Optique™

UiO : University of Oslo

TUHH
Hamburg University of Technology
Department of Telecommunications

SAPIENZA
UNIVERSITÀ DI ROMA

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

National and Kapodistrian
UNIVERSITY OF ATHENS

fluid
Operations

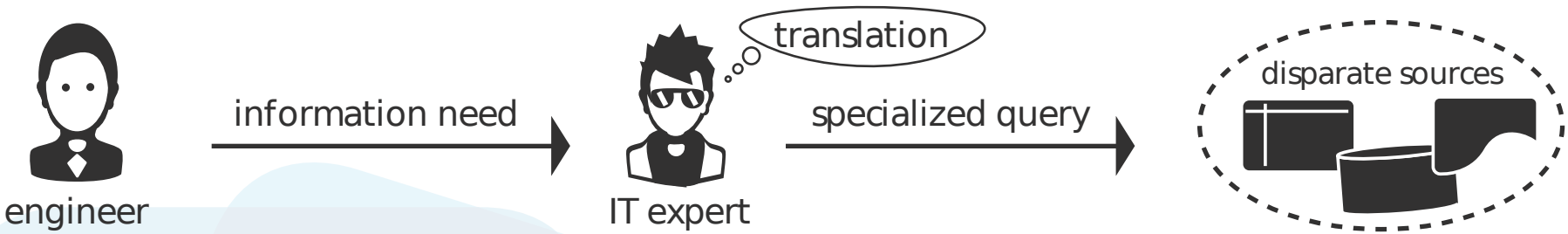
SIEMENS

Statoil

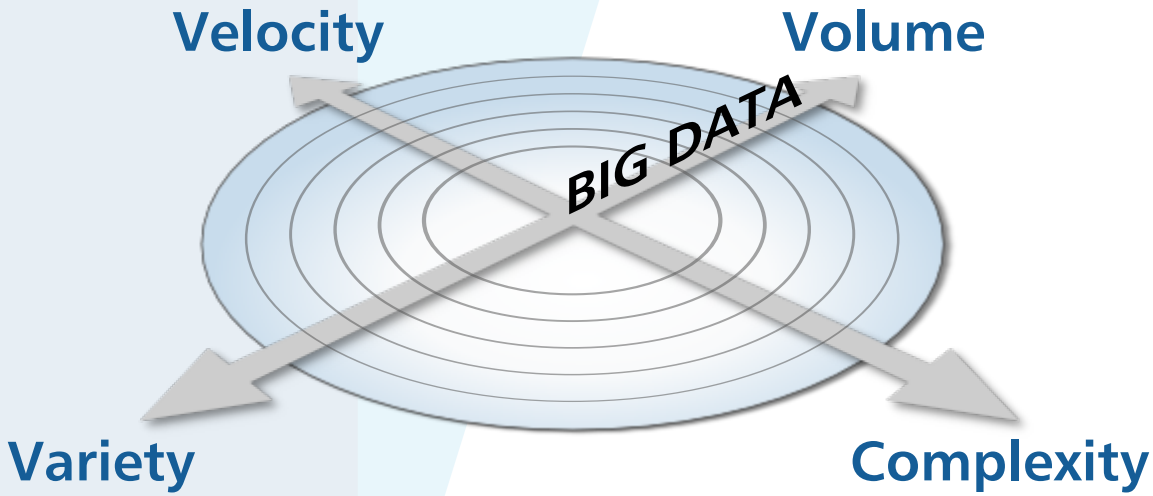
Facts about Optique

- Large-scale integrating project (IP) under FP7-ICT Objective 4.4 Intelligent Information Management
  - 155 proposals in total (IP+STREPS)
  - Optique ranked 2nd, 14 of max 15 points
  - 35 IP proposals, 2 funded
- Total budget: 13 800 000 €
- EU contribution: 9 760 000 €
- Duration: 4 years
- Start: November 1 2012
- Annual reviews in December
- Annual plenary meetings in May/June

Scalable End-user Access to Big Data

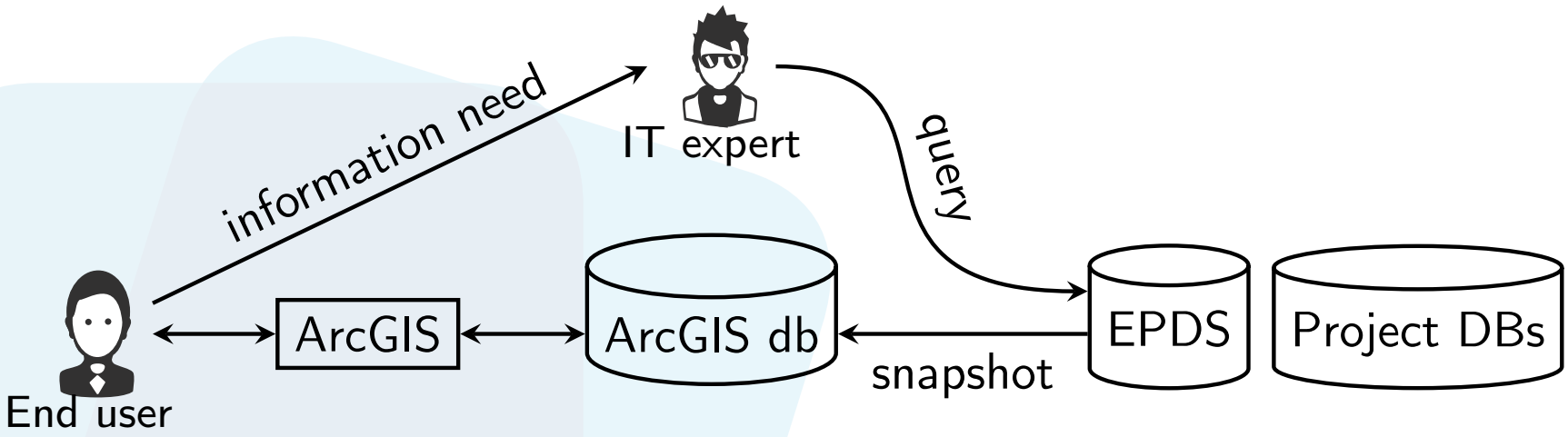


Up to **80%** of experts' time spent accessing Big Data



Optique

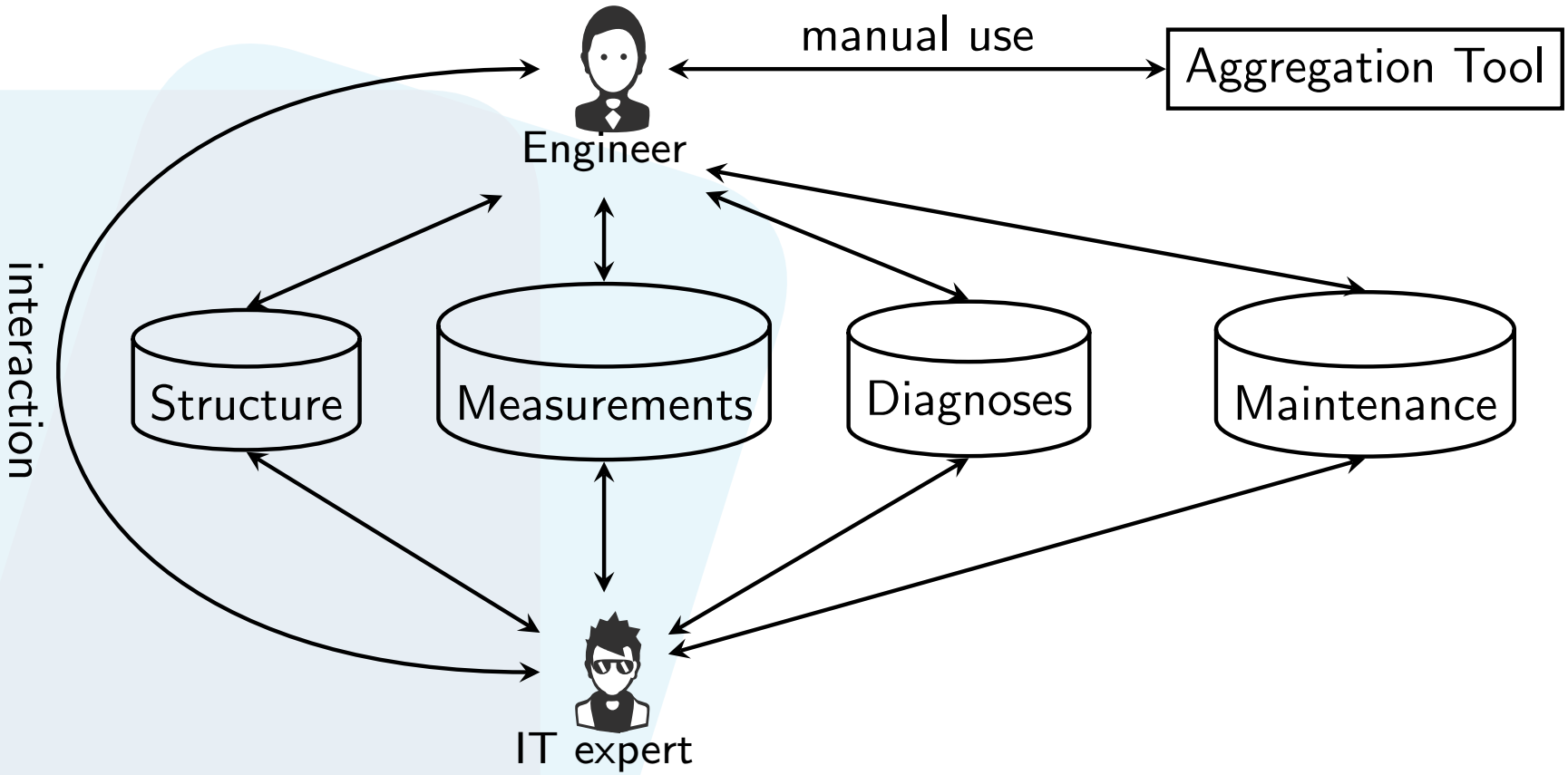
Current situation for Statoil Exploration:



