

Informatik-Doppelstunde: Suchmaschinen

Lehrender:	NN
Mentor:	NN
Ort:	NN
Datum:	27.04.2007
Zeit:	7.45 – 9.20 Uhr
Reihenthema:	Anwendungen und Funktionsweise des Internets
Stundenthema:	Suchmaschinen

1 Betroffene Entscheidungen

1.1 Thematischer Zusammenhang

Nachdem Grundlagen des Internets in den ersten zwei Wochen betrachtet wurden, geht es in der dritten Woche um die Funktionsweise eines wichtigen Werkzeuges für die Internetanwendung World Wide Web. Suchmaschinen wenden das Client-Server-Prinzip an. Die Sammlung der notwendigen Information, die später einem Anwender zur Verfügung gestellt wird, erfolgt mit dem Hypertext Transfer Protocol. Durch die vorangehenden Inhalte kann so ein vertieftes Verständnis der Funktionsweise erreicht werden. Die Kenntnis der Funktionsweise einer Suchmaschine ist Voraussetzung für den systematischen Umgang mit diesem Werkzeug. Es ist jedoch notwendig, auch andere Strategien zur Informationsbeschaffung zu kennen.

Mit dem Wissen über das Domain Name System ergibt sich ein sinnvoller Anknüpfungspunkt für die Bewertung von Webseiten. Gerade in diesem Kontext kann die Information, die im Domainname steckt verwendet werden. In dieser Woche sollen folgende Lernziele thematisiert werden:

Die Schülerinnen und Schüler

- S1 kennen Möglichkeiten und Grenzen bei der Webrecherche mit Suchmaschinen.
- S2 können eine Webrecherche systematisch planen und durchführen.
- S3 können Webseiten bewerten.

1.2 Ziele der Doppelstunde

Die Suchergebnisse einer Webrecherche mit einer Suchmaschine ist abhängig von der verwendeten Datenbasis, der durch die Suchmaschine angelegten Datenbank und den Regeln zur Platzierung einzelner Ergebnisse. Um diese Zusammenhänge verstehen zu können, müssen grundlegende Eigenschaften von Suchmaschinen betrachtet werden. Die Teillernziele für diese Stunde sind daher folgende:

Die Schülerinnen und Schüler

- L1 kennen die Bestandteile einer Suchmaschine Webroboter, Datenbank (Index) und Benutzungsschnittstelle.
- L2 können den Ablauf zur Bearbeitung einer Suchanfrage beschreiben.

1.3 Hausaufgabe

„Informieren Sie sich über den Webroboter, Indexierung von Suchanfragen und Webseiten und Rangierung der Suchergebnisse in der Hilfe zur Wikipedia-Volltextsuche!“

1.4 Geplanter Unterrichtsverlauf

Zeit	Dauer	Phase	Teilziel	Inhalt	Methode	Medien
------	-------	-------	----------	--------	---------	--------

1. Stunde						
7.45	5'			Hausaufgaben: - Gibt es offene Fragen? - Besprechung Teilaufgabe 1c	UG	
7.50	10'	Zielorientierung		- „Wie suchen Sie Material für Referate?“ - Aufzeigen, dass Bücher aus der Bibliothek nicht mit einer Websuchmaschine gefunden werden. (OPAC – Online Public Access Catalog)	UG	
8.00	5'	Erarbeitung		- Analyse eines Aktivitätsdiagramms zum Erstellen eines Stichwortregisters (AB7, Aufgabe 1)	GA	AB
8.05	5'			- Besprechung der Aufgabe und Notieren einer Legende zu Aktivitätsdiagramm (TB)	UG	TA
8.10	10'			Analogie zwischen dem Buchtext und einer Datenbasis, die durch einen Webroboter durchsucht wird, und dem Stichwortregister und einer Datenbank (AB7, Aufgabe 2): - Erarbeitung eines Aktivitätsdiagramms zum Ablauf einer Suchmaschine: <ul style="list-style-type: none"> ○ Identifikation der Objekte ○ Identifikation der Aktivitäten 		

