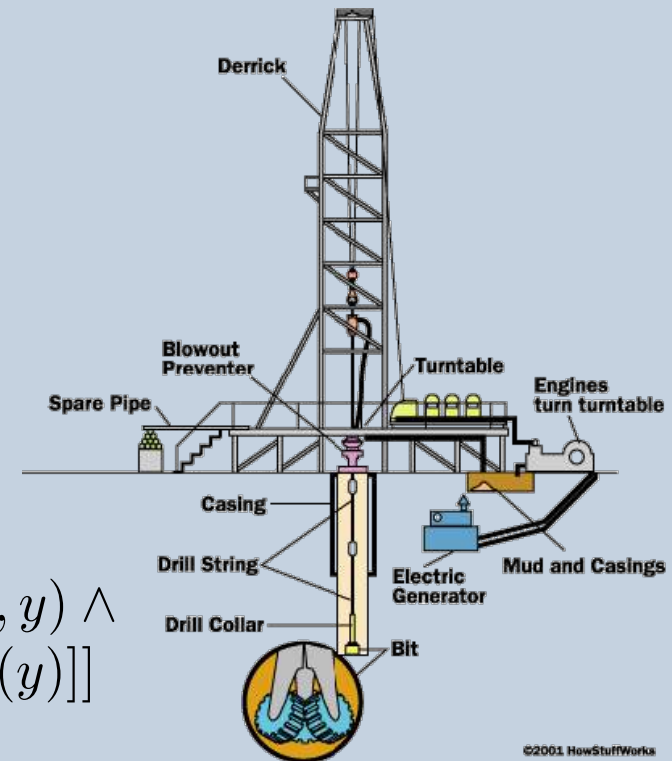
A conceptual model of (some aspect of) the world

- Introduces **vocabulary** relevant to domain
- Specifies **meaning** (semantics) of terms

Oil pipeline is a pipeline from a facility that is an oil facility

- **Formalised** using suitable logic

$$\forall x. [\text{OilPipeline}(x) \rightarrow \text{Pipeline}(x) \wedge \exists y. [\text{fromFacility}(x, y) \wedge \text{OilFacility}(y)]]$$



©2001 HowStuffWorks

What is an Ontology?

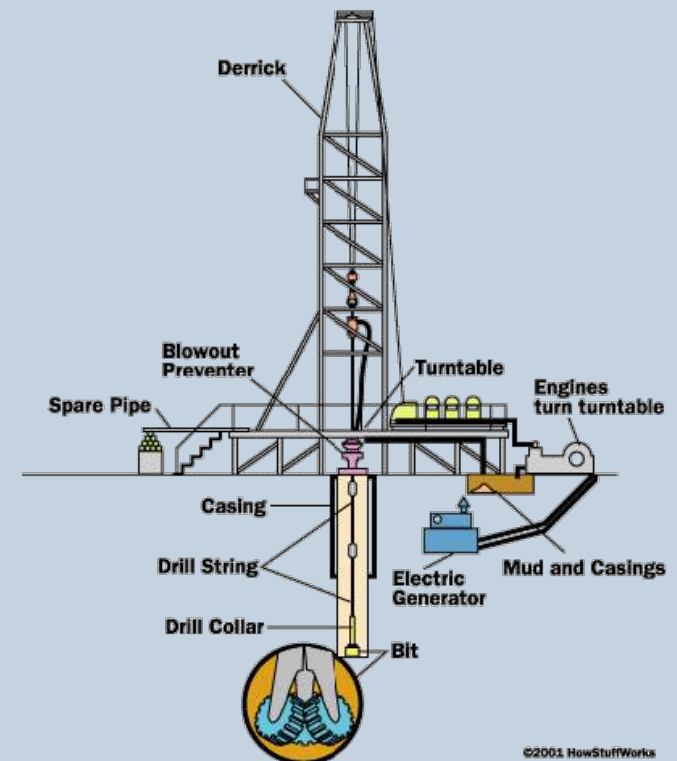
A conceptual model of (some aspect of) the world

- Introduces **vocabulary** relevant to domain
- Specifies **meaning** (semantics) of terms

Oil pipeline is a pipeline from a facility that is an oil facility

- **Formalised** using suitable logic

OilPipeline \sqsubseteq Pipeline \sqcap
 \exists from Facility.OilFacility



©2001 HowStuffWorks

Applications: HCLS

- **OBO foundry** includes more than 100 biological and biomedical ontologies
- **Siemens** “actively building OWL based clinical solutions”
- **SNOMED-CT** (Clinical Terms) ontology
 - used in healthcare systems of more than 15 countries, including Australia, Canada, Denmark, Spain, Sweden and the UK
 - also used by major US providers, e.g., Kaiser Permanente
 - ontology provides common vocabulary for recording clinical data

Applications: Energy Supply Industry

- **EDF Energy** offer personalised energy saving advice to every customer
- **OWL ontology** used to model relevant environmental factors
- **Oxford's HermiT reasoner** used to match customer circumstances with relevant pieces of advice



Applications: Intelligent Mobile Platform

- **Samsung** developing Intelligent Mobile Platform to support context-aware applications
- IMP monitors environment via **sensor data** (GPS, compass, accelerometer, ...)
- **OWL ontology** used to model environment and **infer context** (e.g., coffee with friends)
- Applications exploit context to enable more **intelligent behaviour**



Applications: Oil and Gas Industry

- **Statoil** use data to inform production and exploration management
Large and complex data sets are difficult and time consuming to use
- Using ontologies to **improve access** to relevant data



Barriers to Wider Adoption

- **Ontology engineering** (methodological)
 - Extensive tool support (editing, reasoning, explanation, modularity)
 - Carefully designed “upper ontologies”
 - But it’s still hard

