

# ROSES

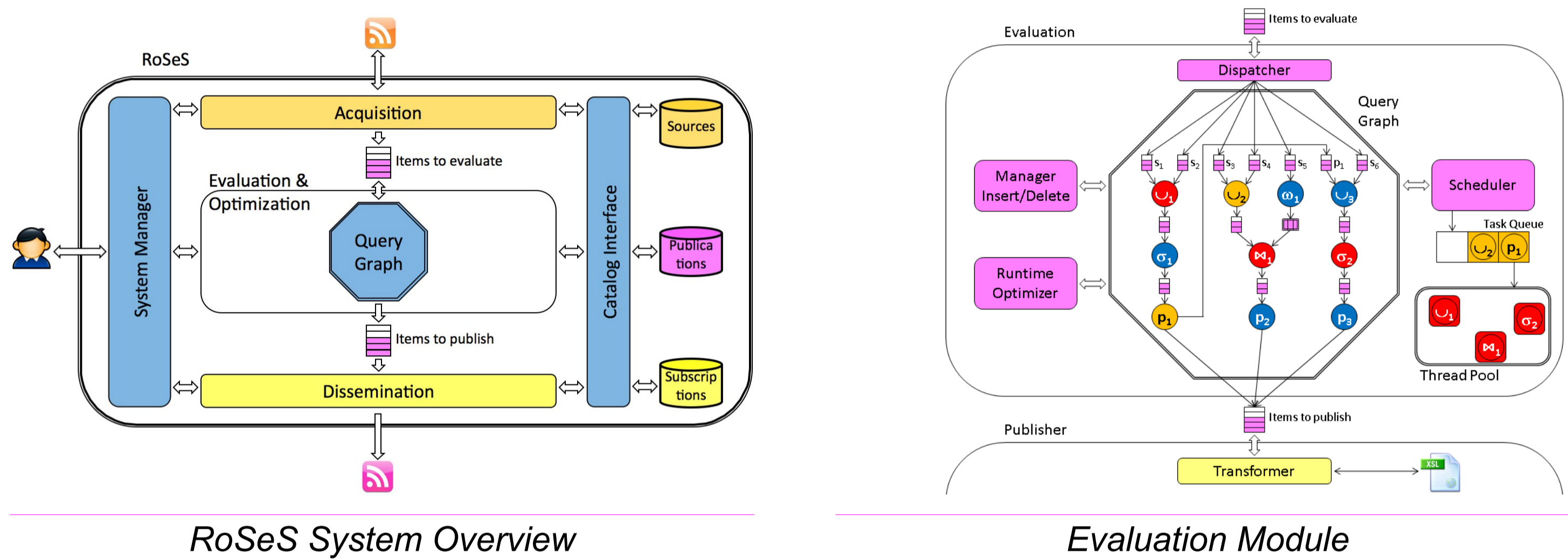
## Create and Personalize your own RSS Feeds

### RoSeS Project

The ROSES project (ANR-07-MDCO-011) aims at defining a set of web resource syndication services and tools for localizing, querying, generating, composing and personalizing RSS feeds available on the Web.

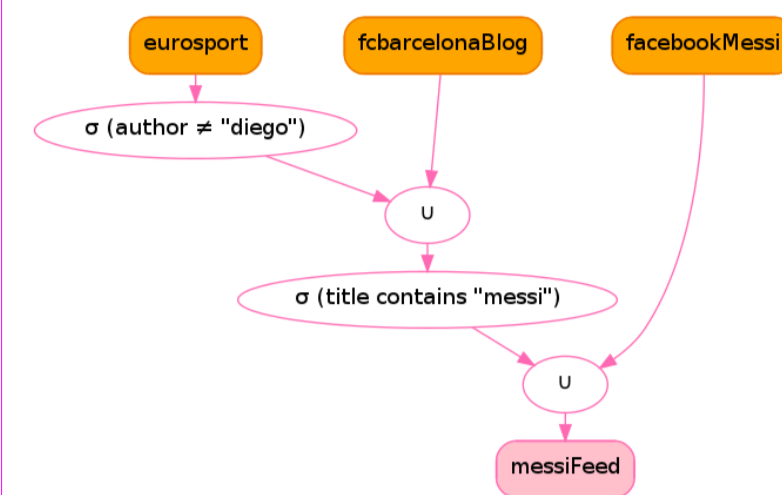
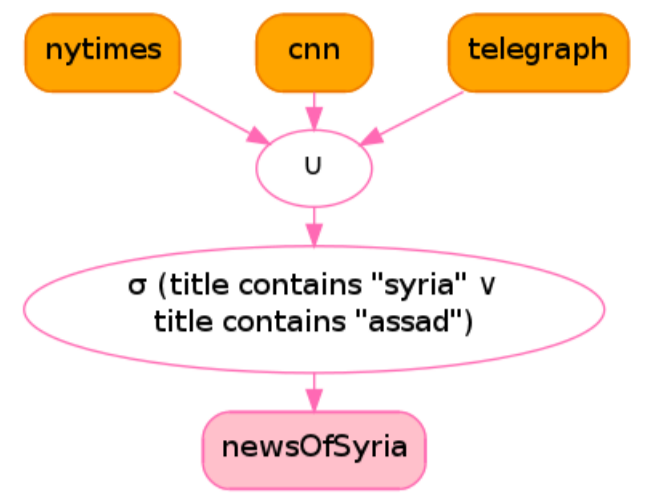
The proposed approach is based on the observation that web content syndication can be considered as a particular large-scale distributed data management problem that might be solved by combining peer-to-peer data sharing infrastructures, XML data management and continuous query processing.