- Query development takes several days
- ProSource turns out to be difficult in use
- Including other sources in queries is difficult/impossible

Optique

Current situation for Siemens Turbines



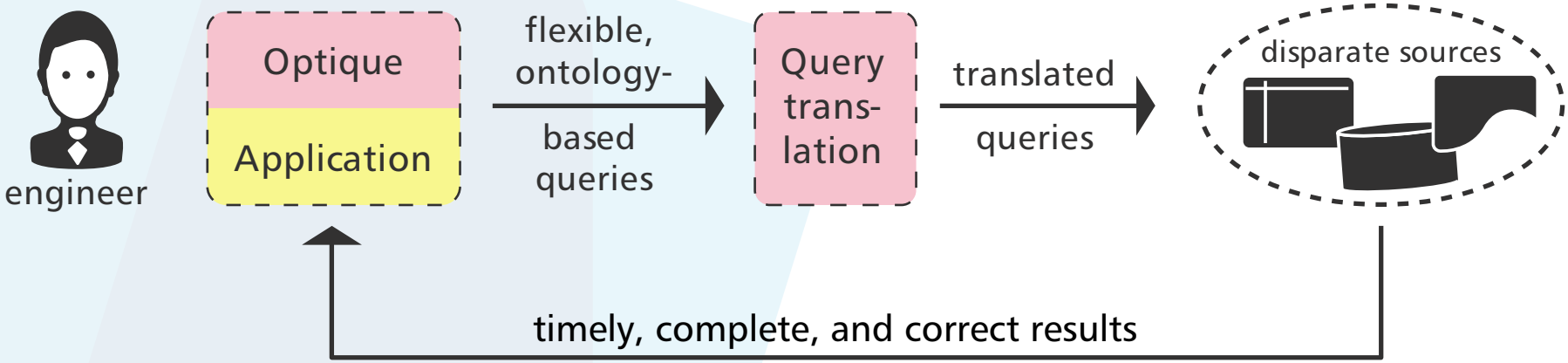
Optique

Challenging Questions for Siemens Turbines

- List all sites where a component of type "C" in version "3.2" is used, which has not been serviced within the last 6 months. *(e.g. for service planning)*
- Determine all turbines showing an error of type "X" defined by at least 5% decrease of quality "Q1" measured at component "C1" within the last month, followed by a statistically significant increase of quality "Q2".*(e.g. for turbine diagnostics)*
- List all turbines for which one of the last five services comprised action "A" followed by action "B", but not action "C".*(e.g. for maintenance quality check)*

Main Innovation: End-to-End Connection

Intelligent, integrated, scalable system:



Optique

Turnaround times: weeks → hours

SIEMENS

Up to 80% time
on data access

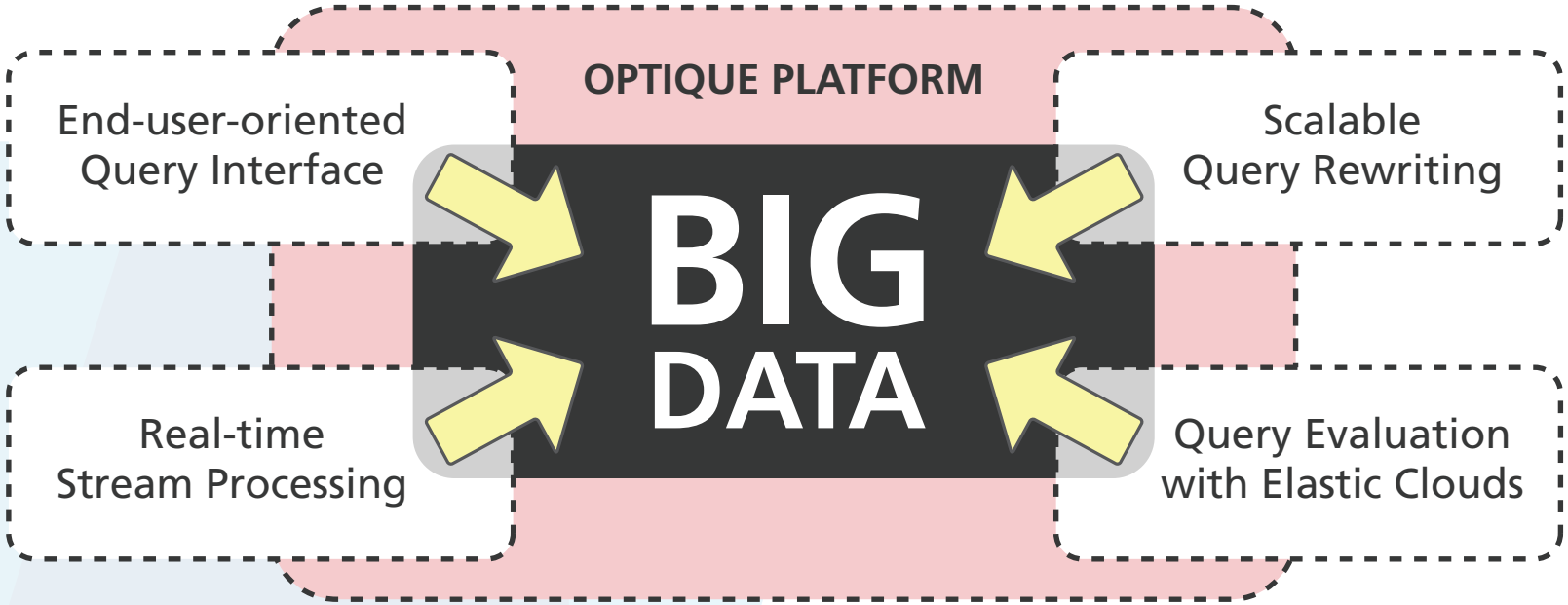
Statoil

Up to 70% time
on data access

Time that should be freed for core tasks!