Barriers to Wider Adoption

- **Ontology engineering** (methodological)
 - Extensive tool support (editing, reasoning, explanation, modularity)
 - Carefully designed “upper ontologies”
 - But it’s still hard
- **Ontology reasoning** (computational)
 - Highly optimised OWL reasoners
 - OWL profiles and specialised reasoners
 - But it’s still hard

Ontology Engineering: Case Study

SNOMED is **BIG** – over **400,000 concepts**

Ontology Engineering: Case Study

SNOMED is **BIG** – over 400,000 concepts

The screenshot shows the CliniClue 2006 SNOMED CT interface. The main window displays the concept 'TB - Pulmonary tuberculosis' (Concept Id: 154283005, Description Id: 1784750013). The interface is annotated with several callout boxes:

- Pulmonary Tuberculosis**: Points to the main concept name.
- pneumonitis**: Points to the 'is a' relationship with 'pneumonitis' in the hierarchy.
- inflammatory disorder of lower respiratory tract**: Points to the 'is a' relationship with 'inflammatory disorder of lower respiratory tract' in the hierarchy.
- Pulmonary disease due to Mycobacteria**: Points to the 'is a' relationship with 'pulmonary disease due to Mycobacteria' in the hierarchy.
- found in lung structure**: Points to the 'finding site' relationship with 'lung structure'.

The hierarchy on the left shows the following structure:

- 205237003 pneumonitis
- 56717001 tuberculosis
- 84353005 pulmonary disease due to Mycobacteria
 - 154283005 pulmonary tuberculosis
 - 428697002 inactive tuberculosis of lung
 - 186175002 infiltrative lung tuberculosis
 - 186188004 isolated tracheal or bronchial tuberculosis
 - 77668003 associated tracheal tuberculosis
 - 80602006 nodular tuberculosis of lung
 - 186192006 respiratory tuberculosis, bacteriologically and histologically confirmed
 - 186202007 respiratory tuberculosis, not confirmed bacteriologically
 - 186177005 tuberculosis of lung with cavitation
 - 81554001 tuberculosis of lung with involvement of bronchus
 - 186194007 tuberculosis of lung, confirmed by culture only
 - 186193001 tuberculosis of lung, confirmed by sputum microscopy
 - 186195008 tuberculosis of lung, confirmed histologically
 - 23022004 tuberculous bronchiectasis
 - 90117007 tuberculous fibrosis of lung

The definitions on the right include:

- Descriptions**: pulmonary tuberculosis (disorder), pulmonary tuberculosis, PTB - Pulmonary tuberculosis, TB - Pulmonary tuberculosis.
- Definition: Fully defined by...**
- is a**: pneumonitis, inflammatory disorder of lower respiratory tract, disorder of lung, inflammation of specific body organs, tuberculosis, pulmonary disease due to Mycobacteria, infectious disease of lung, bacterial lower respiratory infection, mycobacteriosis.
- causative agent**: Mycobacterium tuberculosis complex.
- Group**: associated morphology, finding site (lung structure).
- Qualifiers**: severity (severities), episodicity (episodicities), clinical course (courses).
- Codes**: Original SnomedId: R-F46B3.

Ontology Engineering: Case Study

- **Kaiser Permanente** extending SNOMED to express, e.g.:
 - *non-viral pneumonia* (negation)
 - *infectious pneumonia* is caused by a *virus* or a *bacterium* (disjunction)
 - *double pneumonia* occurs in *two lungs* (cardinalities)
- This is easy in **SNOMED-OWL**
 - but reasoner failed to find expected subsumptions, e.g., that *bacterial pneumonia* is a kind of *non-viral pneumonia*
- Ontology highly **under-constrained**: need to add disjointness axioms (at least)
 - *virus* and *bacterium* must be disjoint

Ontology Engineering: Case Study

- Adding disjointness led to **surprising results**
 - many classes become inconsistent, e.g., *percutaneous embolization of hepatic artery using fluoroscopy guidance*
- Cause of **inconsistencies** identified as class *groin*
 - *groin* asserted to be subclass of both *abdomen* and *leg*
 - *abdomen* and *leg* are disjoint
 - modelling of *groin* (and other similar “junction” regions) identified as incorrect

Ontology Engineering: Case Study

- Correct modelling of groin is quite complex, e.g.:
 - groin has a part that is part of the abdomen, and has a part that is part of the leg (*inverse properties*)

$\text{Groin} \sqsubseteq \exists \text{hasPart} . (\exists \text{isPartOf} . \text{Abdomen})$

$\text{Groin} \sqsubseteq \exists \text{hasPart} . (\exists \text{isPartOf} . \text{Leg})$

$\text{hasPart} \equiv \text{isPartOf}^{-}$

- all parts of the groin are part of the abdomen or the leg (*disjunction*)

$\text{Groin} \sqsubseteq \forall \text{hasPart} . (\exists \text{isPartOf} . (\text{Abdomen} \sqcup \text{Leg}))$

- ...

Ontology Engineering: Case Study

What we learned:

- Ontology engineering is **error prone**
 - errors of omission (e.g., disjointness) and commission (e.g., modelling of groin)
- **Expressive features** of OWL are sometimes needed
- Sophisticated tool support is **essential**
 - handling ontologies of this size is challenging
 - domain experts (and logicians!) often need help to understand the (root) cause of both inconsistencies and non-subsumptions
 - surprising and unexplained (non-) inferences are frustrating for users and may cause them to lose faith in the ontology and/or reasoner



Ontology Deployment: OBDA

Ontology Deployment: **OBDA**

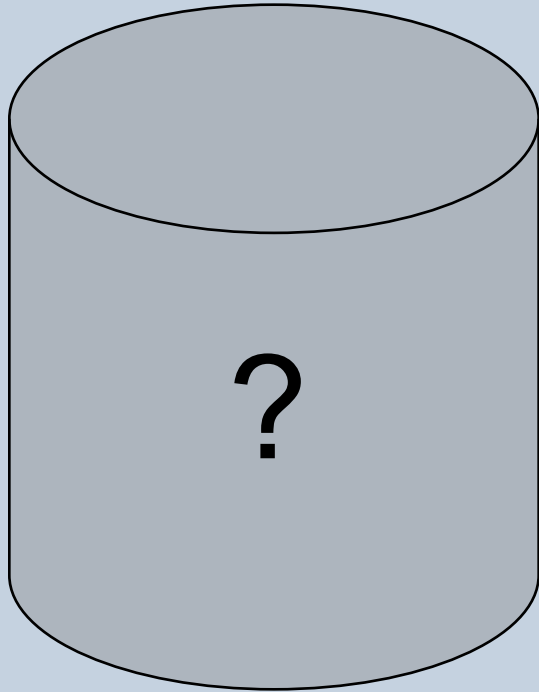


Ontology Deployment: **OBDA**

Pipelines from
oil facilities?