8.20	10'		L1, L2	<p>Besprechung des Ablaufs einer Suchanfrage und Besprechung der Skizze – Zeit zum Abschreiben lassen – unter Berücksichtigung der Aspekte:</p> <ul style="list-style-type: none"> - zeitliche Abhängigkeiten (Erstellung des Indexes unabhängig von Suchanfrage) - Dokumente Auffinden durch Webroboter (Start URL's und weiterverfolgen der Links) → Frage vom Beginn der Stunde beantworten - Indizierung: Stoppwortelimination, Stemming - Rangierungsprinzipien <p>Erklärung Arbeitsauftrag für nächste Stunde</p> <p>Raumwechsel</p>	UG	OHP
------	-----	--	--------	--	----	-----

2. Stunde						
8.35	10'	Ergebnissicherung		- Analyse des Ablaufs von der Untersuchung des Suchraumes bis zum Suchergebnis mit Soekia (AB7, Aufgabe 3)	EA	AB, CO
8.45	20'			<p>Aufgabe zur Erstellung des Indexes und zur Rangierung in Gruppen (Aufgaben 4 u. 5)</p> <p>Die Gruppen, die schon früher fertig sind, bekommen den Auftrag, das Quiz auf der Webseite</p> <p>http://www.swisseduc.ch/informatik/grundlagen_internet/internet_recherche/test/ zu bearbeiten</p>	GA	AB, CO
9.05	15'			Besprechung der Aufgaben und Präsentation der Gruppenergebnisse (Es muss auch deutlich gemacht werden, wie die Lernenden zu den Ergebnissen gekommen sind!)	SV	CP
<i>voraussichtliches Ende der Unterrichtsstunde</i>						
	10'			Arbeitsauftrag: „Informieren Sie sich über den Webroboter, Indexierung von Suchanfragen und Webseiten und Rangierung der Suchergebnisse in der Hilfe zur Wikipedia-Volltextsuche!“	EA	

Methoden:

DE: Demonstrationsexperiment

SE: Schülerexperiment

EA: Einzelarbeit

SV: Schülervortrag

GA: Gruppenarbeit

UG: Unterrichtsgespräch

LV: Lehrervortrag

Medien:

AB: Arbeitsblatt

OHP: Overhead-Projektor

CO: Computer

TA: Tafel

CP: Computerprojektion

TB: Tafelbild

FO: Folien

2 Begründung zentraler didaktischer Entscheidungen

2.1 Lerngruppe

Der Kurs besteht aus 3 Schülerinnen und 21 Schülern. In diesem Schuljahr hat der Informatikkurs begonnen. Bisher wurde objektorientierte Programmierung mit Java und dem Konzept zu „Stiften und Mäusen“ und BlueJ im Unterricht behandelt. Dazu wurden bereits Klassendiagramme mit verschiedenen Beziehungen behandelt. Außerdem wurde eine Unterrichtsreihe zum Thema endliche Automaten durchgeführt. In einer Stunde wurden zudem Kenntnisse zum Binärsystem wiederholt.

2.2 Begründung der Inhalte

„Die technologische Plattform für weltweit verfügbares Wissen ist vorhanden: das WWW [...] wird zunehmend auch als Plattform für die Praxis des weltweiten Wissensangebotes genutzt. Das WWW entwickelt sich in diesem Sinne zur Basis der Globalisierung von Wissen, zur universellen, weltweit verfügbaren Bibliothek. Das ist vorteilhaft, aber produziert zwei Hauptprobleme: erstens das Problem der Geltung des zur Verfügung gestellten Wissens, das sich insbesondere in der Zurechenbarkeit zu einem Autor bzw. zu einer Autorität äußert, sowie zweitens das Problem der Zuordnung von vorhandenem Wissen zu einer Weiterbildungsaufgabe, zu einem Problem am Arbeitsplatz, das sich insbesondere als Suchproblem mit lost-in-hyperspace Effekten äußert“ [Meder, 2000, S. 401].

