Enabling Knowledge-Based Complex Event Processing

Kia Teymourian

Research Assistant at Free University Berlin

http://www.teymourian.de

VLDB 2011, Ph.D Workshop

29 Aug - Sep 3, 2011, Seattle

AG Corporate Semantic Web

Freie Universität Berlin

http://www.inf.fu-berlin.de/groups/ag-csw/

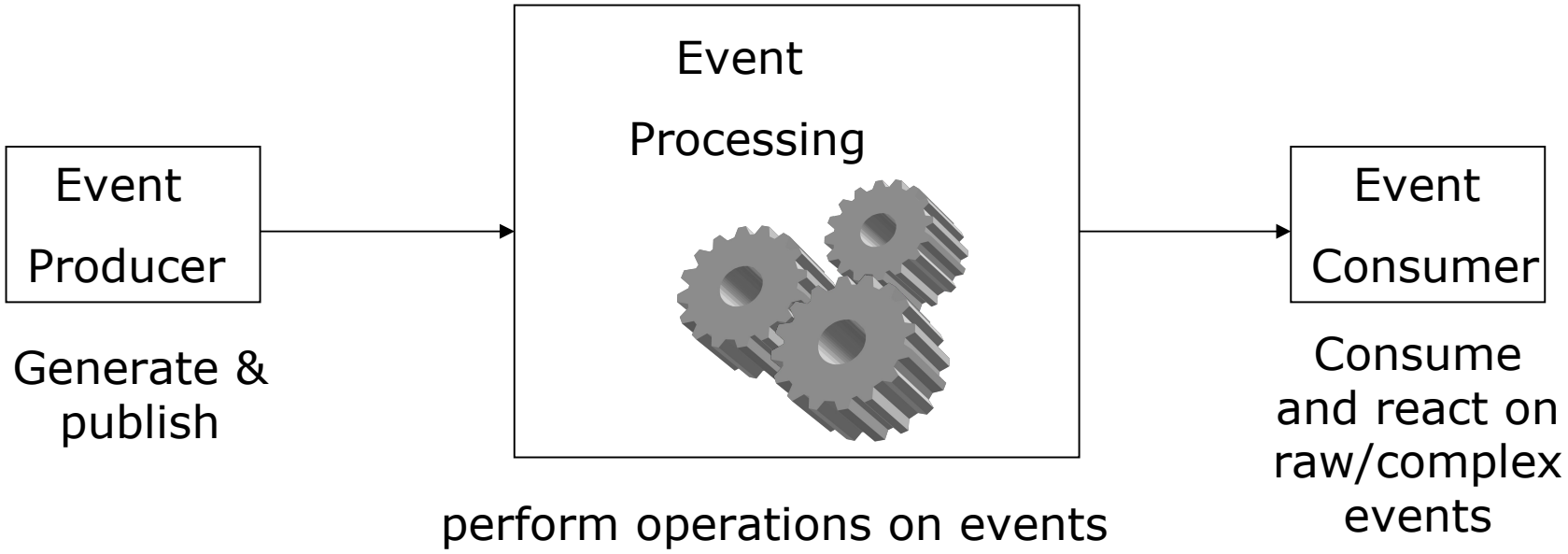
Outline

- Complex Event Processing (CEP)
- Why Semantics + CEP?
- Semantic CEP (SCEP)
  - Knowledge Representation for Events & Event Patterns
  - Real-Time Semantic Event Processing

