DEPARTMENT OF COMPUTER SCIENCE

UNIVERSITY OF OXFORD

OBDA: Theory and Practice

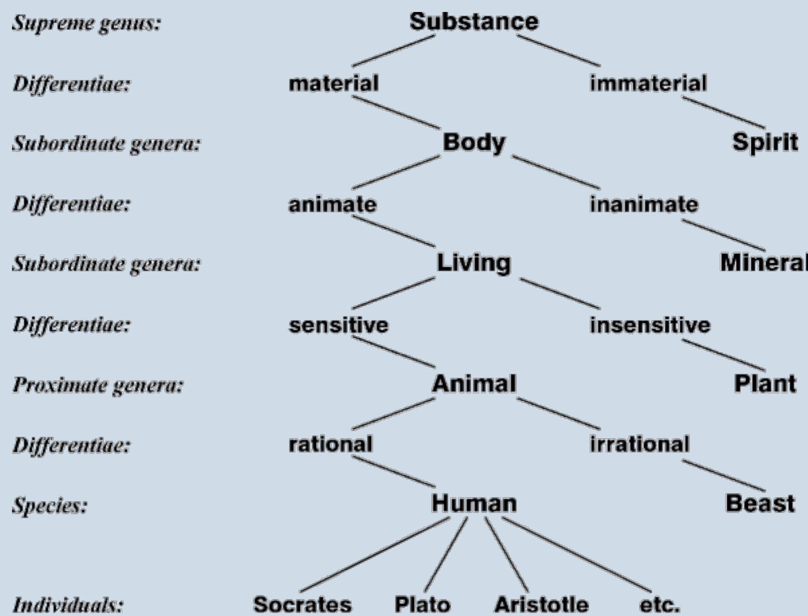
Ian Horrocks
Information Systems Group

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



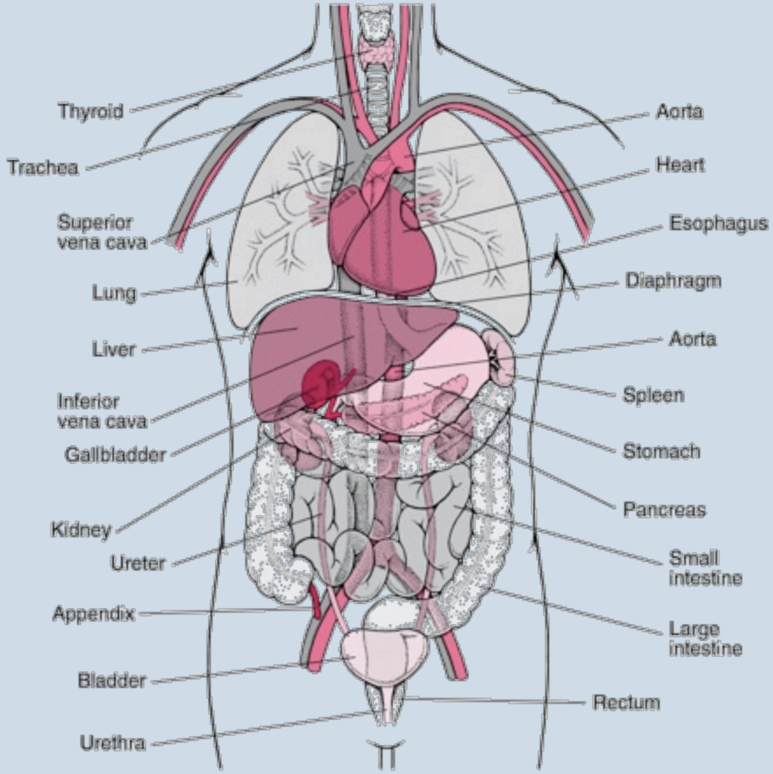
What is an Ontology?

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

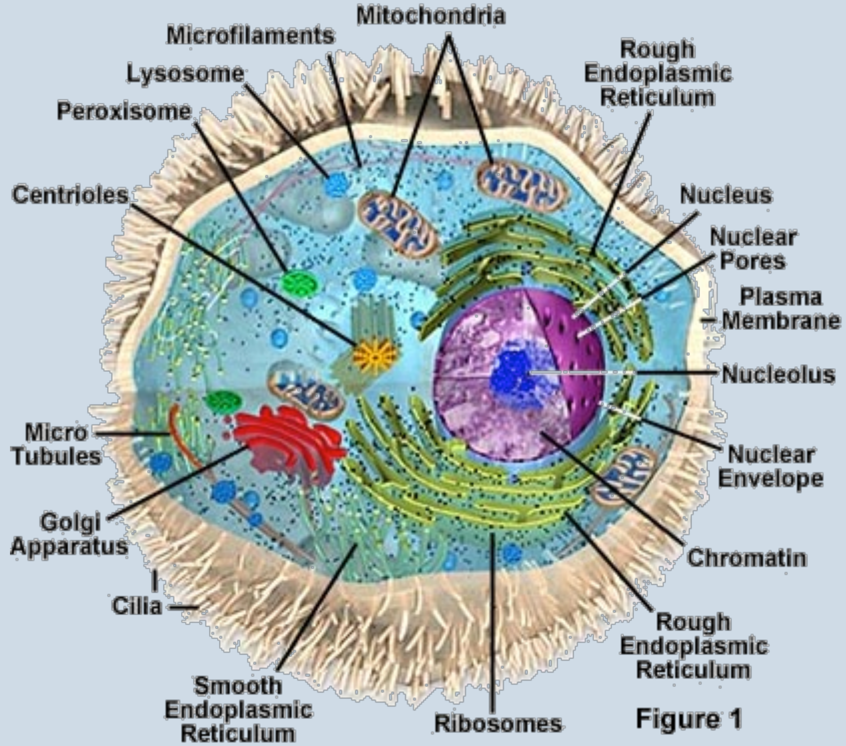


Figure 1

UNIVERSITY OF OXFORD DEPARTMENT OF COMPUTER SCIENCE

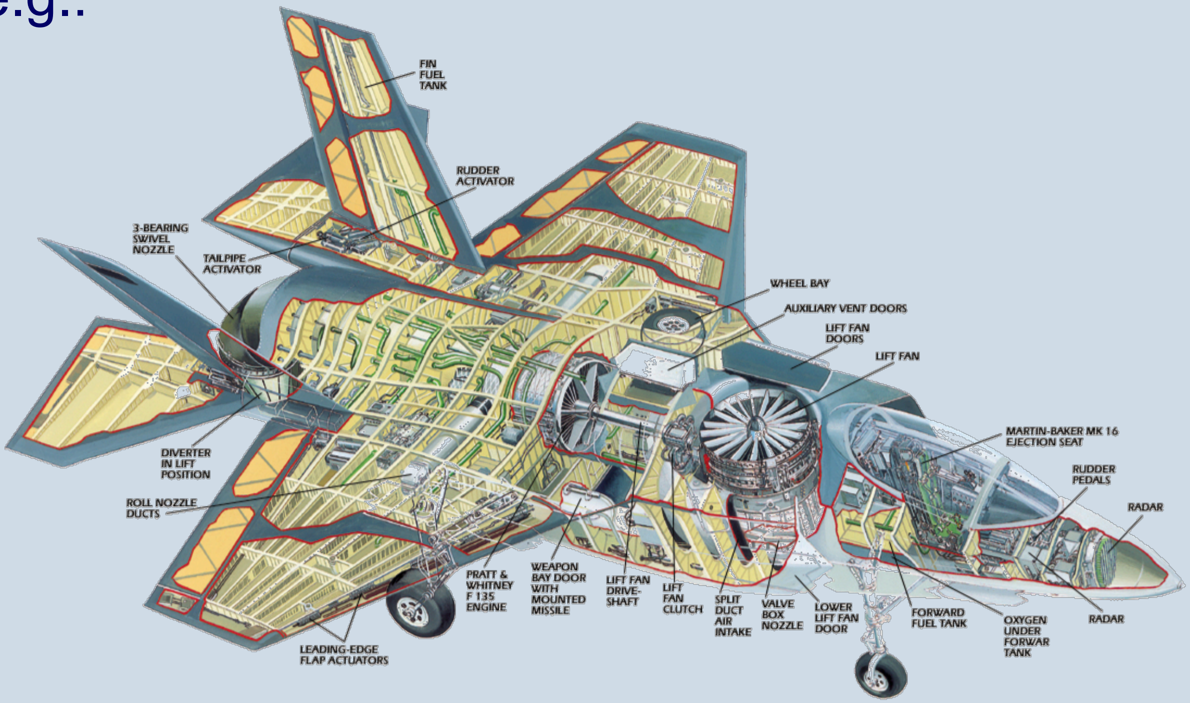
Optique

SEVENTH FRAMEWORK PROGRAMME

What is an Ontology?

