

Proseminar Wissenschaftliches Arbeiten, 181.081, SS 2005, 2h

DATE: 7.3.2005

# Web Information Extraction



Inhalt: Proseminar zu aktuellen Forschungsthemen im Bereich der intelligenten Webdatenextraktion

Lehrende:

- Prof. Georg Gottlob
- Dr. Marcus Herzog
- Dipl.-Ing. Wolfgang Gatterbauer
- Dipl.-Ing. Bernhard Krüpl

Ort: Seminarräume 184/2 und 184/3 (Favoritenstraße 9-11, 3.Stock)

Webpage: [www.dbai.tuwien.ac.at/education/wie/](http://www.dbai.tuwien.ac.at/education/wie/)

Kontakt: [gatter @ dbai.tuwien.ac.at](mailto:gatter@dbai.tuwien.ac.at) (Wolfgang Gatterbauer)

Verwandte LV:

- Webdatenextraktion und -integration (181.130): WS
- Semistrukturierte Daten 2 (181.139): SS

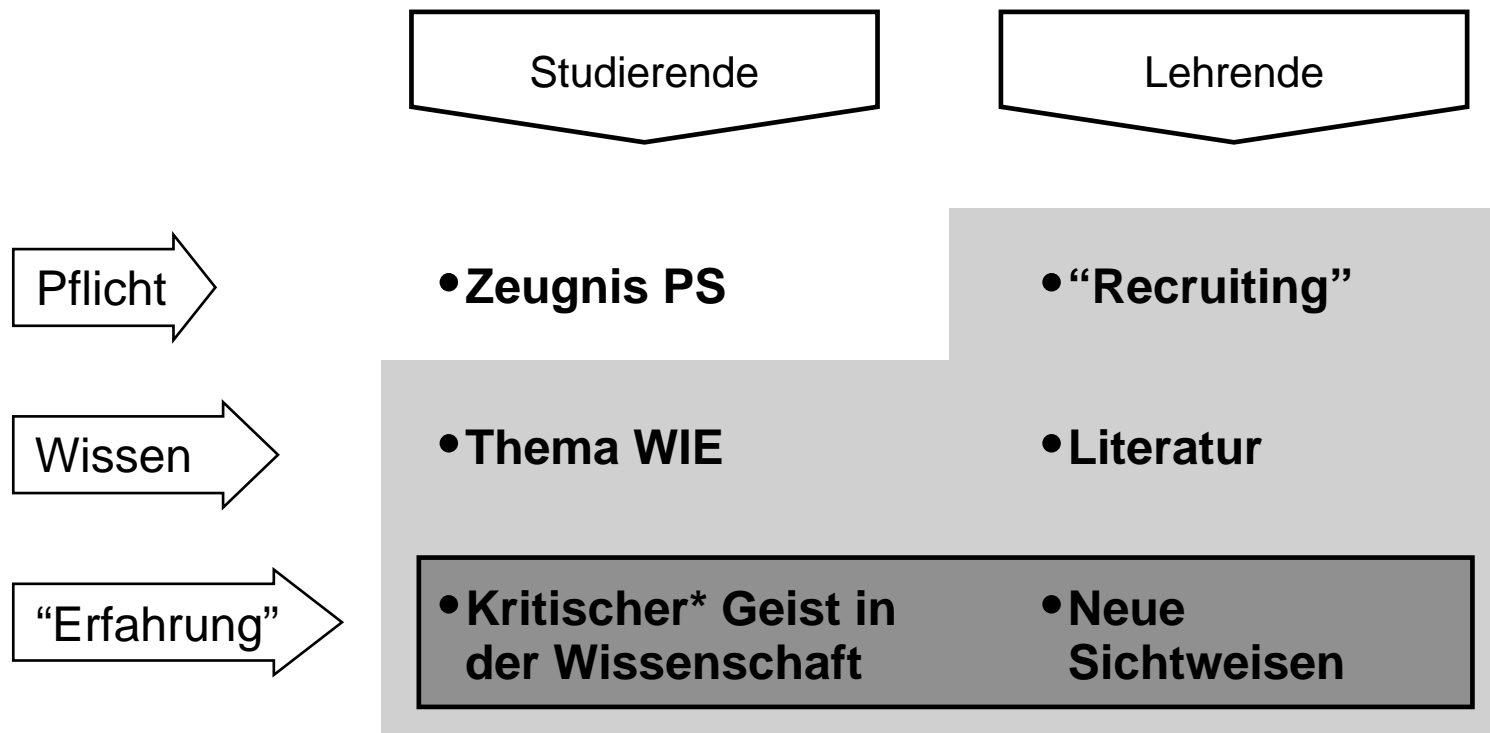
- Pflichterfüllung PS laut Studienplan
- Einführung bzw. Vertiefung der speziellen Thematik
- Einführung in das wissenschaftliche Arbeiten