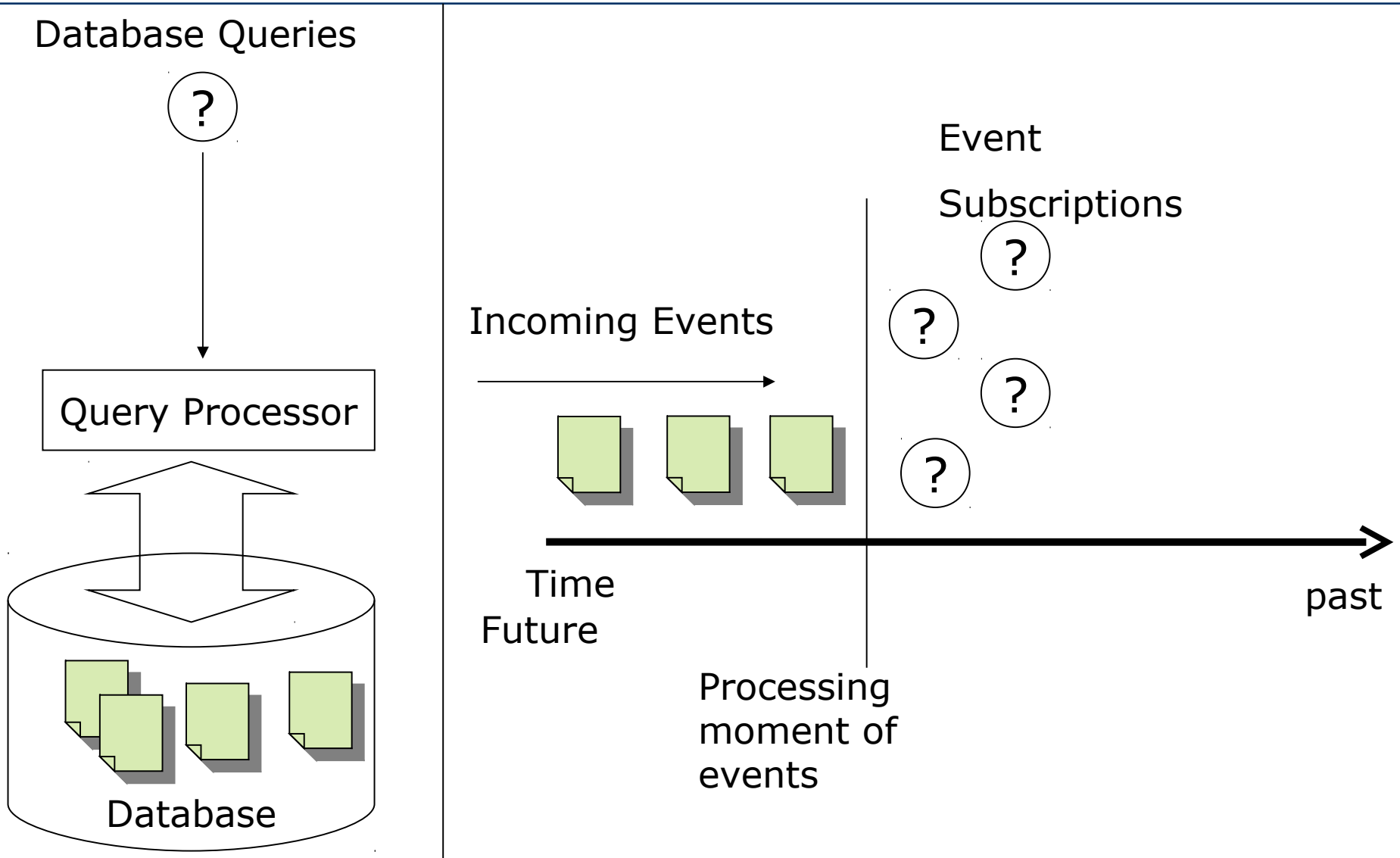
Events

- Anything that happens, or is contemplated as happening
- A Notification is a message that contains information about an event that has occurred.
- Content-based data and filter model:
  - Tuples
  - Structured records
  - *Name/value pair* (n, v) with name n and value v
- Samples:
  - *{(type, StockQuote), (name, "Siemens"), (price,45)}*