In der Fülle der Daten, die im Internet zur Verfügung stehen, kann es schwierig sein die Daten zu finden, die gerade benötigt werden. Mit Suchmaschinen und Katalogen stehen zwei wichtige Werkzeuge zur Verfügung. Für eine effektive Nutzung dieser Werkzeuge ist ein grundlegendes Verständnis der Funktionsweise erforderlich. Ein Beispiel dafür ist die Suchmaschine, die nach Wörtern sucht, ohne diesen eine Semantik zuzuordnen. Zu dem Stichwort „Golf“ erhält man daher Suchergebnisse präsentiert, die eine unterschiedliche Wortbedeutung zugrunde legen (Autotyp, Sportart oder Meeresbucht). Zudem gibt es Strategien von Webseitengestaltern, den Algorithmus für die Bewertung einer Seite zu einem Suchbegriff geschickt auszunutzen, und damit die Ergebnisliste einer Suchmaschine zu beeinflussen. Es ist zudem notwendig, zu wissen, welche Daten überhaupt von Suchmaschinen erfasst werden. Dazu muss die Funktionsweise eines Webroboters berücksichtigt werden. Der Aufbau eines Suchindex bestimmt, welche Begriffe für eine Suchanfrage genutzt werden können.

2.3 Begründung des Lernweges

Zum Einstieg wird ein Demonstrationsexperiment durchgeführt. Eine Suche nach dem Wort „Suche“ wird mit einem Werkzeug für PDF-Dateien in der digitalen Fassung des Buches von Hartmann et al. [Hartmann et al., 2000], mit einer Websuchmaschine (z.B. www.google.de) und mit der Volltextsuche einer Website (z.B. de.wikipedia.org) durchgeführt. Daran wird deutlich, dass die Suchzeit nicht alleine von der Datenmenge abhängig ist (Google liefert am schnellsten eine Ergebnisliste). Die Frage, die zu beantworten ist, ist also, wie eine Ergebnisliste durch eine Suchmaschine erstellt wird.

Eine Erklärung wird durch das Beispiel des Stichwortregisters in einem Buch aufgezeigt. An diesem Beispiel wird deutlich, was die Datenbasis ist, die durch den Webroboter durchsucht wird und was der Suchindex ist, der eine Vorbereitung zur Beantwortung von Suchanfragen darstellt. Mit der Suchmaschine Soekia untersuchen die Lernenden von diesem Beispiel ausgehend, welche Arbeitsschritte notwendig sind, um einen Suchindex zu erstellen und was bei einer Suchanfrage berücksichtigt wird. Im daran anschließenden Unterrichtsgespräch werden

die Beobachtungen von der Suchmaschine Soekia abstrahiert und in einer Darstellung zur Funktionsweise einer Suchmaschine festgehalten.

Zur Vertiefung beschäftigen sich die Schülerinnen und Schüler entweder mit der Indexierung oder mit der Platzierung von Suchergebnissen.

Literatur

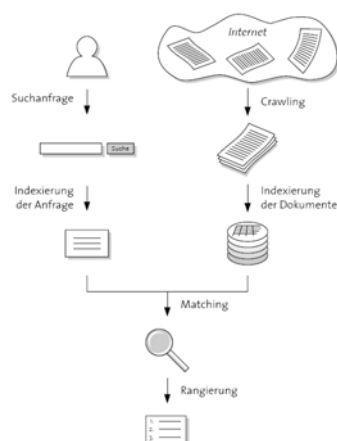
- [Hartmann et al., 2000] Hartmann, W.; Näf, M.; Schäuble, P.: Informationsbeschaffung im Internet. Orell Füssli Verlag, Zürich, 2000, URL: <http://www.internet-kompetenz.ch/infosuche/buch/> (5.10.2006), S. 59-75.
- [Hitz/Kappel, 2003] Hitz, M.; Kappel, G.: UML @ Work. Von der Analyse zur Realisierung. dpunkt.verlag, Heidelberg, 2003.
- [Meder, 2000] Meder, N.: Didaktische Ontologien. In: Ohly, H. P.; Rahmstorf, G.; Sigel, A.: Globalisierung und Wissensorganisation. Proceedings der 6. Tagung der Deutschen Sektion der Internationalen Gesellschaft für Wissensorganisation, S. 401-416. URL: http://cweb.uni-bielefeld.de/agbi/cgi-bin-noauth/cache/VAL_BLOB/167/167/63/did.pdf (06.03.2007).
- [Soekia] – Ein Blick hinter die Kulissen von Suchmaschinen. URL: <http://www.swisseduc.ch/informatik/soekia/> (5.10.2006).

