



OBDA: Theory and Practice

Ian Horrocks
Information Systems Group

What is an Ontology?



DEPARTMENT OF
**COMPUTER
SCIENCE**

Optique



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:

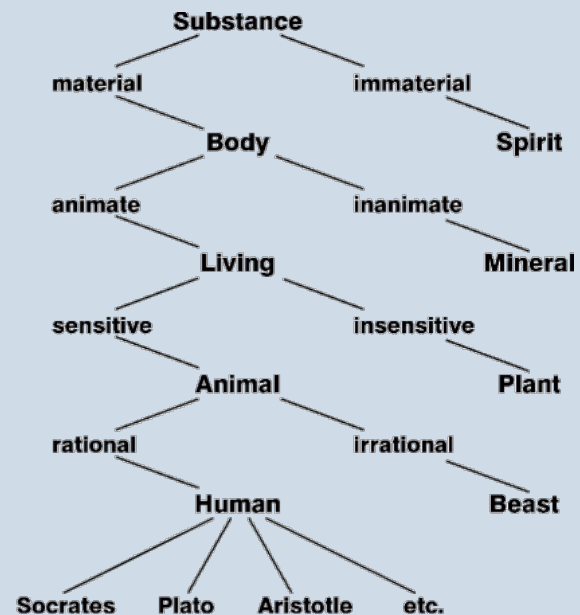
Differentiae:

Proximate genera:

Differentiae:

Species:

Individuals:



What is an Ontology?

A conceptual model of (some aspect of) the world



DEPARTMENT OF
**COMPUTER
SCIENCE**

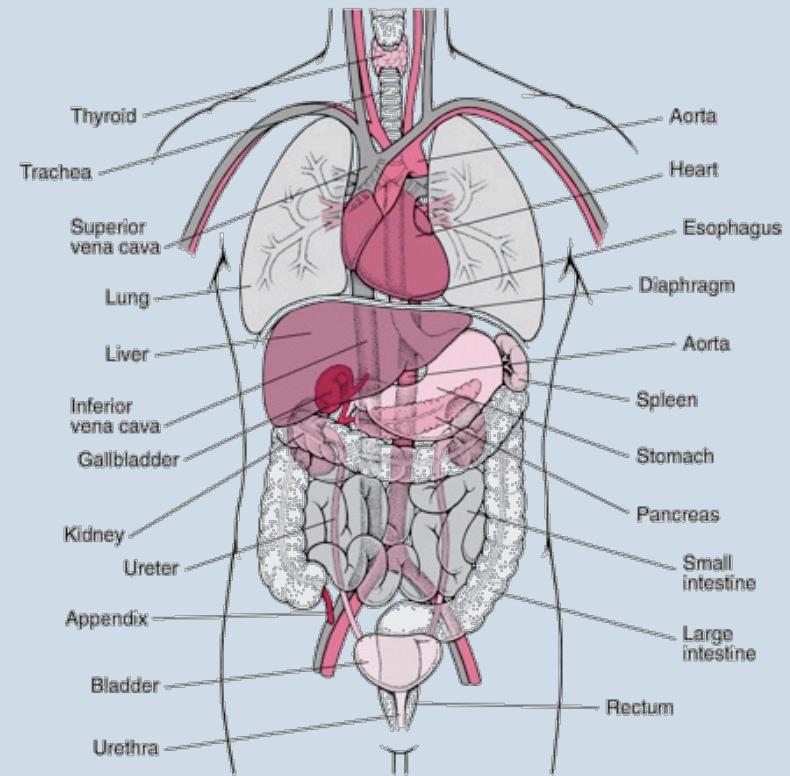
Optique



What is an Ontology?

A conceptual model of (some aspect of) the world

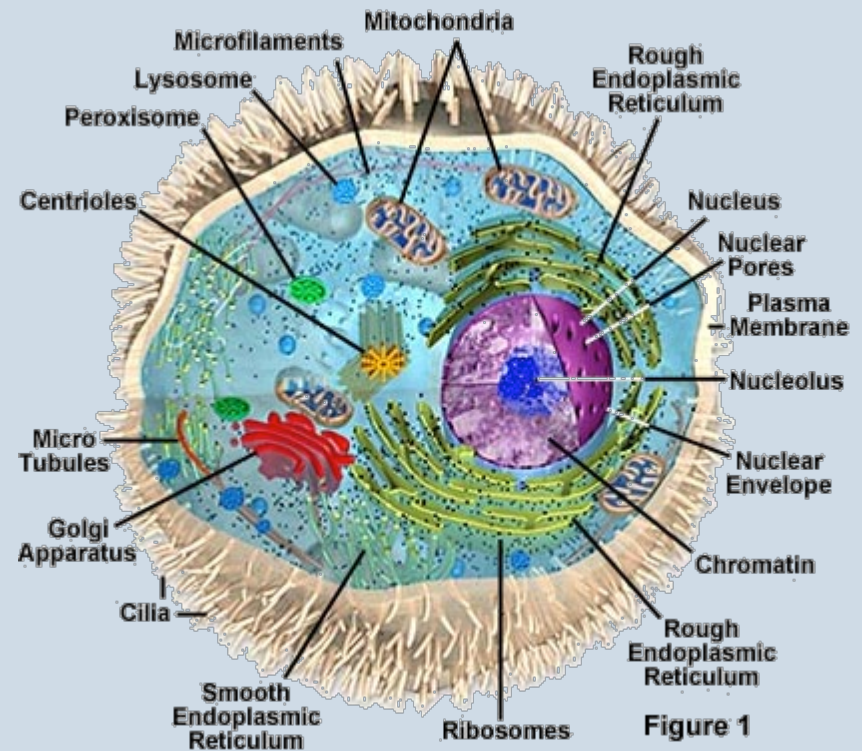
- Introduces **vocabulary** relevant to domain, e.g.:
 - Anatomy



What is an Ontology?

A conceptual model of (some aspect of) the world

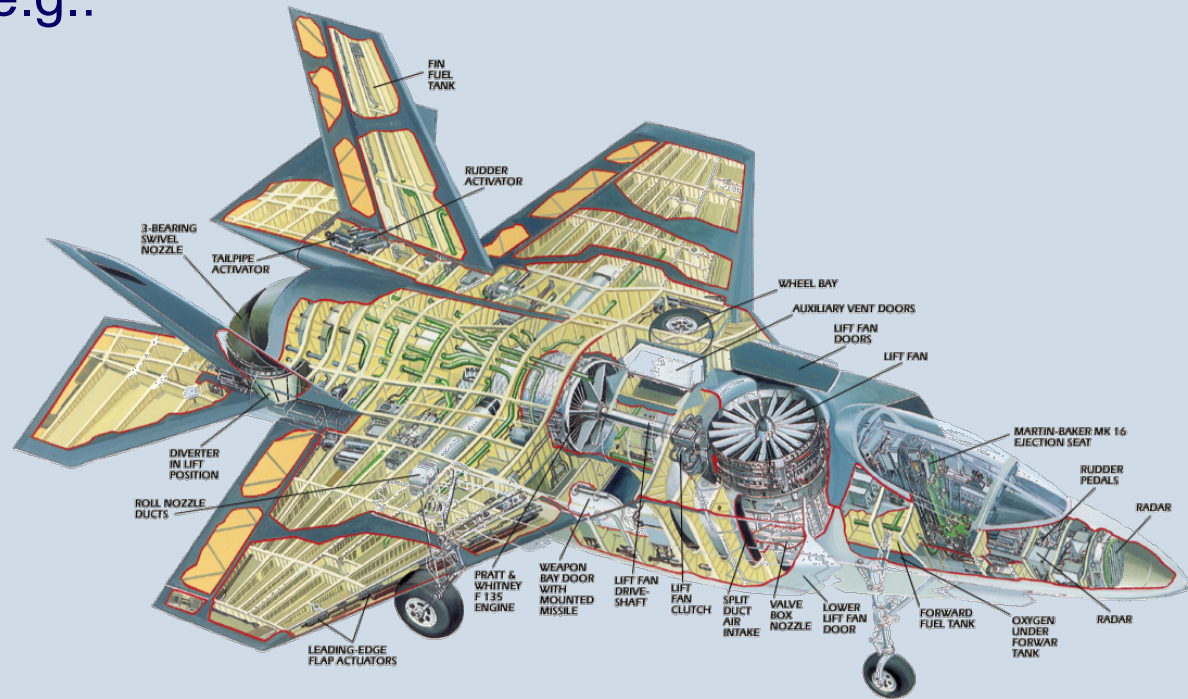
- Introduces **vocabulary** relevant to domain, e.g.:
 - Anatomy
 - Cellular biology



What is an Ontology?

A conceptual model of (some aspect of) the world

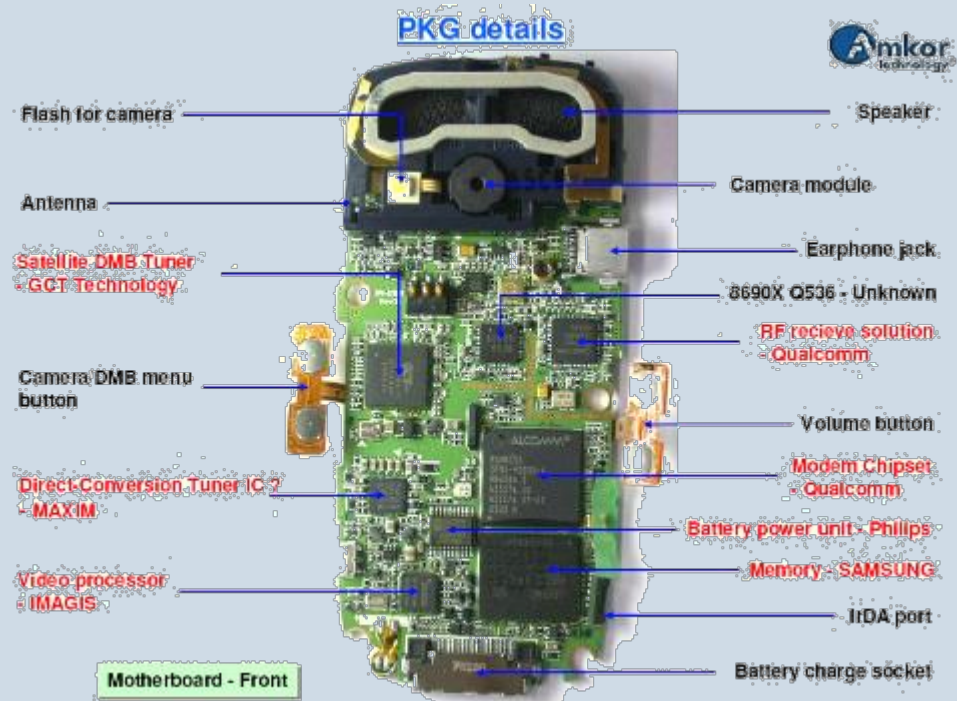
- Introduces **vocabulary** relevant to domain, e.g.:
 - Anatomy
 - Cellular biology
 - Aerospace



What is an Ontology?