Big Picture of Complex Event processing

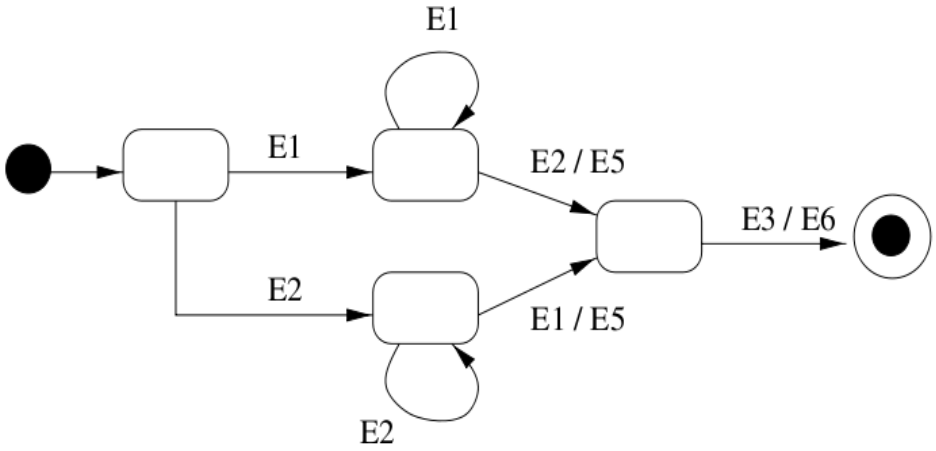


CEP vs. Databases

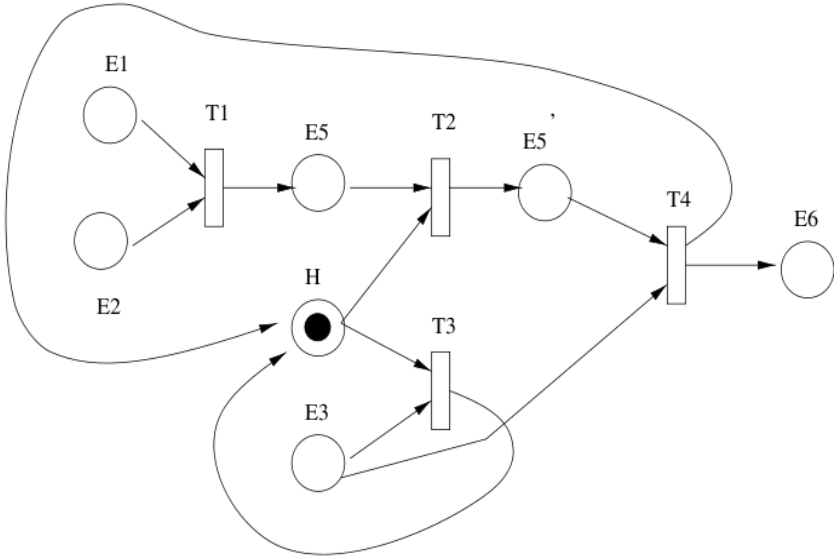


Event Processing Methods

- Syntactic processing of low-level events
- Real-time processing

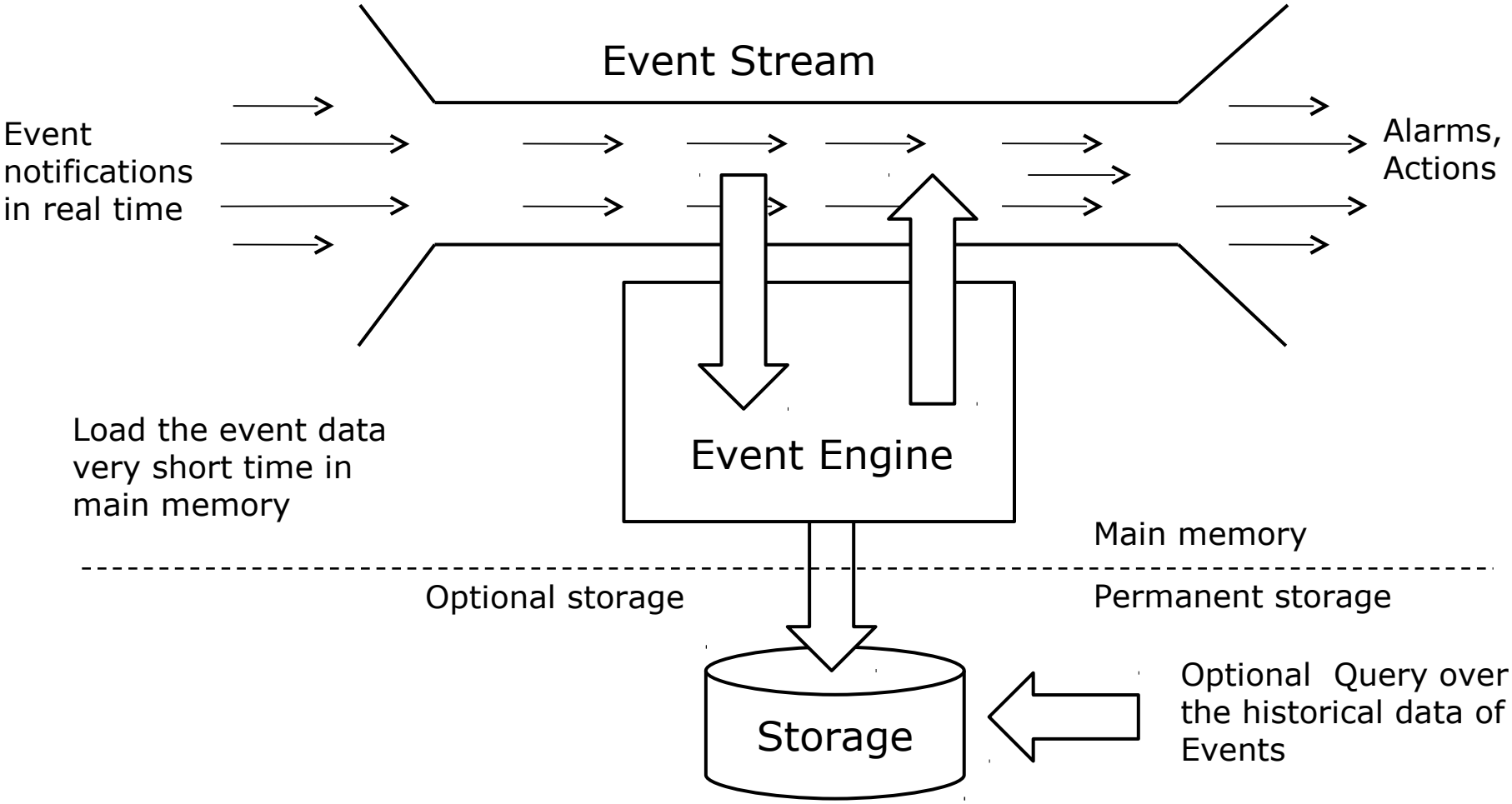


State Machines



Petri Nets

Keep the event data moving!



My Research Challenge

Semantic Complex Event Processing (SCEP)

Research Challenges

Three Main research questions:

1) Can we represent **events and event patterns** based on ontological background knowledge and use it for CEP?
   - Relation to other non-event concepts, e.g. **Situations, Actions, Actors, Processes**, …

2) Is it enough to use **Datalog** as **processing semantic**?

3) Is it possible to **process the events** in timely manner and do **inferencing on a background knowledge**?

Example – Semantic Event Processing

Query:
Buy stocks of companies, who have in **Europe production facilities** *and*
produce products **from iron** *and*
more than **10,000 employees** *and*
are at the moment in **reconstruction phase** *and*
their price/volume **increased stable** in the past 5 minutes.

