

Information Filtering with Semantic Knowledge Base

Ivo Lašek

Faculty of Information Technology
Czech Technical University in Prague
lasekivo@fit.cvut.cz

26.5.2011 – Knowledge Engineering Group Seminar

Outline

- Use Cases
- What? – Targets
- How? – Used Techniques
- Problems
- Future Work

My Questions?

- Does it make sense?
- Tips and improvements? Feature requests?
- Concrete tools?
- Concrete techniques?
- Similar projects?

Use Cases







What?

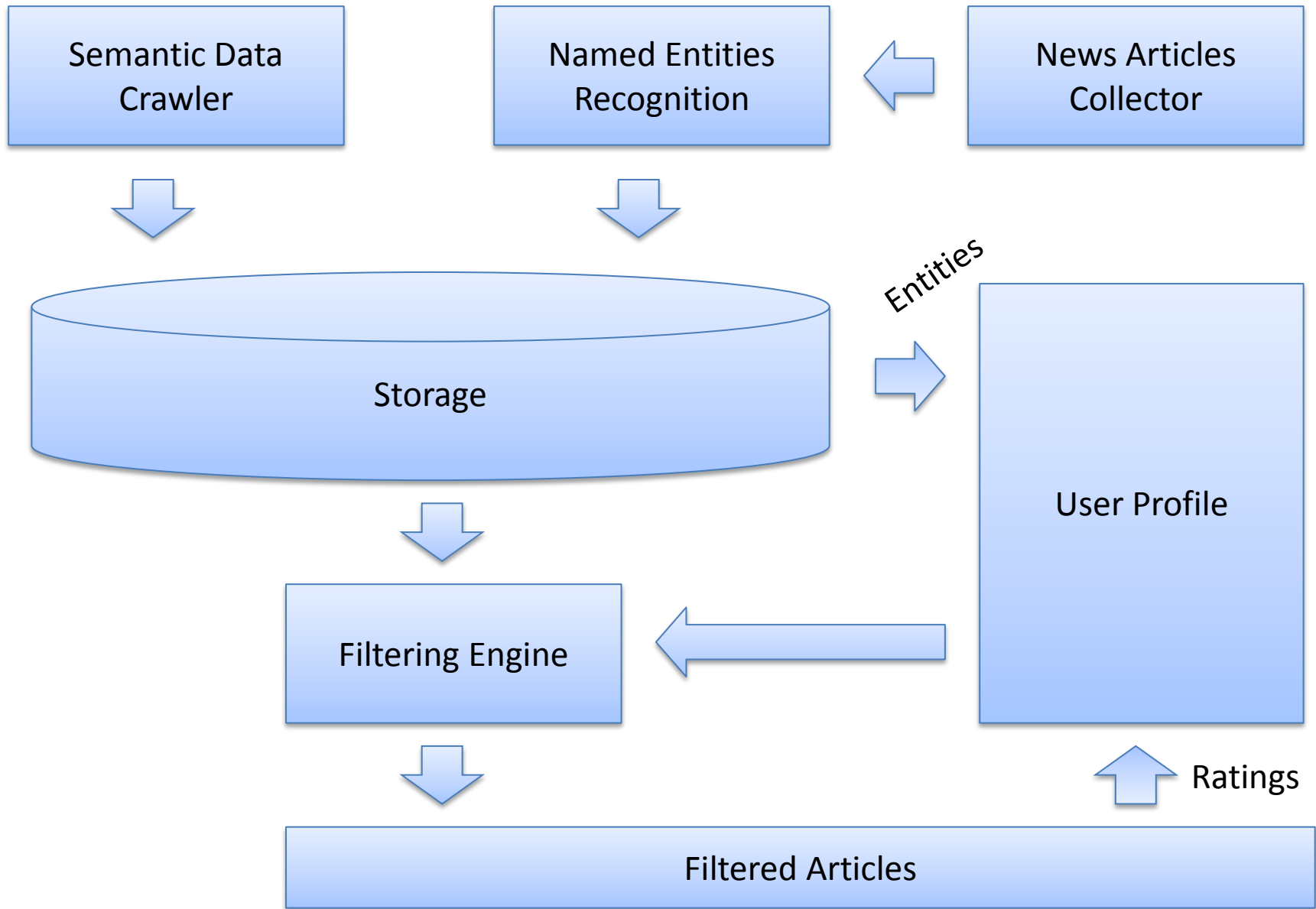
Intelligent RSS reader

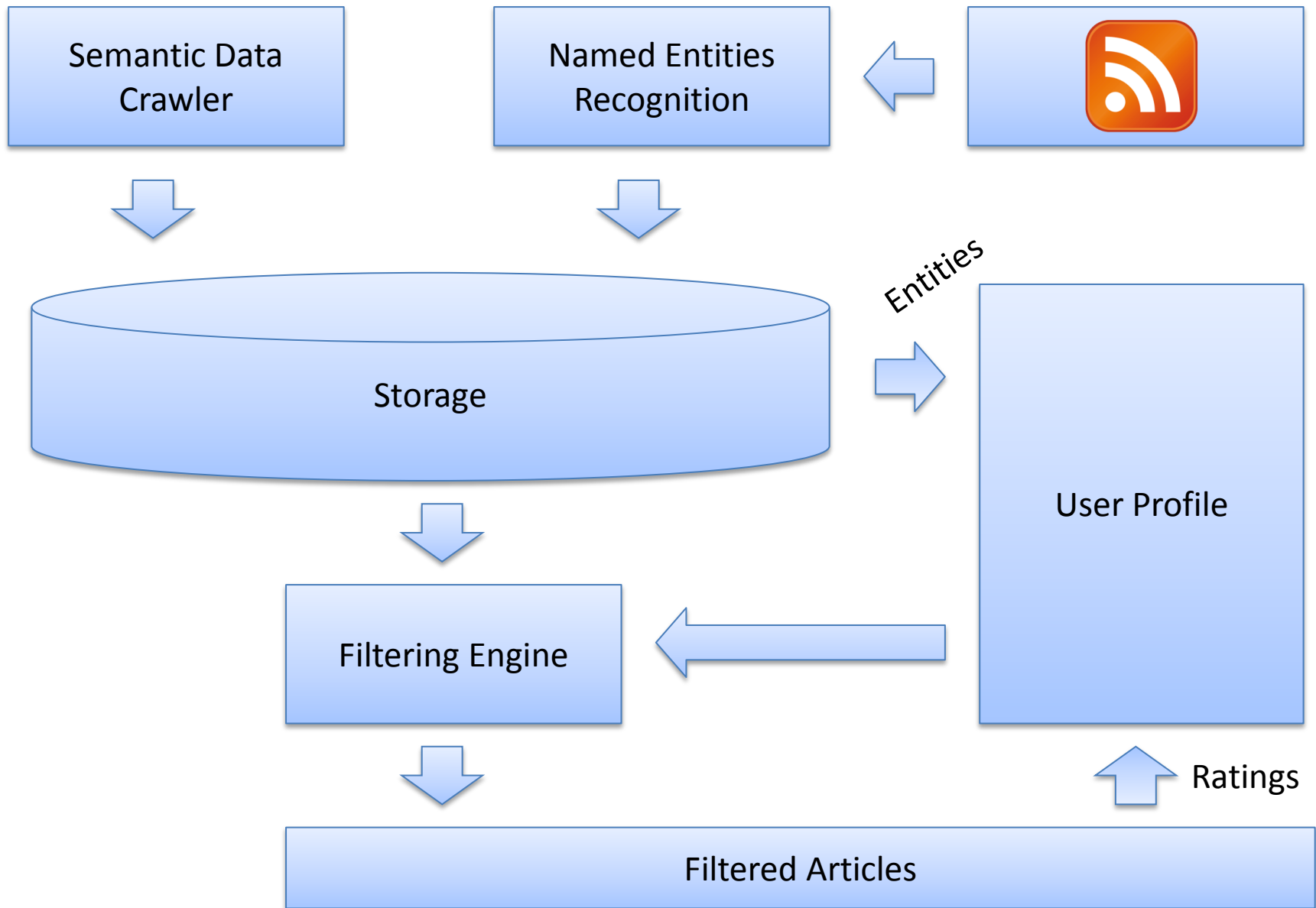
- Add all your favorite information sources
- Define rules (e.g. all events in Prague, news about IT startup companies)
- Rate articles – how well they fit your needs
- Only the articles describing entities fitting your search criteria are delivered, the rest is thrown away