A conceptual model of (some aspect of) the world

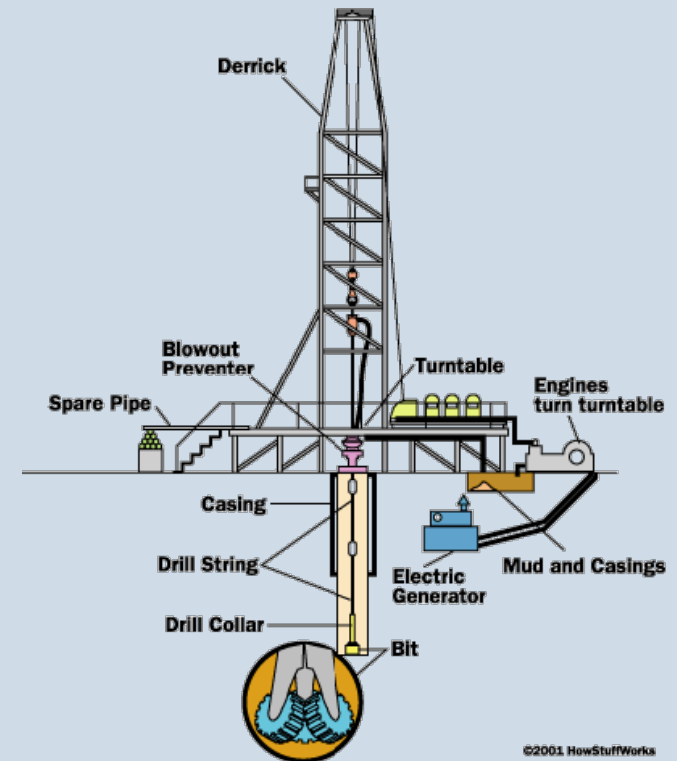
- Introduces **vocabulary** relevant to domain, e.g.:
 - Anatomy
 - Cellular biology
 - Aerospace
 - Cell Phones



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

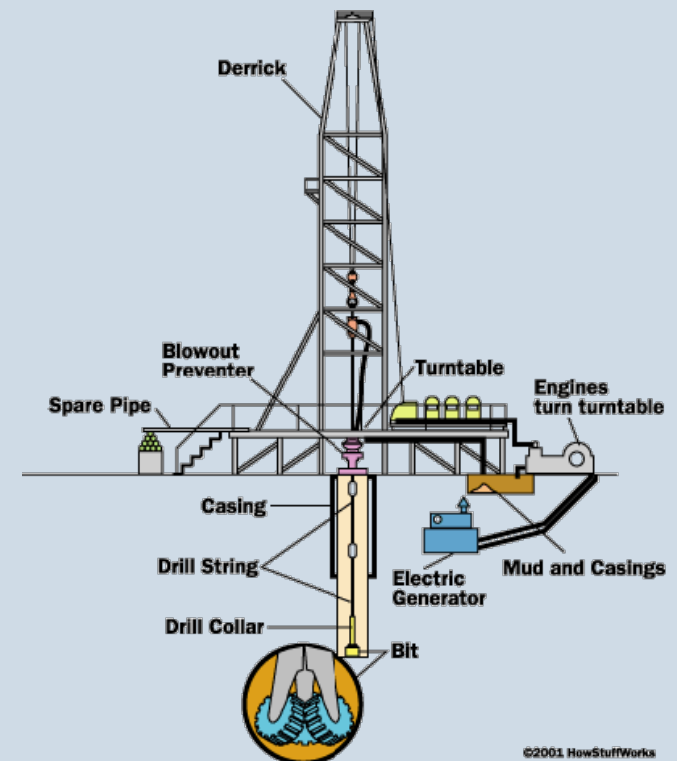


What is an Ontology?

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?

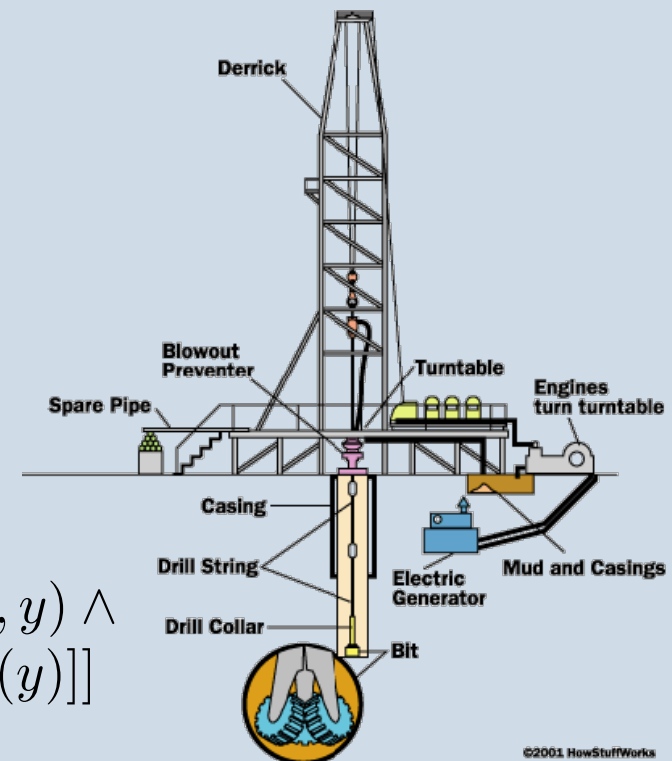
A conceptual model of (some aspect of) the world

- 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

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



What is an Ontology?

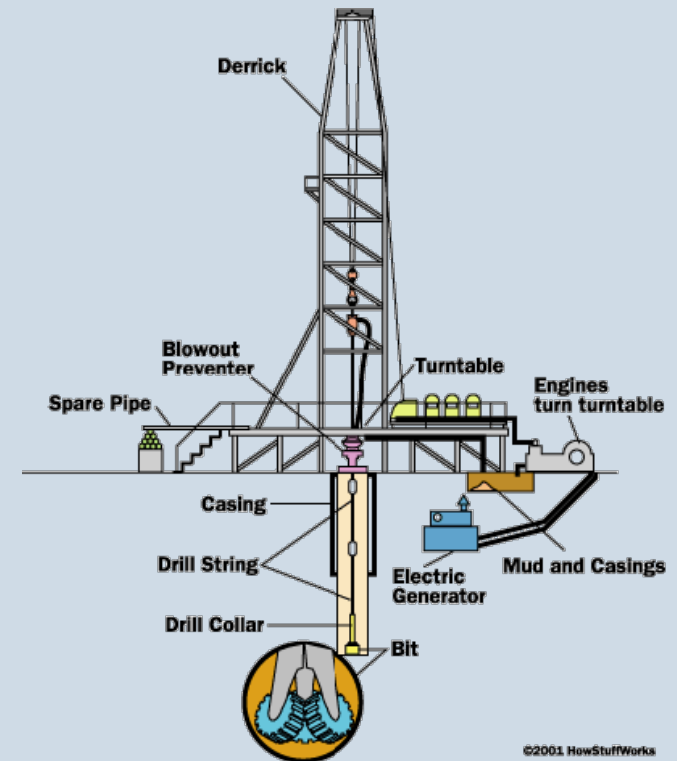
A conceptual model of (some aspect of) the world

