

User Modeling

Saeed Seddighin

Personalized Web Services

- Search
- Page Suggestions

How to collect information

- Questionnaire
- Learning users behavior

Multiple Sources

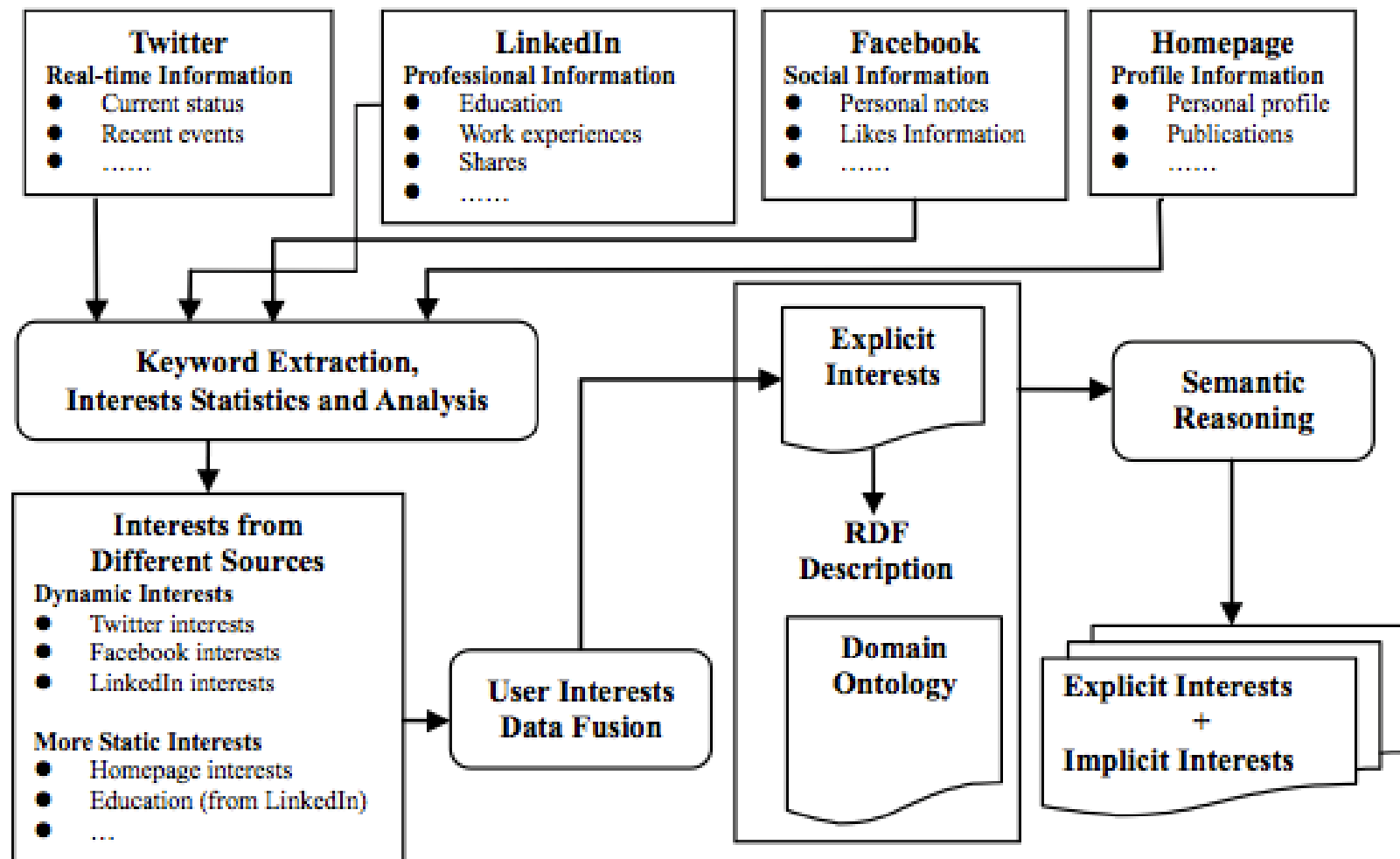
- User modeling based on multisource personal information
- Investigating the semantic relationship between users' interests
- Deriving implicit information based on reasoning

Information

- Explicit / Implicit
- Static / Dynamic

Workflow

- 1- Specifying the sources of information
- 2- Extracting explicit information from each of the sources
- 3- Merging all the information
- 4- Describing information by knowledge representation language
- 5- Semantic reasoning to extract implicit information based on explicit information



```
<foaf:Person rdf:about="http://www.cs.vu.nl/~frankh/">
  <foaf:name>Frank van Harmelen</foaf:name>
  <e-foaf:interest>
    <rdf:Description rdf:about="http://www.wici-lab.org/wici/
      wiki/index.php/Web">
      <dc:title>Web</dc:title>
      <e-foaf:cumulative_interest_value rdf:parseType="Resource">
        <rdf:value rdf:datatype="xsd:number">34.922</rdf:value>
      </e-foaf:cumulative_interest_value>
    </rdf:Description>
  </e-foaf:interest>
  ...
</foaf:Person>
```