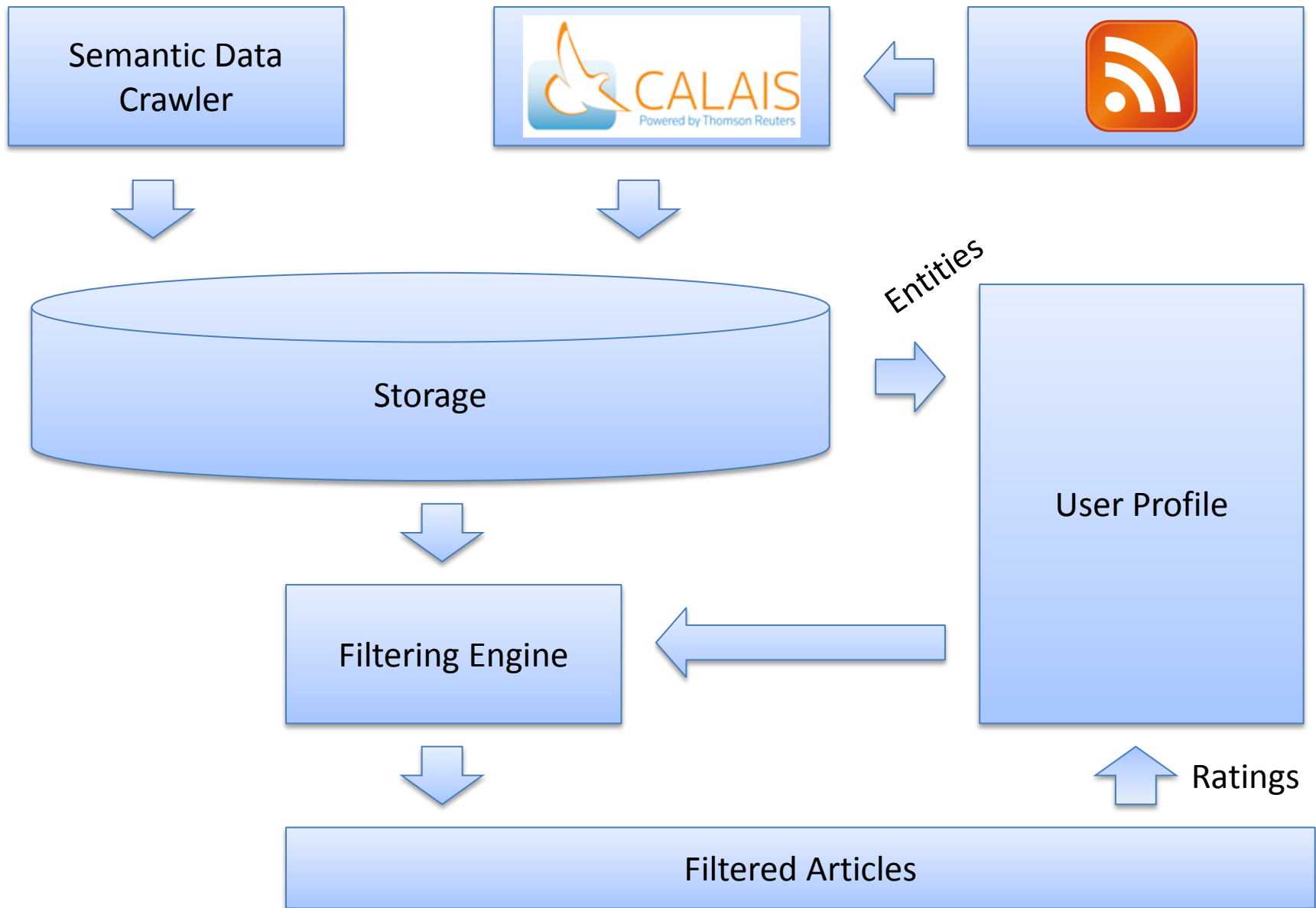
Related Work

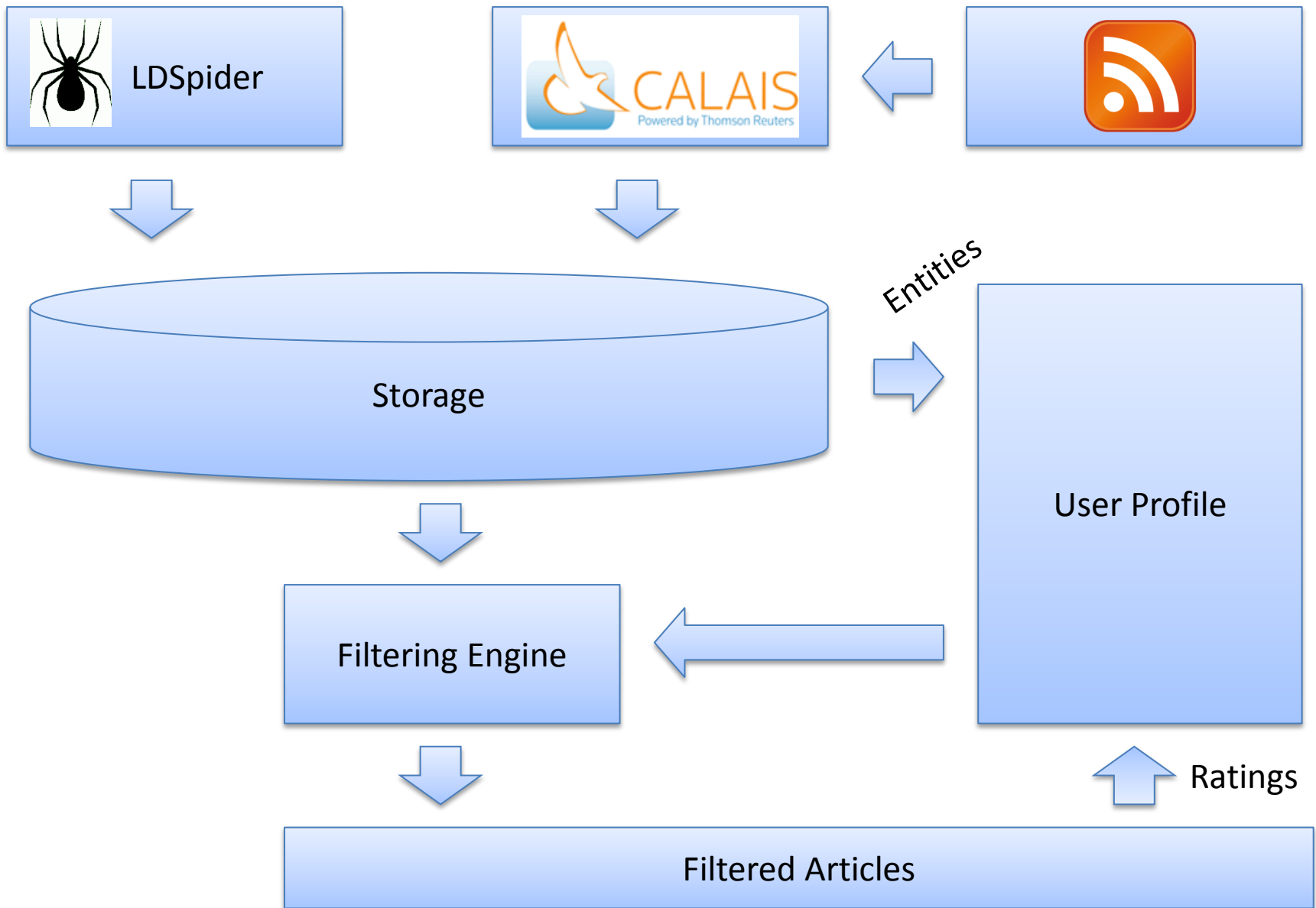
- Collaborative filtering
 - E.g. Google News 2007
- Content based filtering
 - E.g. Google News 2010
 - Text based information retrieval: Language Modeling, Okapi BM25