- 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

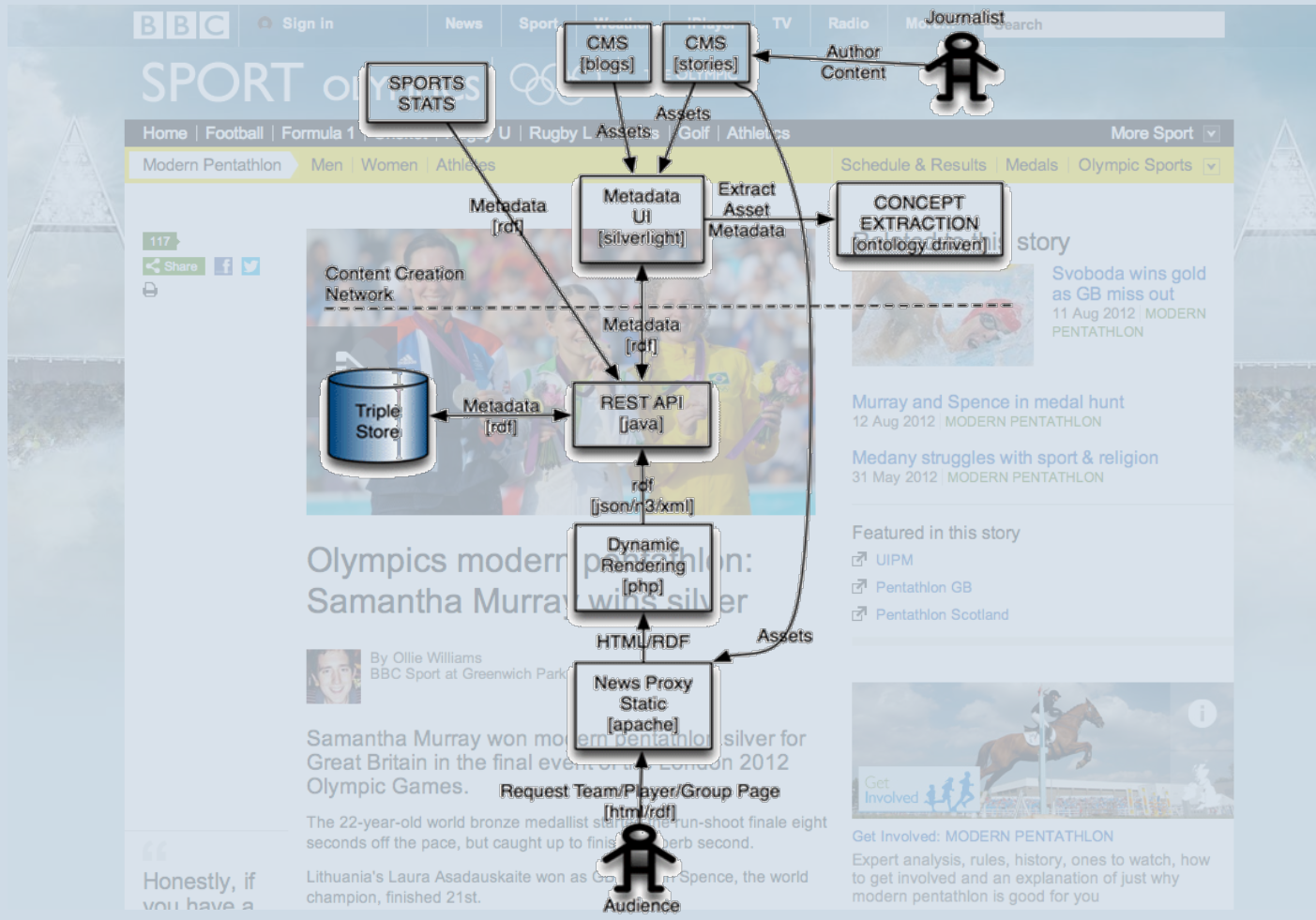
OilPipeline \sqsubseteq Pipeline \sqcap
 \exists fromFacility.OilFacility



Applications: Semantic Web

The screenshot shows the BBC Sport Olympics website. The top navigation bar includes 'BBC', 'Sign in', and various sports categories like News, Sport, Weather, IPlayer, TV, Radio, and More... A search bar is also present. Below the navigation, the main header reads 'SPORT OLYMPICS' with the Olympic rings logo and 'THE OLYMPIC BROADCASTER'. A secondary navigation bar lists sports like Home, Football, Formula 1, Cricket, Rugby U, Rugby L, Tennis, Golf, and Athletics. A yellow bar highlights 'Modern Pentathlon' with sub-links for Men, Women, and Athletes, along with 'Schedule & Results', 'Medals', and 'Olympic Sports'. The main content area features a large photo of three athletes on a podium. To the left of the photo are social media share buttons for 117 users, and a play button icon. Below the photo is the article title 'Olympics modern pentathlon: Samantha Murray wins silver' and the author 'By Ollie Williams, BBC Sport at Greenwich Park'. The article text states that Samantha Murray won silver for Great Britain in the final event of the London 2012 Olympic Games. A quote from the article reads: 'Honestly, if you have a...'. To the right of the article is a 'Related to this story' section with three items: 'Svoboda wins gold as GB miss out' (11 Aug 2012), 'Murray and Spence in medal hunt' (12 Aug 2012), and 'Medany struggles with sport & religion' (31 May 2012). Below this is a 'Featured in this story' section with links to 'UIPM', 'Pentathlon GB', and 'Pentathlon Scotland'. At the bottom right, there is a 'Get Involved' section with a photo of a horse and rider, and text about expert analysis and rules of modern pentathlon.

Applications: Semantic Web

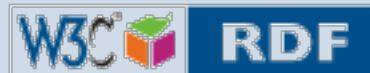


Semantic Technologies

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

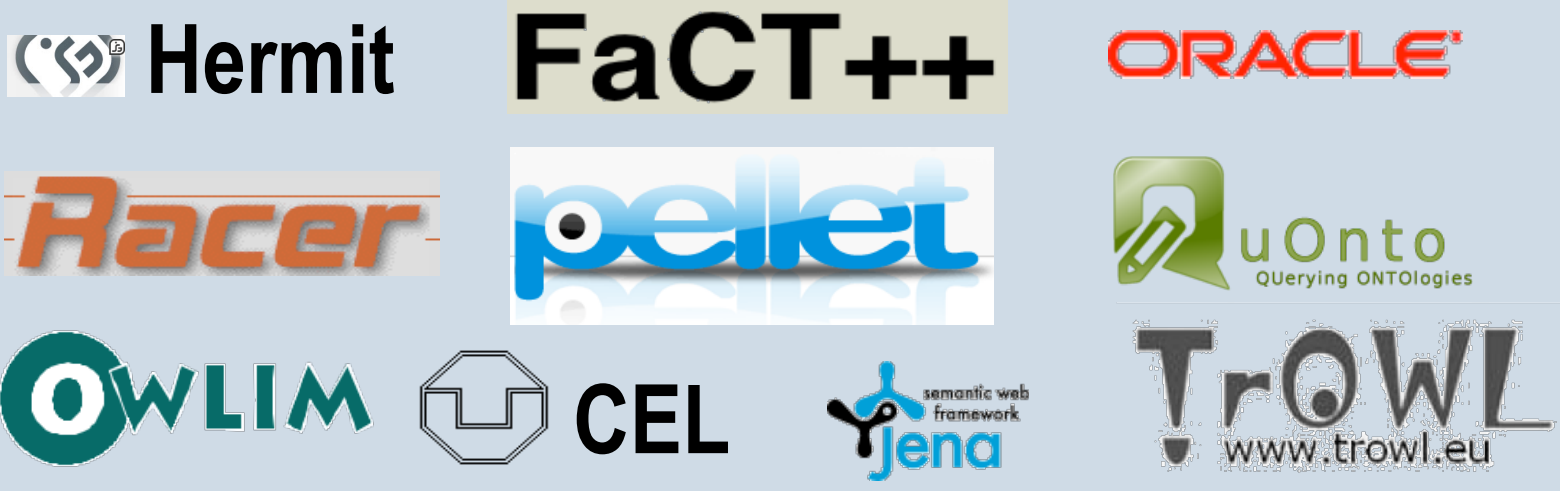
Semantic Technologies

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



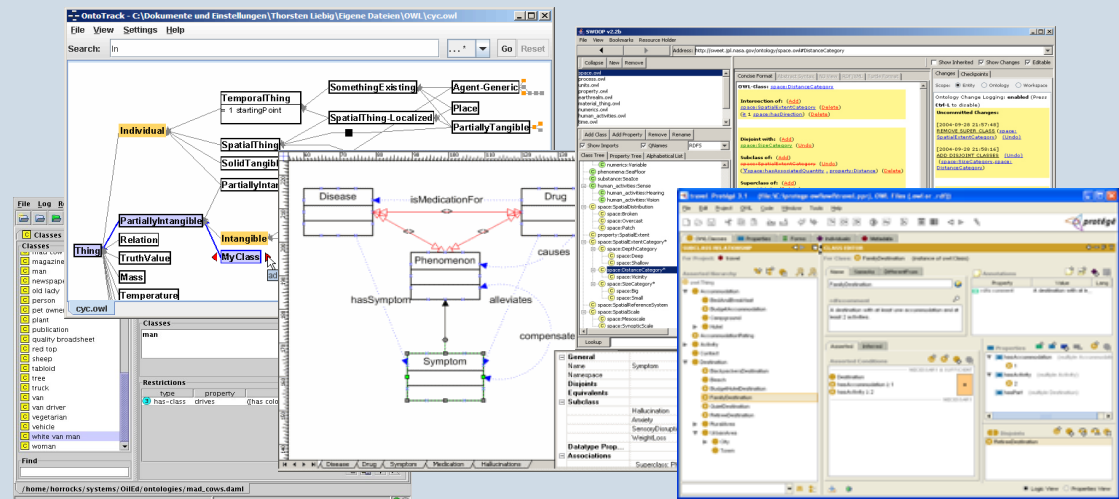
Semantic Technologies

- SemWeb motivated development of **robust infrastructure**:
 - Languages
 - Storage and querying



Semantic Technologies

- SemWeb motivated development of **robust infrastructure**:
 - Languages
 - Storage and querying
 - Development tools



Semantic Technologies

- SemWeb motivated development of **robust infrastructure**:
 - Languages
 - Storage and querying
 - Development tools
- Increasingly used in “Intelligent Information Systems”, and in particular **Ontology Based Data Access (OBDA)**



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: **OBDA**



DEPARTMENT OF
**COMPUTER
SCIENCE**

Optique



Applications: OBDA



DEPARTMENT OF
**COMPUTER
SCIENCE**

Optique

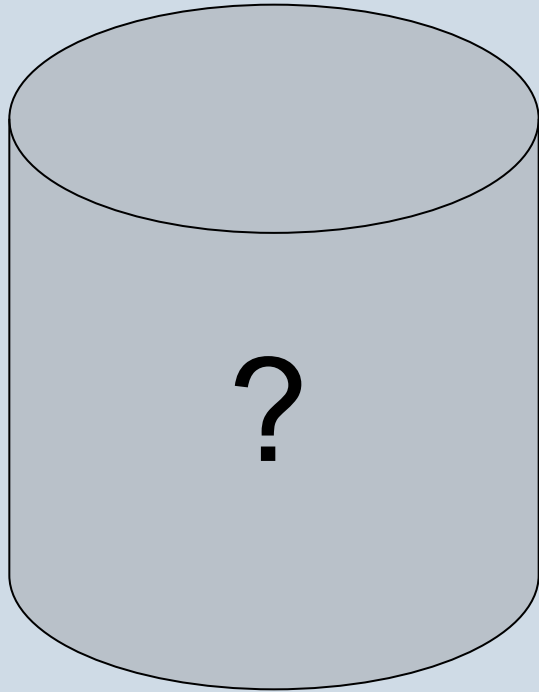


Applications: **OBDA**

Pipelines from
oil facilities?



