

### A

#### **Abstract**

Ein Abstract ist ein Kurzreferat, das kurz und klar den Inhalt eines Dokuments wiedergibt. Zusätzlich ist es ein Verfahren zur Inhaltserschließung. Es ist also eine Reduktion eines Textes auf die inhaltskennzeichnenden Inhalte. Die Suche während des IR-Prozesses erfolgt über den Abstract, was eine schnellere Suche zur Folge hat. Die Erstellung eines Abstracts erfolgt entweder manuell/intellektuell durch einen Experten oder den Autor oder automatisch durch ein computergestütztes System.

#### **Affix removal stemmer**

Ist eine Stemming-Methode. Affix removal stemmer arbeiten regelbasiert, dabei werden Präfixe und/oder Suffixe in einer bestimmten Reihenfolge entfernt. So wird auf den Wortstamm hingearbeitet.

#### **Anfragesprache**

Art und Weise, wie der Benutzer seine Anfrage in einer dem System verständlichen Form formuliert (z.B. formal oder natürlichsprachlich).

#### **Aggregation**

Eine Aggregation ist in der Informatik eine Verbindung zwischen Daten oder Objekten. Anfragen an ein Datenbankmanagementsystem geben meist Mengen von Datensätzen zurück. Jedoch können durch Aggregatfunktionen auch zusammengefasste Informationen ermittelt werden. Ein einfaches Beispiel ist die Errechnung eines Durchschnittsalters aus einer Personentabelle.

#### **Attribut**

Ein Attribut ist die Eigenschaft einer Entität oder einer Relation, z.B. der Titel einer Vorlesung. In einem Entitätstyp existiert immer ein Primärschlüssel, ein Attribut (oder eine Kombination von Attributen), das die einzelnen Entitäten eindeutig identifiziert, z.B. die Matrikelnummer eines Studierenden.

### B

#### **Benutzeroberfläche**

Grafisch-direktmanipulative Benutzungsoberflächen zeichnen sich vor allem dadurch aus, dass durch direkte Manipulation (z.B. Mausklicks) veränderbar sind und ein sofortiges Feedback an den Benutzer ermöglichen (z.B. Bestätigung oder Nachfrage bei der Durchführung eines Arbeitsschrittes wie Dateien löschen). Die Elemente einer BOF z. B. folgende: Menüs, Schaltflächen, Fenster-Technik, Formulare, Icons, Buttons.

### **Boole'sche Retrieval**

Das Boole'sche Retrieval ist ein klassisches Verfahren des Information Retrieval und basiert auf dem Exact Match Paradigma. Es prüft, ob eine wohldefinierte Bedingung erfüllt ist, oder nicht, basierend auf der Boole'schen Algebra. Die Grundidee ist es, Mengenoperationen auf Mengen von Dokumenten oder Objekten anzuwenden, die durch Attributwerte charakterisiert sind. Dabei werden die Terme der Anfragen mit Operatoren verknüpft, übliche Operatoren sind u.a.: AND OR und NOT.

## **C**

### **Clusterin**

Clustering beschreibt eine Art der Klassifizierung von thematisch ähnlichen Objekten und ist ein Verfahren zur Inhaltserschließung. Dokumente, die einem Cluster zugeordnet werden sind sich also ähnlich. Man kann entweder viele Dokumente in wenig Clustern oder wenig Dokumente in vielen Clustern zusammenfassen. Die einzelnen Cluster können sich auch überschneiden, das bedeutet, dass bestimmte Dokumente aufgrund ihres Inhalts zu mehreren Clustern gehören können.

### **Cross Language Information Retrieval (CLIR)**

CLIR ist ein Akronym für Cross Language Information Retrieval und bedeutet, dass über die Sprachgrenzen hinaus nach Informationen gesucht werden kann. So kann beispielsweise eine Suchanfrage auf Englisch in verschiedene Sprachen übersetzt werden. Somit werden dann nicht nur Dokumente auf Englisch gefunden, sondern auch auf anderen Sprachen, die die Antworten für die Suchanfrage liefern.