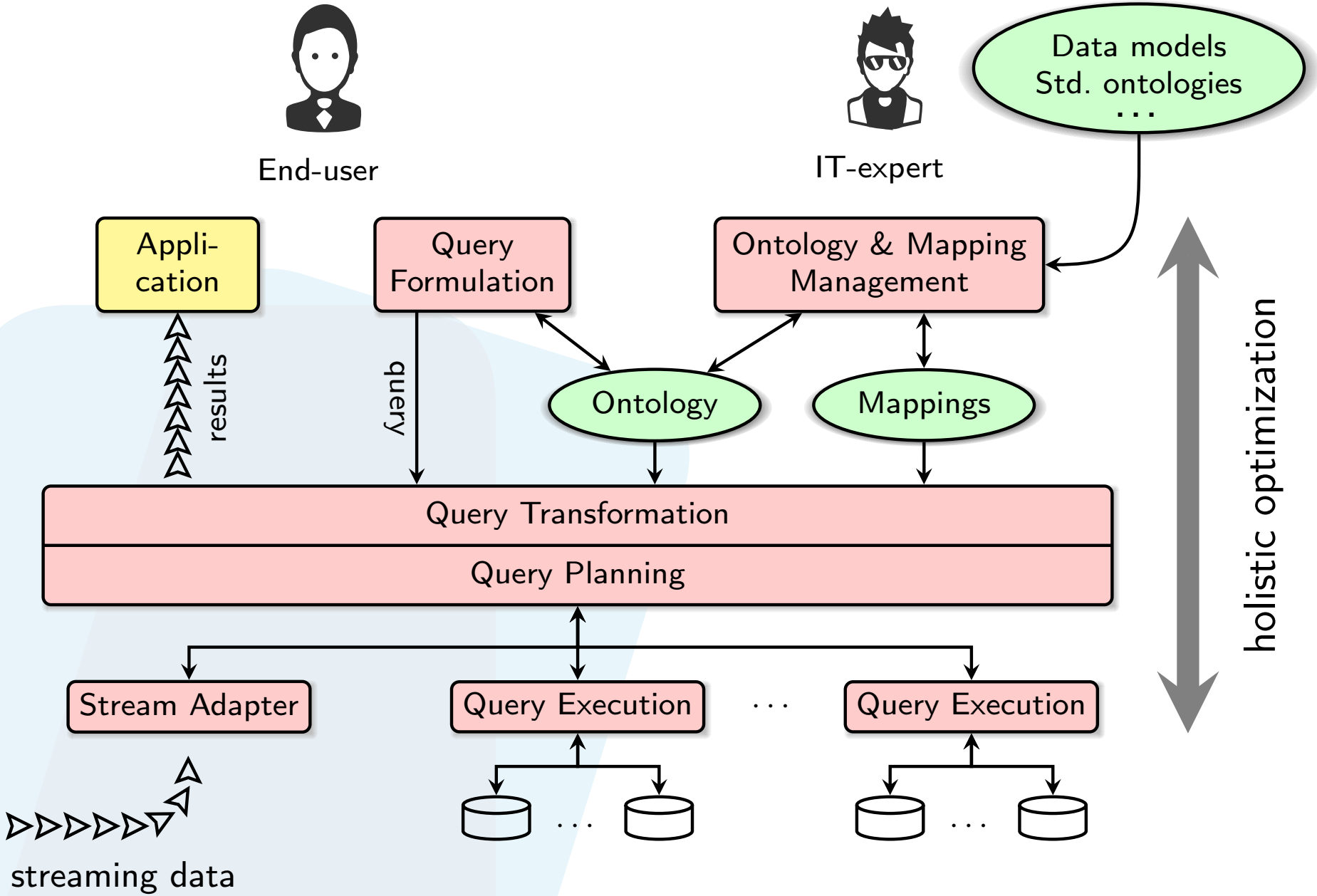
Optique

World leading Consortium

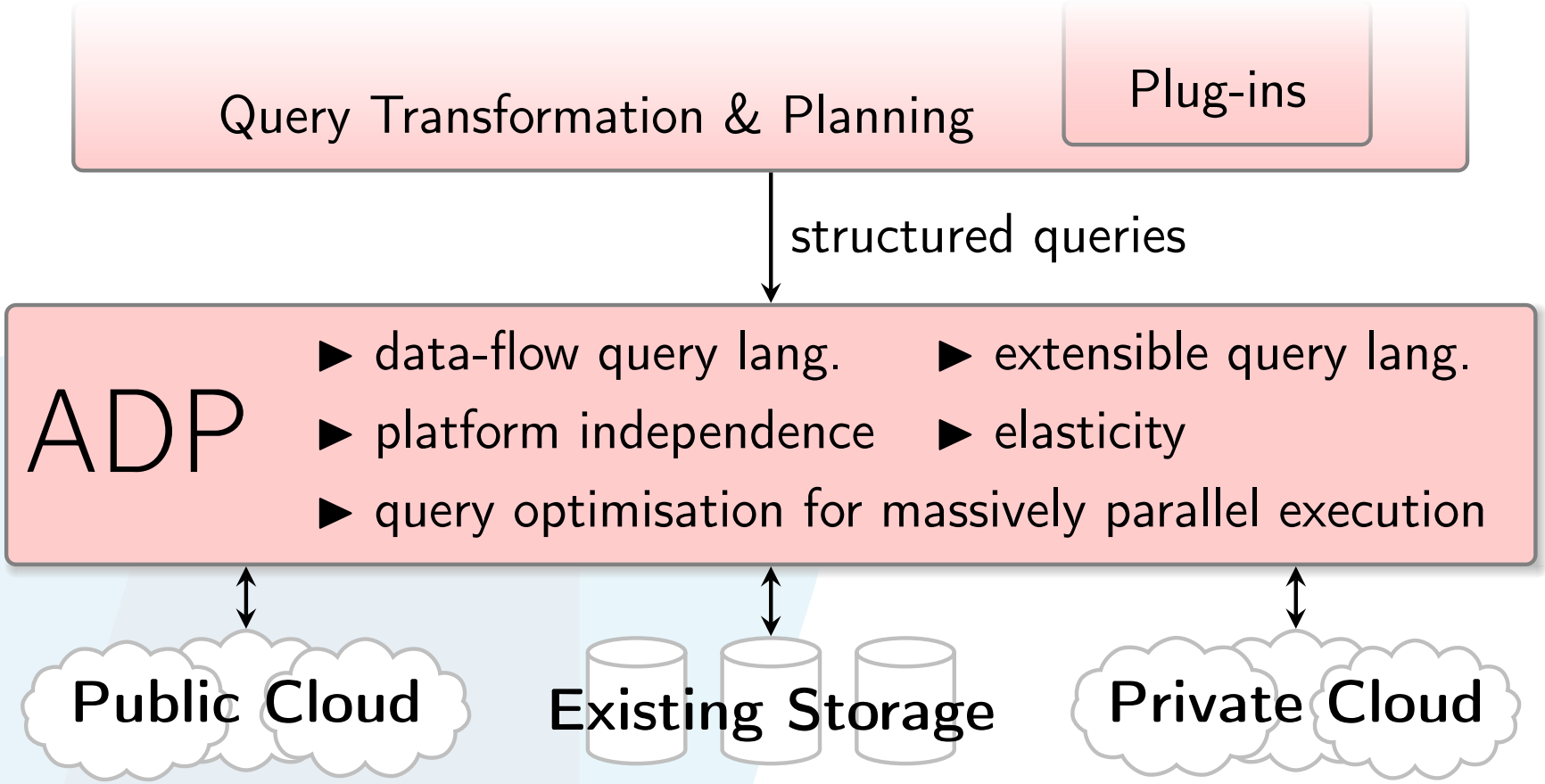


Shaping the future of ICT

Optique


End-user
IT-expert
Data models
Std. ontologies
...
Appli-
cation
Query
Formulation
Ontology & Mapping
Management
results
query
Ontology
Mappings
Query Transformation
Query Planning
holistic optimization
Stream Adapter
Query Execution
...
Query Execution
streaming data

Optique



Optique extensions of ADP:

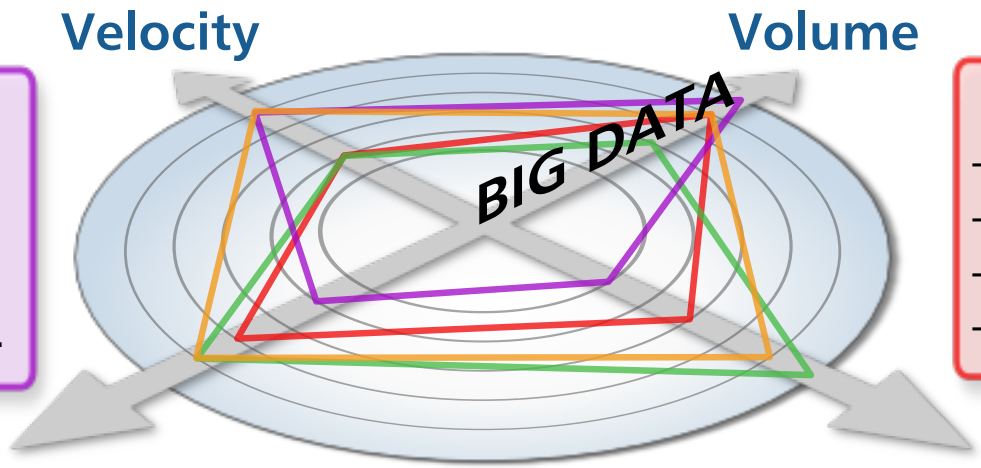
- ▶ federation support
- ▶ temporal queries
- ▶ continuous queries

Optique

Optique

Real-time Stream Processing

Query Evaluation with Elastic Clouds



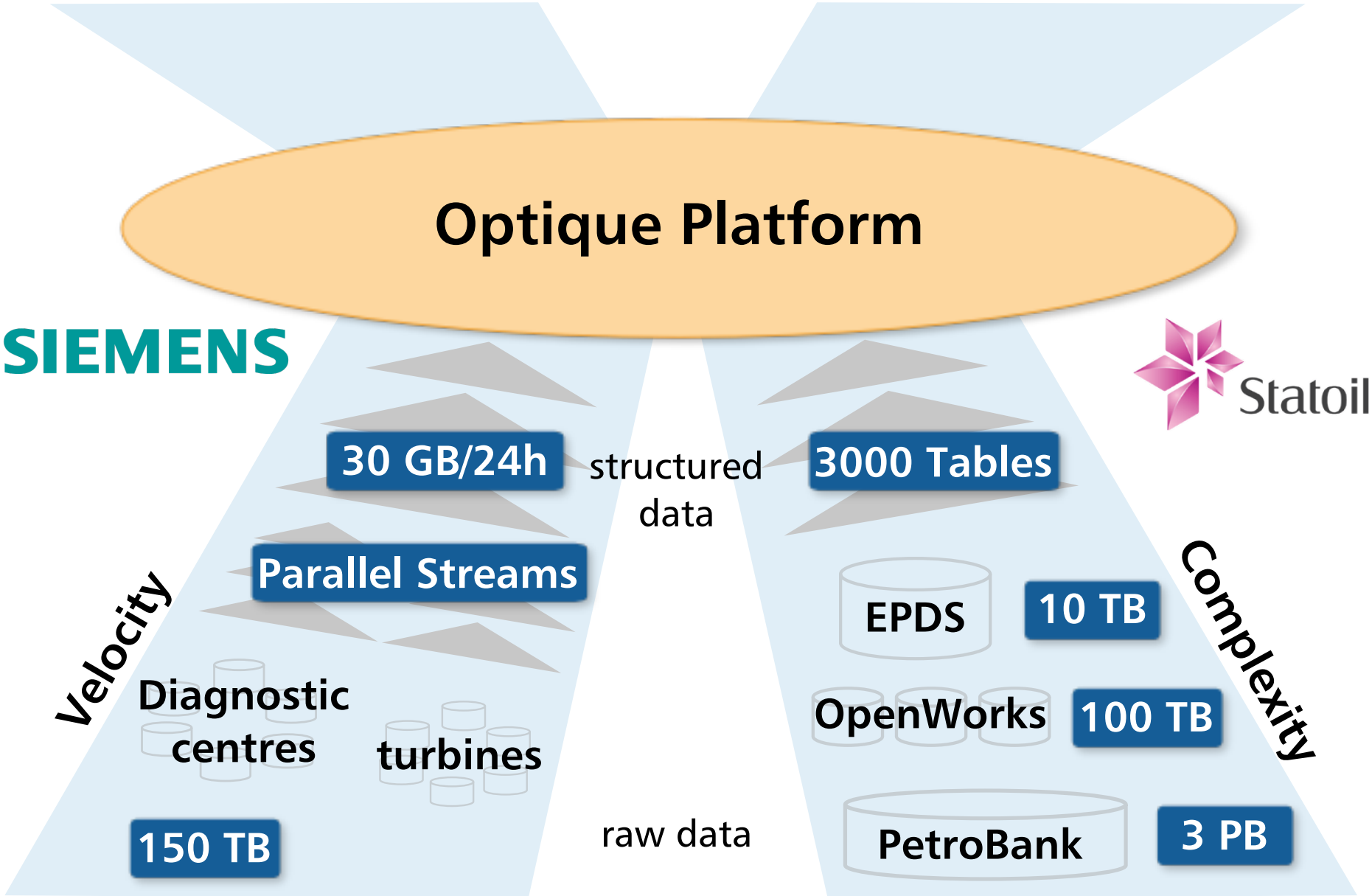
NoSQL:
+scalability
+elasticity
−weak structure
−weak query lang.