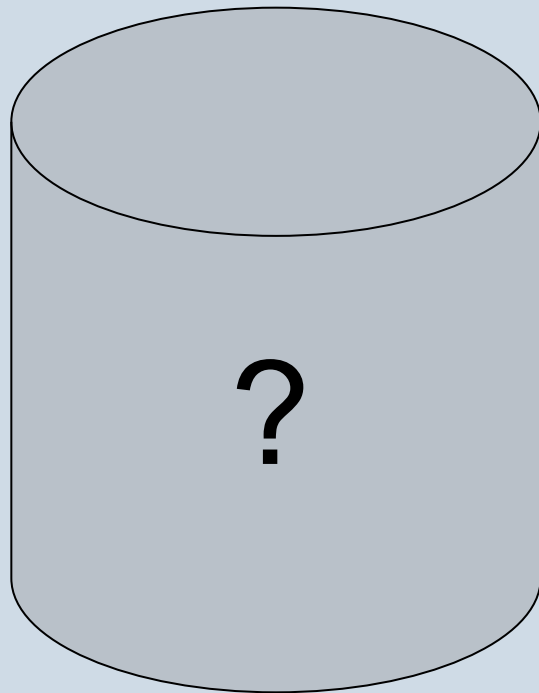
Applications: **OBDA**



Pipelines from
oil facilities?



Applications: OBDA

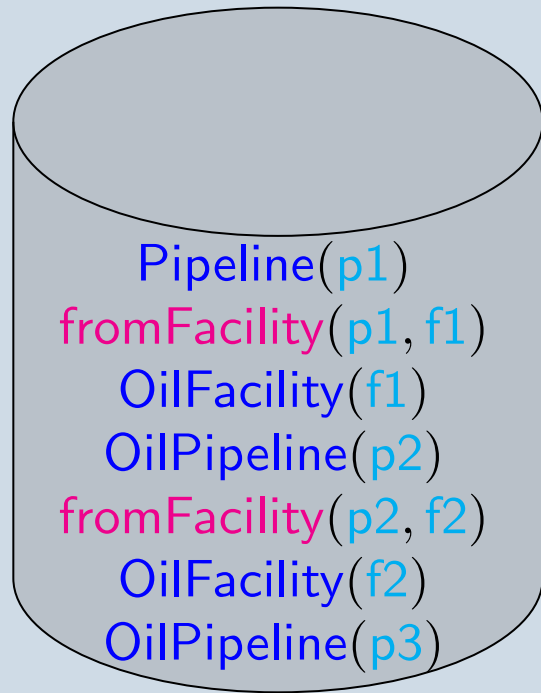


Pipelines from
oil facilities?



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

Applications: OBDA



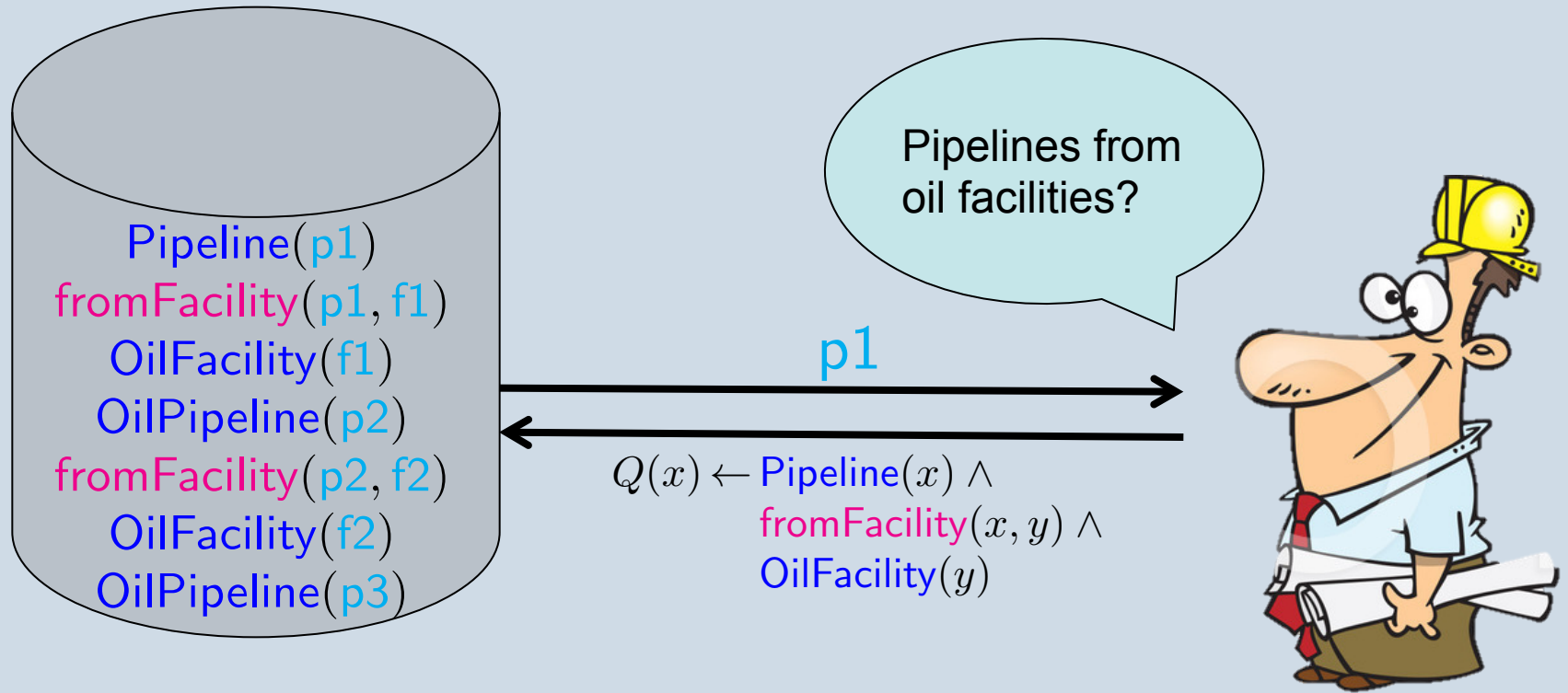
Pipelines from oil facilities?



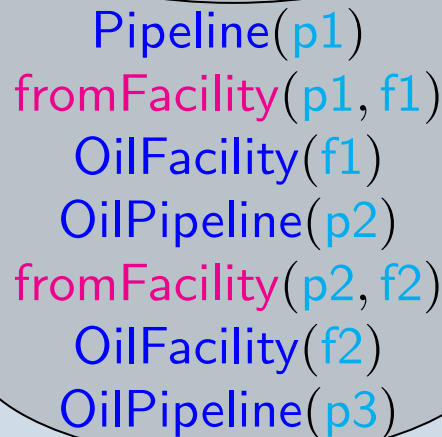
$$Q(x) \leftarrow \text{Pipeline}(x) \wedge \text{fromFacility}(x, y) \wedge \text{OilFacility}(y)$$



Applications: OBDA



Applications: OBDA



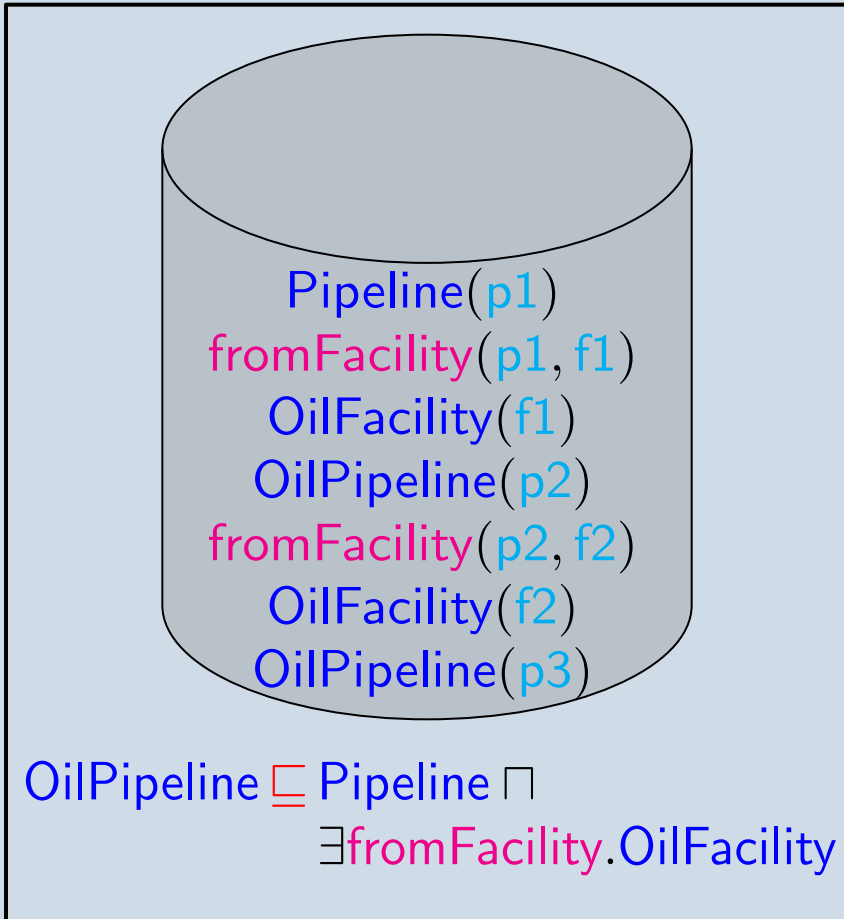
Pipeline(p1)
fromFacility(p1, f1)
OilFacility(f1)
OilPipeline(p2)
fromFacility(p2, f2)
OilFacility(f2)
OilPipeline(p3)

OilPipeline \sqsubseteq Pipeline \sqcap
 \exists fromFacility.OilFacility

Pipelines from
oil facilities?