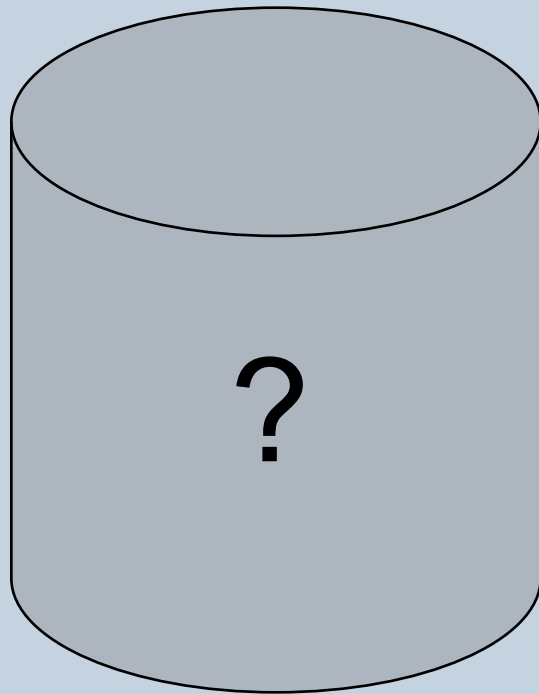
Ontology Deployment: OBDA



Pipelines from
oil facilities?



Ontology Deployment: OBDA

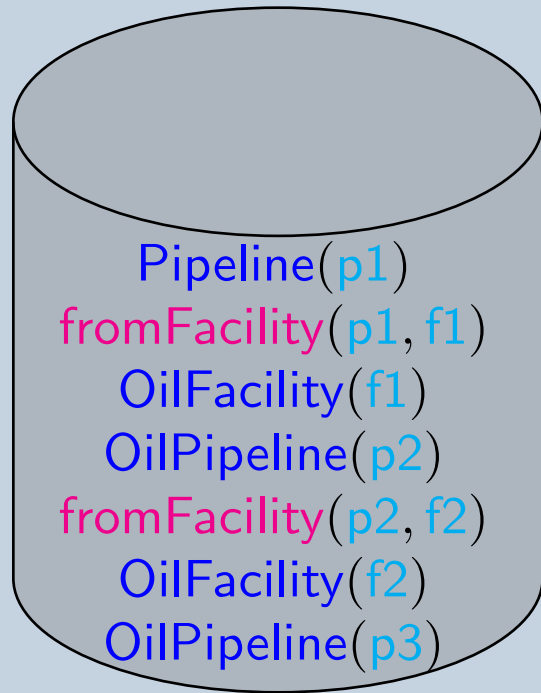


Pipelines from
oil facilities?



$Q(x) \leftarrow \text{Pipeline}(x) \wedge$
 $\text{fromFacility}(x, y) \wedge$
 $\text{OilFacility}(y)$