A conceptual model of (some aspect of) the world

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



DEPARTMENT OF COMPUTER SCIENCE

UNIVERSITY OF OXFORD

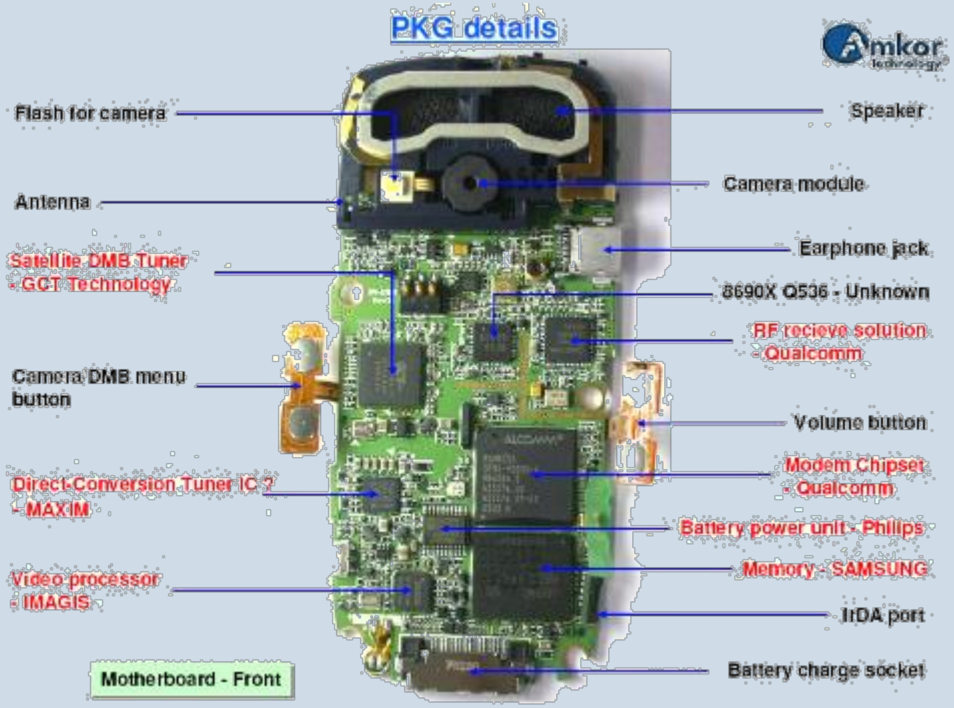
Optique

SEVENTH FRAMEWORK PROGRAMME

What is an Ontology?

A conceptual model of (some aspect of) the world

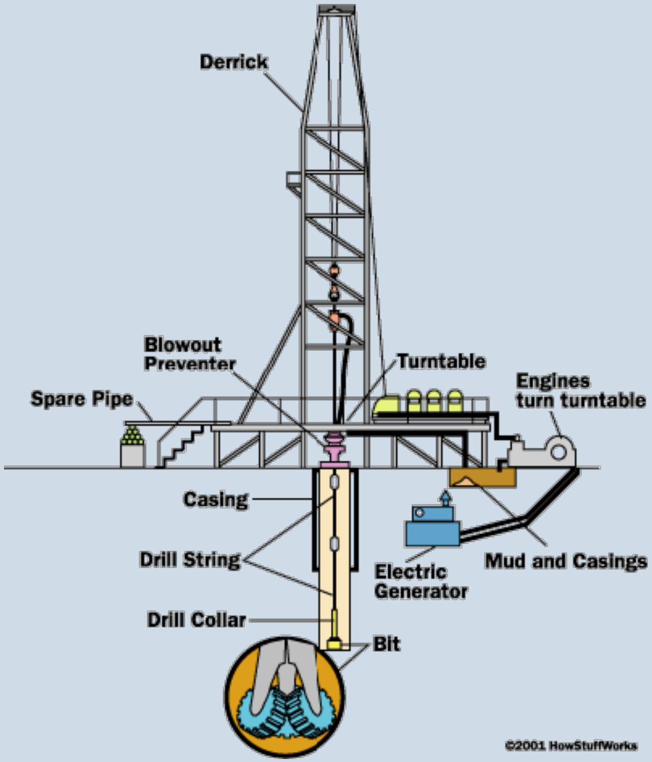
- Introduces **vocabulary** relevant to domain, e.g.:
  - Anatomy
  - Cellular biology
  - Aerospace
  - Cell Phones



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

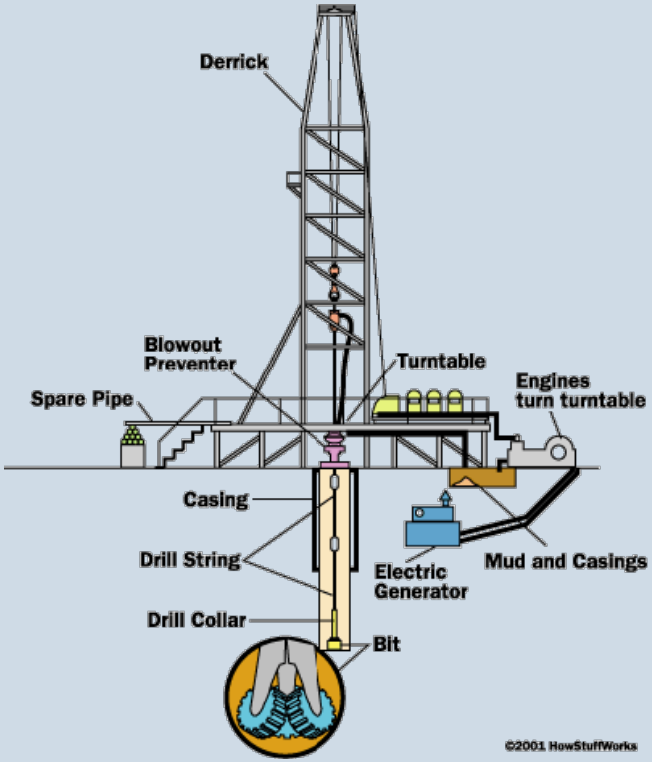


What is an Ontology?

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?

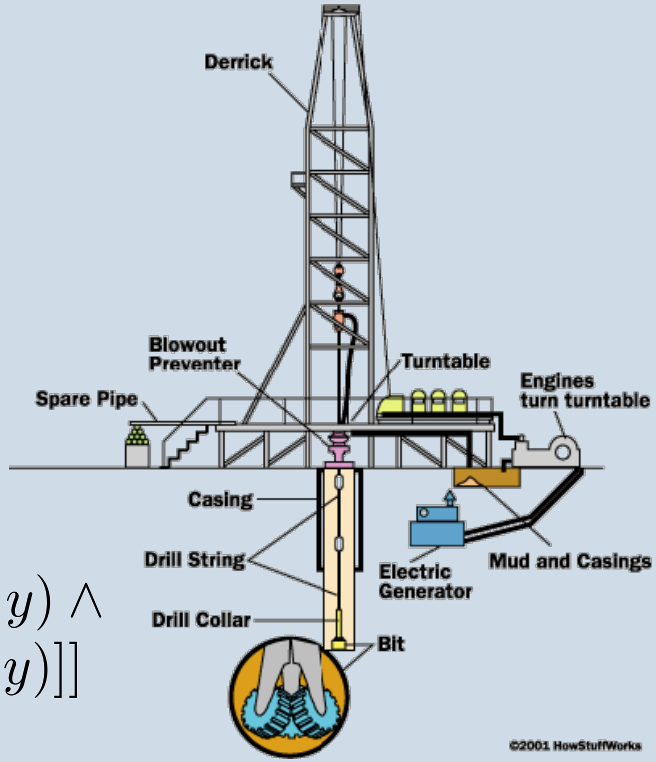
A conceptual model of (some aspect of) the world

- 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

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



What is an Ontology?

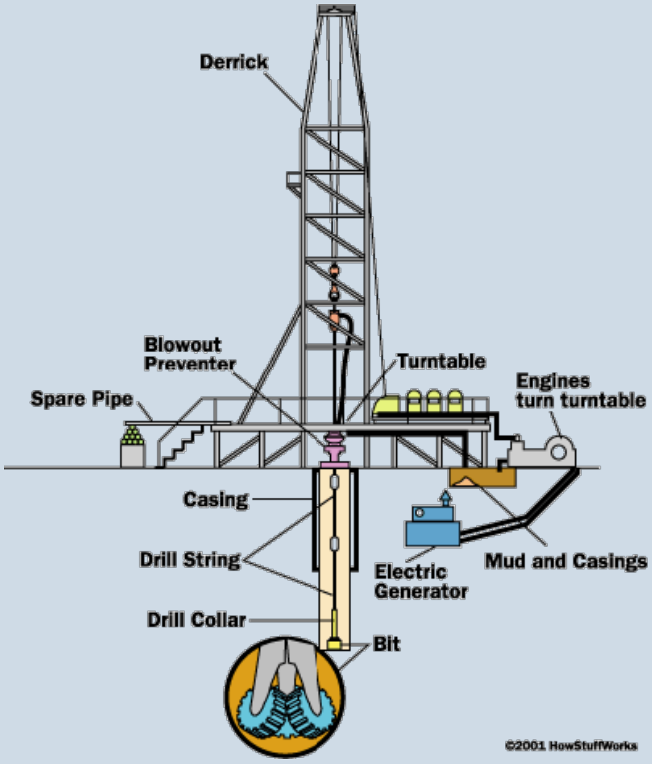
A conceptual model of (some aspect of) the world

- 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

$$\mathrm{OilPipeline} \sqsubseteq \mathrm{Pipeline} \sqcap \exists \mathrm{fromFacility}.\mathrm{OilFacility}$$