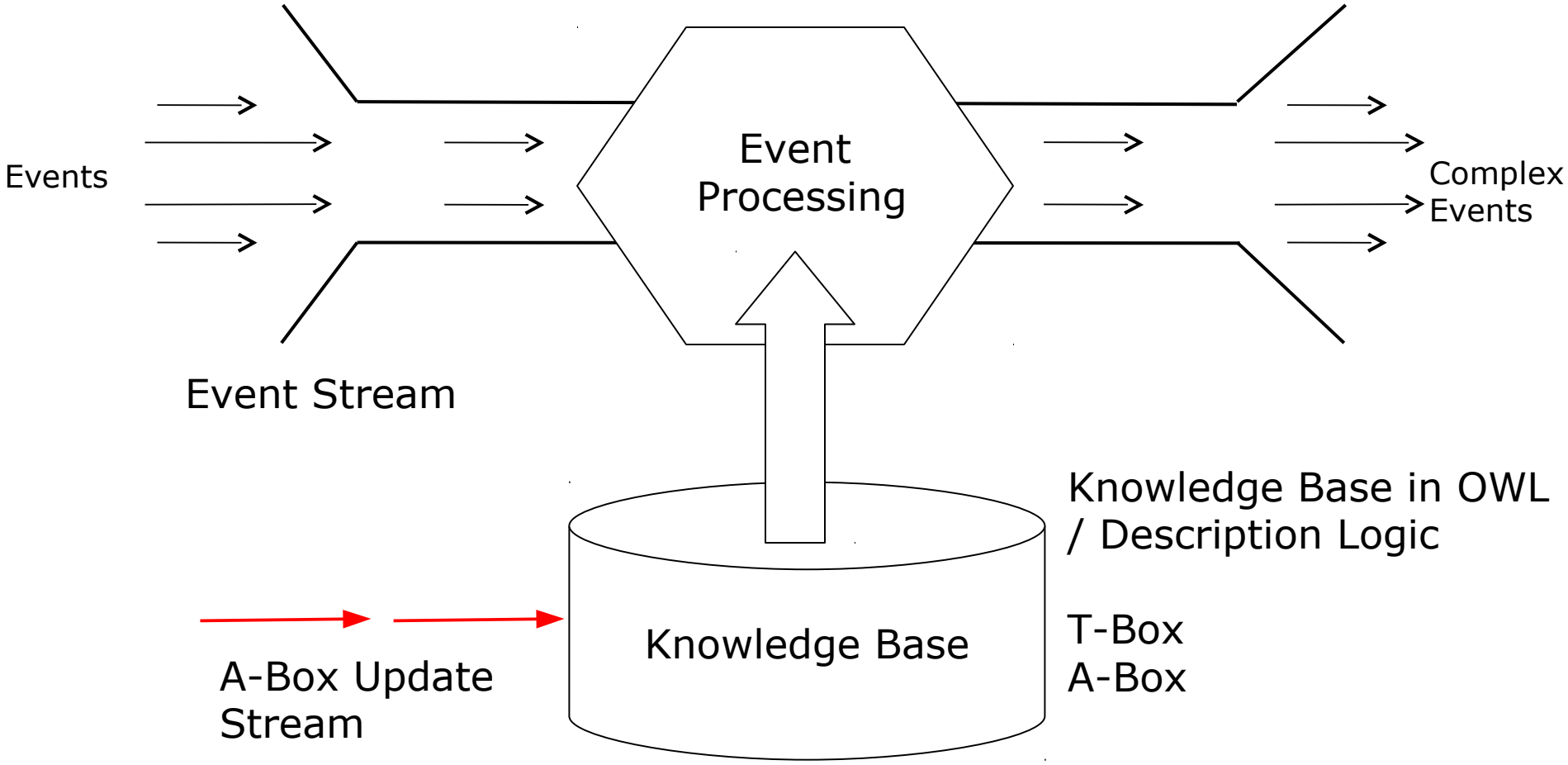
Event Stream

{(Name, "*OPEL*")(Preis, 45)(Volumen, 2000)}
{(Name, "SAP")(Preis, 65)(Volumen, 1000)}

Knowledge Base

{(***OPEL***, is_a, **automobil_company**),
(**automobil_company**, build, **Cars**),
(**Cars**, are_build_from, **Iron**),
(***OPEL***, hat_production_facilities_in, **Germany**),
(**Germany,** is_in, **Europe**)
(**OPEL**, is_a, **Major_corporation**),
(**Major_corporation**, have, over**_10,000_employees**),
(**OPEL**, is_in, **reconstruction_phase**)}

Knowledge-based Event Processing