Ontology Deployment: OBDA

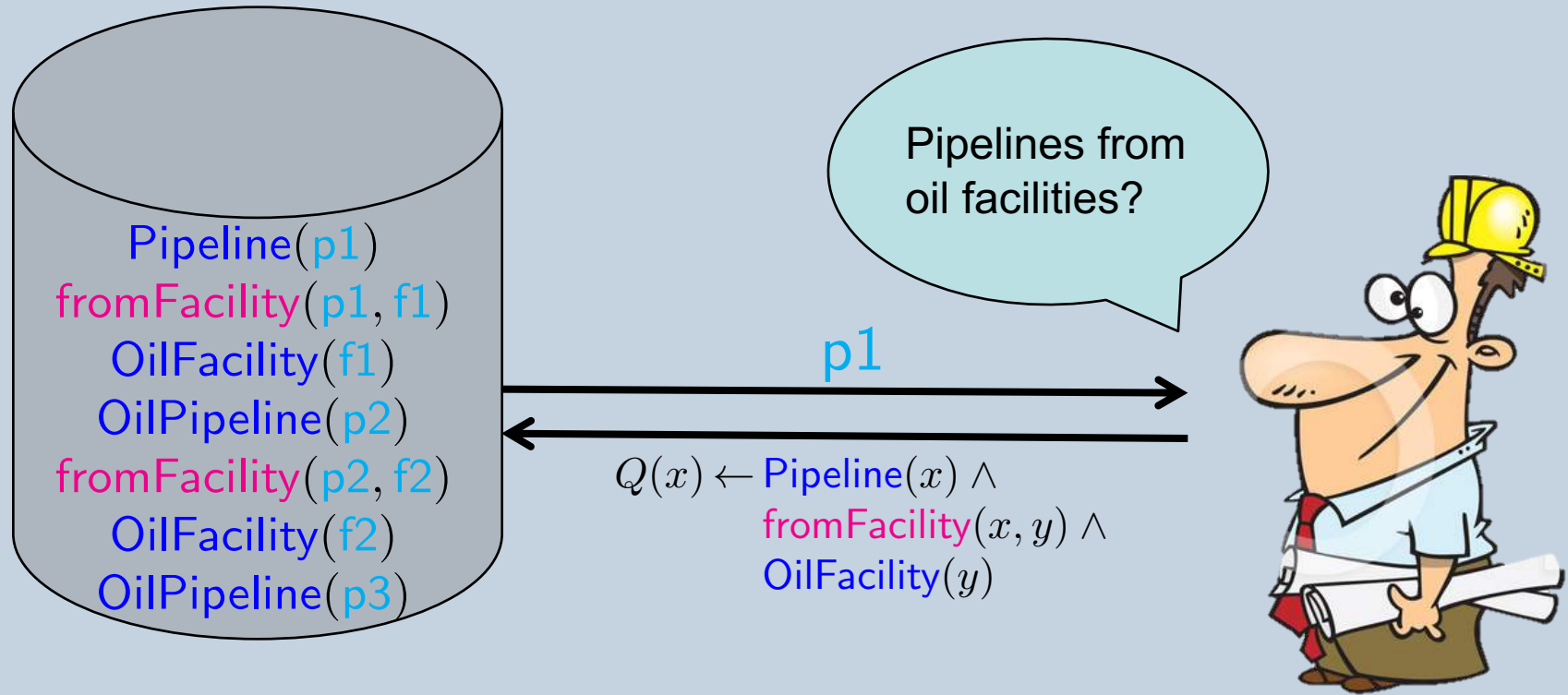


Pipelines from oil facilities?

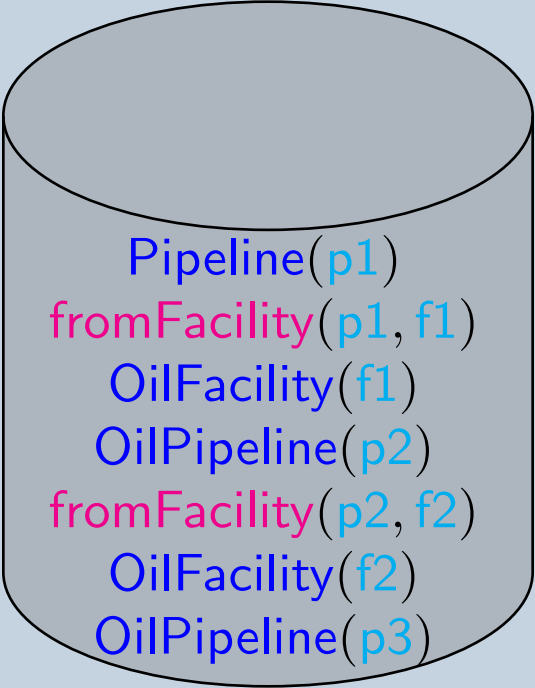


$Q(x) \leftarrow \text{Pipeline}(x) \wedge$
 $\text{fromFacility}(x, y) \wedge$
 $\text{OilFacility}(y)$

Ontology Deployment: OBDA

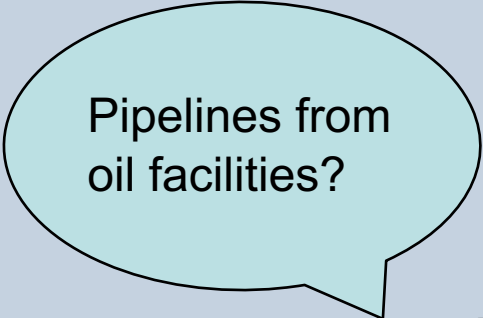


Ontology Deployment: OBDA



Pipeline(p1)
fromFacility(p1, f1)
OilFacility(f1)
OilPipeline(p2)
fromFacility(p2, f2)
OilFacility(f2)
OilPipeline(p3)

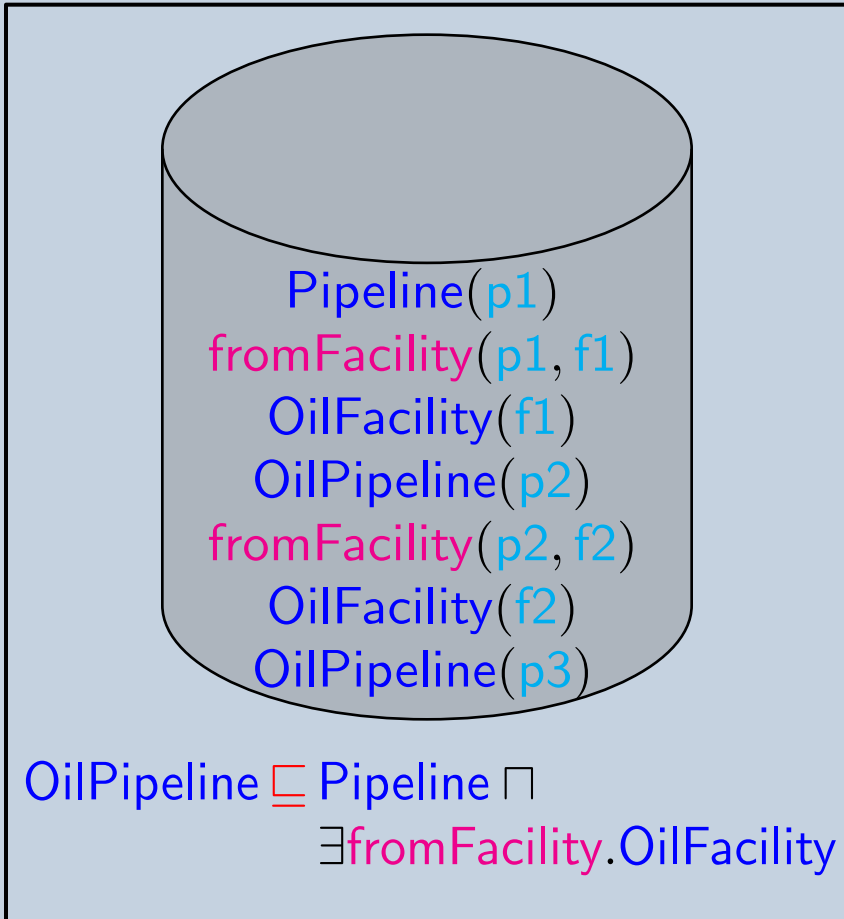
OilPipeline \sqsubseteq Pipeline \sqcap
 \exists fromFacility.OilFacility



Pipelines from
oil facilities?