### RoSeS System Architecture



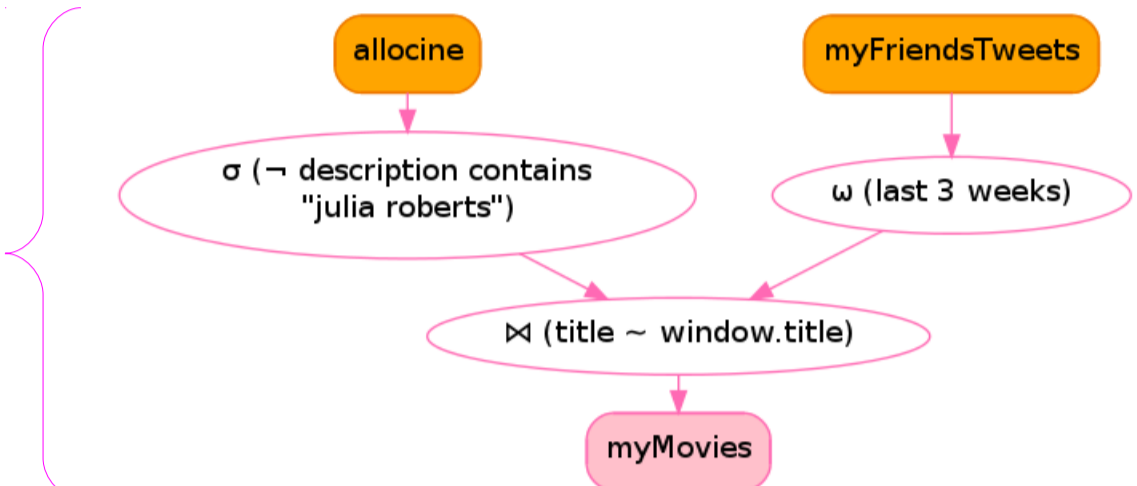
### RoSeS Publication queries

**CREATE FEED newsOfSyria**  
 FROM nytimes | cnn | telegraph  
 WHERE title CONTAINS "syria" OR title CONTAINS "assad"

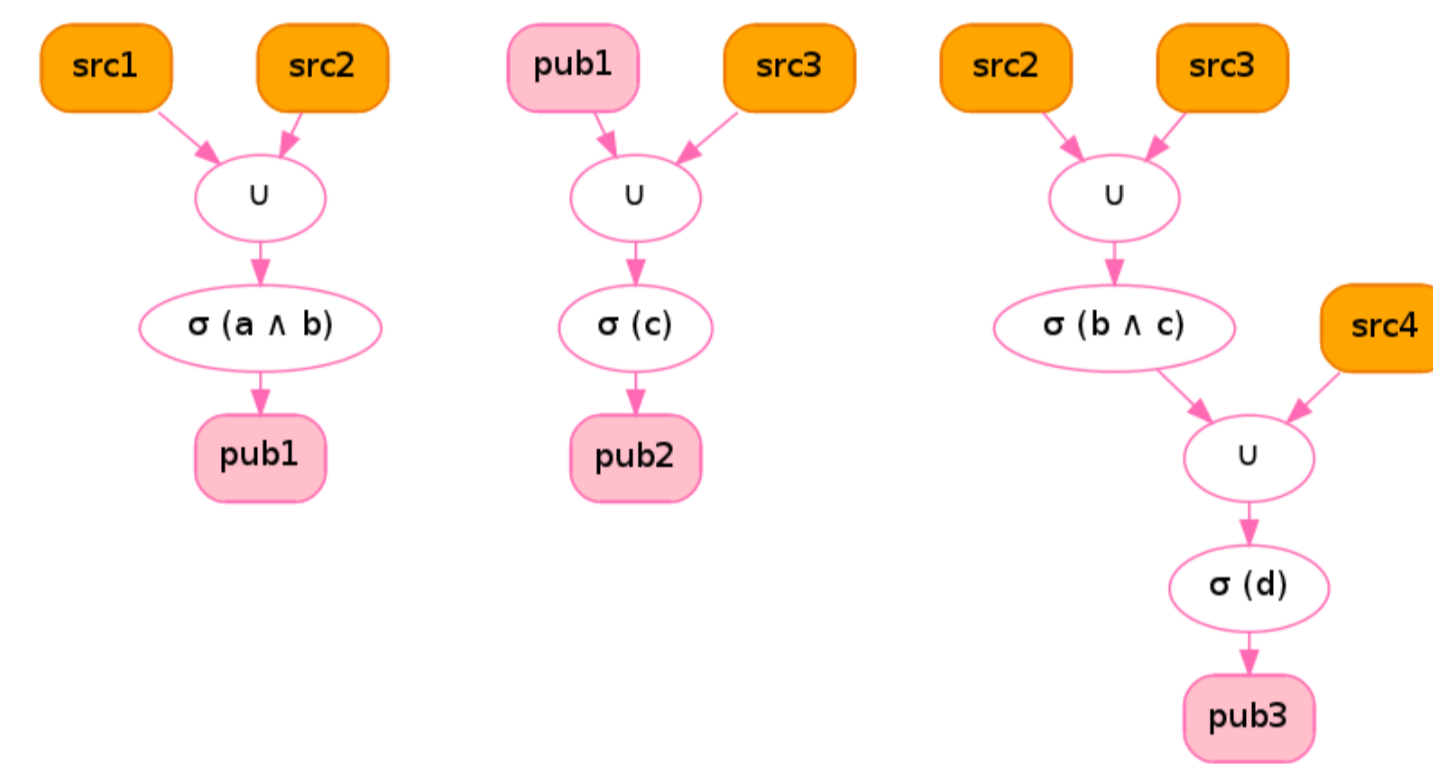


**CREATE FEED messiFeed**  
 FROM (eurosport AS \$e | fcbarselonaBlog) AS \$u | facebookMessi  
 WHERE \$e[author <> "diego"] AND \$u[title CONTAINS "messi"]

**CREATE FEED myMovies**  
 FROM allocine AS \$a  
 JOIN LAST 3 WEEKS ON myFriendsTweets  
 WITH \$a[title SIMILAR WINDOW.title]  
 WHERE \$a[description NOT CONTAINS "julia roberts"]