## **D**

### **Data Control Language (DCL)**

Die Data Control Language, dient der Steuerung von Zugriffsrechten auf die Datenbank. Mit der DCL können somit Privilegien an Benutzer oder auch auf Objekte gezielt gesteuert werden.

### **Data Definition Language (DDL)**

Die Data Definition Language (Datendefinitionssprache) dient zur Einrichtung, Änderung oder Löschung von Datenbankobjekten wie z. B. Tabellen, Indizes usw. Zu den DDL-Anweisungen gehören: CREATE, um Tabellen zu erstellen. ALTER, um die Tabellenstruktur zu ändern, z.B. Hinzufügen einer Spalte. DROP, um Spalten oder auch Tabellen zu löschen.

## **Data Manipulation Language (DML)**

Mit der Data Manipulation Language (Datenmanipulationssprache) können Daten selektiert, geändert und gelöscht werden. Zu den wichtigsten Anweisungen gehören: INSERT UPDATE DELETE

### **Datenbank**

Unter einer Datenbank versteht man eine auf Dauer für flexiblen und sicheren Gebrauch ausgelegte Datenorganisation, die sowohl eine Datenbasis als auch die zugehörige Datenverwaltung umfasst.

### **Datenmodellierung**

Unter Datenmodellierung wird bei Datenbanken das Erstellen des Datenbankschemas verstanden. Das Datenbankschema legt fest, welche Daten in der Datenbank gespeichert werden und welche Beziehungen die Daten zueinander haben. Es gibt drei Arten der Datenmodellierung: Konzeptuelles Datenbankschema (semantisches Datenmodell), logisches Datenbankschema und physisches Datenbankschema.

### **DELETE (SQL-Befehl)**

Mit dem SQL-Befehl DELETE kann man Einträge aus der Datenbank löschen. Mittels DELETE kann man also beliebige Einträge aus einer Tabelle löschen, doch vorsicht: Wenn man die WHERE-Bedingung nicht setzt, werden alle Einträge innerhalb der Tabelle gelöscht.

### **Deskriptoren**

Eine einfache Art der inhaltlichen Beschreibung von (Text-)Dokumenten ist die mit Hilfe von Stich- und Schlagwörtern. Diese werden Deskriptoren genannt, wenn sie im Rahmen eines → Thesaurus (=kontrolliertes Vokabular) zur inhaltlichen Beschreibung vorgesehen werden (= Vorzugsbenennungen). In einem solchen Thesaurus gibt es auch sogenannte "Nichtdeskriptoren", die mit Deskriptoren in einer inhaltlichen Beziehung stehen, die aber aus Gründen der Einheitlichkeit der Indexierung nicht verwendet werden sollen und stattdessen auf einen an ihrer Stelle zu verwendenden Deskriptor verweisen. Deskriptoren werden entweder von einem Indexierer (intellektuelle Indexierung) oder von automatischen oder computergestützten Indexierungssystemen vergeben.

## E

### **Effektivität**

Effektivität bezeichnet die Unterstützung, die das System dem Benutzer bei der Lösung seines Informationsproblems bietet.

### **Effizienz**

Effizienz bezeichnet die Nutzung von Systemressourcen (CPU-Zeit, Ein-/Ausgabe-Operationen, Speicherplatz, Antwortzeit) bei der Bewältigung einer Aufgabe.

### **Elementarparameter**

Die Elementarparameter  $a$ ,  $b$ ,  $c$ ,  $d$  enthalten die Anzahl der Dokumente, die sich in den einzelnen Teilmengen befinden, und dienen zur Berechnung von recall und precision.

### **E-Maß**

Um eine einzige Maßzahl zu gewinnen, mit der IR-Systeme verglichen werden können, müssen die beiden Standardmaße recall und precision miteinander verknüpft werden. Dies geschieht mit Hilfe des e-Maßes.

### **Entität**

Als Entität (auch Informationsobjekt) wird in der Datenmodellierung ein eindeutig zu bestimmendes Objekt bezeichnet, über das Informationen gespeichert oder verarbeitet werden sollen.