Applications: Semantic Web

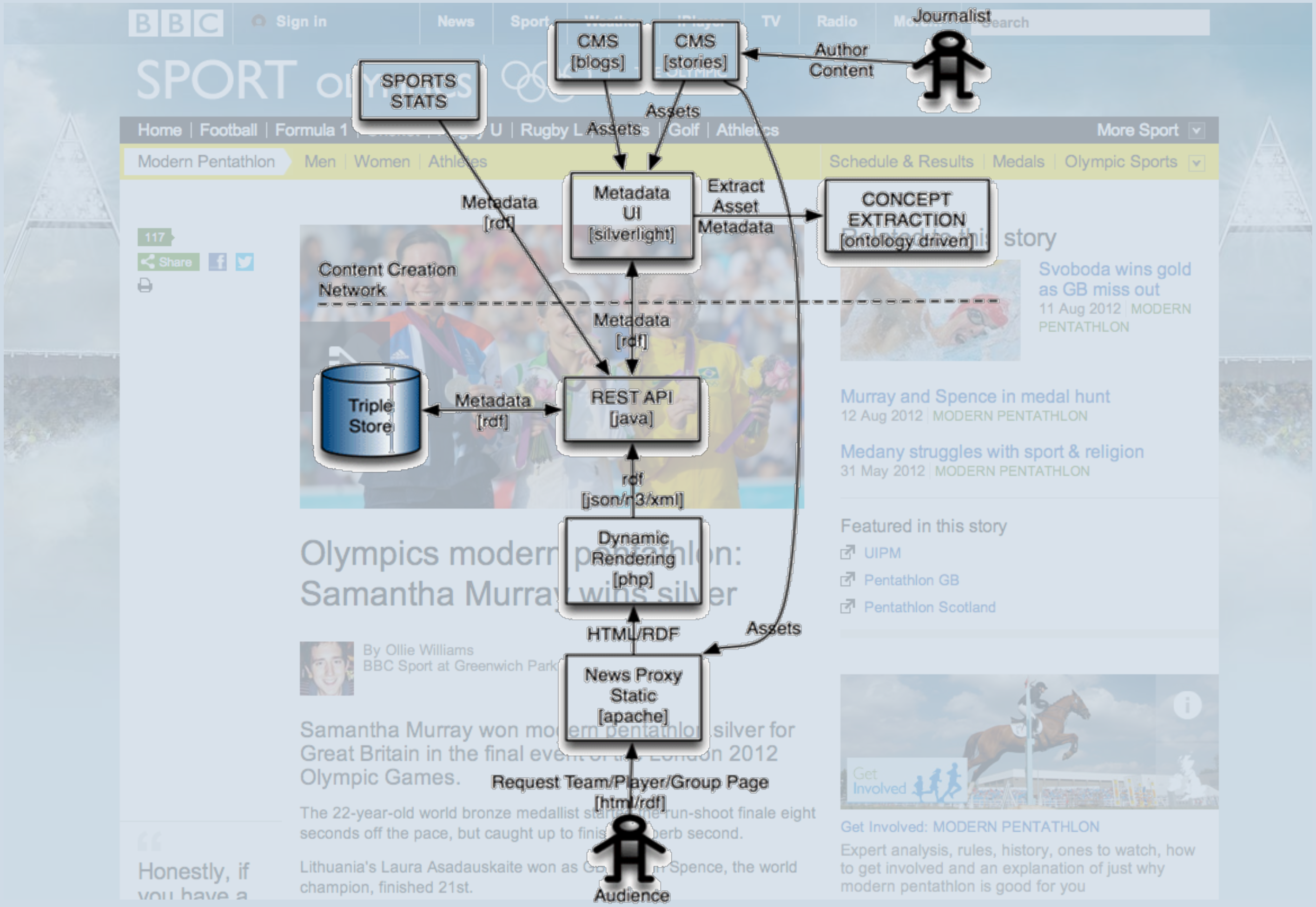
UNIVERSITY OF OXFORD

DEPARTMENT OF COMPUTER SCIENCE

Optique

SEVENTH FRAMEWORK PROGRAMME

Applications: Semantic Web



UNIVERSITY OF OXFORD DEPARTMENT OF COMPUTER SCIENCE

Optique

SEVENTH FRAMEWORK PROGRAMME

Semantic Technologies

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

Semantic Technologies

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

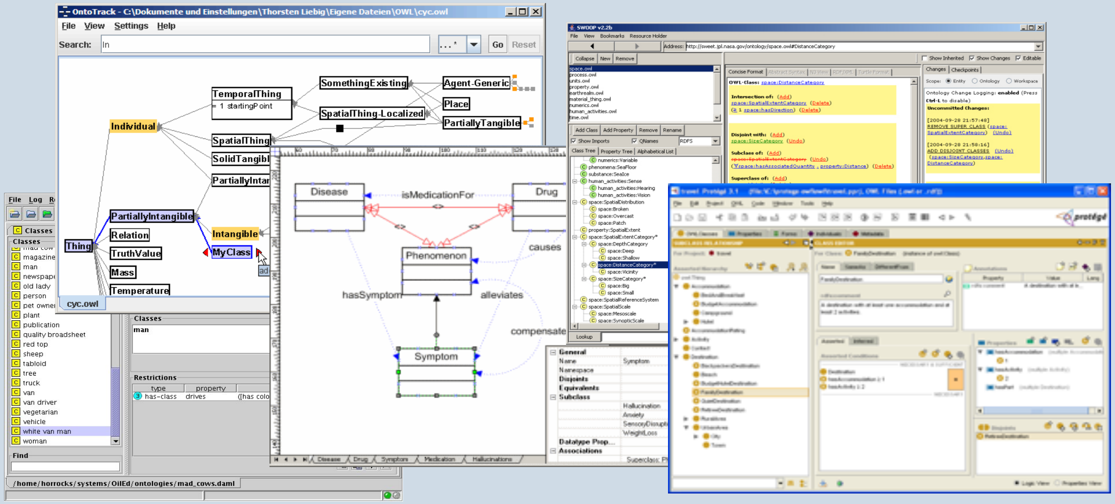
W3C OWL
W3C RDFa
W3C POWDER
W3C SKOS
W3C RDF
W3C SPARQL

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



UNIVERSITY OF OXFORD
DEPARTMENT OF COMPUTER SCIENCE

Optique

SEVENTH FRAMEWORK PROGRAMME

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

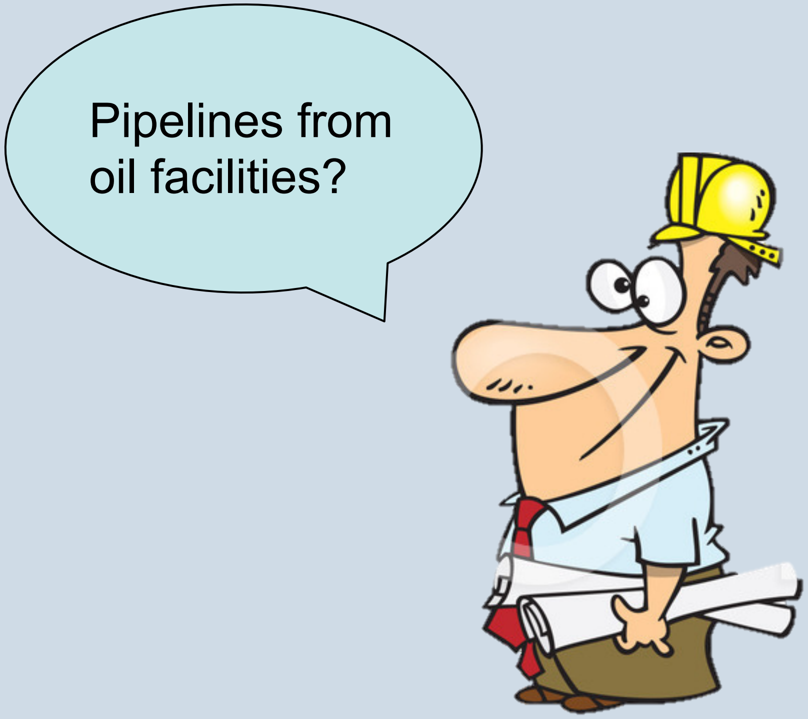
Applications: Intelligent Mobile Platform

- **Samsung** developing Intelligent Moblile 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**