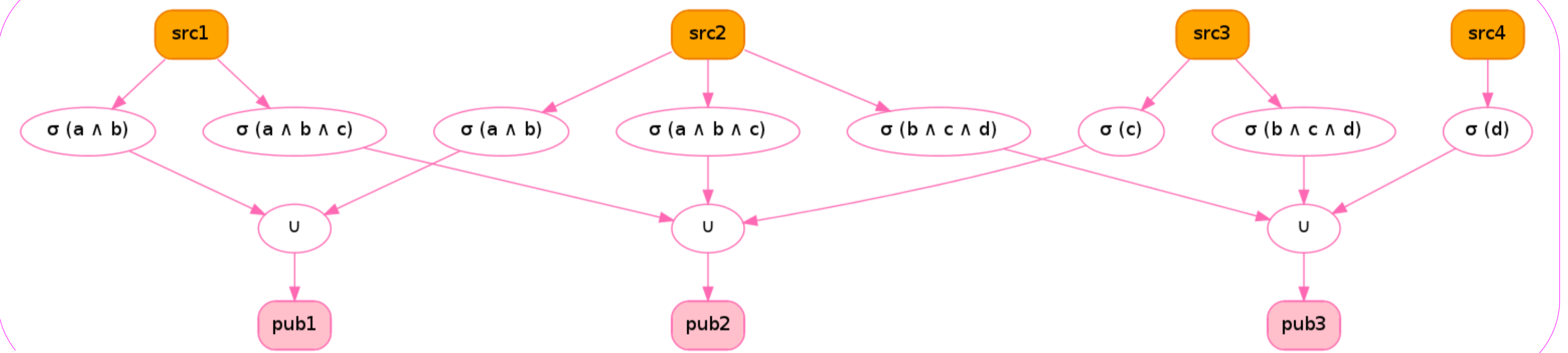


### Multi-query Optimization



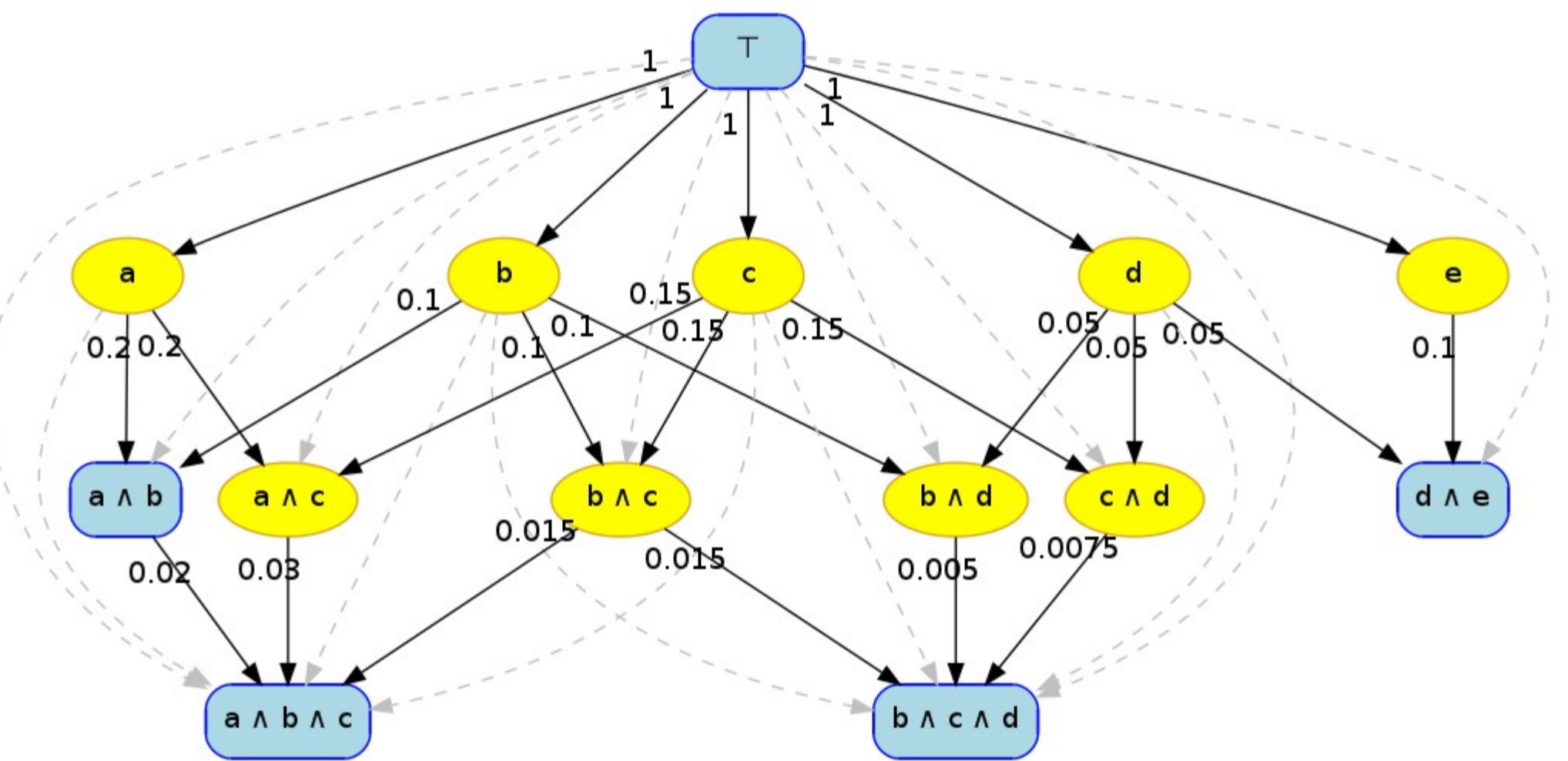
- Normalization**
- Push selection operators towards sources
  - Decompose publications
  - etc.