Representation of Events and Event Patterns

Event Query Representation

- SQL – Like:
  - Esper , Event Processing Language http://esper.codehaus.org/
  - XchangeEQ , (LMU, Munich)
  - Cayuga Event Language (CEL), Cornel University

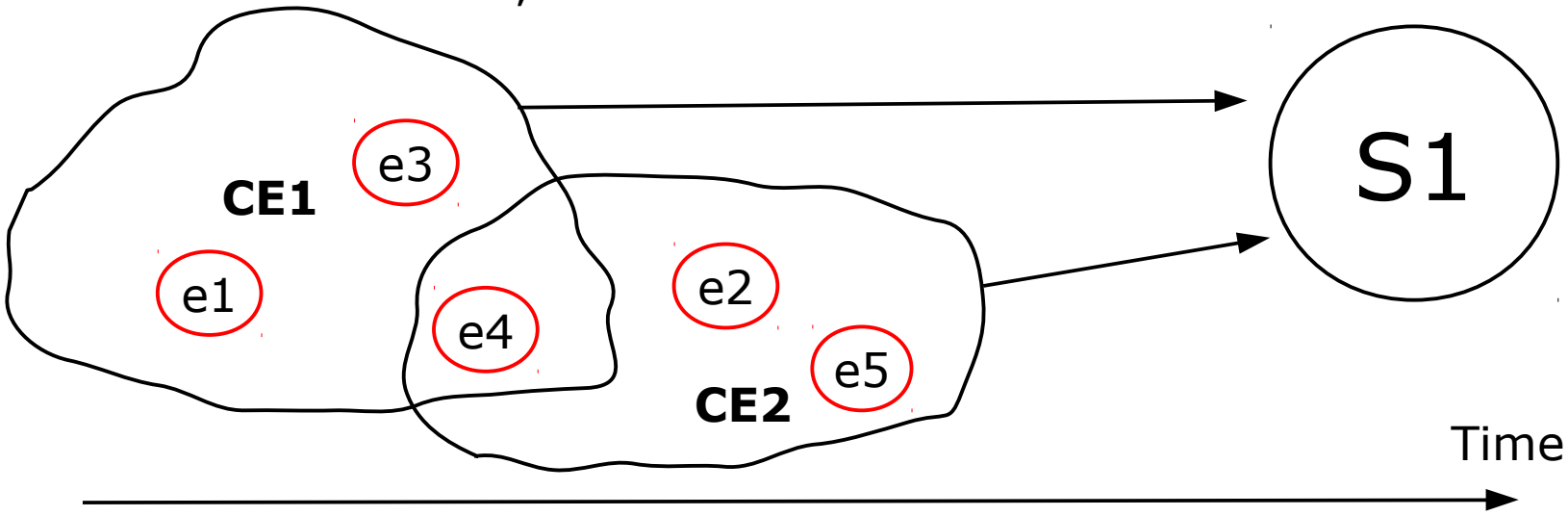
- Declarate Language
  - Prolog
  - Drools Fusion , http://www.jboss.org/drools/drools-fusion.html
  - Rule Core, XML-based rule language http://rulecore.com
  - ETALIS http://code.google.com/p/etalis/ using Prolog
  - Prova http://www.prova.ws Prolog + Java + MAS