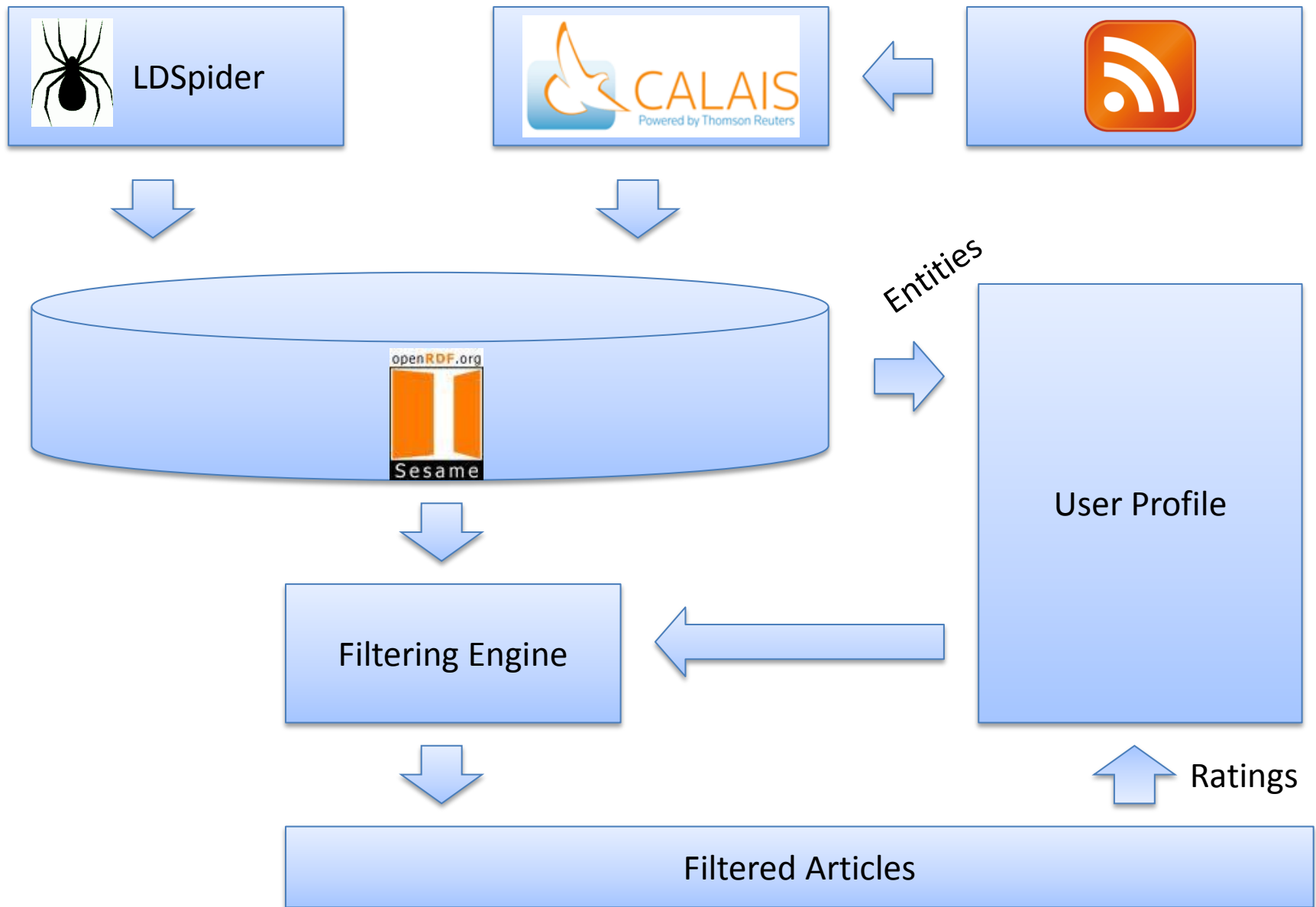
How?



