

# Comparison



Stefan Bischof  
Stefan Rümmele



## Compared Papers

### **[AKM+03]**

Automatic Ontology-Based Knowledge  
Extraction from Web Documents (#15)

### **[ETL03]**

Automating the Extraction of Data from  
HTML Tables with Unknown Structure (#16)



# Differences (1)

## Input

[AKM+03] unstructured text

[ETL03] structured text (tables)

## Output

[AKM+03] unstructured text; Narrative

[ETL03] structured text

## Interface

[AKM+03] human users

[ETL03] interface for further processing



## Differences (2)

### Internal data source

**[AKM+03]** Knowledge base (KB)

**[ETL03]** not specified

### Project structure

**[AKM+03]** combines several tools

**[ETL03]** standalone project



# Similarities

## **Extraction ontology**

- key aspect of the extraction process
- automates wrapper creation

## **Narrow domain of interest**

- thus usable for specific applications
- no general information extraction tool

## **Manual ontology creation**

- time-consuming adaptation to other domains

## **Domain knowledge**

- needed for ontology creation



**Almost done ...**

**Thanks for your attention**

**Any Questions?**

**HAPPY WEEKEND !!!**



JT.

© Jolyon Troscianko  
jolyon.troscianko@lineone.net  
From 'Consciousness' S. Blackmore