Ontology Deployment: OBDA

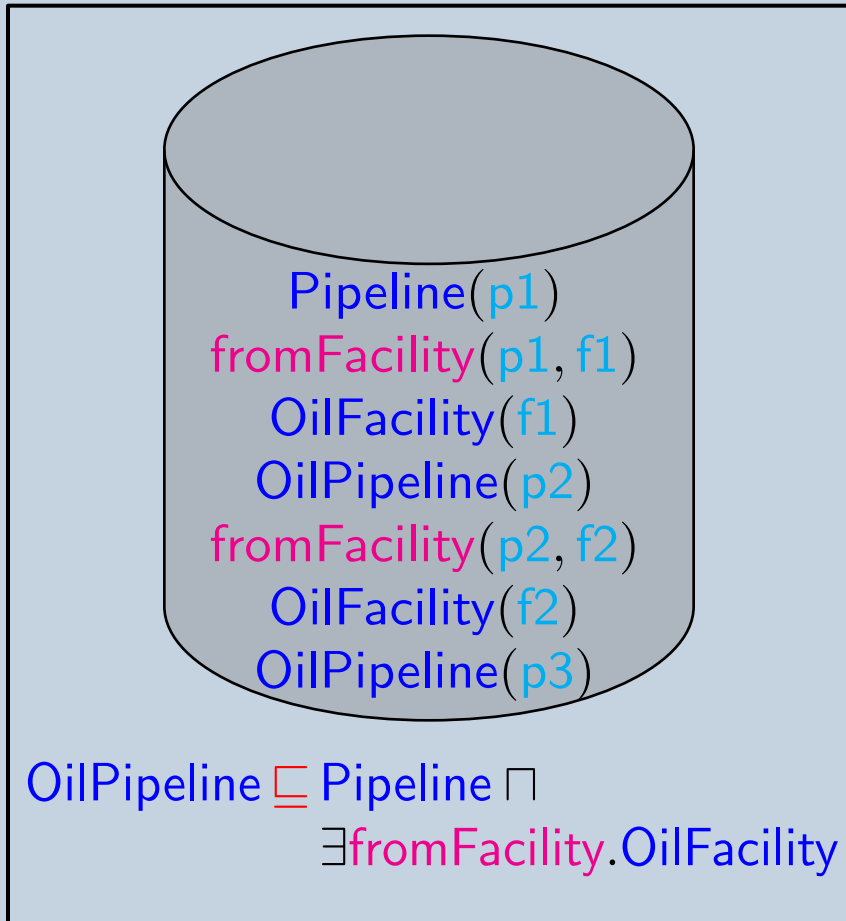


Pipelines from oil facilities?

$Q(x) \leftarrow \text{Pipeline}(x) \wedge \text{fromFacility}(x, y) \wedge \text{OilFacility}(y)$

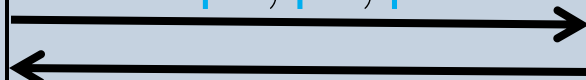


Ontology Deployment: OBDA

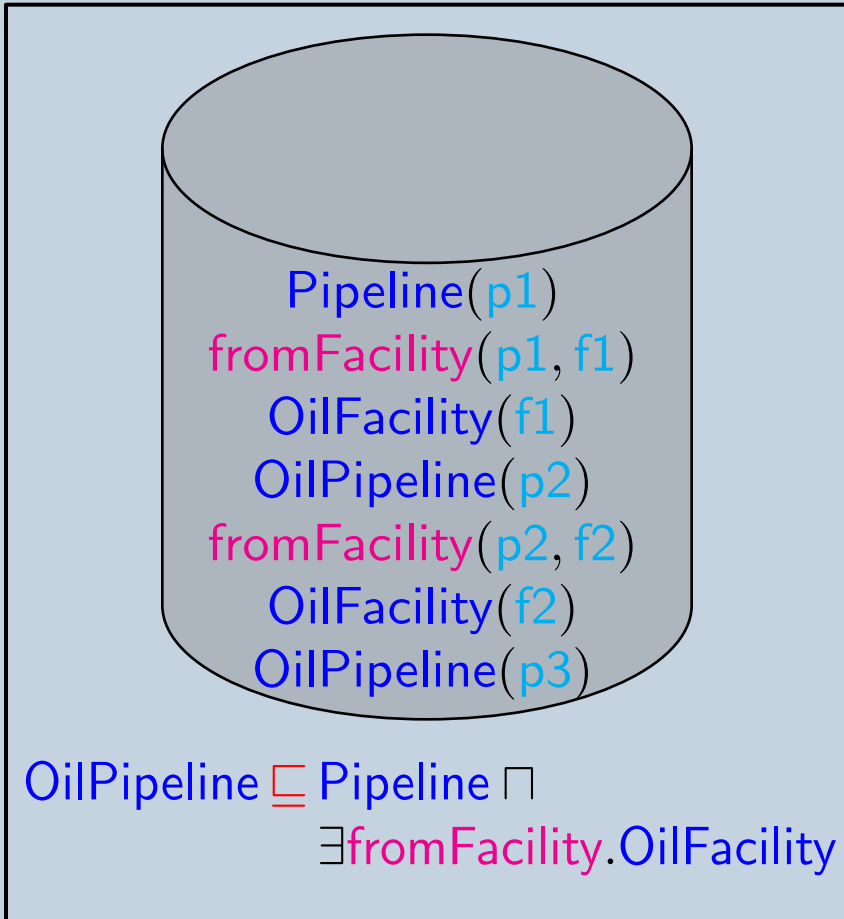


Pipelines from oil facilities?