Improving Search Quality

- Defining a model to capture users interest
- Extracting users' interest based on his click history
- Improving search quality based on the extracting information

User Modeling

Definition 1 (Topic Preference Vector) A user's topic preference vector is defined as an m -tuple $\mathbf{T} = [T(1), \dots, T(m)]$, in which m is the number of topics in consideration and $T(i)$ represents the user's degree of interest in the i^{th} topic (say, "Computers"). The vector $\mathbf{T}$ is normalized such that $\sum_{i=1}^m T(i) = 1$. □

Topic-Sensitive Page Rank

$$PR(v) = d * \sum_{u \in \mathcal{A}_v} PR(u)/l_u + (1 - d) * E(i)$$

$$E_t(i) = \begin{cases} 1/n_t & \text{if page } i \text{ is related to topic } t \\ 0 & \text{otherwise,} \end{cases}$$

Creating Hierarchical Profiles

- Allowing applications to access documents on users desktop
- Using Wikipedia concepts to tag files
- Creating hierarchical profiles based on the tags

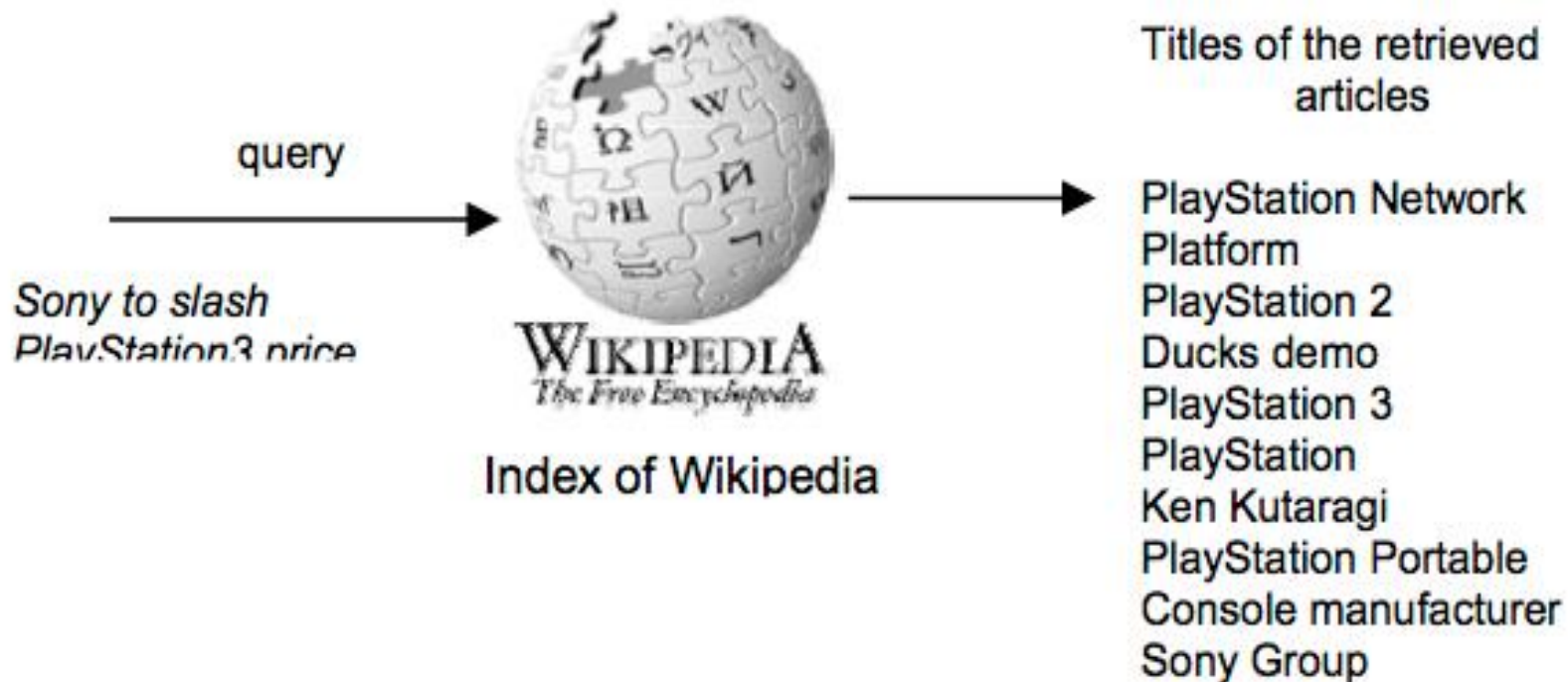
How to tag files?