### Global Normal Query Graph

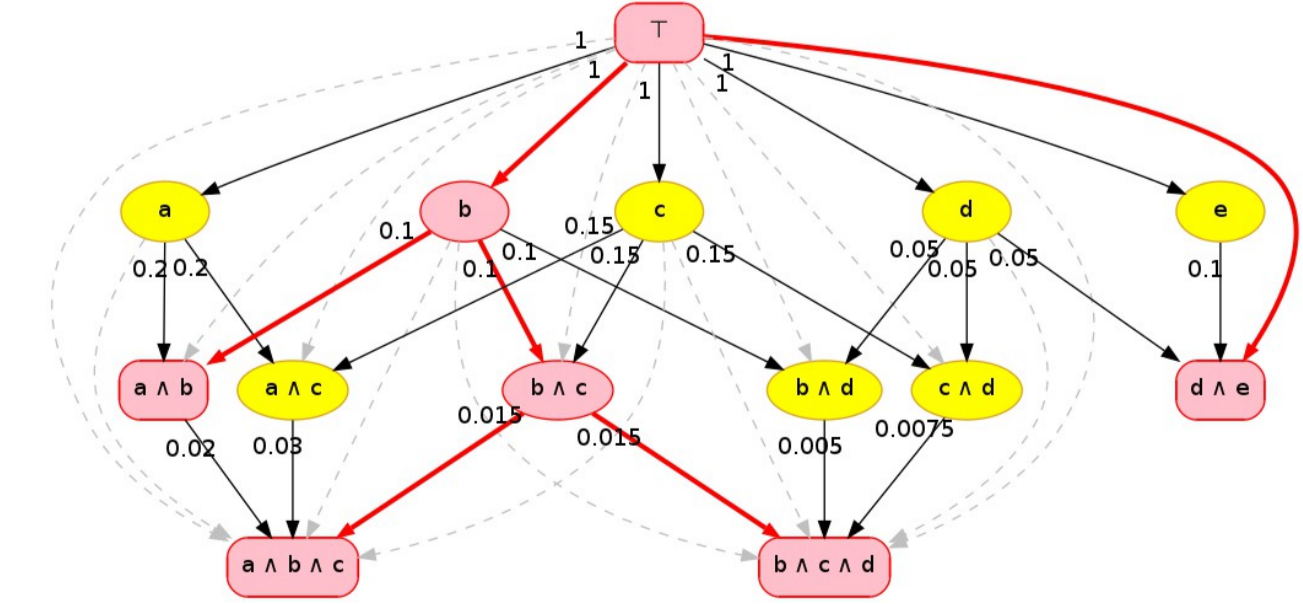


- For each Source:
- Generate a *subsumption graph*
  - Populate it with processing costs
  - Seek a *Steiner tree* on the subsumption graph (Steiner tree: *Minimum tree* spanning at least a given subset of vertices called *terminal vertices*)

### Subsumption Graph (source 2)



### Steiner Tree

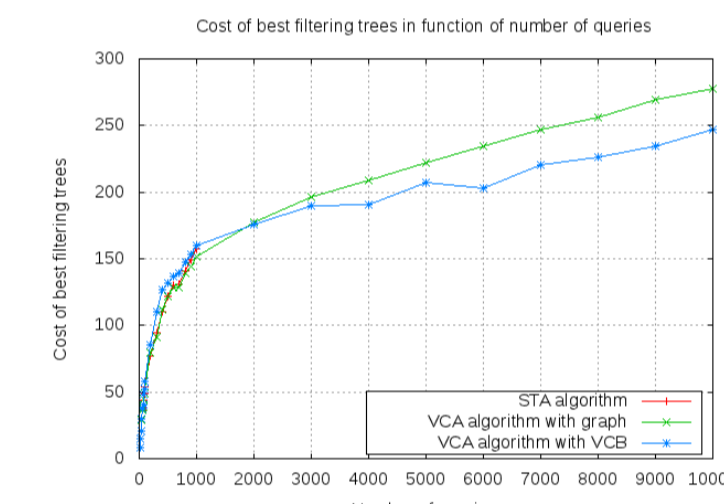
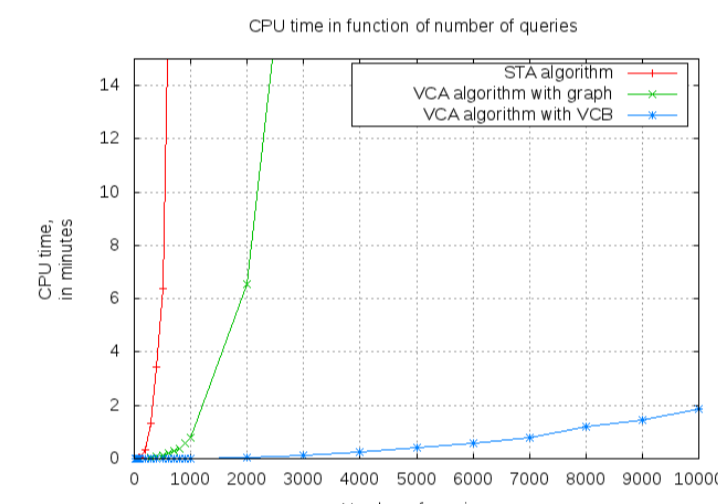


