

WAIS

Mario Nenno, 28.1.93, WAIS.TXT

Was bedeutet "WAIS"?

WAIS = **Wide Area Information Servers**, sprich: "ways"

Was ist WAIS?

Eine Netzwerk-Anwendung, bei der man **verteilte Datenbanken** im Netz über ein Protokoll abfragen kann. Der Inhalt kann als Text oder als Binär-Dateien vorliegen (Bilder, Sound usw.).

WAIS ist ein Projekt von Brewster Kahle bei Thinking Machines Corp. in Zusammenarbeit mit Apple Computer und Dow Jones. Der Zweck war die Entwicklung eines Retrieval-Systems das für Leute die nicht professionell mit Datenbanken arbeiten einfach und so natürlich wie möglich zu bedienen ist.

WAIS hat drei wesentliche Bestandteile: **Client** mit Userinterface, **Server** mit Indizierungsprogramm und Abfragebearbeitung und dem **Protokoll** (Z39.50) zwischen Client und Server.

Nach der ersten Entwicklung wurden von Thinking Machines Beispiel-Clients und -Server geschrieben die dann 1991 veröffentlicht wurden. Dadurch war es auch Anderen möglich entsprechende Programme zu schreiben.

Da WAIS im **Client-Server** Modell arbeitet, kann man mit einem beliebigen WAIS-Client einen/mehrere WAIS-Server abfragen.

Mittlerweile gibt es eine Vielzahl von Client-Programme für verschiedenste Rechner-Typen vom PC mit DOS, Macintosh, NeXT, UNIX und VMS, wodurch es sich momentan eine sehr schnell Verbreitung findet.

Eine typische WAIS-Abfrage verläuft folgendermaßen:

Der Anwender startet lokal seinen WAIS-Client und wählt eine Datenbank aus (bei WAIS "Source" genannt) und gibt seine Suchabfrage ein. Der WAIS-Client verbindet sich automatisch mit dem Rechner/WAIS-Server der ausgewählten Datenbank, um an ihn die Abfrage weiterzuleiten.

Nach kurzer Suchzeit schickt der Server zunächst eine Liste von gefundenen Datensätzen an den Client zurück. Dabei hat sich der Client um die Darstellung der Liste zu kümmern. In dieser Liste wird für jeden gefundenen Datensatz eine Überschrift und eine Gewichtung (1000 bis 1) angegeben und der Anwender entscheidet anhand der Liste und der Gewichtung einzelner Einträge welche er kopieren möchte.

Der Client holt den ausgewählten Eintrag vom Server, zeigt den Inhalt an und gibt dem Anwender die Möglichkeit den Inhalt lokal abzuspeichern.

Wichtige Begriffe

Directory-of-Servers

Auf dem Rechner think.com wird eine Liste aller 'offiziellen' weltweit WAIS-Datenbanken geführt. Diese Liste wird Directory-of-Servers genannt.

Source

Der Begriff hat zwei Bedeutungen. (1) Source ist ein Synonym für WAIS-Datenbank. (2) Source ist eine Text-Datei die alle wesentlichen Angaben einer Wais-Datenbank enthält, z.B. Rechneradresse, Portnummer, Datenbank-Name und Beschreibung der Datenbank.

Relevance feedback

Er erlaubt Suchergebnisse, die man zuvor auf eine Abfrage erhalten hat, weiter einzuschränken. Dabei werden alle Wörter der Suchergebnisse als neue Suchabfrage verwendet.

Dieser Mechanismus, ist noch nicht bei allen WAIS-Clients verfügbar.

Vorteile von WAIS?

- Einfaches Abfragen durch Eingabe von Suchwörtern
- Keine spezielle Retrieval(Abfrage)-Sprache, einfaches Englisch
- Abfragen mehrerer Datenbanken
- Durch Gewichtung werden auch "ähnliche" Begriffe gefunden
- Suchergebnisse lassen sich als Abfrage für weitere Abfragen verwenden
- Kostenlos
- besonders geeignet für Freitext-Suche

Nachteile von WAIS?

- offizielle Version hat keine logischen Operatoren (AND, OR)
- noch in der Entwicklung (Zukunft ?)

Welche Datenbanken sind erreichbar?

Derzeit sind etwa 350 Datenbanken oder "Sources" abrufbar. Die Datenbanken enthalten z.B. Informationen über das Internet, News-Archive, FAQ-Listen, CIA-Factbook, Liedtexte, Gedichte, Bibel und Koran-Texte, Inhaltsverzeichnisse von Zeitschriften, DNA-/Protein-Sequenzen, Satelliten-Bildern und vieles mehr.

Wais zum Ausprobieren:

Zum Testen kann man sich per Telnet bei **quake.think.com** unter der Login: **wais** einloggen und dort einen sehr einfachen WAIS-Client (swais) ausprobieren. Die Geschwindigkeit ist im Vergleich zu einem lokalen Client sehr langsam.

WAIS in Gopher

Von der Möglichkeit in Gopher direkt einen WAIS Client einzubauen wird bereits häufig Gebrauch gemacht. Der Nachteil gegenüber einem normalen WAIS-Client ist allerdings man kann nur eine Datenbank auf einmal abfragen, den 'relevant feedback'-Mechanismus nicht nutzen, die Fragen können nicht abgespeichert werden und man bekommt die Gewichtung meist nicht angezeigt.

Wo bekommt man WAIS?

Per anonymous ftp für folgende Systeme:

UNIX: think.com, cd wais, wais-8-b5.1.tar.Z

Mac: think.com, cd wais, WAISation-0-63.sit.hqx

PC/DOS: ftp.oit.unc.edu, cd pub/wais/UNC/DOS, pcdist.zip

Weitere Information zu WAIS:

- /wais/bibliography.txt und andere *.txt@think.com, anonymous ftp
- directory-of-servers auf think.com ist eine Liste der weltweiten Datenbanken mit einer kurzen Beschreibung, /wais/wais-sources.tar.Z@think.com, anonymous ftp
- News: comp.infosystems.wais

Erweiterung

Don Gilbert von der University of Indiana hat WAIS modifiziert, so daß es möglich ist auch Abfragen mit **logischen Verknüpfungen** zu erstellen. (/util/wais/iubio-wais-8b5.tar.Z@ftp.bio.indiana.edu). Diese Möglichkeit wird bisher nur von wenigen WAIS-Servern unterstützt (z.B. IUBIO-xxx WAIS-Datenbanken).

Literatur:

comp.infosystems.wais Frequently Asked Questions [FAQ], 29 Dec 1992

Lehvaslaiho, H., Harper, R. (1992) BioBit 21, pub/bio/biobit/biobit.arc@ftp.uni-kl.de

Marshall, Peter. (1992) WAIS: The Wide Area Information Server or Anonymous What???. University of Western Ontario pub/internet/wais-paper.ps.Z@ftp.uni-kl.de

Ward, Lynn (1992) Shopping for Information on the Internet with WAIS (INTERNET TREASURES). University of Illinois Publication Newsletter UIUCnet. doc/net/uiucnet/vol5no6.txt@ftp.cso.uiuc.edu