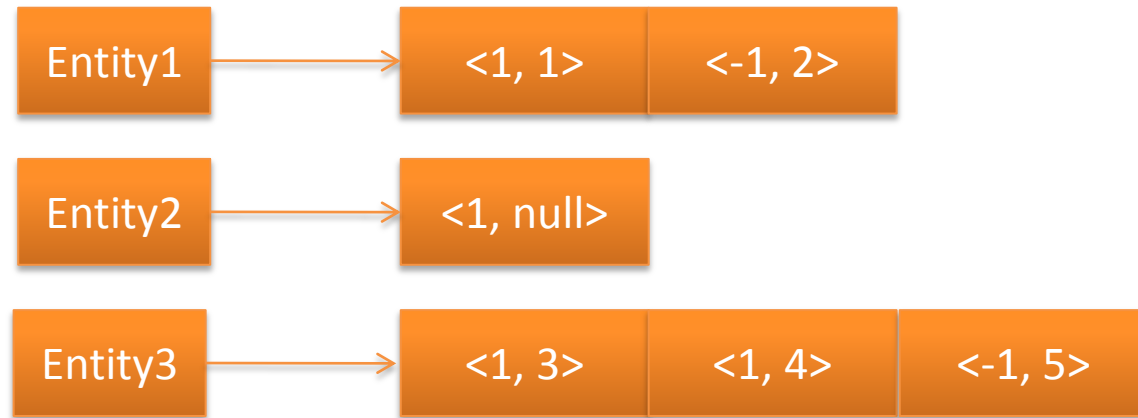








User Profile



EF-IDEF

- Proposed ef-idef metric similar to the tf-idf metric known in Information Retrieval
- Entity Frequency – Inverse Document Entity Frequency

$$ef_{i,j} = \frac{e_{i,j}}{\sum_k e_{k,j}}$$

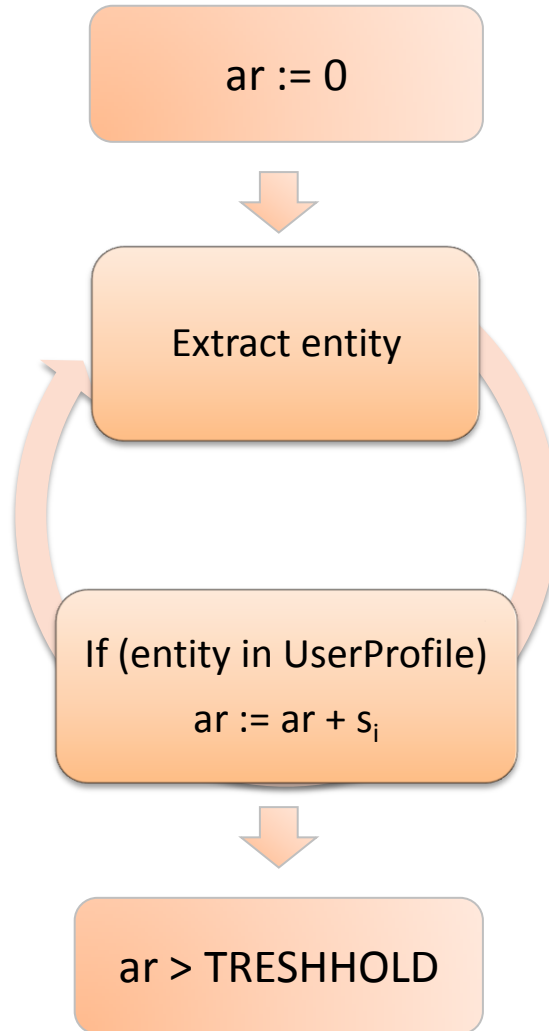
$$idef_i = \log \frac{|D|}{|\{d : e_i \in d\}|}$$