### Experimental Evaluation

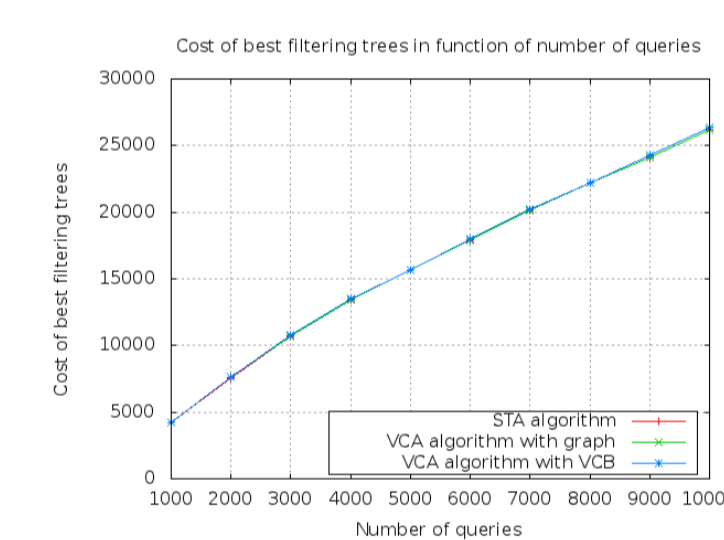
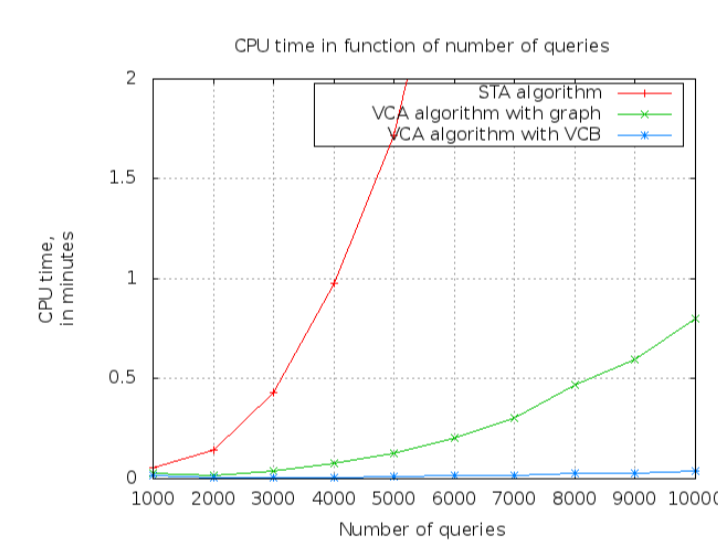
- Benchmark Dataset:**
- 1300 real-world sources (2 weeks period)
  - 5200 keywords
  - 4 keywords/source average

Synthetic Query Generator

**CREATE FEED publicationName**  
 FROM (src<sub>1</sub> | src<sub>2</sub> | src<sub>3</sub> | ... | src<sub>n</sub>) AS \$var  
 WHERE \$var[predicate\_in\_CNF]

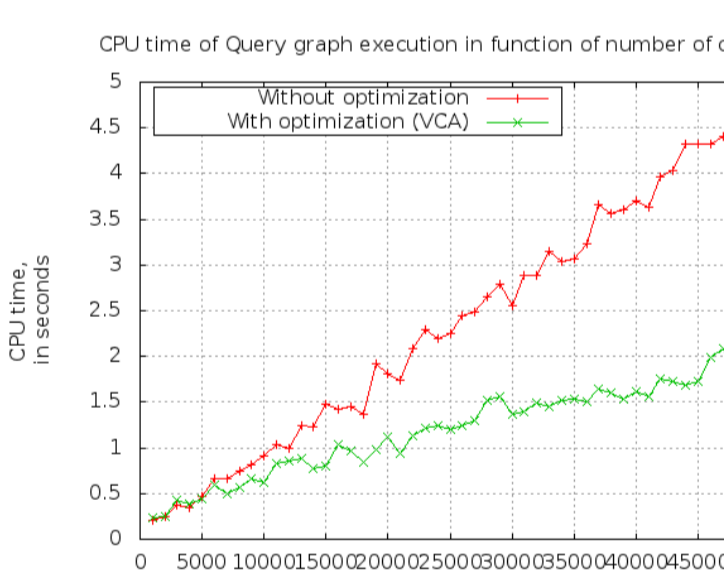
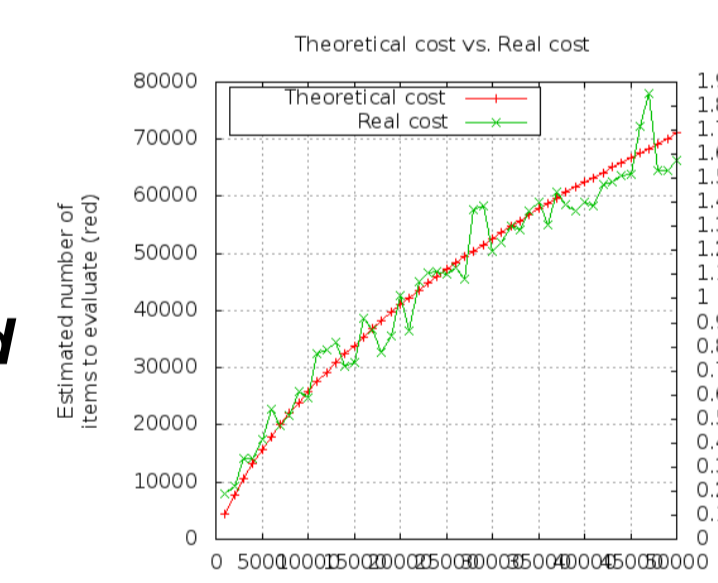