p1, p2, p3


$$Q(x) \leftarrow \text{Pipeline}(x) \wedge \text{fromFacility}(x, y) \wedge \text{OilFacility}(y)$$


Ontology Deployment: OBDA

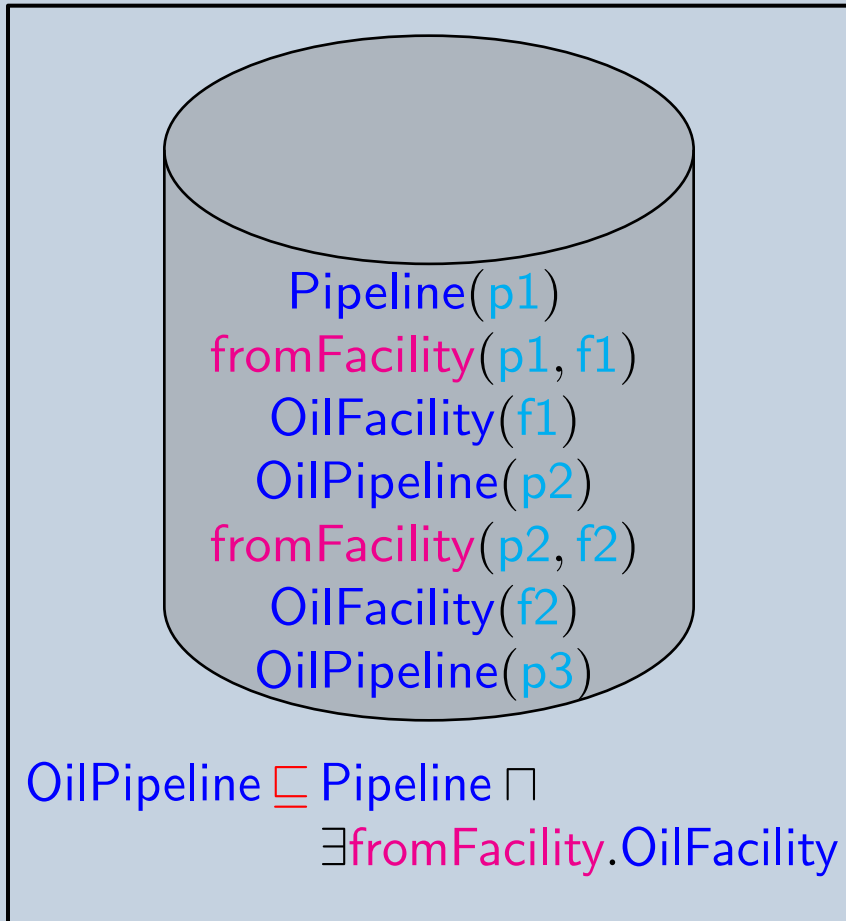


Is an oil pipeline a kind of pipeline?

$\forall x. [\text{Pipeline}(x) \rightarrow \text{OilPipeline}(x)]?$



Ontology Deployment: OBDA



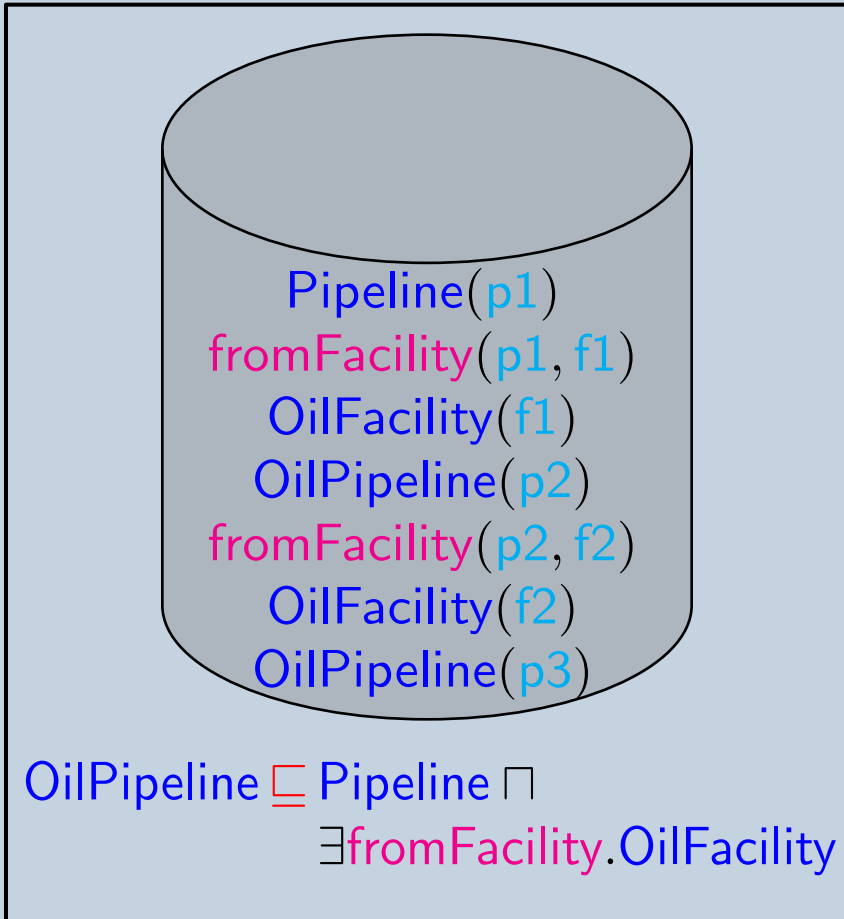
Is an oil pipeline a kind of pipeline?