Parallel SQL:
+parallel analytics
+in-DB MapReduce
−integration difficult
−no streaming

Scalable Query Rewriting

Semantic Web:
+strong reasoning
+ont.-based integration
−no streaming
−poor scalability

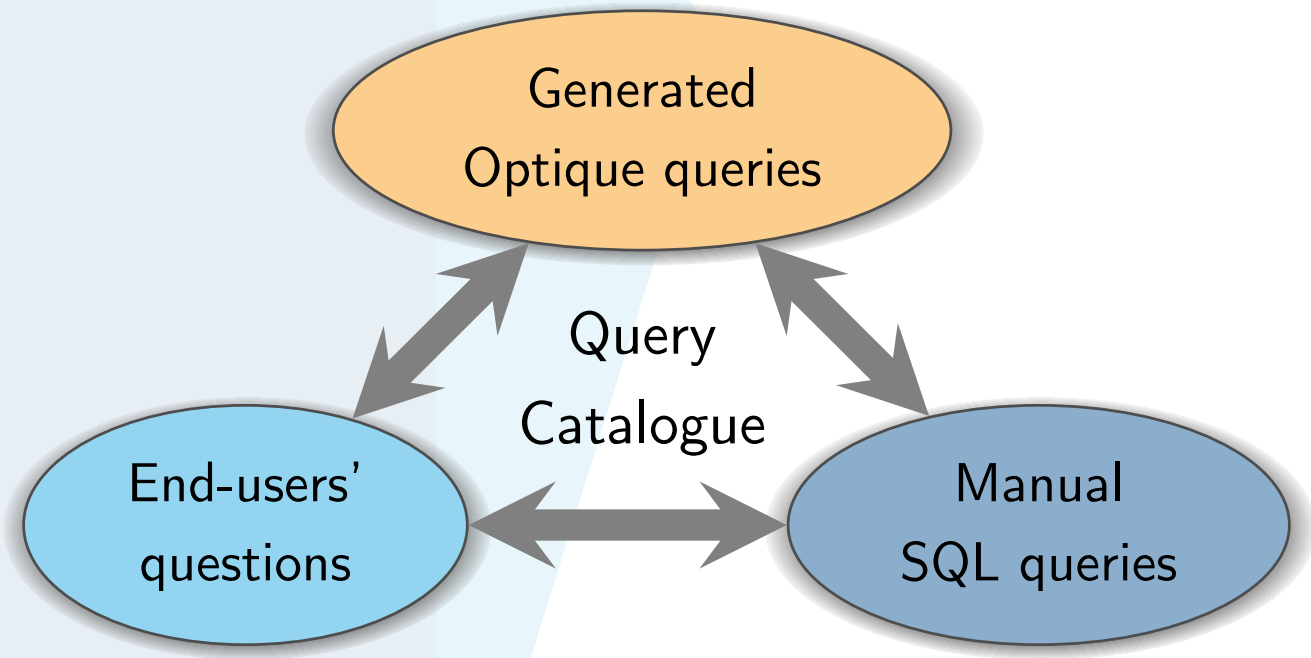
End-user-oriented Query Interface


Optique Platform
SIEMENS
Statoil
30 GB/24h
structured data
3000 Tables
Parallel Streams
Velocity
Complexity
EPDS
10 TB
Diagnostic centres
turbines
OpenWorks
100 TB
150 TB
raw data
PetroBank
3 PB

Optique

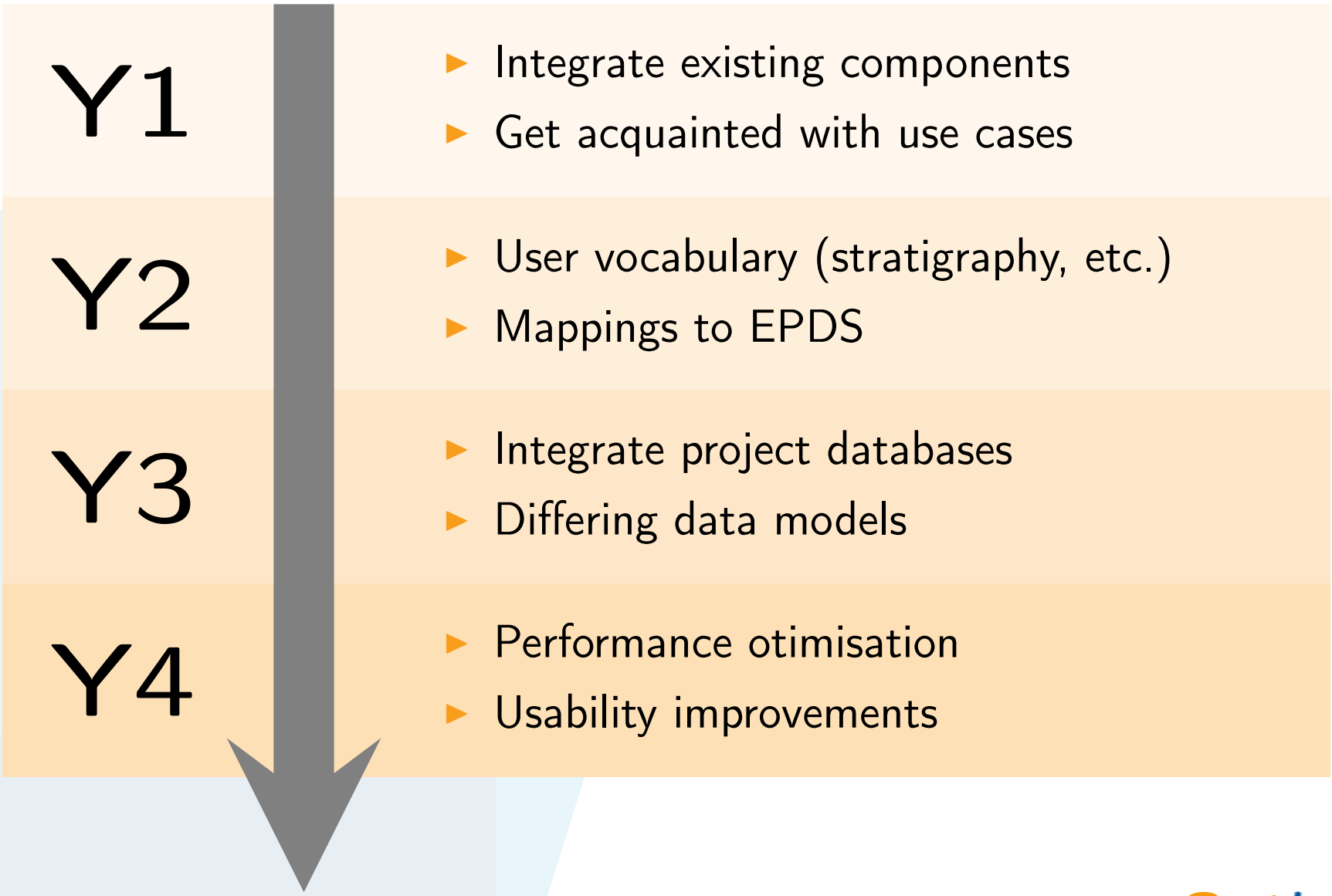
Query Catalogue for Requirements and Evaluation

Query formulation	▶	90% success
Turnaround time	▶	from weeks to hours
Data volume	▶	from TB to PB