Example: A Semantic Query Language

```
@prefix fin:<http://csw.fu-berlin.de/fin#>.

ACTION{  buy(?S1); }

STREAM{  e1:($S1,  $P1, $V1),
         e2:($S2,  $P2, $V2)

}WHERE{

  (?X1, fin:company, $S1),
  (?X2, fin:company, $S2),
  (?X1, fin:produce, ?Z),
  (?Z,  fin:buildfrom, mtr:metal),
  (?X1, fin:facilitiesin, geo:Europe),
  (?X1, fin:employees, 12,000),
  (?X1, fin:is_in, fin:reconstruction),
  (?X2, fin:oilconsume, 2.000.000),

}ON{
  e2 AFTER e1 ,
  ($P1 * $V1 >= 20000)
}WITHIN{ 10 min }
```

Complex Events vs. Situations

- What is a **complex event**?
  - An event that is an abstraction of other events called its members (EPTS Glossary)
- What is a **Situation**?
  - Is the same as **complex event**? Or is the **result** of Complex Event.
  - Situation calculus, Event calclus?



Event Processing Methods

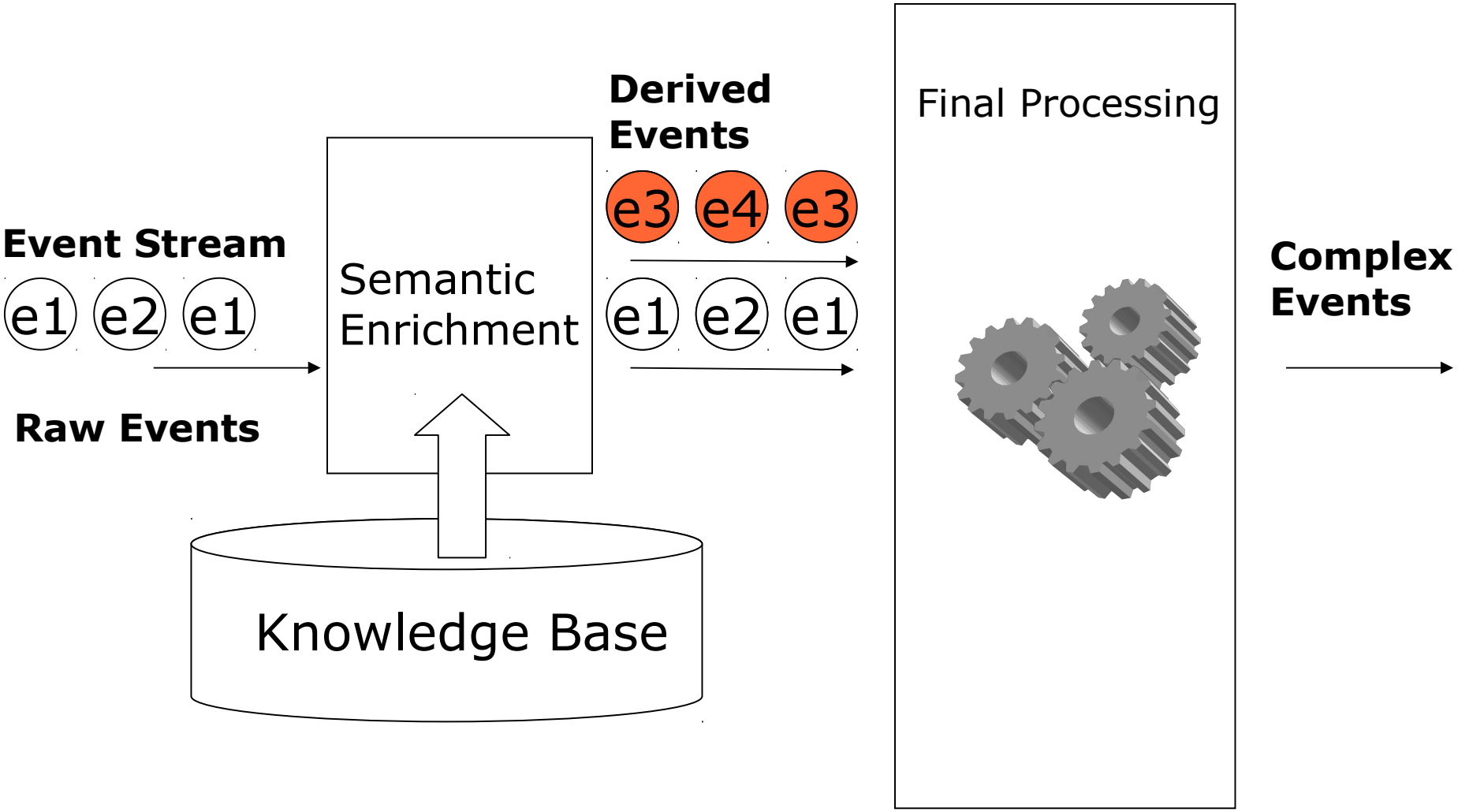
Semantic CEP Requirments

- General CEP Requirements:
  - Timely Processing (real time, near real time)
  - Scalability
    - High **throughput** of events
    - **Number of Processing Rules**
- Special SCEP Requirements:
  - Scalability
    - **Size of background knowledge**
    - **Level of reasoning** on KB
    - Frequency of **KB updates**

Processing Methods

1) Storage-based
2) Central rule engine
3) **Semantic Enrichment of Event Stream (SEES)**
4) **Event Query Pre-Processing (EQPP)**