YES

$\forall x. [\text{Pipeline}(x) \rightarrow \text{OilPipeline}(x)]?$



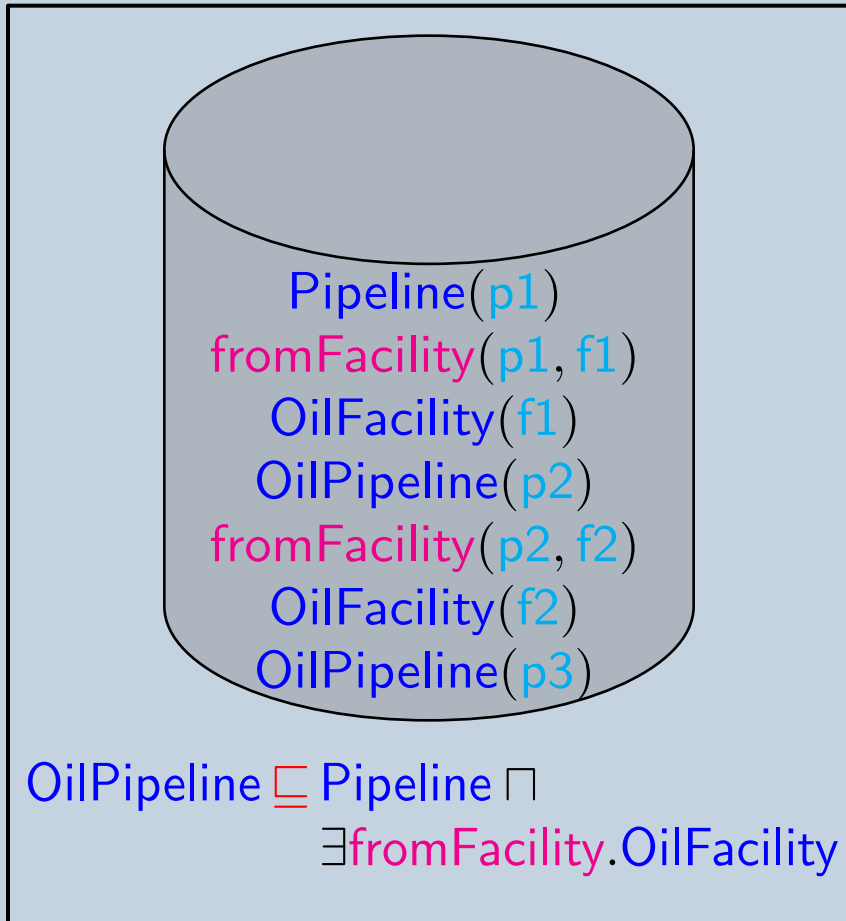
Ontology Deployment: OBDA



Why?



Ontology Deployment: OBDA



Why?

$\text{OilPipeline} \subseteq \text{Pipeline} \dots$

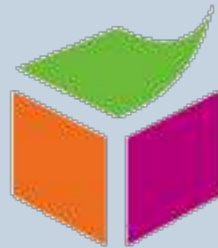


Semantic Technologies

- SemWeb motivated development of **robust infrastructure**:

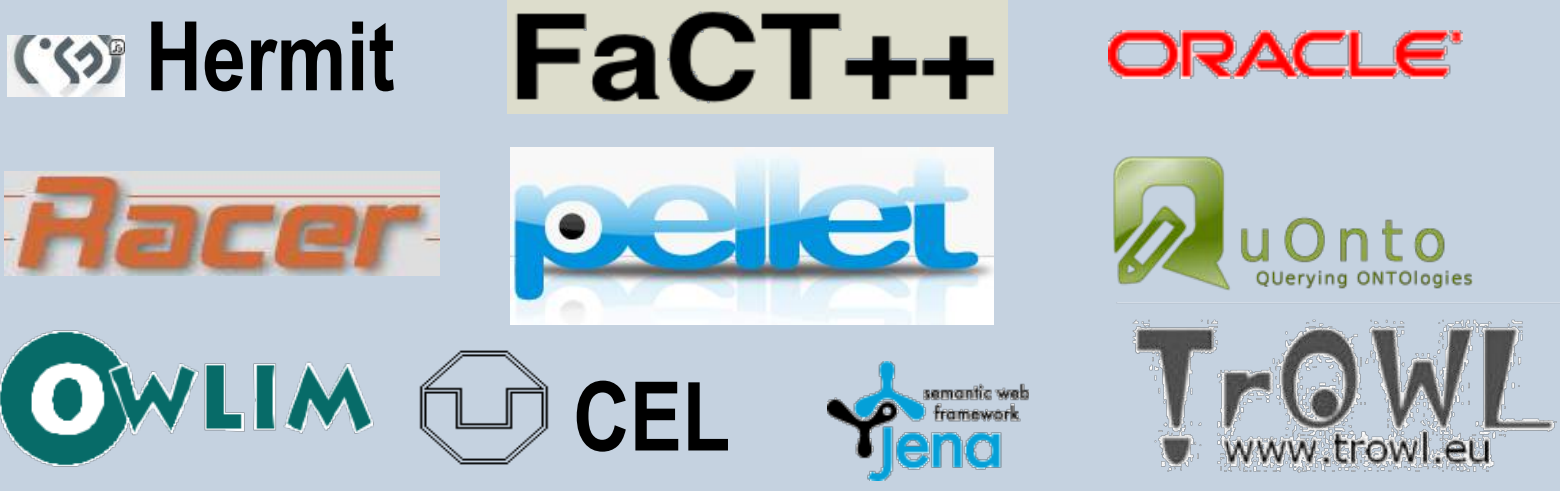
Semantic Technologies

- SemWeb motivated development of **robust infrastructure**:
 - Languages



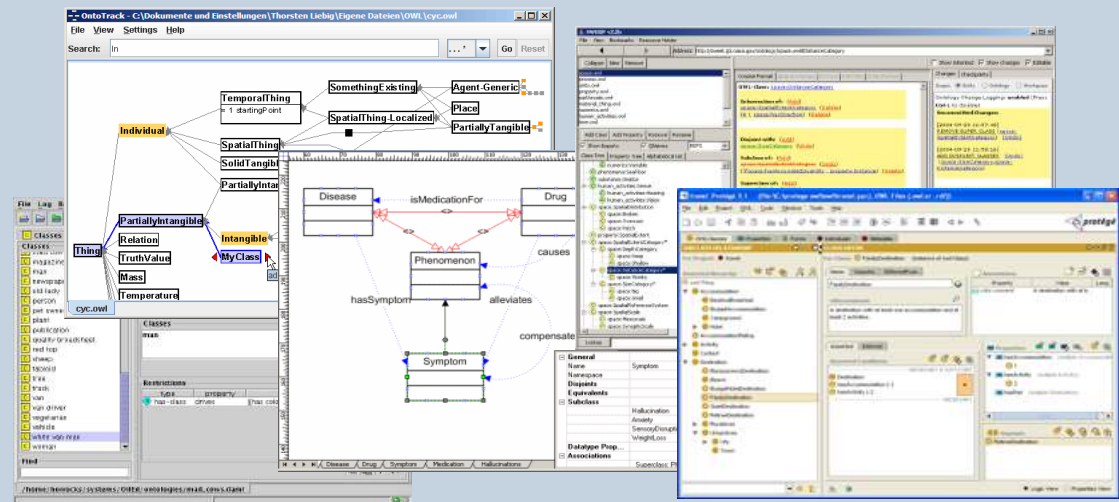
Semantic Technologies

- SemWeb motivated development of **robust infrastructure**:
 - Languages
 - Storage reasoning and query answering



Semantic Technologies

- SemWeb motivated development of **robust infrastructure**:
 - Languages
 - Storage reasoning and query answering
 - Development tools



Semantic Technologies

- SemWeb motivated development of **robust infrastructure**:
 - Languages
 - Storage reasoning and query answering
 - Development tools
- Increasingly used as part of “Intelligent Information Systems



DEPARTMENT OF
**COMPUTER
SCIENCE**

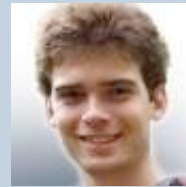
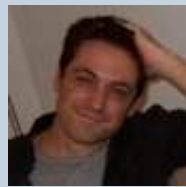
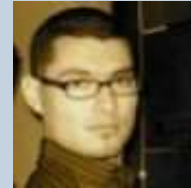
Information Systems Group



Ongoing Research

- Query answering
 - [Motik et al], [Kontchakov et al], [Konev et al], [Baader et al], ...
- Diagnosis and repair
 - [Horridge et al], [Peñaloza et al], ...
- Extensions
 - [Motik et al], [Artale et al], ...
- Optimisation/Profiles
 - [Kazakov], [Glimm et al], [Faddoul et al], [Savo et al], ...
- Parallelization/Distribution/HPC
 - [Motik et al], [Gurajada et al], [Neumann et al], [Wu et al], ...

Acknowledgements



Engineering and Physical Sciences
Research Council



SIXTH FRAMEWORK
PROGRAMME



SEVENTH FRAMEWORK
PROGRAMME

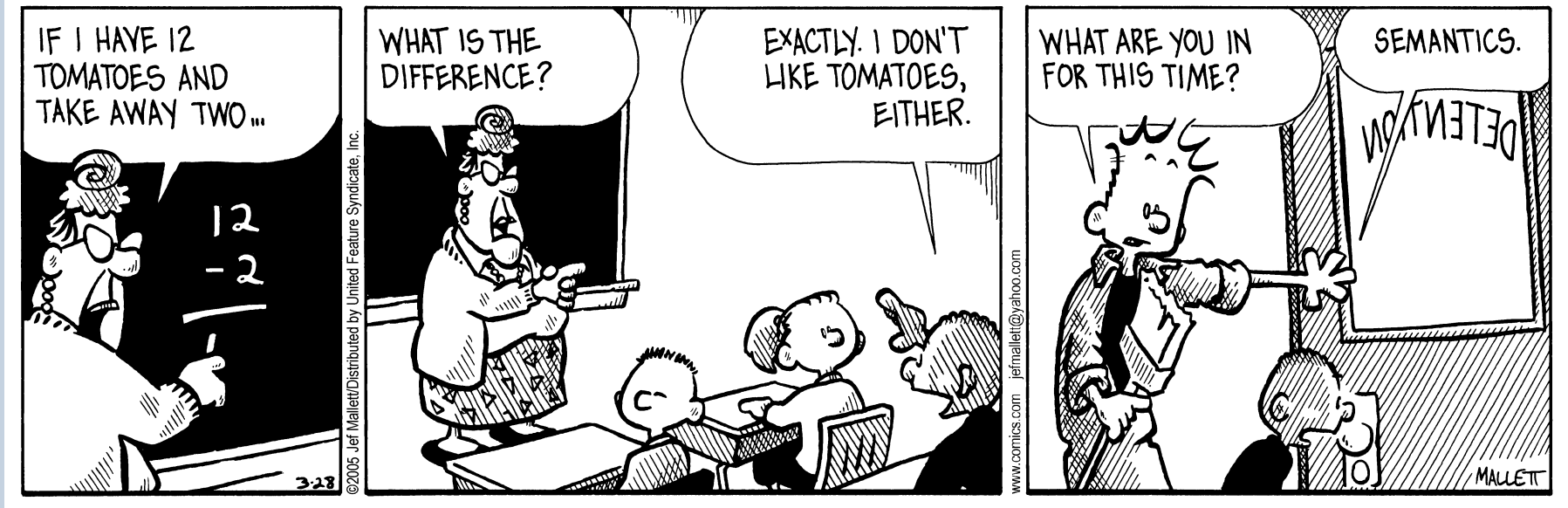


DEPARTMENT OF
**COMPUTER
SCIENCE**

Information Systems Group



Thank you for listening



FRAZZ: © Jeff Mallett/Dist. by United Feature Syndicate, Inc.

Any questions?