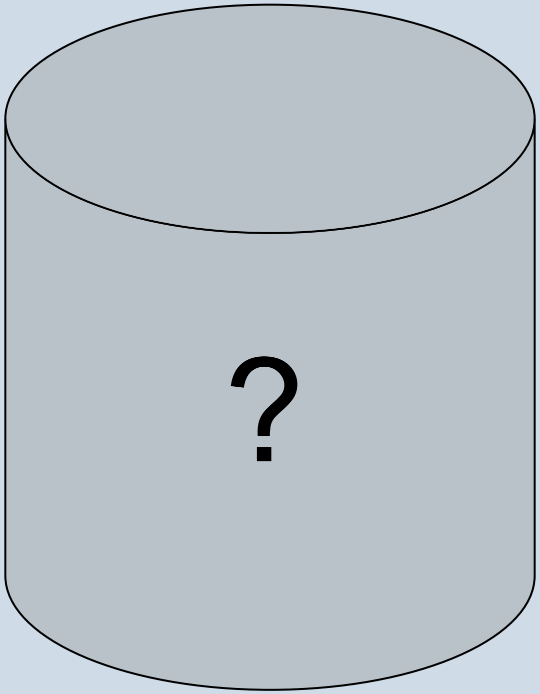
Applications: **OBDA**

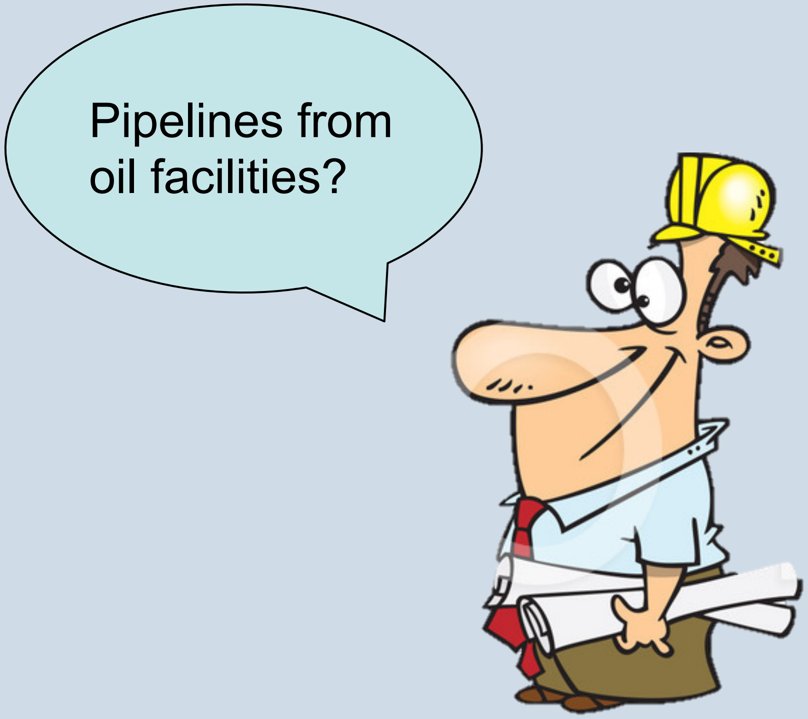
Applications: **OBDA**

Applications: **OBDA**

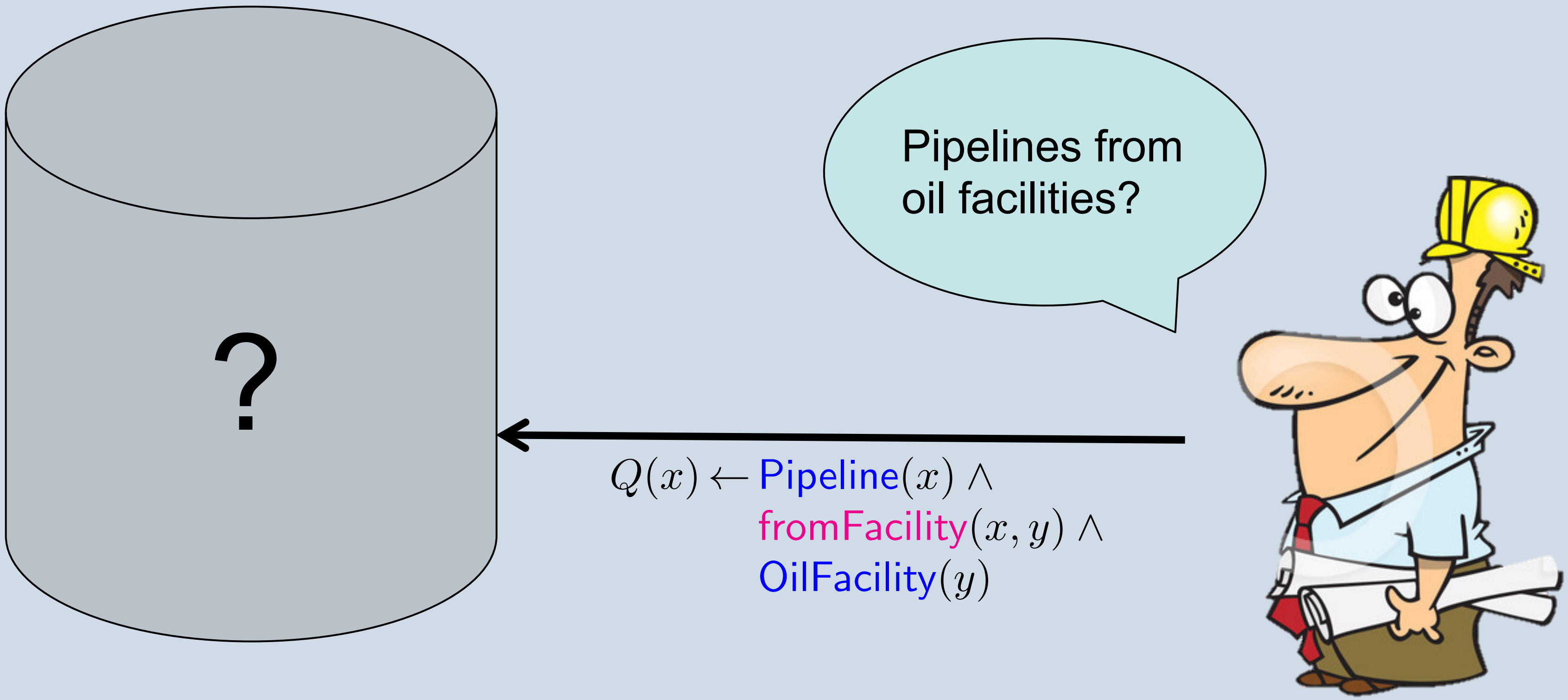


Applications: **OBDA**





Applications: **OBDA**



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

Applications: **OBDA**

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

Pipelines from oil facilities?

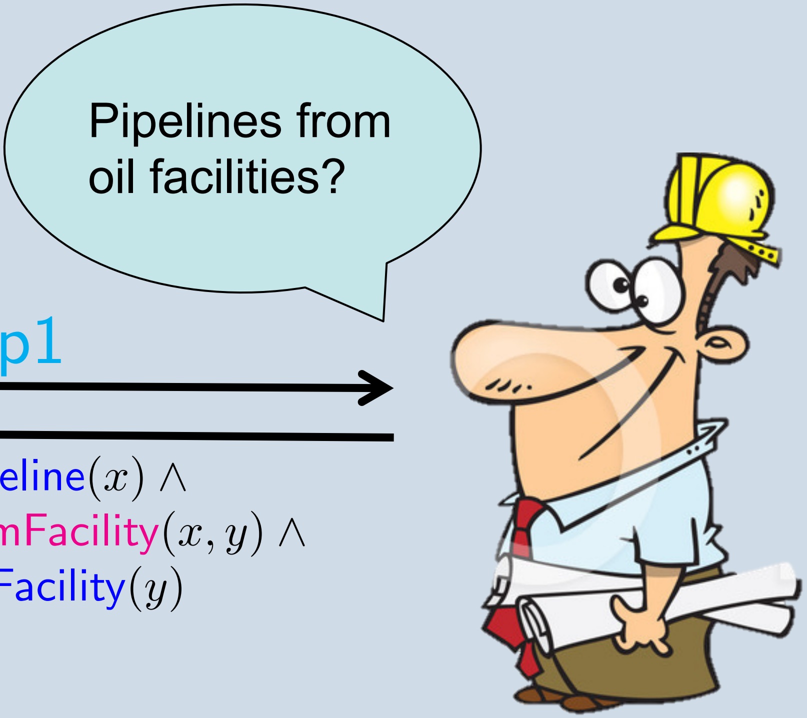
$$Q(x) \leftarrow \mathsf{Pipeline}(x) \wedge \mathsf{fromFacility}(x, y) \wedge \mathsf{OilFacility}(y)$$

Applications: **OBDA**

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

p1

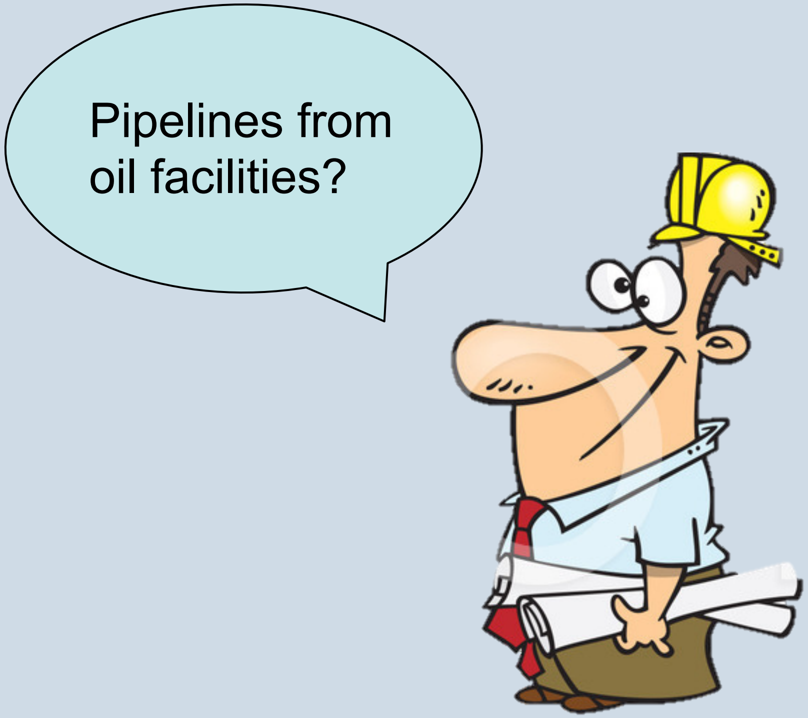
$$Q(x) \leftarrow \mathsf{Pipeline}(x) \wedge \mathsf{fromFacility}(x, y) \wedge \mathsf{OilFacility}(y)$$



Applications: **OBDA**

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

$$\mathsf{OilPipeline} \sqsubseteq \mathsf{Pipeline} \sqcap \exists \mathsf{fromFacility}.\mathsf{OilFacility}$$



Applications: **OBDA**

$\mathsf{Pipeline}(\mathsf{p1})$
$\mathsf{fromFacility}(\mathsf{p1}, \mathsf{f1})$
$\mathsf{OilFacility}(\mathsf{f1})$
$\mathsf{OilPipeline}(\mathsf{p2})$
$\mathsf{fromFacility}(\mathsf{p2}, \mathsf{f2})$
$\mathsf{OilFacility}(\mathsf{f2})$
$\mathsf{OilPipeline}(\mathsf{p3})$

$$\mathsf{OilPipeline} \sqsubseteq \mathsf{Pipeline} \sqcap \exists \mathsf{fromFacility}.\mathsf{OilFacility}$$

Pipelines from oil facilities?

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

UNIVERSITY OF OXFORD DEPARTMENT OF COMPUTER SCIENCE

Optique

SEVENTH FRAMEWORK PROGRAMME

Applications: **OBDA**

$\mathsf{Pipeline}(\mathsf{p1})$
$\mathsf{fromFacility}(\mathsf{p1}, \mathsf{f1})$
$\mathsf{OilFacility}(\mathsf{f1})$
$\mathsf{OilPipeline}(\mathsf{p2})$
$\mathsf{fromFacility}(\mathsf{p2}, \mathsf{f2})$
$\mathsf{OilFacility}(\mathsf{f2})$
$\mathsf{OilPipeline}(\mathsf{p3})$

$$\mathsf{OilPipeline} \sqsubseteq \mathsf{Pipeline} \sqcap \exists \mathsf{fromFacility}.\mathsf{OilFacility}$$

Pipelines from oil facilities?