Semantic Enrichment of Event Stream

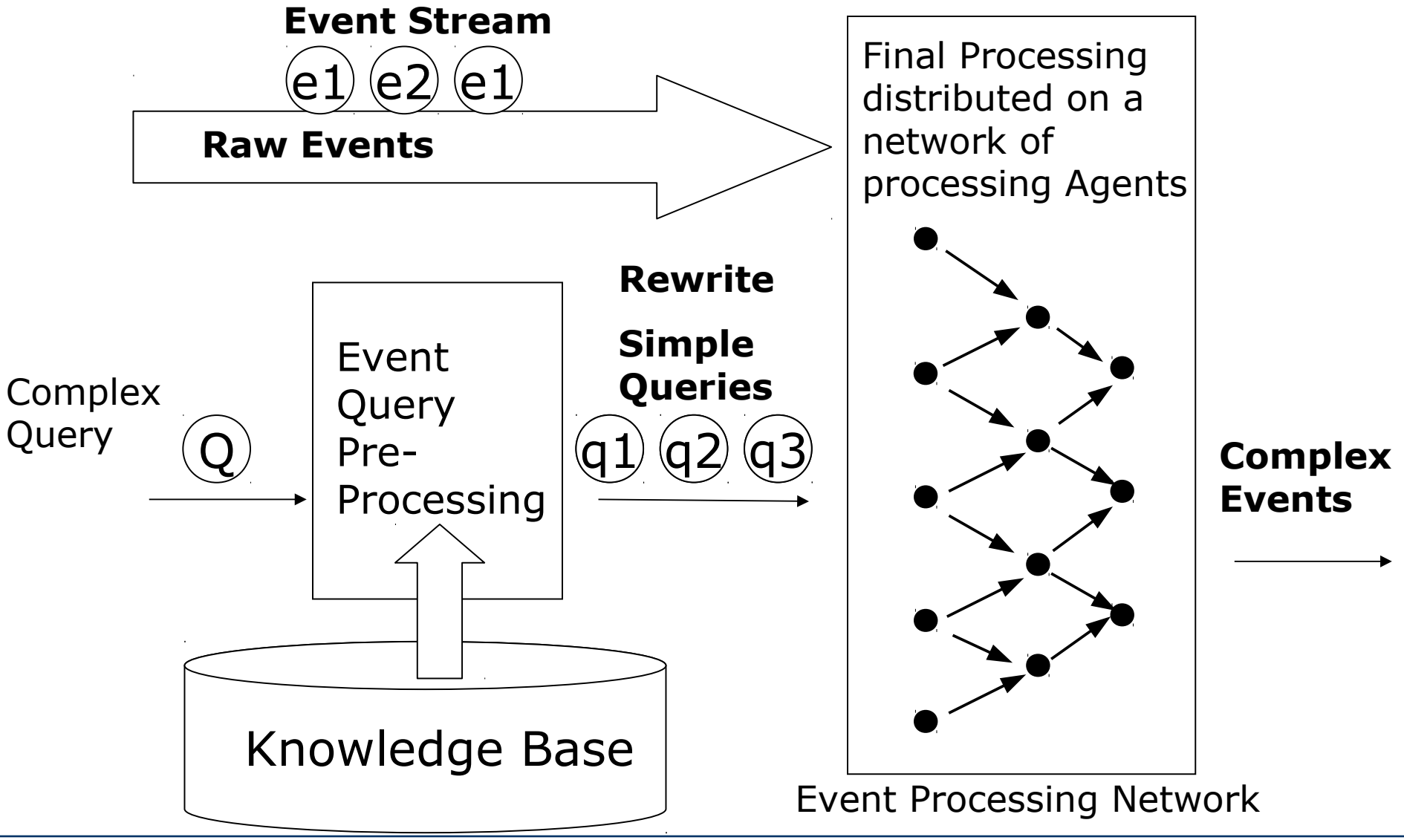


Event-Query Pre-Processing

- The **complex *query*** is pre-processed and rewritten in several **simple queries**.
- **Simple query** is a query which can be processed **without** the external KB
- New ***simple queries*** can be ***generated using the knowledge base.***

  Q U KB → q1 , q2 , q3, q4, …
- Simple queries can be in conjunction and disjunction
- Queries are processed by several Event Processing Agents
- Results are jointed together by EPAs

Event Query Pre-Processing



Comparison of Methods

	DB-Based	Rule Engine	SEES	EQPP
Performance	low	high	limited	high
Scalability	limited	limited	limited	high
Elasticity	no	no	high	high
Reasoning on KB	No/limited	No/limited	high	high

Future Work

- Representation of Event Patterns/event query
- Algorithms for rewriting complex event query
- Prove of concept implementation
- Evaluation

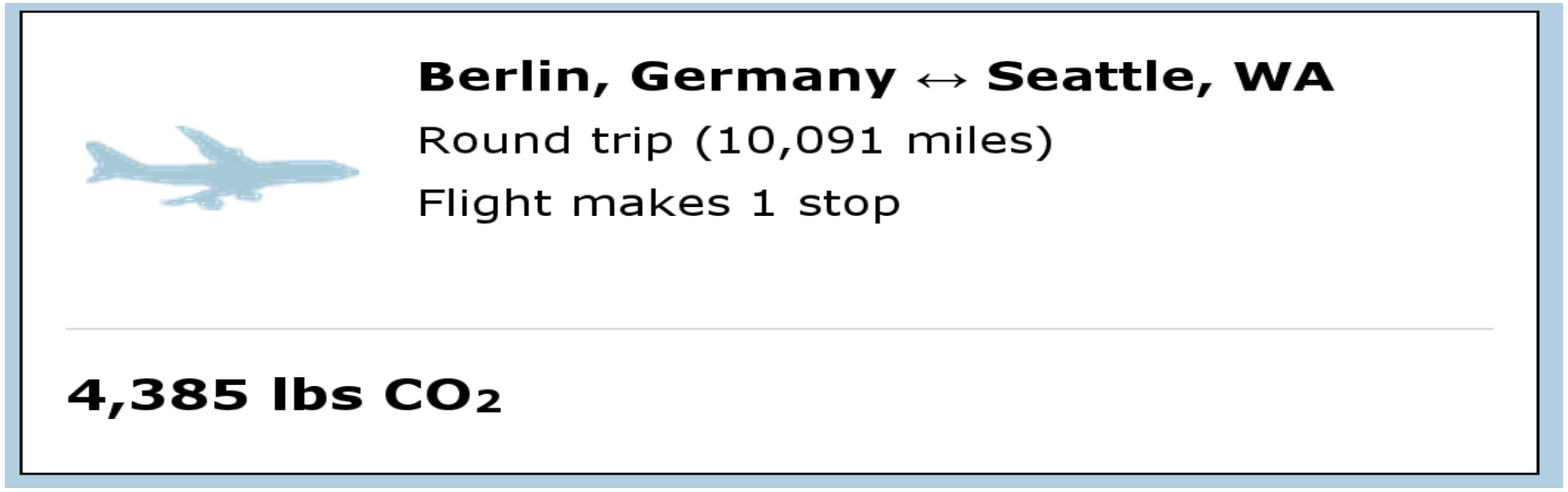
Excluded, but related Subjects

- Noisy event stream
- Uncertain events stream
- Event pattern mining

Thank You!

Please give me feedback!

I am here in Seattle until Sunday ...



Thank you!

http://www.corporate-semantic-web.de

AG Corporate Semantic Web
Freie Universität Berlin
http://www.inf.fu-berlin.de/groups/ag-csw/