Applications: OBDA

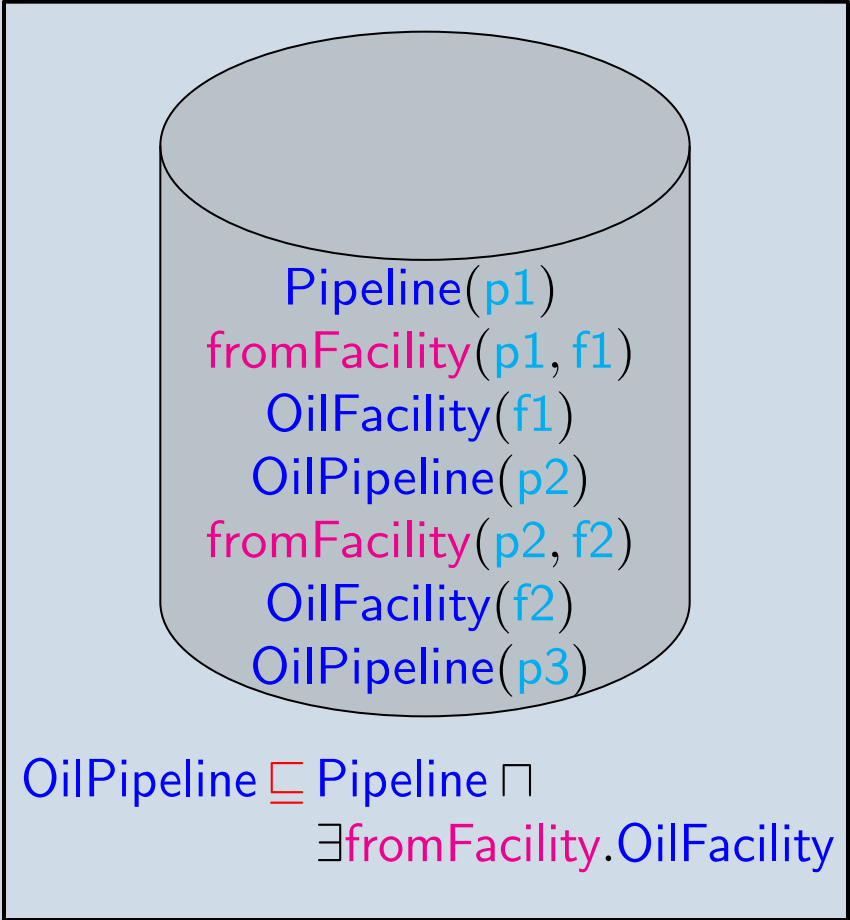


Pipelines from oil facilities?

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

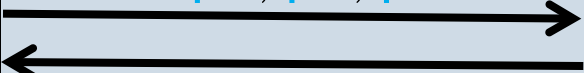


Applications: OBDA



Pipelines from oil facilities?

p1, p2, p3



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



OBDA: Practical Issues

- Scalability is always an issue, particularly when we have **large and heterogeneous data sources**
- Can be a large (startup) cost for **building and maintaining ontologies (and mappings)**
- Users may still have difficulty **formulating queries**



Optique Objectives

- Provide **semantic end-to-end connection** between users and data sources



DEPARTMENT OF
**COMPUTER
SCIENCE**

Optique



Optique Objectives

- Provide **semantic end-to-end connection** between users and data sources
- Enable users to rapidly formulate **intuitive queries** using familiar vocabularies and conceptualisations

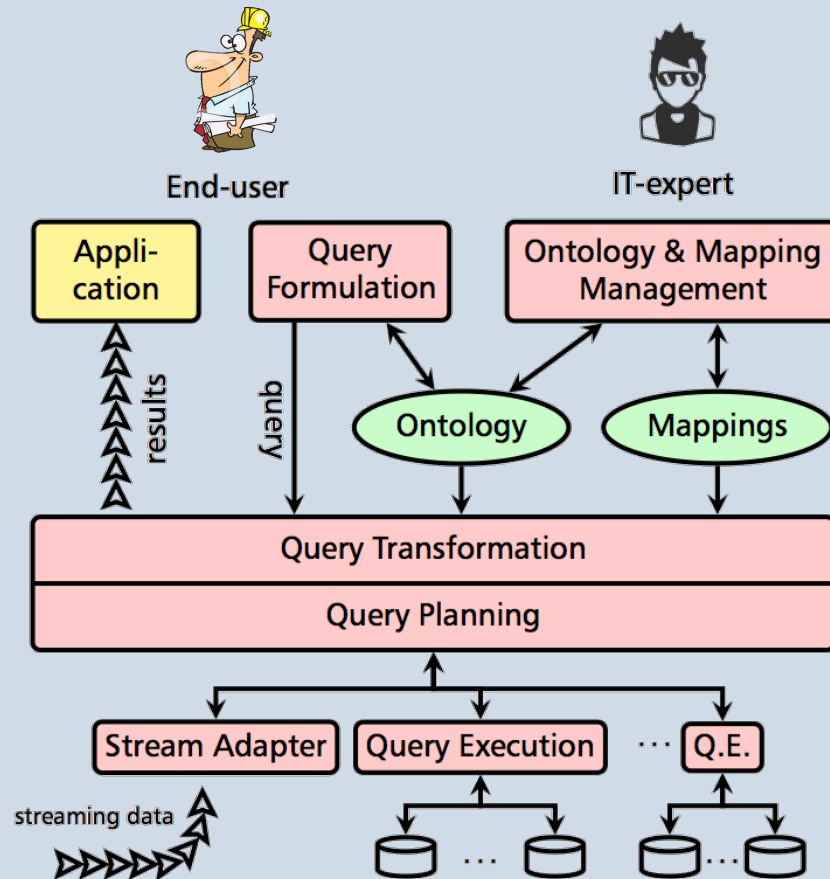


Optique Objectives

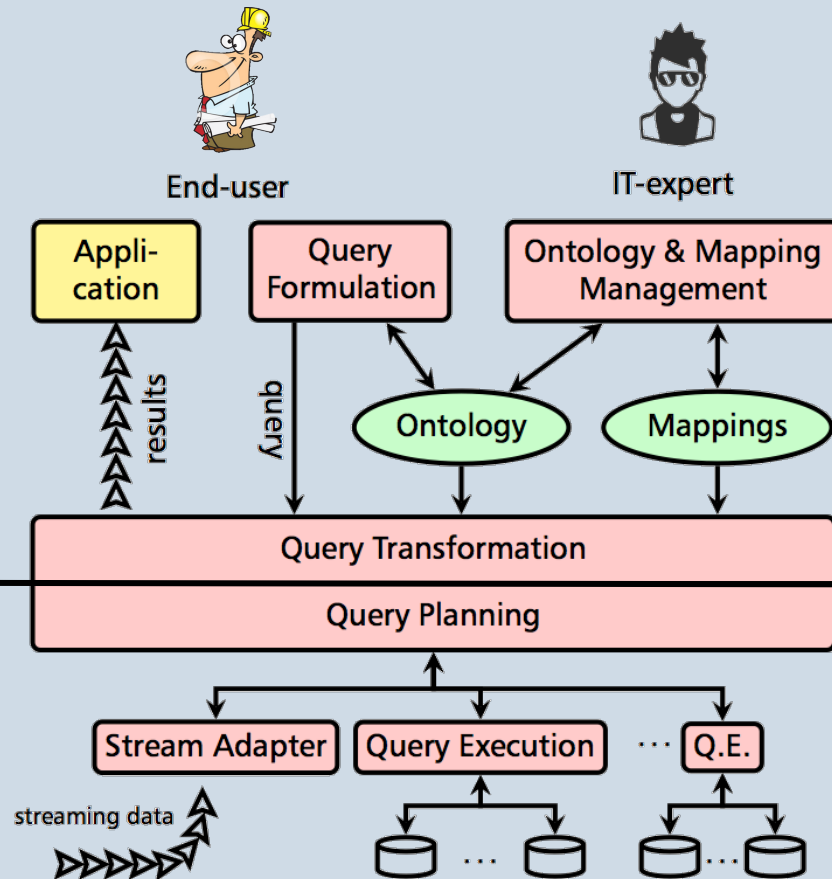
- Provide **semantic end-to-end connection** between users and data sources
- Enable users to rapidly formulate **intuitive queries** using familiar vocabularies and conceptualisations
- Return **timely answers** from large scale and heterogeneous data sources



Optique[™] Solution



Optique[®] Solution

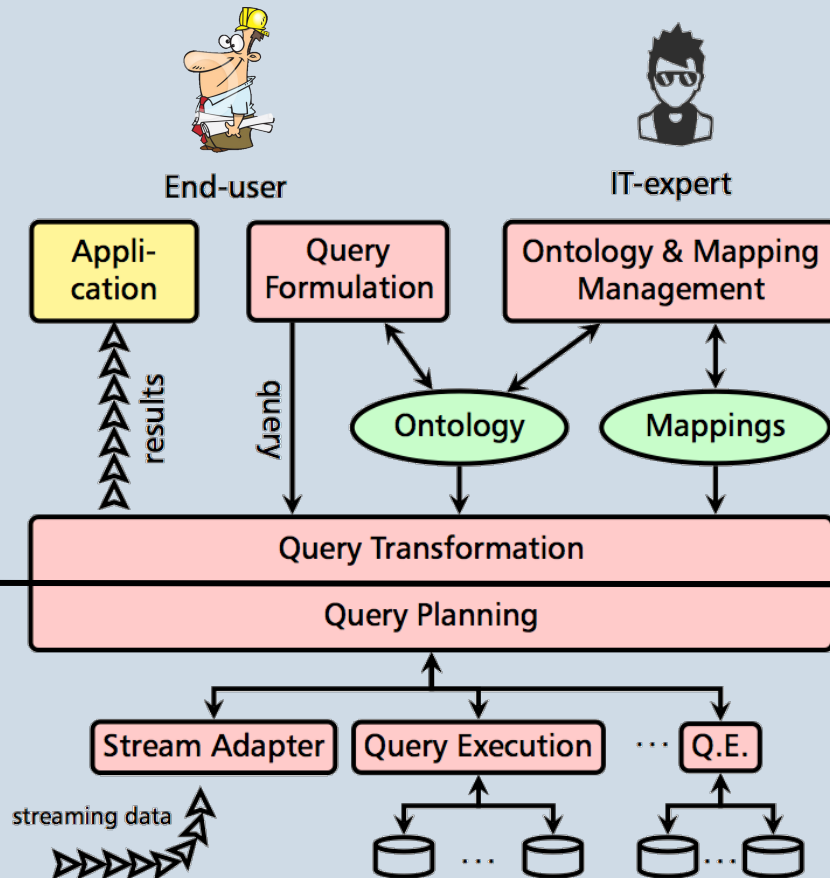


Optique[®] Solution



Query rewriting:

- uses ontology & mappings
- computationally hard
- ontology & mappings small



Optique Solution



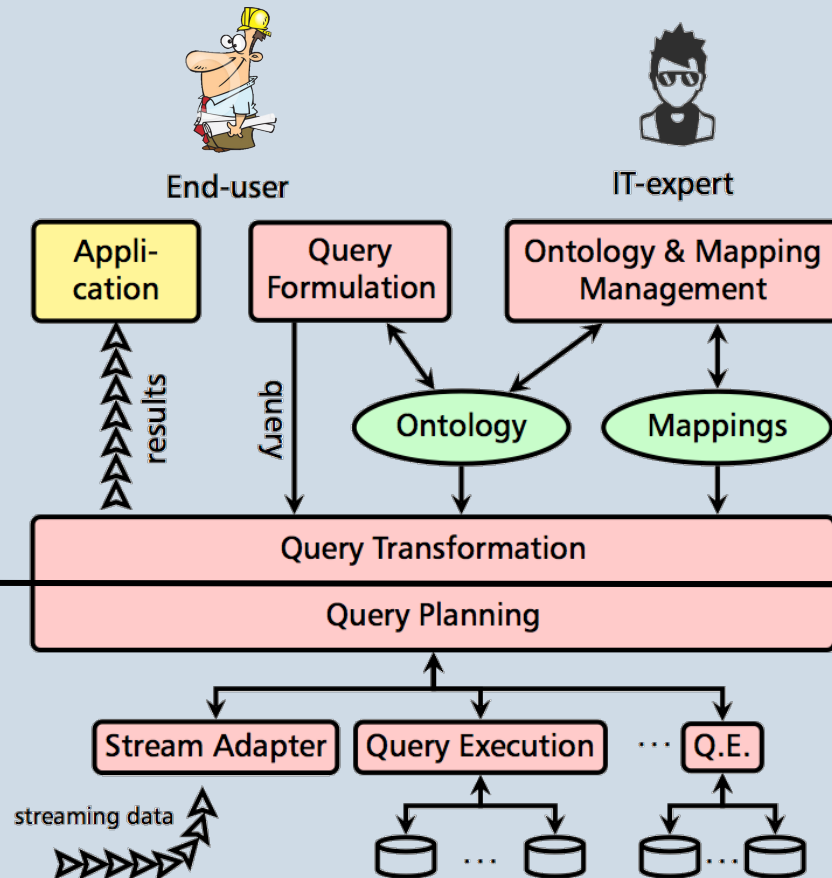
Query rewriting:

- uses ontology & mappings
- computationally hard
- ontology & mappings small



Query evaluation:

- ind. of ontology & mappings
- computationally tractable
- data sets very large



Query Rewriting

Given ontology \mathcal{O} query Q and mappings \mathcal{M} :



DEPARTMENT OF
**COMPUTER
SCIENCE**

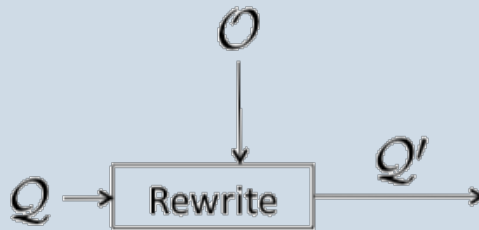
Optique



Query Rewriting

Given ontology \mathcal{O} query Q and mappings \mathcal{M} :

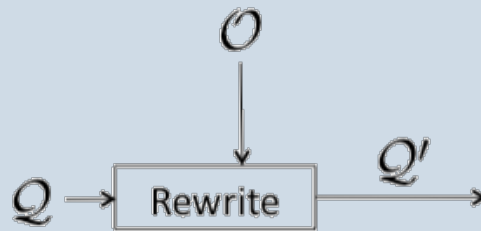
- **Rewrite** $Q \rightarrow Q'$ s.t. answering Q' without \mathcal{O} equivalent to answering Q w.r.t. \mathcal{O} *for any dataset*



Query Rewriting

Given ontology \mathcal{O} query Q and mappings \mathcal{M} :

- **Rewrite** $Q \rightarrow Q'$ s.t. answering Q' without \mathcal{O} equivalent to answering Q w.r.t. \mathcal{O} *for any dataset*

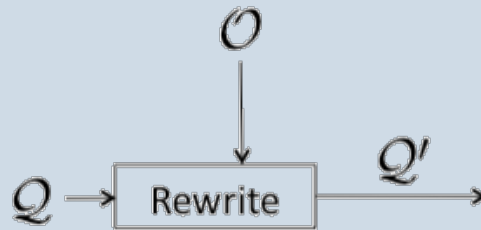


$$Q(x) \leftarrow \text{Pipeline}(x) \wedge \text{fromFacility}(x, y) \wedge \text{OilFacility}(y)$$

Query Rewriting

Given ontology \mathcal{O} query Q and mappings \mathcal{M} :

- **Rewrite** $Q \rightarrow Q'$ s.t. answering Q' without \mathcal{O} equivalent to answering Q w.r.t. \mathcal{O} *for any dataset*

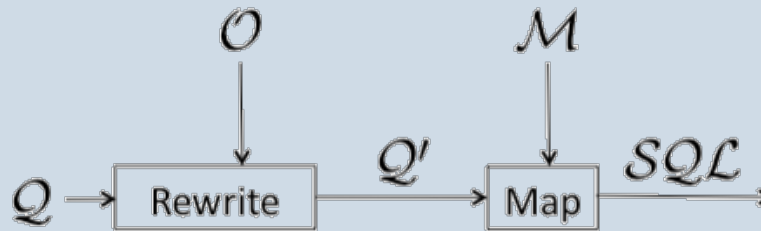


$$Q(x) \leftarrow \text{Pipeline}(x) \wedge \text{fromFacility}(x, y) \wedge \text{OilFacility}(y) \\ \vee \text{OilPipeline}(x)$$

Query Rewriting

Given ontology \mathcal{O} query Q and mappings \mathcal{M} :

- **Rewrite** $Q \rightarrow Q'$ s.t. answering Q' without \mathcal{O} equivalent to answering Q w.r.t. \mathcal{O} *for any dataset*
- **Map** ontology queries \rightarrow DB queries (typically SQL) using mappings \mathcal{M} to rewrite Q' into a DB query

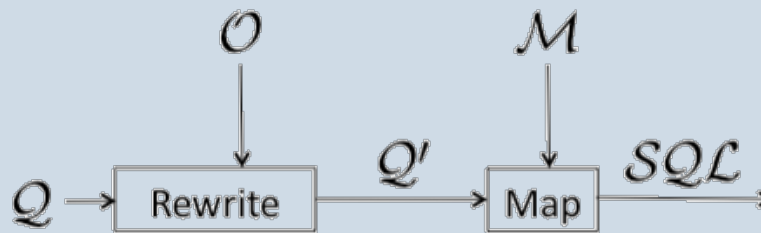


$$Q(x) \leftarrow \text{Pipeline}(x) \wedge \text{fromFacility}(x, y) \wedge \text{OilFacility}(y) \\ \vee \text{OilPipeline}(x)$$

Query Rewriting

Given ontology \mathcal{O} query Q and mappings \mathcal{M} :

- **Rewrite** $Q \rightarrow Q'$ s.t. answering Q' without \mathcal{O} equivalent to answering Q w.r.t. \mathcal{O} *for any dataset*
- **Map** ontology queries \rightarrow DB queries (typically SQL) using mappings \mathcal{M} to rewrite Q' into a DB query

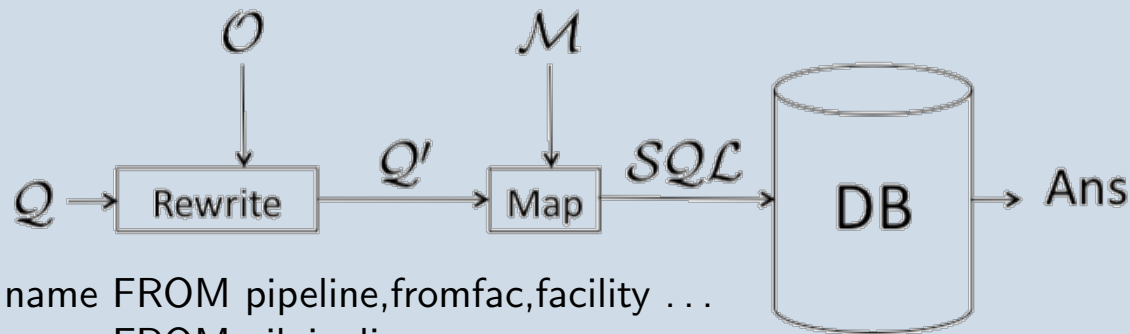


```
SELECT name FROM pipeline,fromfac,facility ...  
UNION SELECT name FROM oilpipeline
```

Query Rewriting

Given ontology \mathcal{O} query Q and mappings \mathcal{M} :

- **Rewrite** $Q \rightarrow Q'$ s.t. answering Q' without \mathcal{O} equivalent to answering Q w.r.t. \mathcal{O} *for any dataset*
- **Map** ontology queries \rightarrow DB queries (typically SQL) using mappings \mathcal{M} to rewrite Q' into a DB query
- **Evaluate** (SQL) query against DB



SELECT name FROM pipeline,fromfac,facility ...
UNION SELECT name FROM oilpipeline