$\mathsf{p1}, \mathsf{p2}, \mathsf{p3}$

$$Q(x) \leftarrow \mathsf{Pipeline}(x) \land \mathsf{fromFacility}(x, y) \land \mathsf{OilFacility}(y)$$

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

Optique Objectives

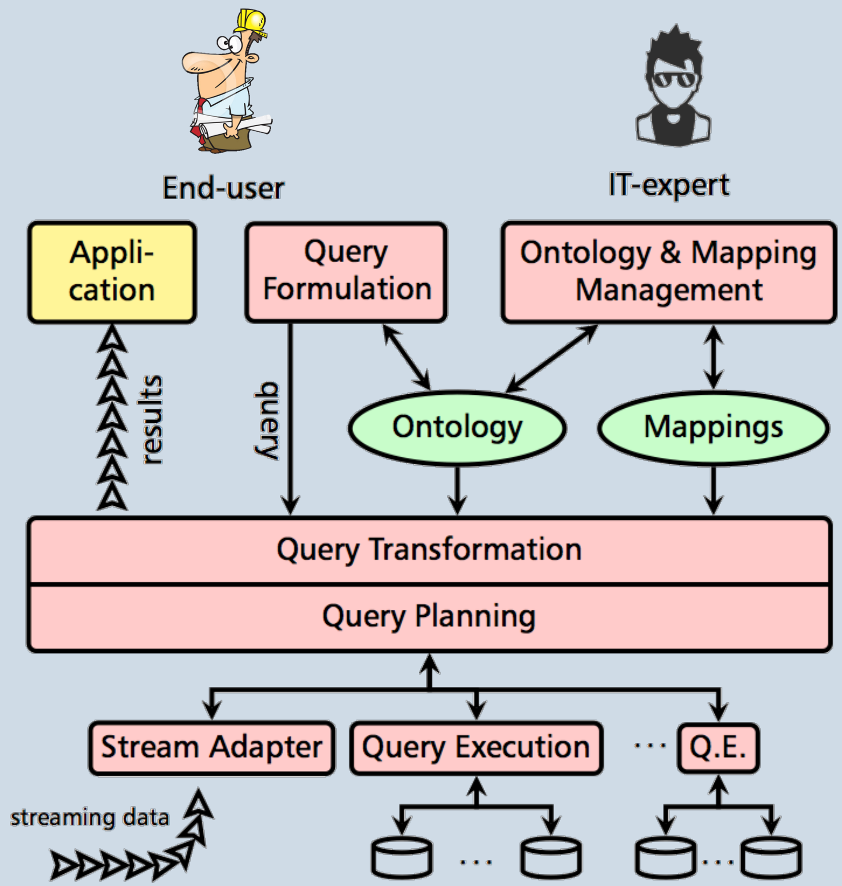
- Provide **semantic end-to-end connection** between users and data sources

- Enable users to rapidly formulate **intuitive queries** using familiar vocabularies and conceptualisations

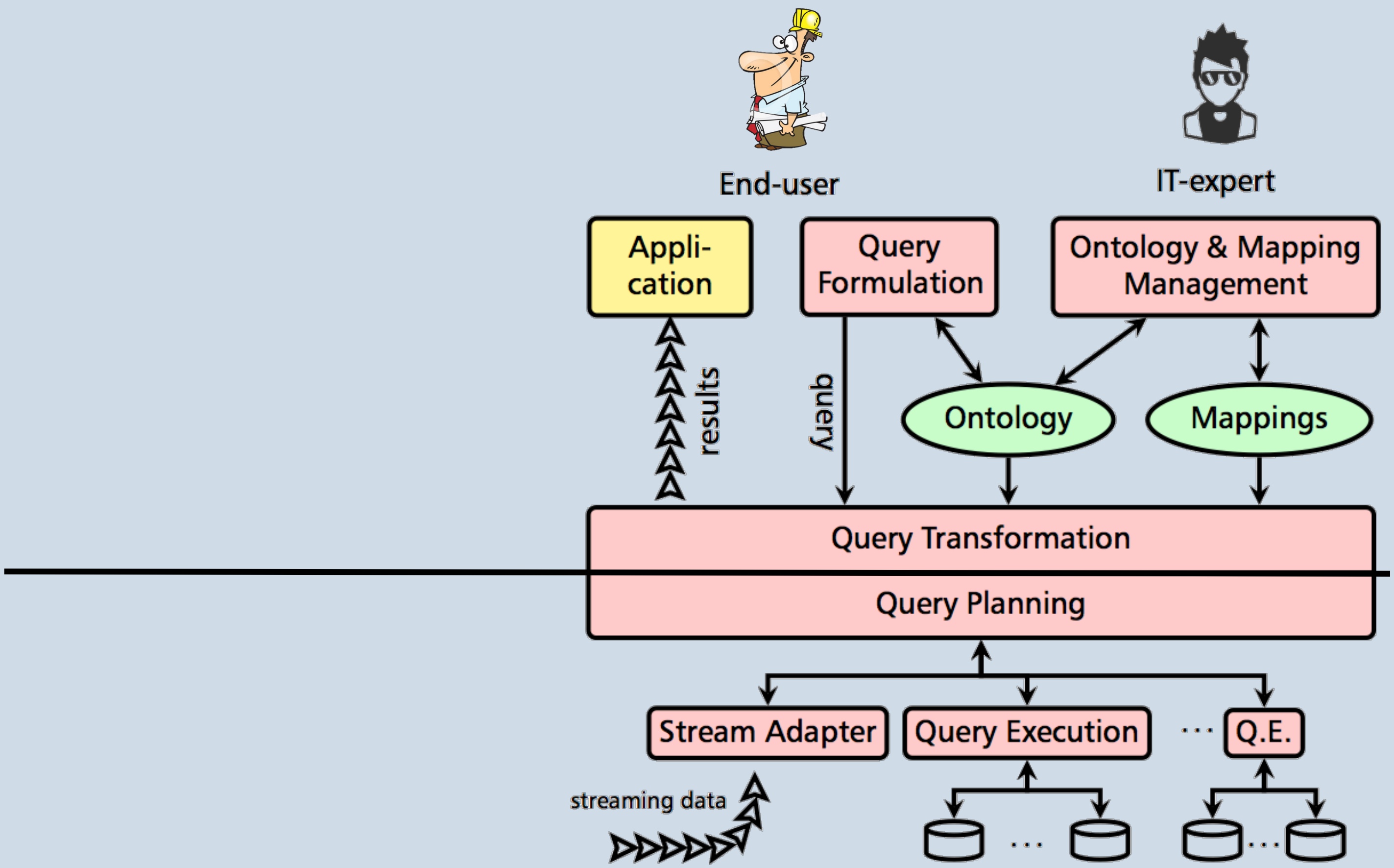
Optique Objectives

- 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



DEPARTMENT OF COMPUTER SCIENCE
UNIVERSITY OF OXFORD

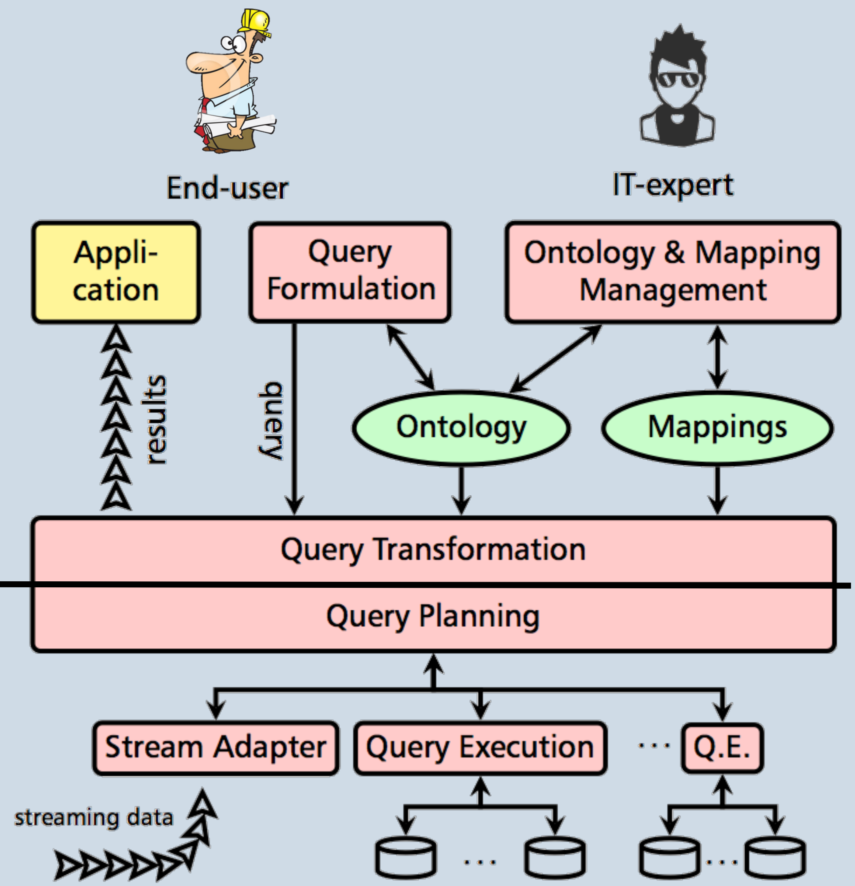
Optique

SEVENTH FRAMEWORK PROGRAMME

Optique Solution

Query rewriting:

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



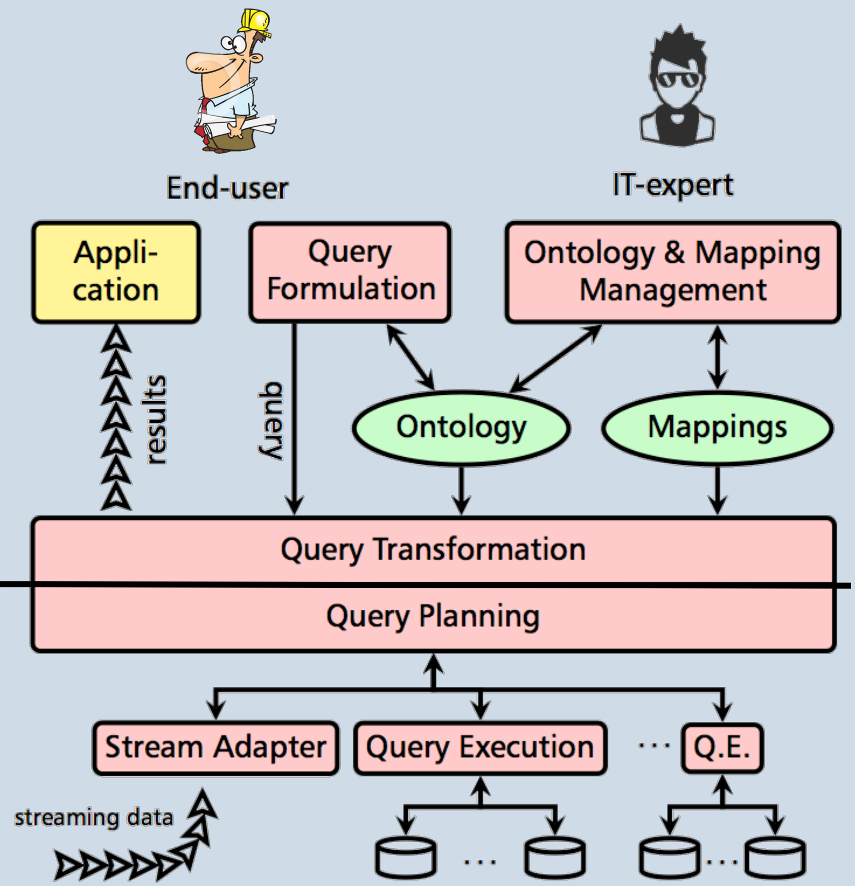
Optique Solution

Query rewriting:

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

Query evaluation:

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



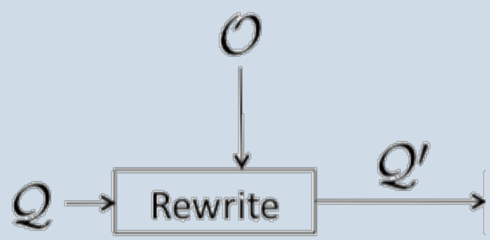
Query Rewriting

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

Query Rewriting

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

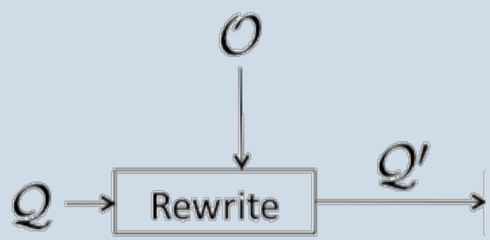
- **Rewrite** $\mathcal{Q} \rightarrow \mathcal{Q}'$ s.t. answering $\mathcal{Q}'$ without $\mathcal{O}$ equivalent to answering $\mathcal{Q}$ w.r.t. $\mathcal{O}$ *for any dataset*



Query Rewriting

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

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

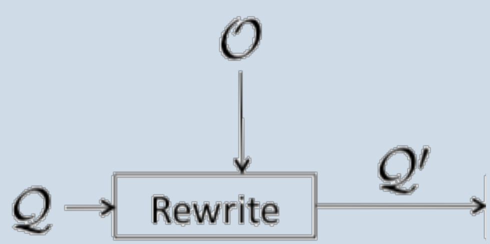


$$Q(x) \leftarrow \mathsf{Pipeline}(x) \wedge \mathsf{fromFacility}(x, y) \wedge \mathsf{OilFacility}(y)$$

Query Rewriting

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

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

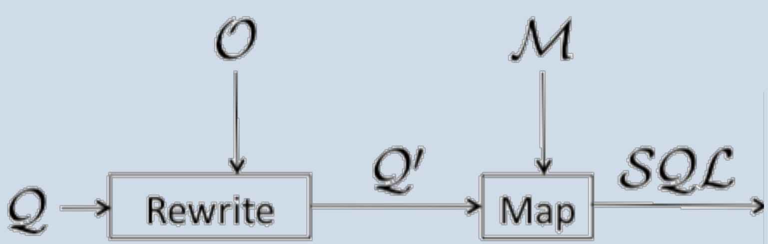


$$Q(x) \leftarrow \mathsf{Pipeline}(x) \wedge \mathsf{fromFacility}(x,y) \wedge \mathsf{OilFacility}(y)$$
$$\vee\ \mathsf{OilPipeline}(x)$$

Query Rewriting

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

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

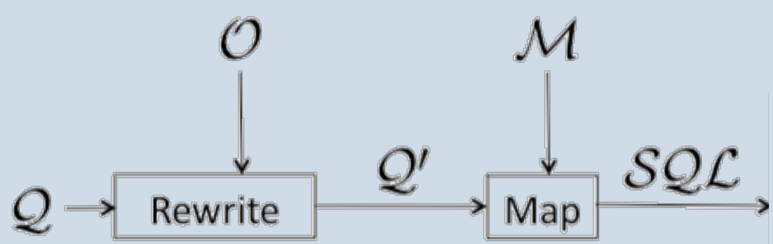


$$Q(x) \leftarrow \mathsf{Pipeline}(x) \wedge \mathsf{fromFacility}(x, y) \wedge \mathsf{OilFacility}(y)$$
$$\vee\ \mathsf{OilPipeline}(x)$$

Query Rewriting

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

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

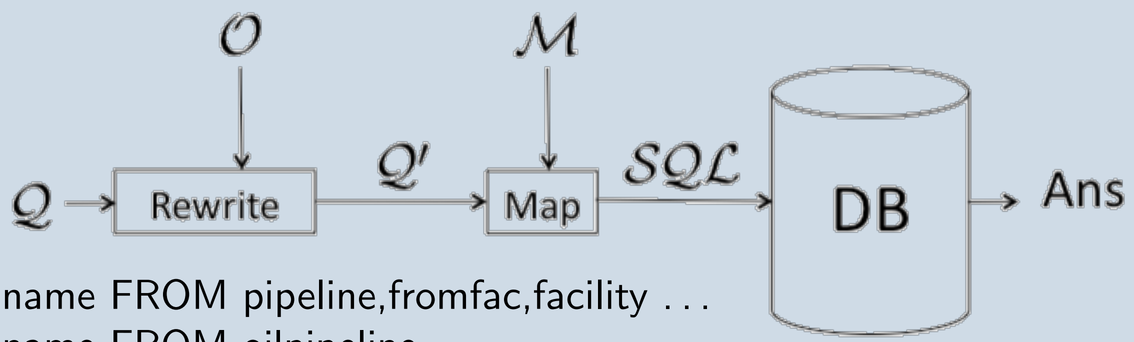


SELECT name FROM pipeline,fromfac,facility . . .
UNION SELECT name FROM oilpipeline

Query Rewriting

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

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



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

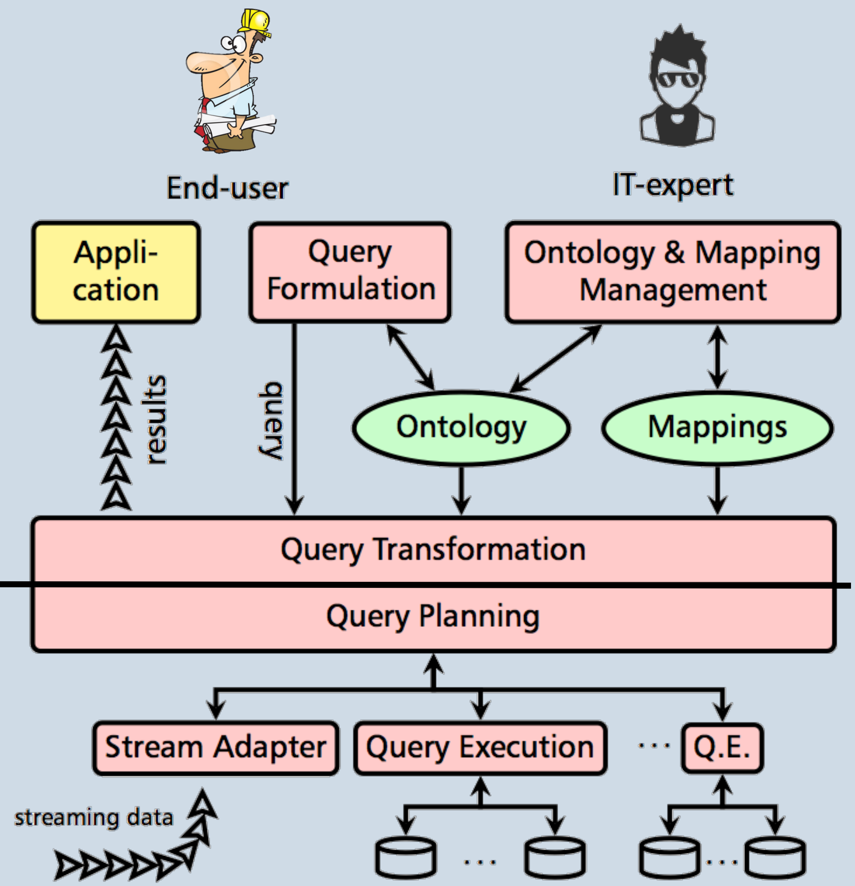
Optique Solution

Query rewriting:

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

Query evaluation:

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

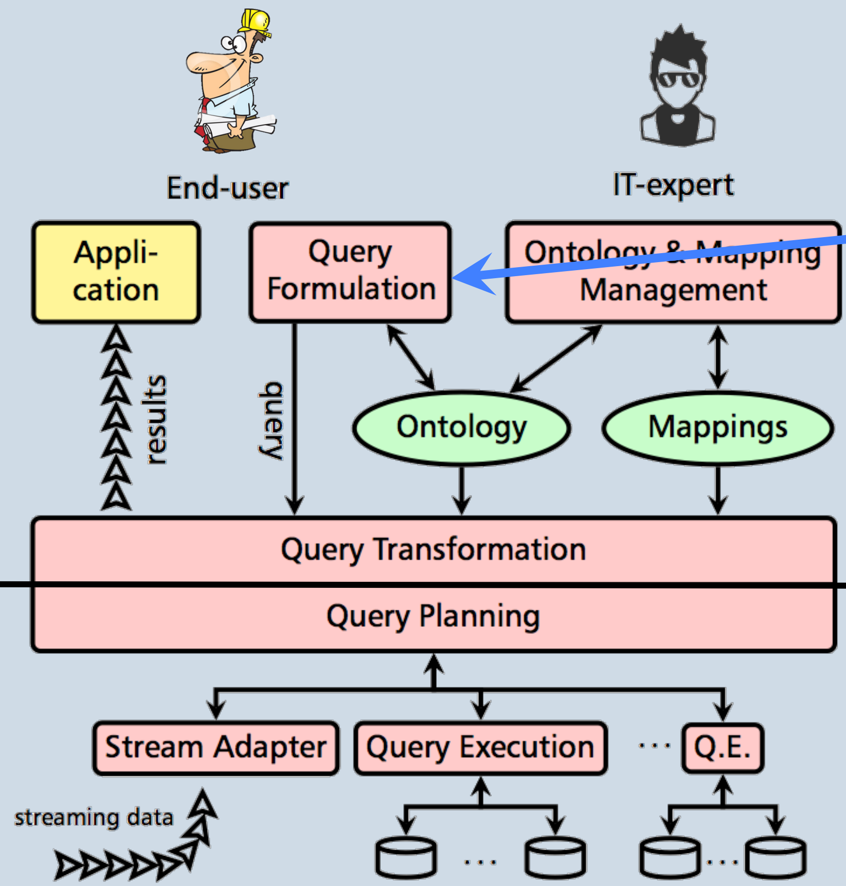


UNIVERSITY OF OXFORD
DEPARTMENT OF COMPUTER SCIENCE

Optique

SEVENTH FRAMEWORK PROGRAMME

Optique Solution



Other features:

support for query formulation

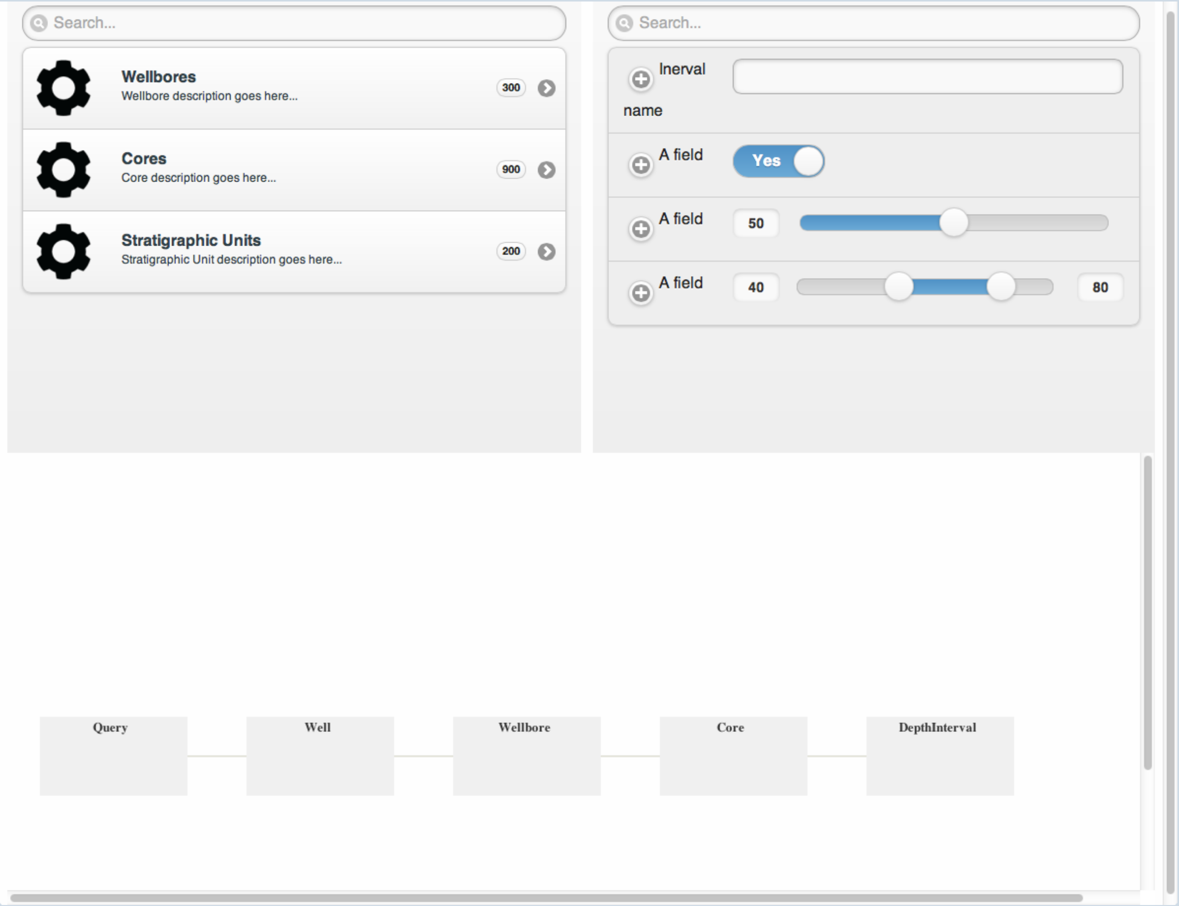
Query rewriting:

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

Query evaluation:

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

Query Formulation



UNIVERSITY OF OXFORD
DEPARTMENT OF COMPUTER SCIENCE

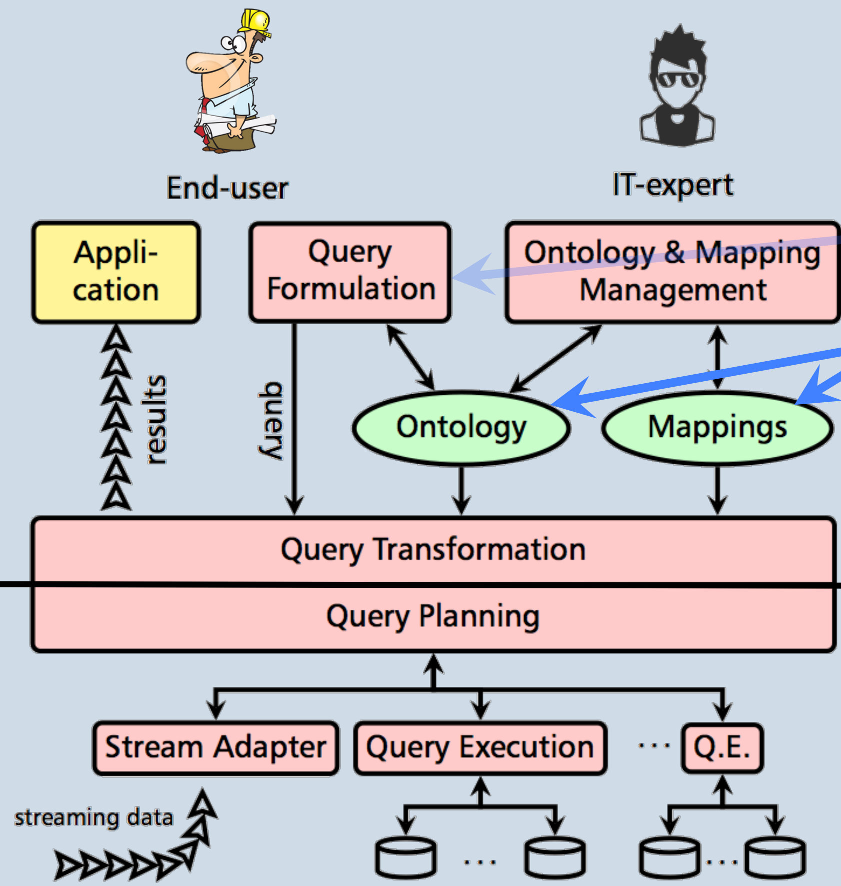
Optique

SEVENTH FRAMEWORK PROGRAMME

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"