Entity Significance

- Measure of the user interest in a particular entity

$$s_i^t = \sum_j (r_{i,j} \times ef_{i,j}^t \times idef_i^t)$$

Filtering Criteria



Problems

Crawling

- Breath first search
- Many irrelevant data sources (hundreds of thousands of foaf profiles of completely unimportant people)
- Lack of focus
- Semantic site maps may help – crawl primarily well managed data sources (Freebase, DBPedia)
- Requests from the Entity Extractor

Storage

- Sesame is extremely slow for certain queries (data from about 25 000 structured documents)
- Sesame index optimization?
- Another repository?
- Identification of most critical query types => custom made indexes?

Entity Matching

- Calais contains links to other Linked Data sources
- Not for all entities – possible automatic matching?

Future Work

Annotated Articles

- Primary direction – information filtering
 - Simple model so far
 - In the future possible data mining techniques based on properties of identified entities
- Timelines
- Sentiment analysis & connection to share price development
- Article recommendation – similar to Seevl.net (<http://www.seevl.net/>)

Nirvana was an American rock band that was formed by singer/guitarist Kurt Cobain and bassist Krist Novoselic in Aberdeen, Washington in 1987. Nirvana went through a succession of drummers, the longest-lasting ...

[read more](#)

📺 Videos



provozováno na 

📄 Fact sheet

Activity end(s): 1994

Activity start(s): 1987

Collaboration(s): [Fecal Matter](#)

Genre(s): [Alternative rock](#)
[Grunge](#)
[Punk rock](#)
[Rock music](#)

Label(s): [DGC Records](#)
[Geffen Records](#)
[Sub Pop](#)

Band member(s): [Aaron Burckhard](#)
[Chad Channing](#)
[Dale Crover](#)

Fecal Matter

[Show the relations !](#)

Fecal Matter was a short-lived band from Aberdeen, Washington, that formed in late 1985. It was the first band conceived by Kurt Cobain, who would become famous as the frontman and guitarist of the seminal grunge band, Nirvana. The band recorded only one demo tape, Illiteracy Will Prevail, which remains ...

[more info](#) ▶

Foo Fighters

[Show the relations !](#)

- Both have the same **origin**: [United States](#)
- Both **collaborated with**:
 - [Dave Grohl](#)
 - [Nukeateen](#)
- Both have similar **topic**:
 - [1990s music groups](#)
 - [Grammy Award winners](#)
 - [American alternative rock groups](#)
 - [BRIT Award winners](#)
- Both play the same **genre**:
 - [Punk rock](#)
 - [Rock music](#)
 - [Alternative rock](#)
- Both have common **members**:
 - [Dave Grohl](#)
 - [Pat Smear](#)

The Foo Fighters are an American rock band from Seattle, Washington. Former Nirvana drummer Dave Grohl formed the group as a one-man project in 1994 after the dissolution of his previous band. Prior to the release of the Foo Fighters' 1995 debut album Foo Fighters, which featured Grohl as the ...

[more info](#) ▶

Questions?

My Questions?

- Does it make sense?
- Tips and improvements? Feature requests?
- Concrete tools?
- Concrete techniques?
- Similar projects?