Optique Solution



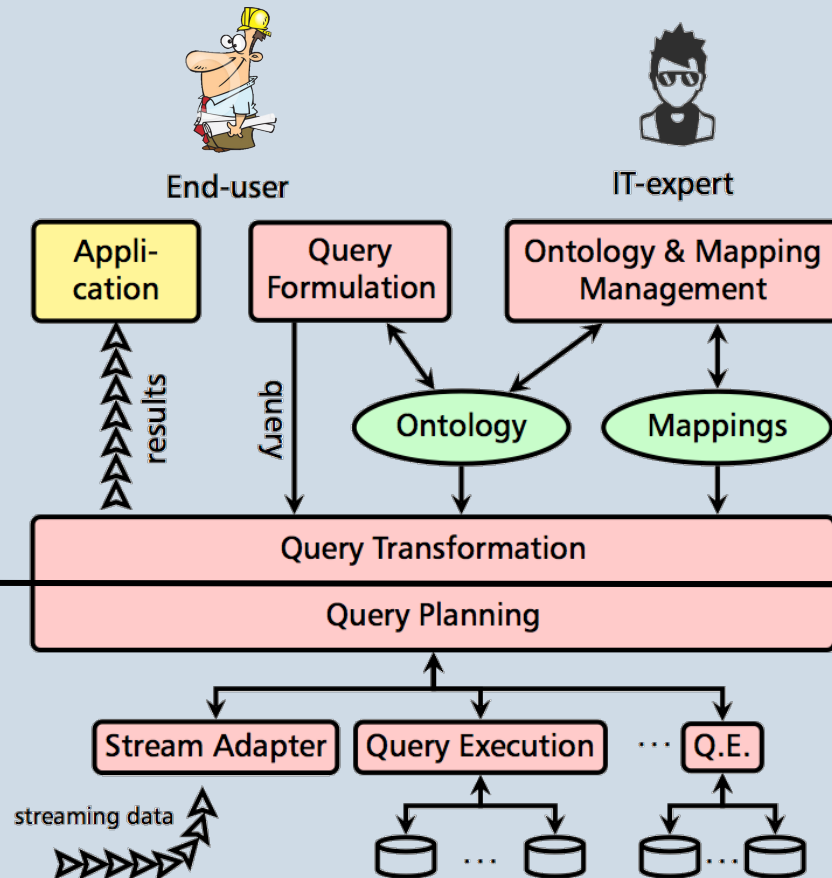
Query rewriting:

- uses ontology & mappings
- computationally hard
- ontology & mappings small



Query evaluation:

- ind. of ontology & mappings
- computationally tractable
- data sets very large



Optique Solution



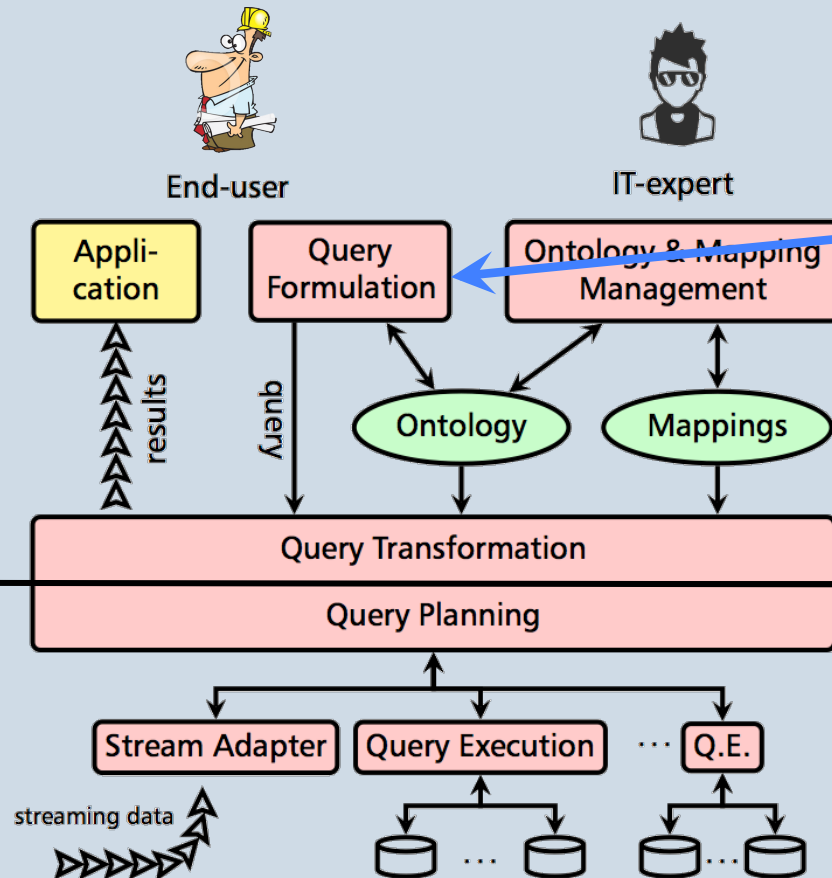
Query rewriting:

- uses ontology & mappings
- computationally hard
- ontology & mappings small



Query evaluation:

- ind. of ontology & mappings
- computationally tractable
- data sets very large



Other features:

support for query formulation

Query Formulation

The screenshot displays the Optique query formulation interface. On the left, a sidebar lists three data sources: Wellbores (300 items), Cores (900 items), and Stratigraphic Units (200 items). The main panel features a search bar and a filter section with three fields: 'Interval name' (text input), 'A field' (toggle set to 'Yes'), and two 'A field' sliders (one with a value of 50, the other with values 40 and 80). Below the interface is a horizontal flow diagram showing the sequence of entities: Query -> Well -> Wellbore -> Core -> DepthInterval.

Optique Solution



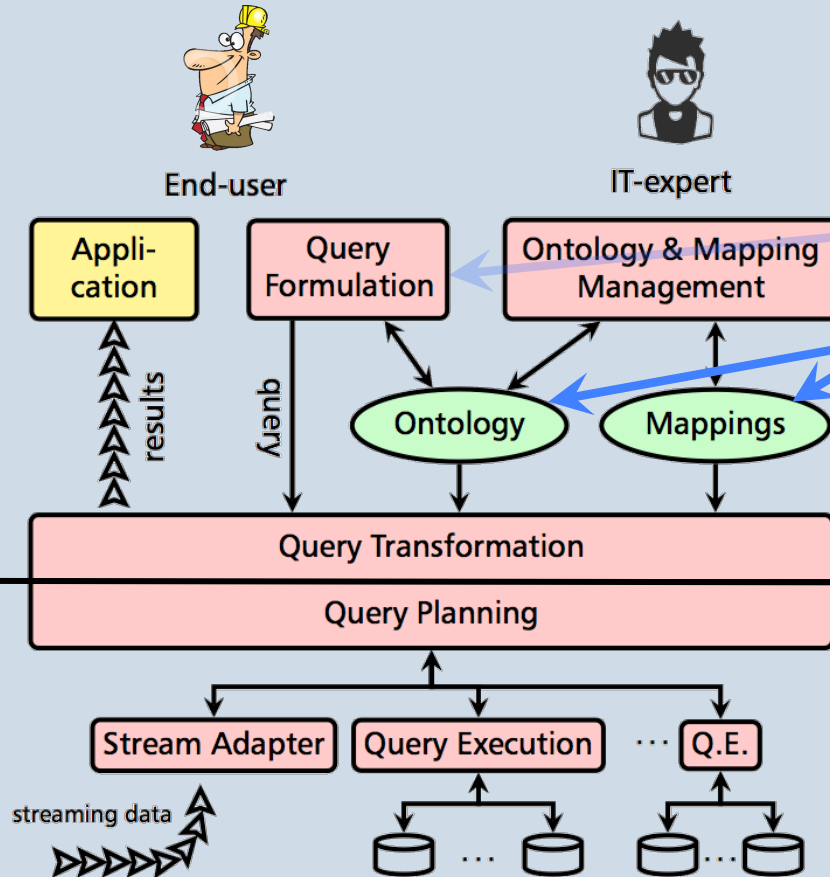
Query rewriting:

- uses ontology & mappings
- computationally hard
- ontology & mappings small



Query evaluation:

- ind. of ontology & mappings
- computationally tractable
- data sets very large



Other features:

support for query formulation

Ontology & mappings extended "on the fly"

Optique Solution



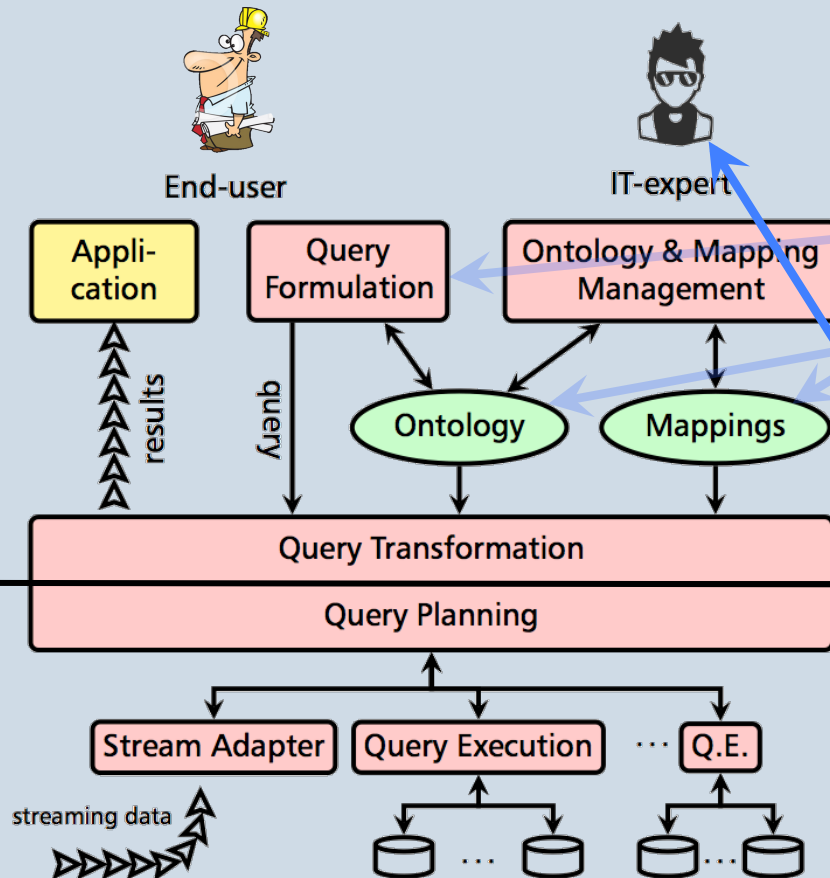
Query rewriting:

- uses ontology & mappings
- computationally hard
- ontology & mappings small



Query evaluation:

- ind. of ontology & mappings
- computationally tractable
- data sets very large



Other features:

- support for query formulation
- Ontology & mappings extended “on the fly”
- IT-expert oversees O&M management

Optique Solution



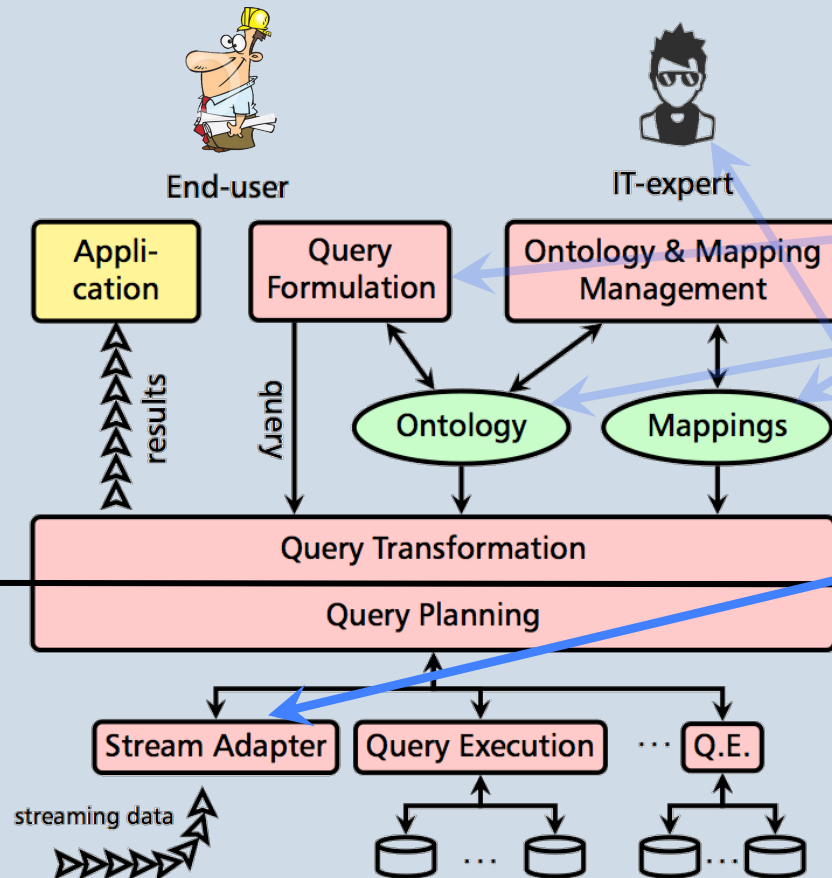
Query rewriting:

- uses ontology & mappings
- computationally hard
- ontology & mappings small



Query evaluation:

- ind. of ontology & mappings
- computationally tractable
- data sets very large



Other features:

- support for query formulation
- Ontology & mappings extended “on the fly”
- IT-expert oversees O&M management
- Adapter to support streaming data

Optique Solution



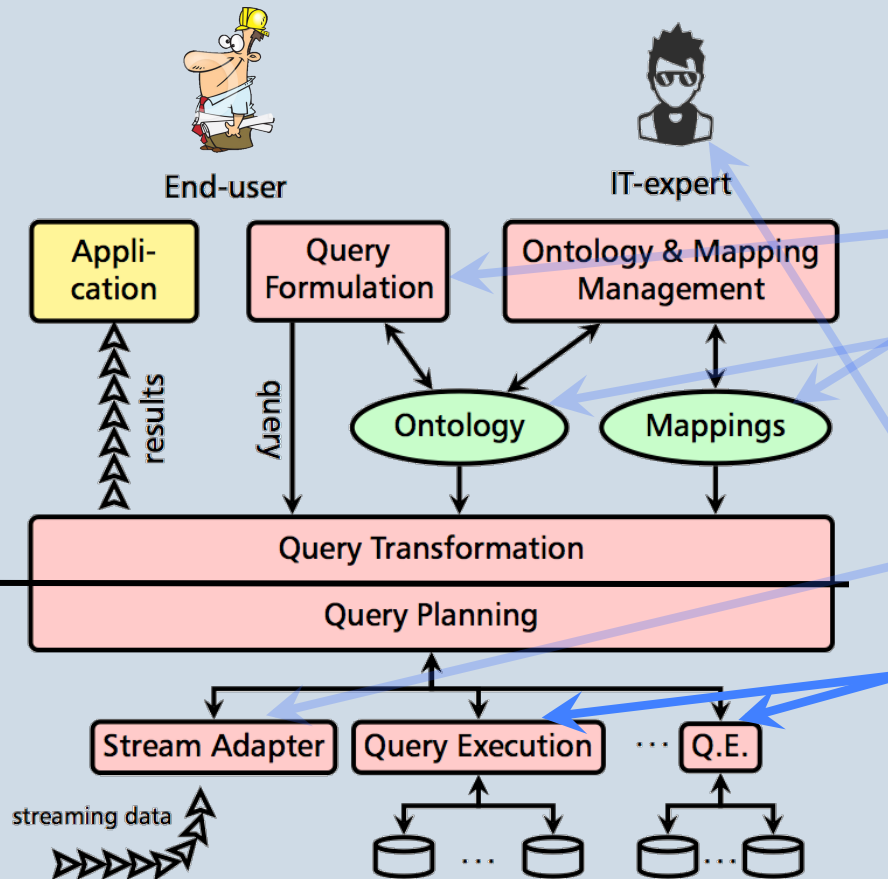
Query rewriting:

- uses ontology & mappings
- computationally hard
- ontology & mappings small



Query evaluation:

- ind. of ontology & mappings
- computationally tractable
- data sets very large



Other features:

- support for query formulation
- Ontology & mappings extended “on the fly”
- IT-expert oversees O&M management
- Adapter to support streaming data
- Distributed query execution



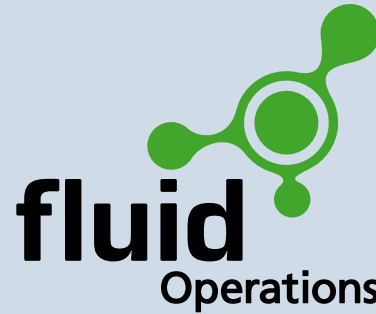
UiO : University of Oslo



UNIVERSITY OF
OXFORD

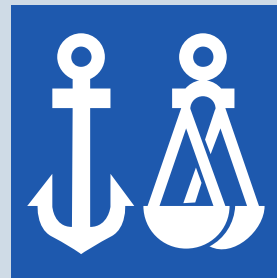


Statoil



National and Kapodistrian
UNIVERSITY OF ATHENS

SIEMENS



SAPIENZA
UNIVERSITÀ DI ROMA

TUHH

Hamburg University of Technology
Department of Telecommunications



FREIE UNIVERSITÄT BOZEN
LIBERA UNIVERSITÀ DI BOLZANO
FREE UNIVERSITY OF BOZEN · BOLZANO

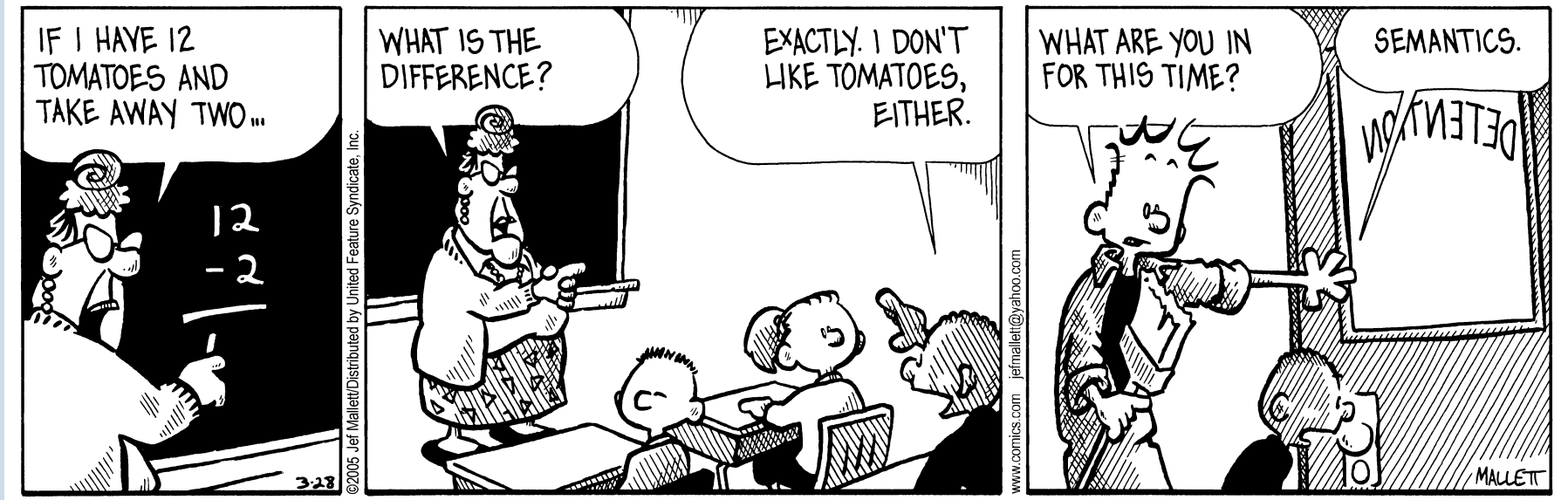


DEPARTMENT OF
COMPUTER
SCIENCE

Optique



Thank you for listening



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

Any questions?



DEPARTMENT OF
**COMPUTER
SCIENCE**

Optique