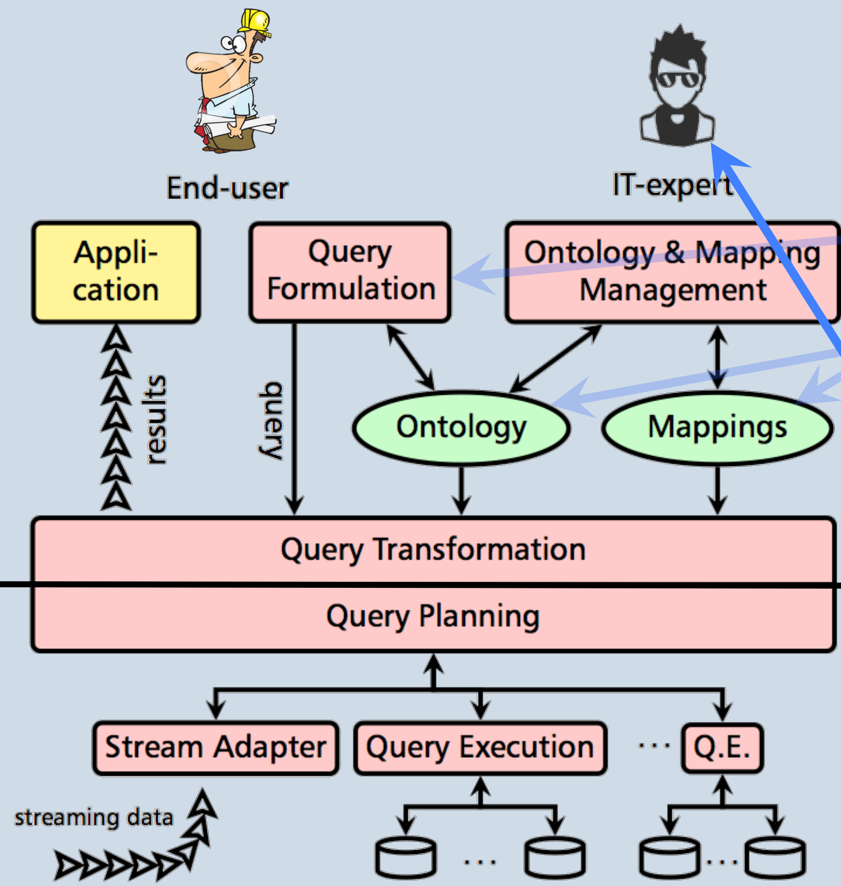
Optique Solution

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

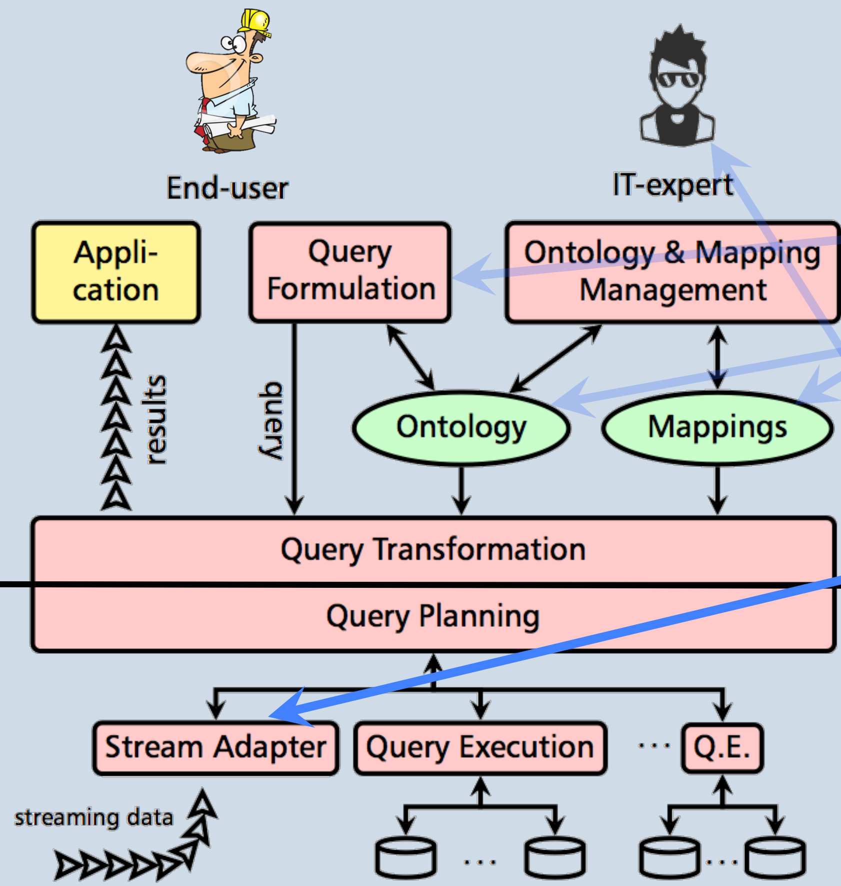
Optique Solution

Query rewriting:

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

Query evaluation:

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



Other features:

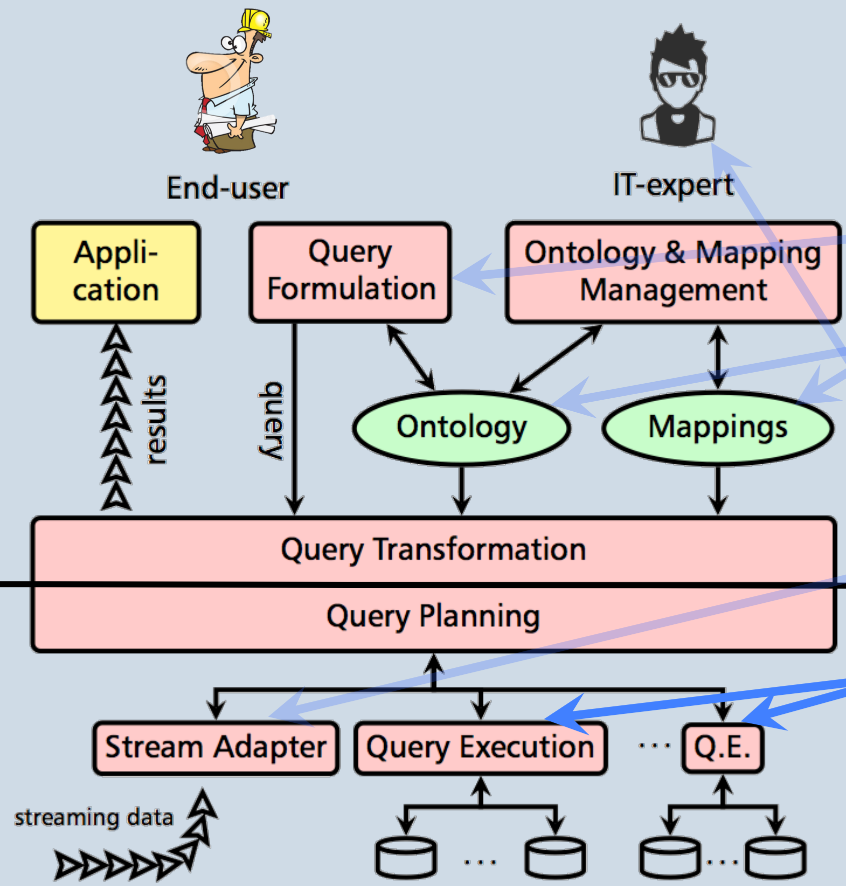
support for query formulation

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

fluid Operations

National and Kapodistrian UNIVERSITY OF ATHENS

SIEMENS

SAPIENZA UNIVERSITÀ DI ROMA

TUHH

Hamburg University of Technology
Department of Telecommunications

DNV

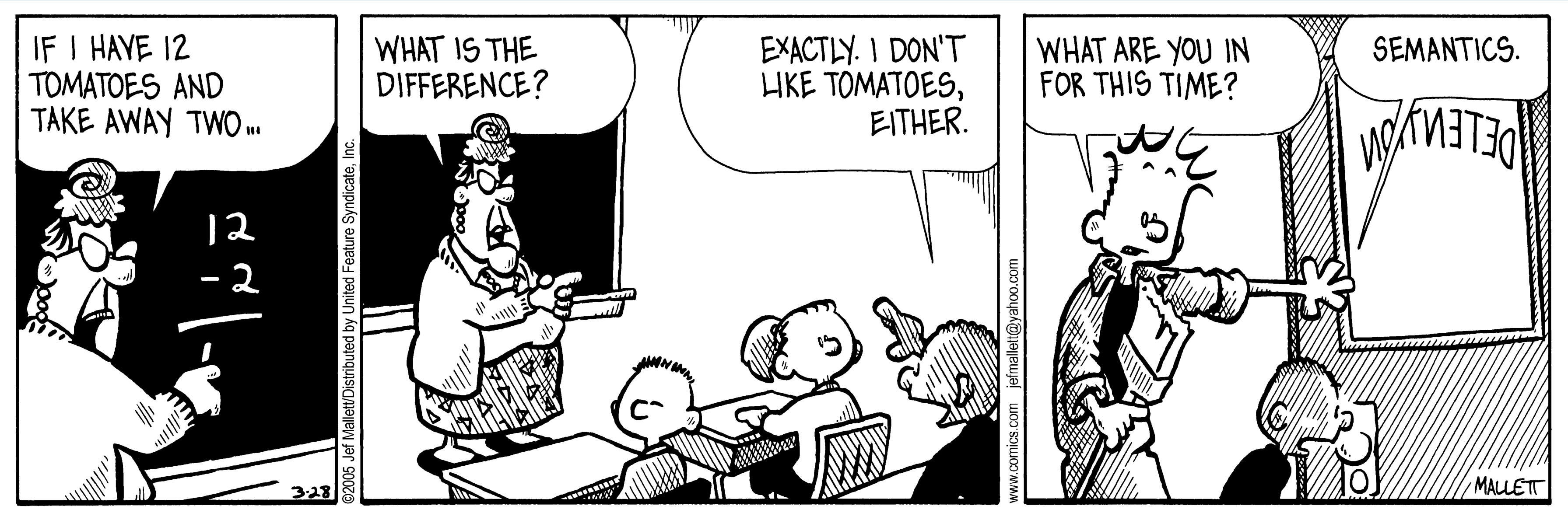
Freie Universität Bozen
Libera Università di Bolzano
Free University of Bozen · Bolzano

DEPARTMENT OF COMPUTER SCIENCE
UNIVERSITY OF OXFORD

Optique

SEVENTH FRAMEWORK PROGRAMME

Thank you for listening



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

Any questions?

DEPARTMENT OF COMPUTER SCIENCE
UNIVERSITY OF OXFORD

Optique

SEVENTH FRAMEWORK PROGRAMME