**Mono-source queries with conjunctive filter predicates of 1 to 3 atomic predicates**

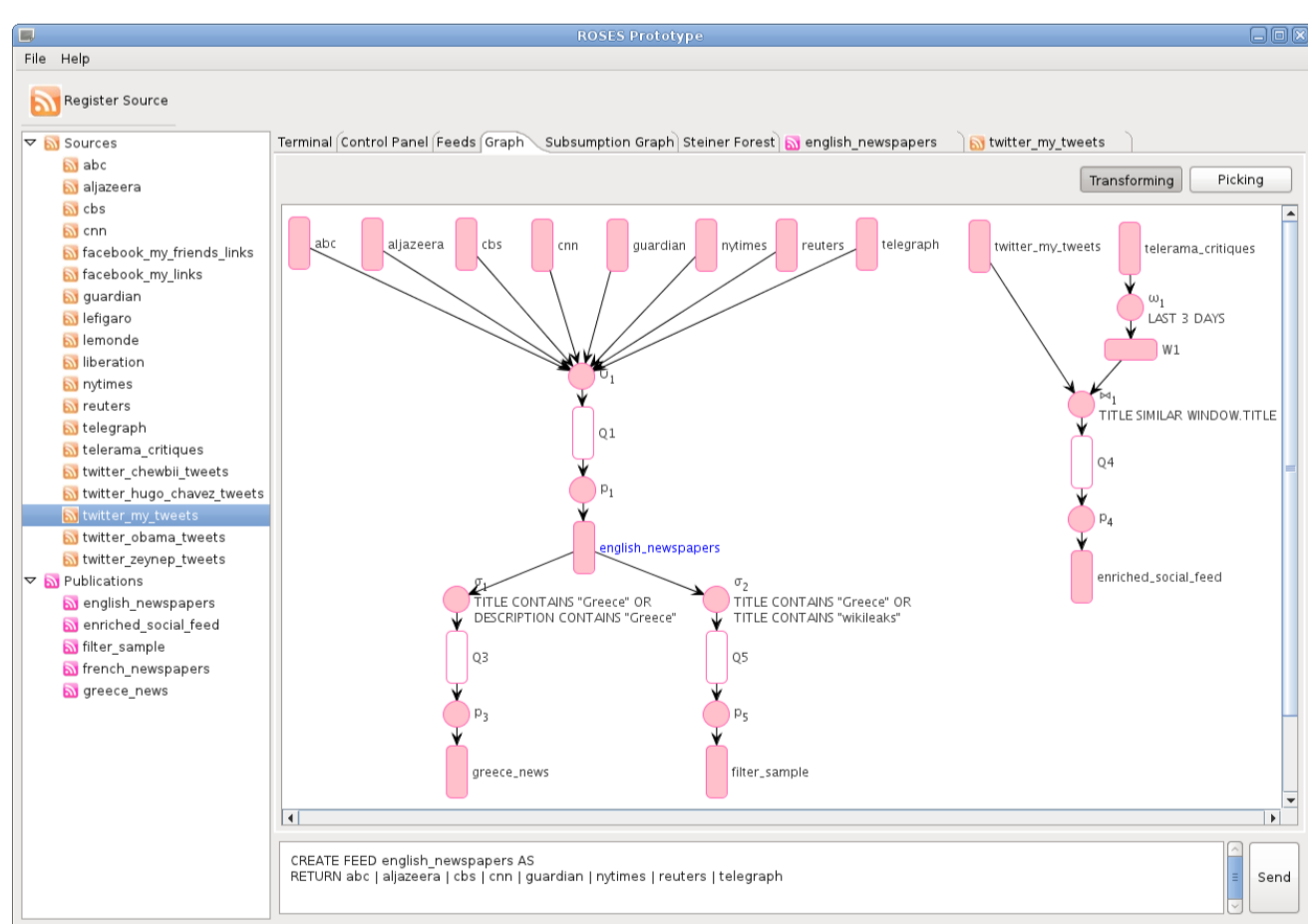


**Multi-source queries (1 to 10 sources per query) with conjunctive filter predicates of 1 to 3 atomic predicates**

**Cost model validation and Query graph evaluation with and without optimization**



### RoSeS Prototype Screenshots

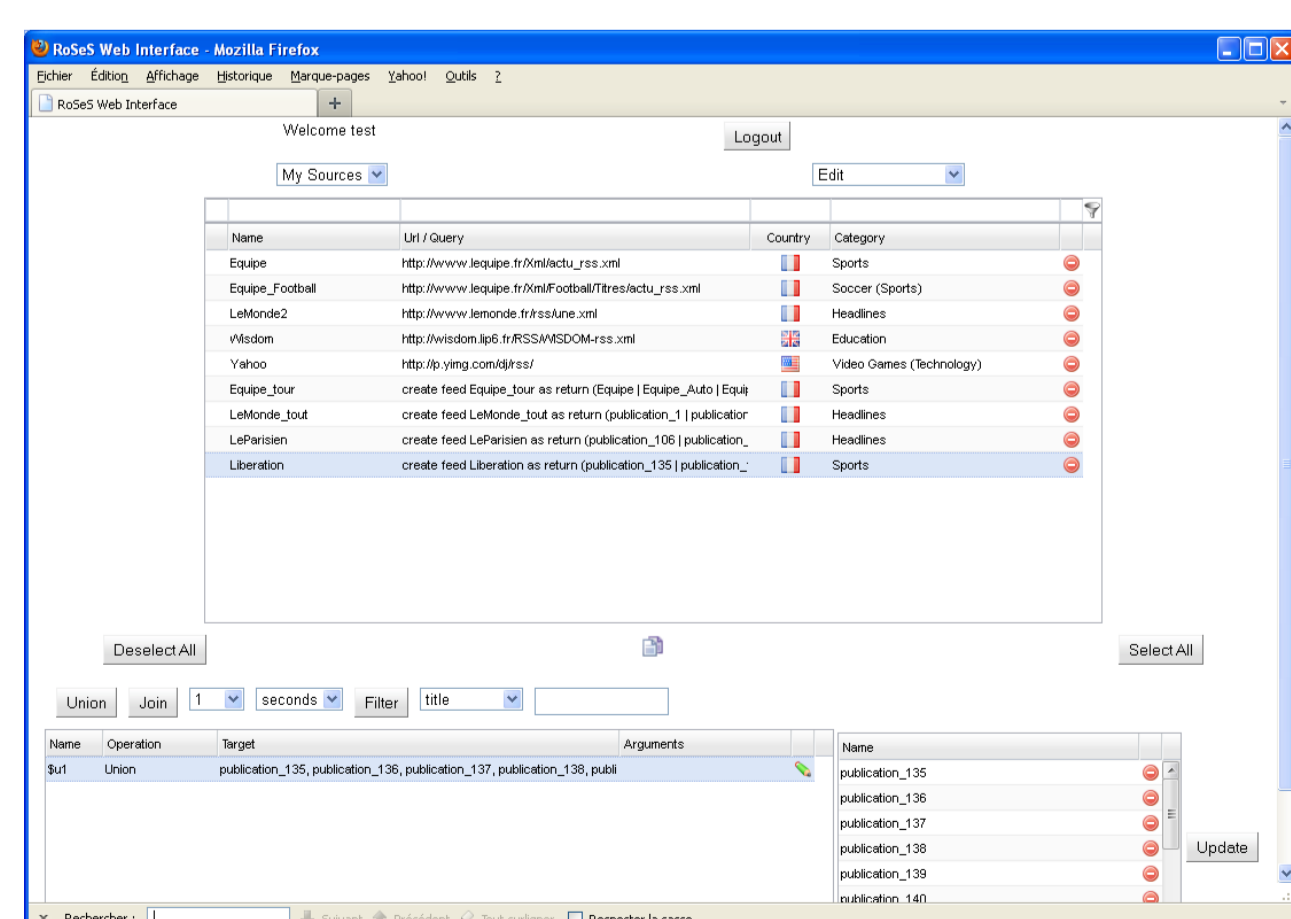


RoSeS web-client

- Features:
- Visual Query-builder

### RoSeS java Prototype

- Features:
- RSS/Atom feed crawling
  - Multi-thread Continuous query-engine
  - Publication output formatter
  - Source registering
  - Publication query insertion
  - Physical query graph Visualizer
  - Query graph Optimization



<http://www-bd.lip6.fr/roses>