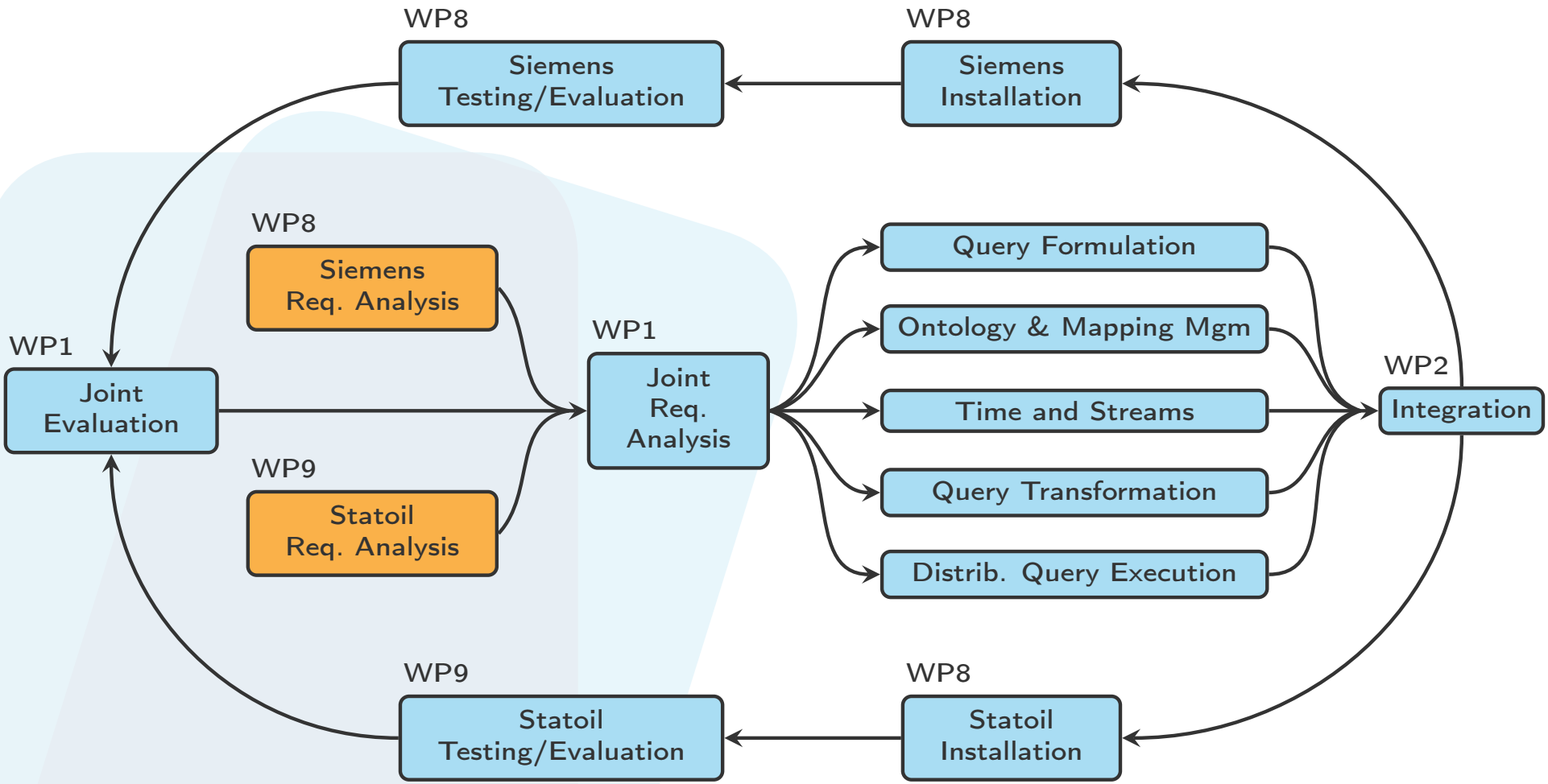
Optique

4 Project phases



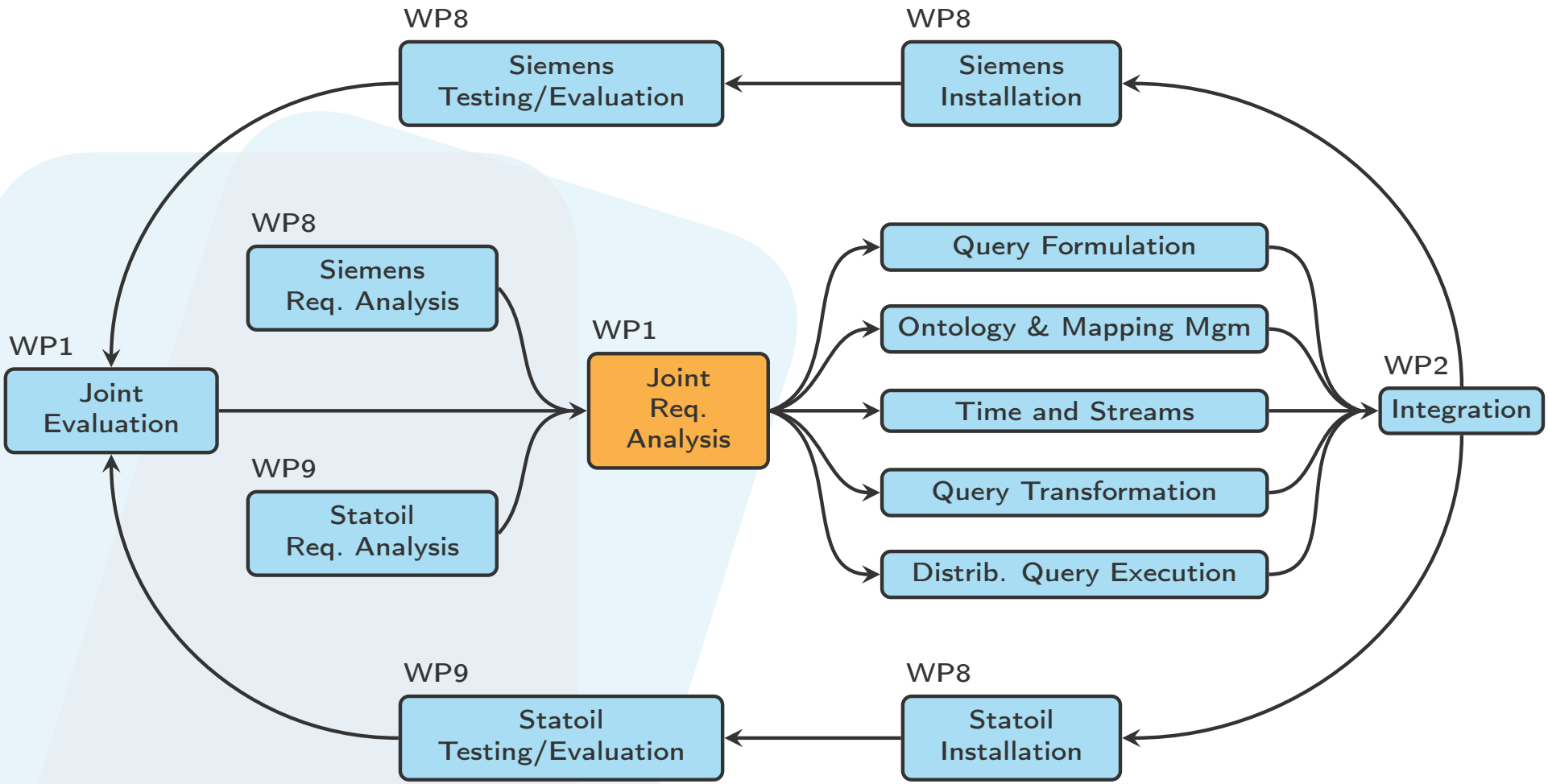
Optique

Yearly development cycle



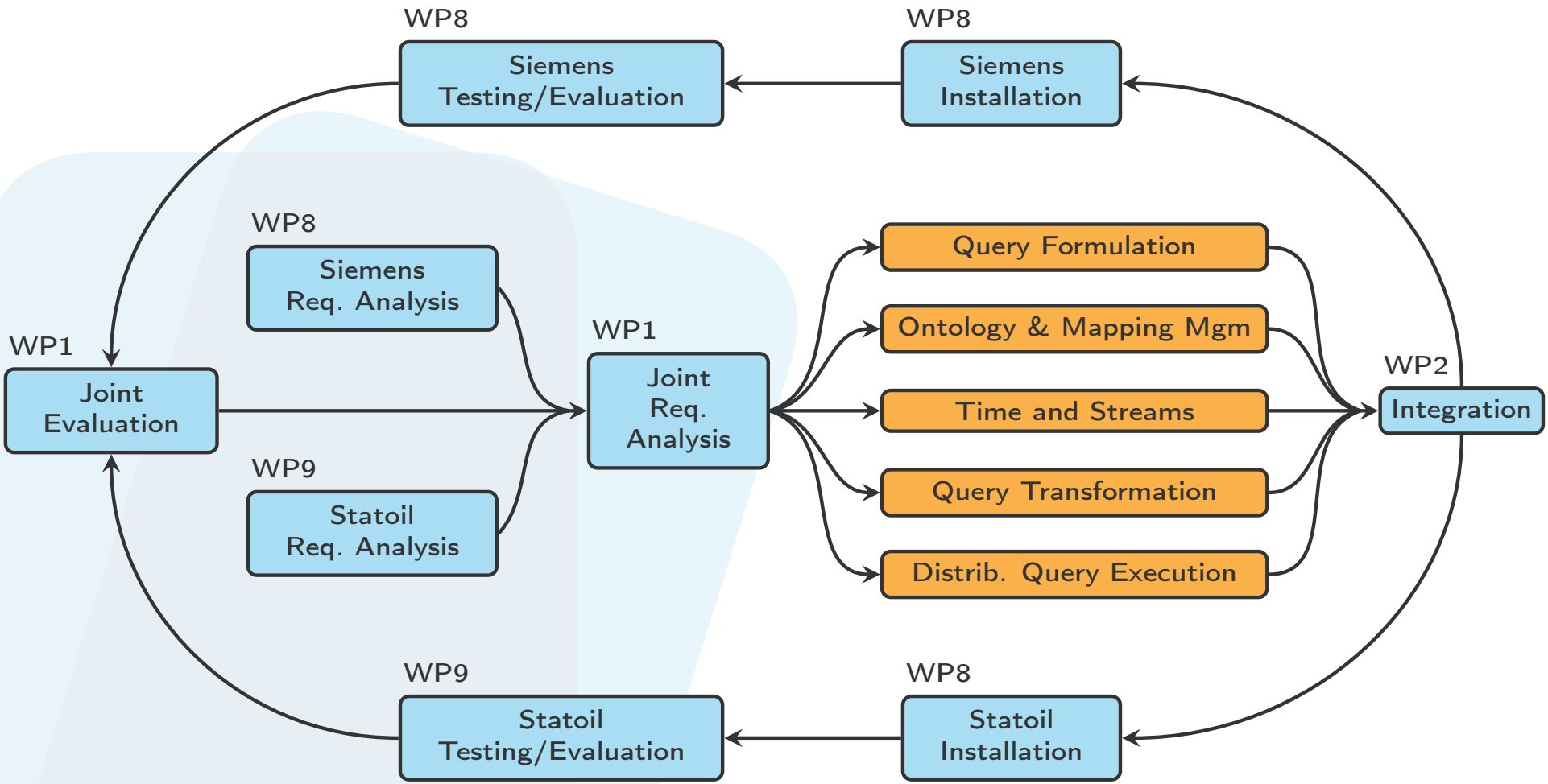
Optique

Yearly development cycle



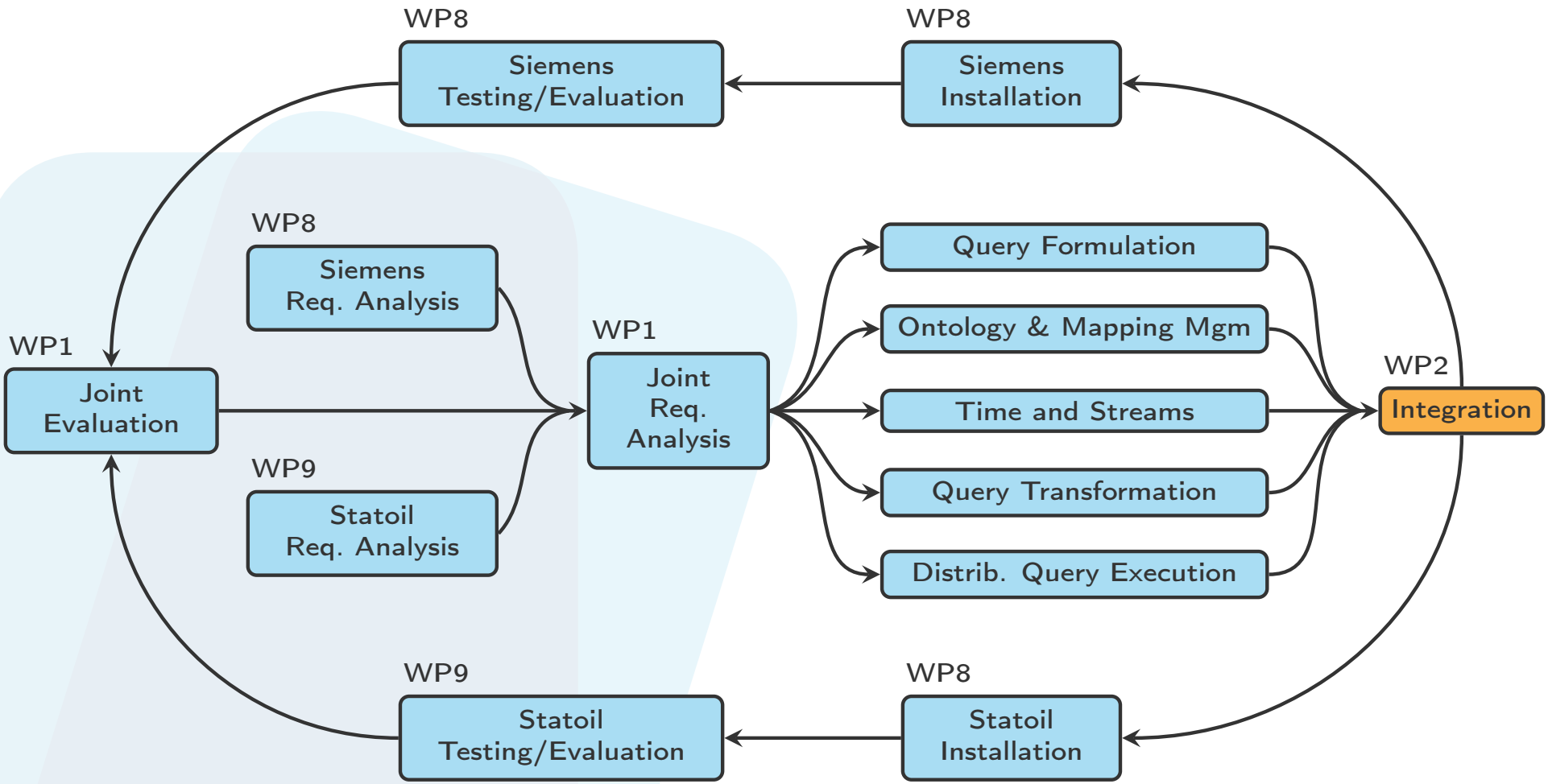
Optique

Yearly development cycle



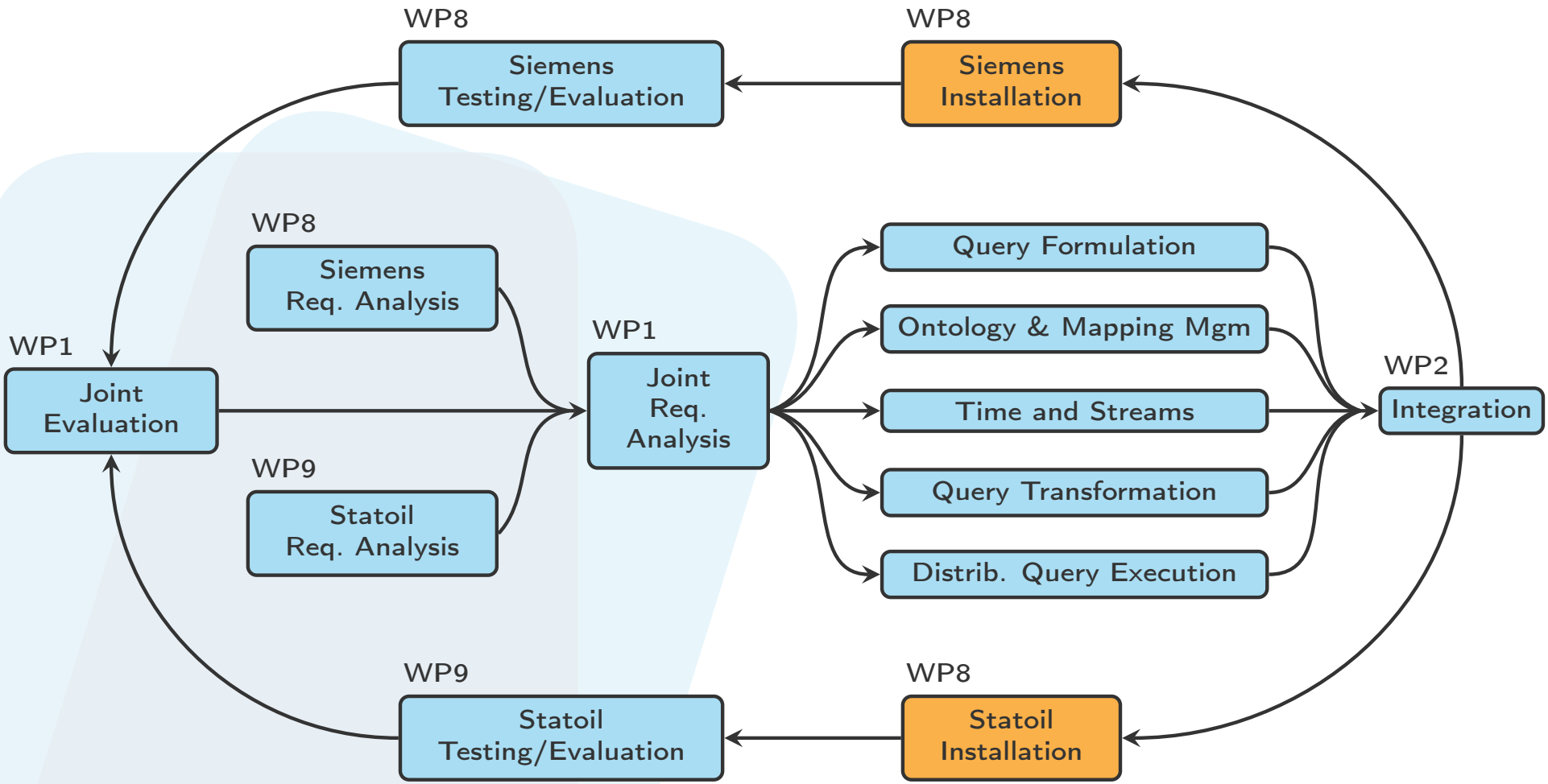
Optique

Yearly development cycle



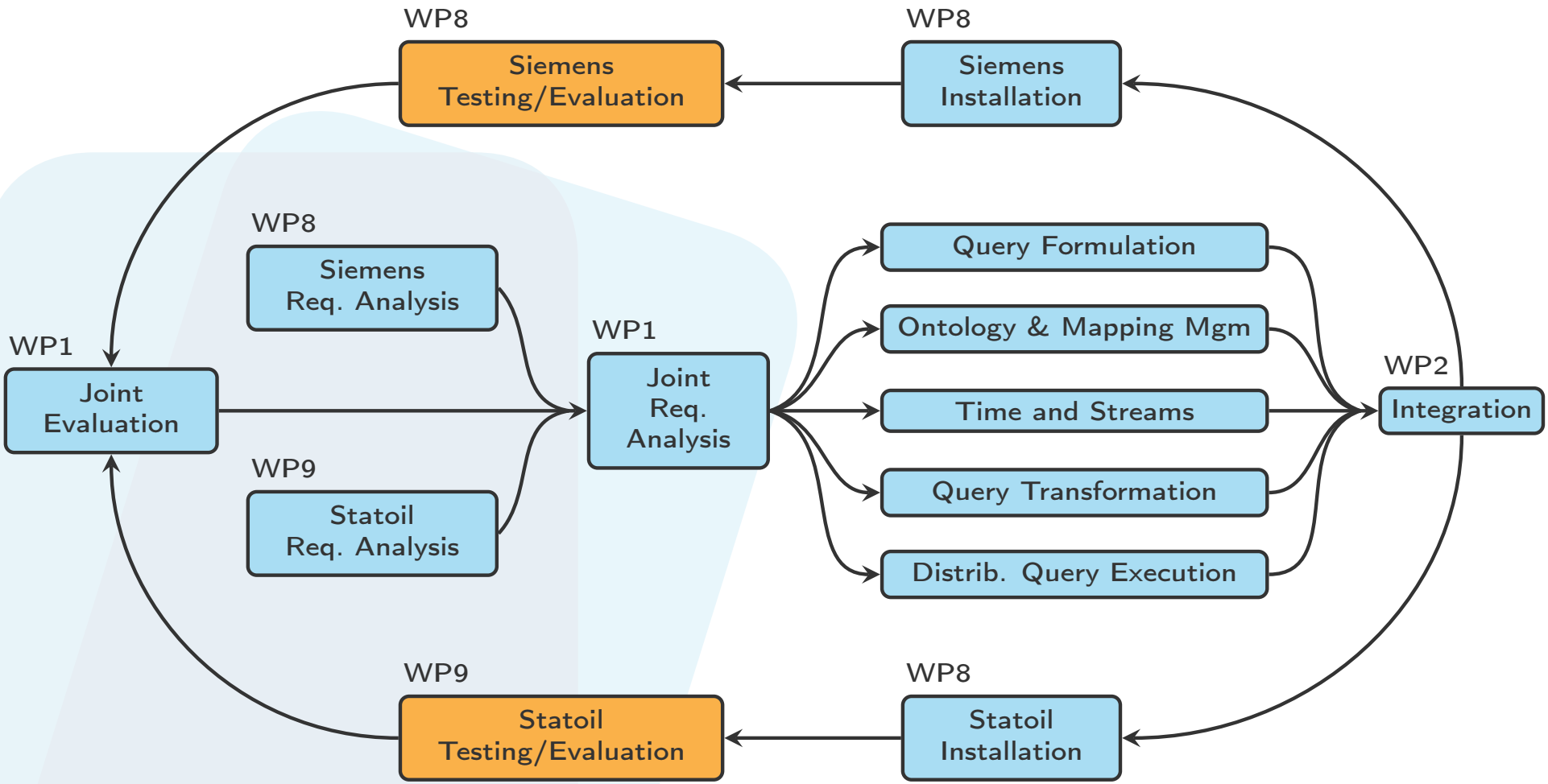
Optique

Yearly development cycle



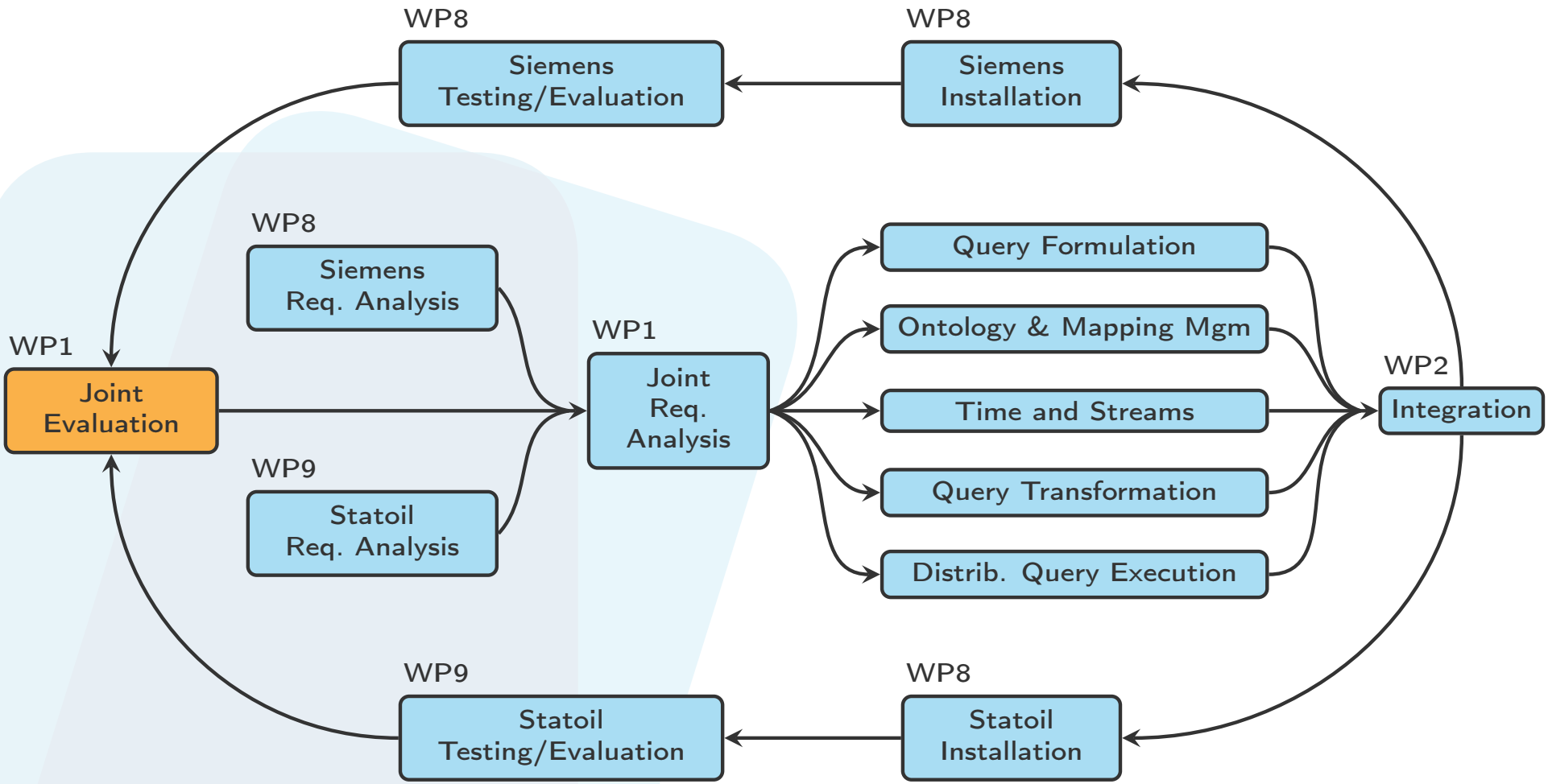
Optique

Yearly development cycle



Optique

Yearly development cycle



Optique

WP3: Query Formulation

What: Front-end component to help non-technical users formalize their information needs

How:

- Query by Navigation
  - Query by Example?
  - Faceted Querying?
- Direct Editing for experts
- Query-driven ontology construction

Who: Oxford University, University of Oslo

WP4: Ontology and Mapping Management

What: Creation, evolution, transformation & management of

- ontologies
- mappings

How:

- "Bootstrapping" of ontologies and mappings from existing models using ontology alignment and approximation techniques
- Bottom-up refinements on demand
- New framework/metamodel for analysing and manipulating sets of mappings

Who: Sapienza Univ. of Rome, most of the partners

WP5: Time and Streams

What: Semantics, Queries, and Query Rewriting for

- time-stamped (historical) data
- streaming (real-time) data

How:

- Extended query language (C-SPARQL?)
- Optimization of query rewriting for temporal data
- Data mining and query log analysis (caching)

Who: TU Hamburg-Harburg, Siemens, Univ. of Athens

WP6: Query Transformation

What: Rewriting user queries to data source queries

- efficient rewriting
- efficient queries output

How:

- Established OBDA methodology
- + utilize existing structure in sources
- + storing auxiliary information in data sources

Who: Free Univ. Bolzaono, Sapienza Univ. of Rome

WP7: Distributed Query Execution

What: Query planning and execution

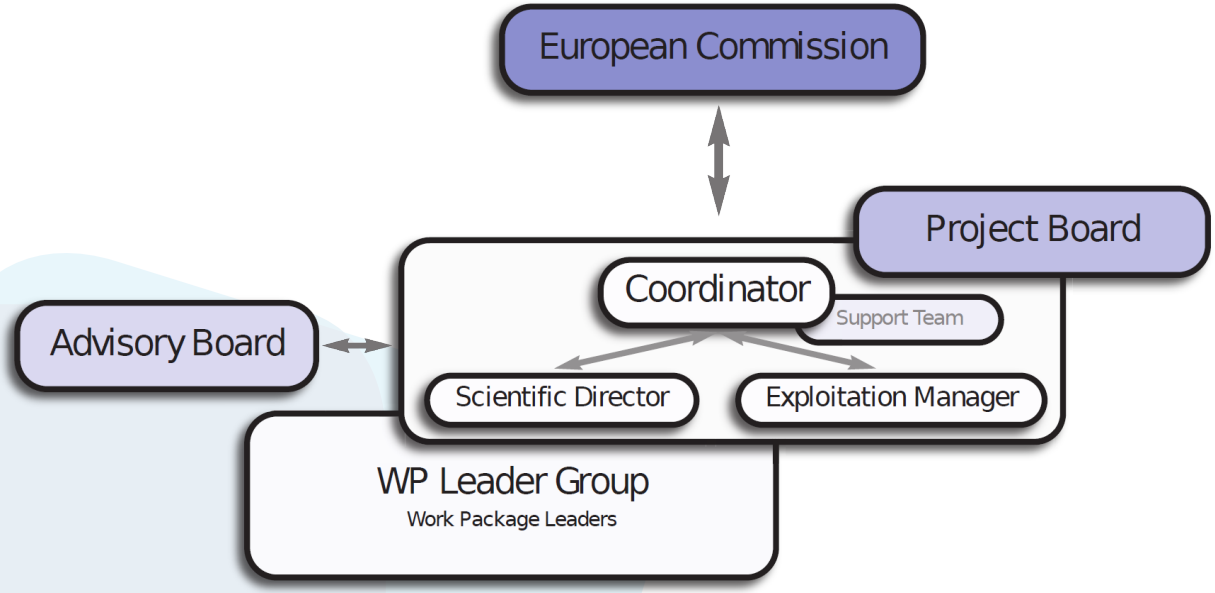
- "standard"/temporal/continuous
- distribution to disparate sources (federation)
- highly scalable