- Ansporn für Auseinandersetzung mit Literatur
- Neue Sichtweisen der Studenten
- Begeisterung der Studenten für Thema

# PS WIE: Gegenüberstellung der Motivationen

DATE: 7.3.2005

- Haupthebel
- Mögl. Resultat



\* kritisch = hinterfragend := positiv (!)

Source: [www.dbai.tuwien.ac.at/education/wie](http://www.dbai.tuwien.ac.at/education/wie)

## PS WIE: Organisatorisches (1/2) – Ablauf

DATE: 7.3.2005

- Vorbereitung
  - Je 2 Studierende (=Team) wählen ein Paper-Paar
  - Jeder der beiden Studierenden bearbeitet zuerst ein Paper alleine
  
- Zwischenbesprechung
  - Bis zur Zwischenbesprechung überlegt sich jeder Studierende die Hauptaussage seines Papers -> Zeichnen der Hauptaussage auf A4 im Querformat
  - In der Zwischenbesprechung wird die Hauptaussage jedes Papers von den Studierenden kurz vorgestellt (je max. 30 sec) und anschließend gemeinsam das Paper zu anderen in Beziehung gesetzt
  
- Berichte
  - Verfassen von 3 Berichten
    - 2 individuelle Berichte der beiden Studierenden zu Ihren Papers
    - 1 gemeinsame, kritisch vergleichende Stellungnahme
  - Länge nicht entscheidend (max. 8 Seiten/Team), Inhalt zählt
  
- Vorträge
  - Max. 60 min/Team inkl. Diskussion
  - Ebenfalls 3 Vorträge analog zu Berichten
  - Anschließend Feedback der Zuhörer

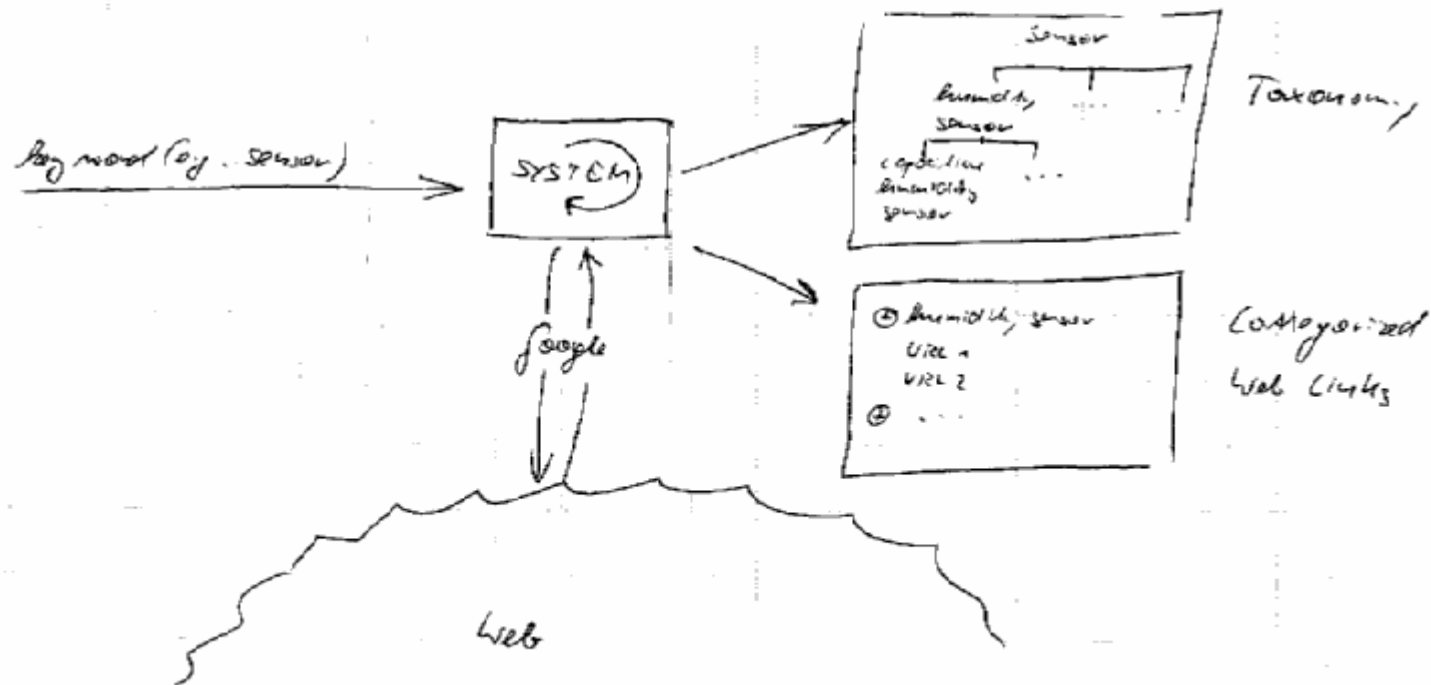
## PS WIE: Organisatorisches (2/2)