Anhang A Aktivitätsdiagramm zum Ablauf der Verarbeitung von Suchanfragen durch eine Suchmaschine

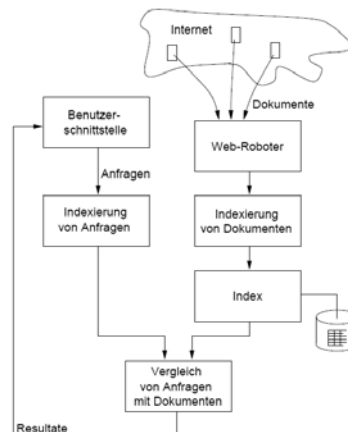
Hartmann et al. beschreiben den Ablauf zur Datengewinnung einer Suchmaschine und zur Verarbeitung einer Suchanfrage, wie rechts dargestellt. Die Bezeichner der Grafikelemente bezeichnen dabei Objekte oder auch Aktionen. Außerdem wird die Rangierung in dieser Darstellung nicht berücksichtigt. Im Begleitmaterial zu Soekia wird der Ablauf durch die linke Abbildung beschrieben. In dieser Darstellung wird der Ablauf genauer beschrieben. Durch die Pfeile werden die Aktivitäten beschrieben. Die Grafiken, die die Ergebnisse der einzelnen Schritte repräsentieren sind jedoch nicht eindeutig. Daher habe ich die formale Darstellung mit einem Aktivitätsdiagramm gewählt. Hitz und Kappel führen als Einsatzzwecke unter anderem einen Anwendungsfall und das Zusammenspiel mehrerer Anwendungsfälle an.

„Aktivitätsdiagramme werden in UML mit einer ähnlichen Intention wie Datenflussdiagramme in OMT [Object-Modeling Technique, SF] eingesetzt. Aktivitätsdiagramme ermöglichen die Beschreibung eines Ablaufs, wobei spezifiziert werden kann,

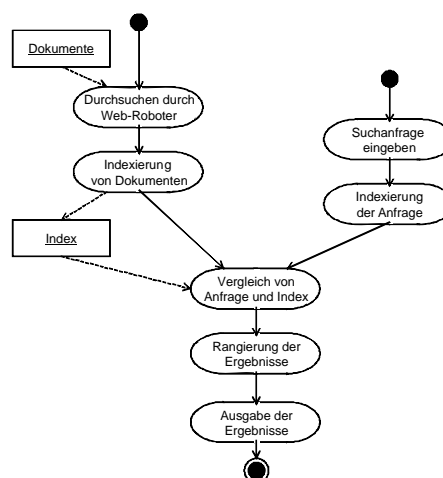
- was die einzelnen Schritte des Ablaufs tun (durch Angabe einer Bezeichnung, eines Schritts und der durch ihn manipulierten Objekte),
- in welcher Reihenfolge sie ausgeführt werden und
- wer für sie verantwortlich zeichnet (optional)“ [Hitz/Kappel, 2003, S. 160].



Quelle: [Soekia]



Quelle: [Hartmann et al., 2000, S. 22]



Gestaltung: S. Freischlad