How:

- existing ADP framework for cloud storage/querying
- + federation support
- + temporal queries
- + continuous queries

Who: Univ. of Athens, Sapienza Univ. of Rome

Management structure



Management:

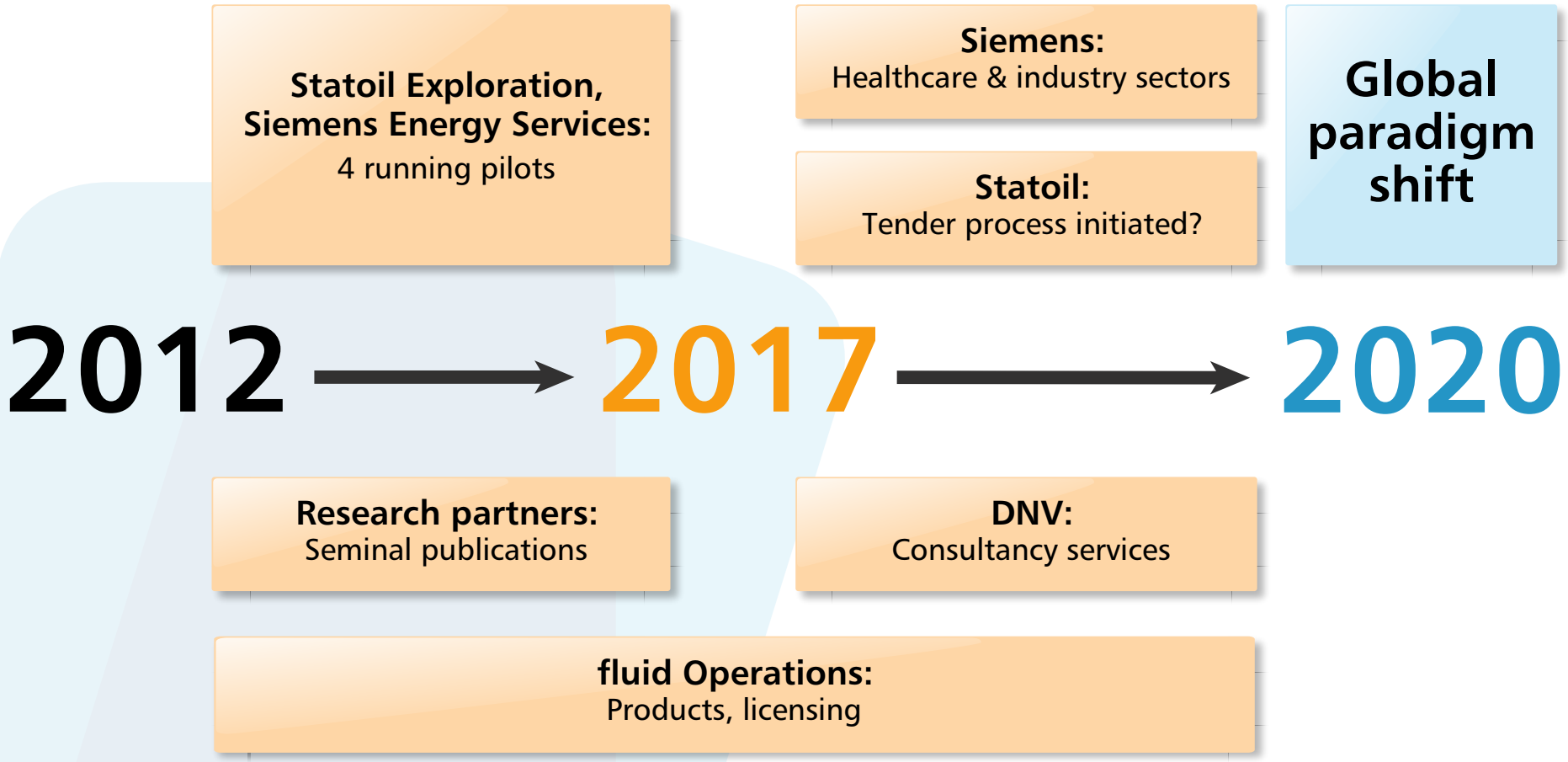
- One from each partner in WP leader group and in Board
- Board meetings twice a year at plenary meetings
- WP leader meetings at least every 3rd week, virtual
- Advisory Board meets once a year

Optique

Advisory Board

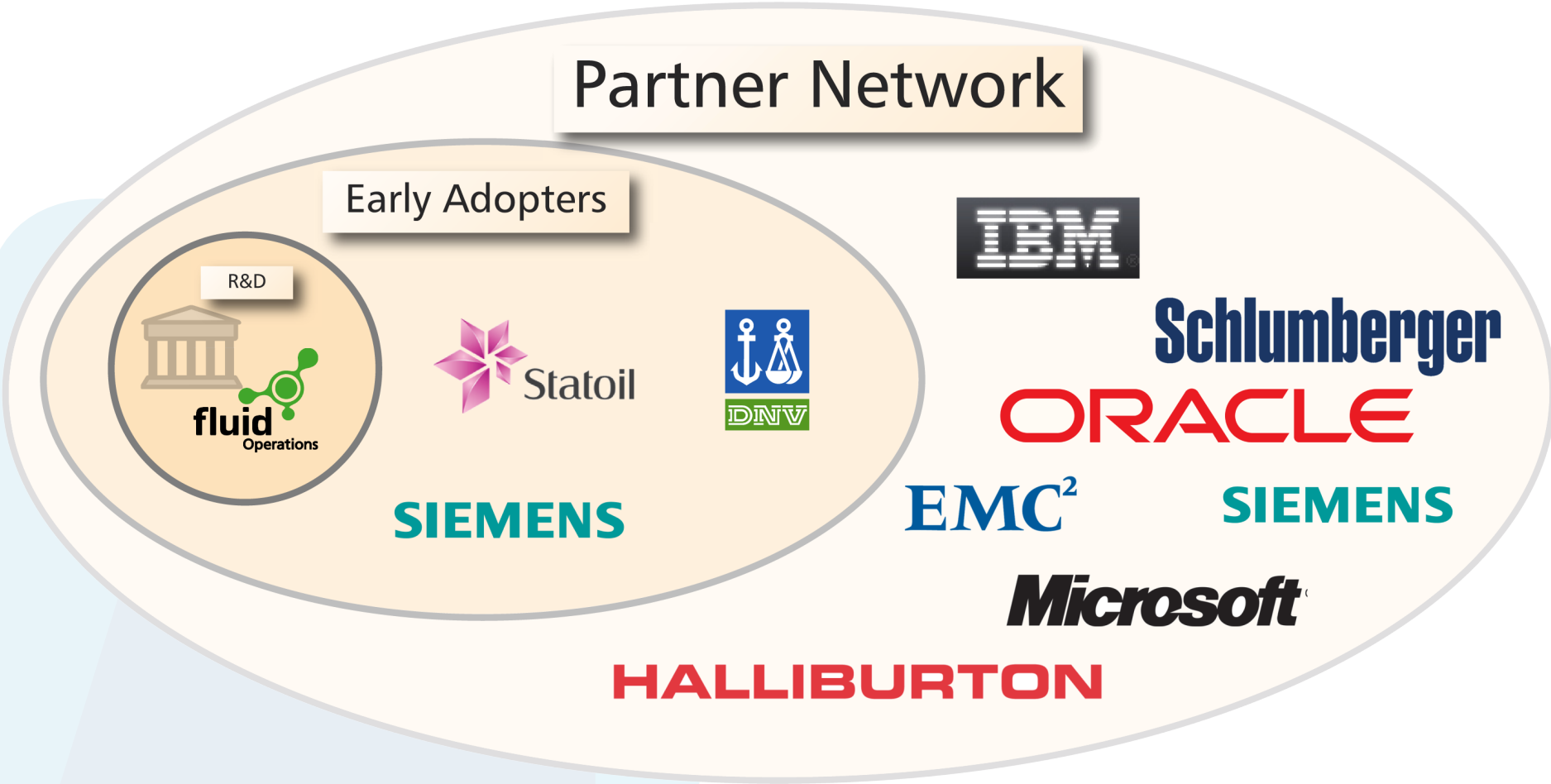
- Morten Thorkildsen, Country general Manager, IBM Norway
- Joe King, Director of Strategy, Landmark R&D
- Bjørn Olstad, General Manager for Search and Enterprise Content Management, Microsoft Development Center Norway
- Rune Syverinsen, Technical Director, Oracle Nordics
- Vasu Guruswamy, Vice President Schlumberger Information Solutions

Towards a Global Paradigm Shift



Optique

Exploitation through the Optique Ecosystem



Optique

Optique™