DATE: 7.3.2005

- Beurteilung
  - 30%: Mitarbeit (inkl. Intermediate Hand-in)
  - 40%: Präsentation (und anschließende Diskussion)
  - 30%: Bericht
  
- Abgabeformate:
  - Zwischenbericht:
    - A4 im Querformat, Medium beliebig (z.B. Zeichnung -> Sekretariat, Fax, Scan + Email, PowerPoint, ...)
    - Seite enthält Titel der Publikation & Name des Studierenden
  - Vortrag:
    - PPT bzw. OS Äquivalent
  - Bericht:
    - DOC bzw. OS Äquivalent
  
- Anwesenheitspflicht, außer Einheit im Juni
- Sprache: schriftlich (Berichte & Präsentationsfolien): Englisch; mündlich: nach Belieben
- Bericht und Vortrag werden am Ende des SS ins Web gestellt, es sei denn von jdm. explizit nicht gewünscht; Default: ja

aim 2 process

- To build a system that
  - automatically builds a taxonomy of terms and
  - accordingly classifies web sites



# PS WIE: Feedback Bogen

DATE: 7.3.2005

Evaluation	Speaker: _____	Date: _____
1. How do you evaluate the overall <b>presentation</b> regarding	<b>content/</b>	<b>delivery?</b>
	<small>very poor</small> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <small>very good</small>	<small>very poor</small> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <small>very good</small>
2. How do you evaluate the <b>speaker</b> regarding	<b>competence/</b>	<b>execution?</b>
	<small>very poor</small> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <small>very good</small>	<small>very poor</small> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <small>very good</small>
3. What should the speaker <b>maintain (+)</b> ? How could he <b>improve (!)</b> ? [e.g. <i>transition between slides, intonation and pronunciation, gesticulation, speed of the talk, motivation of audience to ask questions, fun factor, ...</i> ]	<hr/> <hr/> <hr/> <hr/> <hr/>	
4. How do you rate the <b>value of the presentation for your future</b>	<small>very poor</small> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <small>very good</small>	
5. What are the <b>main points</b> you have taken away from this talk?	<hr/> <hr/> <hr/> <hr/> <hr/>	

# PS WIE: Zeitplan SS 2005

VORLÄUFIG, DATE: 7.3.2005

- Anwesenheitspflicht
- Einreichtermin

<b>Wann</b>		<b>Was</b>	<b>Ort</b>	<b>To Do:</b>
Mo	7.3. 13:30 – 14:00	Vorbesprechung		
Mo	14.3. 15:00 – 15:30	Papers liegen auf bzw. stehen online	SE 184/2	
Fr	18.3. 23:59	Auswahl: Team + Paper-Paar		Email
Fr	8.4. 23:59	Intermediate Hand-in: Hauptaussage des Papers auf 1 Seite		1 Seite A4 quer
Di	12.4. Zwischen 14:00 und 17:00	Zwischenbesprechung, Dauer ca. 1 h	SE 184/2	
Fr	15.4. 23:59	Einreichen Bericht		Email DOC
Mo	18.4. 10:00	Berichte Studierende online		
Mi	20.4. 14:00 – 18:00	1.Vortragstag	SE 184/3	Vorträge (3 Teile / Team)
Do	21.4. 14:00 – 18:00	2.Vortragstag		
Fr	22.4. 09:00 – 12:00	3.Vortragstag		
Fr	22.4. 23:59	Einreichen Präsentation		Email PPT
	~Juni	Aktuelles Forschungsbeispiel am DBAI; Anschließend gemeinsames Mittagessen		[freiwillig]