- Word-Based approach is not efficient
 - Many different words have the same meanings
 - A word may have different meanings
 - A document may contain many irrelevant words.
 - Size of the profile grows by the number of words

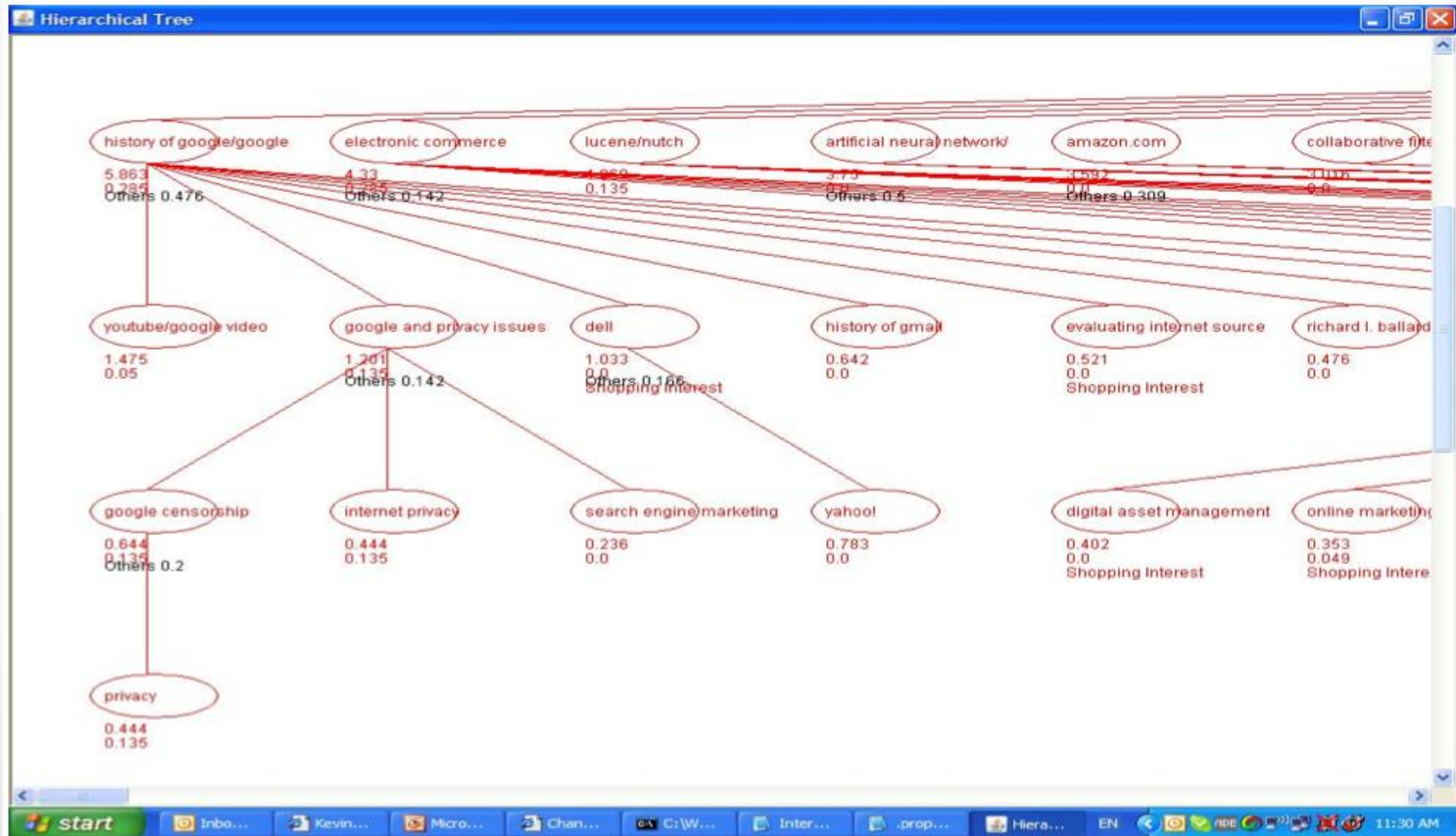
How to tag files?

- Wikipedia Concepts
 - We can tag each document with a set of Wikipedia concepts
 - Concepts of Wikipedia have more information than words:
 - tree (data structure)
 - tree (graph theory)
 - tree (set theory)

How to tag files?



Hierarchical Profiles



THANK
YOU