### **Entity-Relationship-Modell (ERM)**

ERM ist ein Akronym für das Entity-Relationship-Modell. Dieses ist ein konzeptionelles Datenmodell zur Erstellung eines Datenbauplans für ein Informationssystem.

## F

### **Fachinformationen**

Fachinformationen sind aufbereitete und zur Verfügung gehaltene Informationen über verschiedene Wissensbereiche, speziell Fachwissen.

### **Fakteninformationssystem (FIS)**

Die Basis von Fakteninformationssystemen bilden Datenbanken. Um Informationen in Datenbanken verwalten zu können, muss zunächst ein 'Datenbauplan' als Implementierungsgrundlage erstellt werden. Dies geschieht mit den Instrumenten der Datenmodellierung. Der Zugriff auf die Daten einer Datenbank erfolgt mit der Datenbanksprache SQL.

### **Fremdschlüssel**

Ein Fremdschlüssel ist ein Attribut oder eine Attributkombination einer Relation, welches auf einen Primärschlüssel (bzw. Schlüsselkandidaten) einer anderen oder der gleichen Relation verweist.

### **Funktion**

Mit SQL Funktionen können bestimmte Werte einer Spalte berechnet oder bestimmte Werte mittels der SQL Funktionen im Nachhinein manipuliert werden. Es existieren zwei Arten von SQL Funktionen: SQL Aggregat Funktionen und SQL Skalar Funktionen.

## G

### **Get (HTTP Request)**

Unter einem HTTP-Request versteht man die Anfrage eines HTTP-Clients (z.B. Browser) an den HTTP-Server. Jeder Request wird durch die Angabe einer Methode eingeleitet, um dem Server zu sagen, was er mit dem Request machen soll. Die GET-Methode wird verwendet, um Informationen aus dem gegebenen Server mit Hilfe einem gegebenen URI abzurufen. Anfragen mit Hilfe von GET sollten nur Daten abrufen und sonst keine weiteren Auswirkungen auf die Daten haben.

### **Gewichtung im Information Retrieval**

Die Gewichtung im Information Retrieval berechnet die Gewichtung der Relevanz von Dokumenten im Exact Match Paradigm. Ein wichtiger Bestandteil der Gewichtung ist die IDF. Ein weiterer wichtiger Bestandteil der Gewichtung ist die normalized IDF. Die Harmonisierung von Termgewichten kann mit der Harmonic Sum durchgeführt werden.

### **Grundformreduction**

Für viele Anwendungen der Automatischen Sprachverarbeitung ist es hilfreich,

Wortformen, d.h. in verschiedenen Formen auftretende Varianten einer Grundform (z.B. Hauses, Häuser, Häusern zu Haus) zu deren Grundformen zu reduzieren: Dadurch werden Informationen, die über Wortformen vorliegen, an der Grundform akkumuliert, was zu mehr und so statistisch sichereren Daten führt.

### **GitHub**

Eine Website, auf der Quellcode in Git gehostet wird. GitHub ist die beliebteste Anlaufstelle, um Open-Source-Projekte zu teilen und gemeinsam an diesen zu arbeiten. Die Website kann auch genutzt werden, um private „Datenlager“ von Unternehmen zu hosten.

## **H**

### **Homonymie**

Homonymie bezeichnet die lautliche Übereinstimmung von Wörtern mit verschiedener Bedeutung. Zum Beispiel die „Bank“ als Geldinstitut und die „Bank“ als Sitzgelegenheit.

### **Host**

Zentraler Rechner, auf dem Datenbanken verwaltet werden und der Anwendungsprogramme zur Verfügung stellt.

### **HTML**

HTML steht für Hypertext Markup Language und bezeichnet eine Markup-Sprache, in der WWW-Dokumente erstellt werden. HTML gilt als Standard, der auf SGML aufbaut.

### **HTTP**

Die Abkürzung HTTP steht für Hypertext Transfer Protocol und bezeichnet ein zustandsloses Protokoll, mit dem sich Daten in einem IP-Netzwerk übertragen lassen. Die wichtigste Anwendung des Hypertext Transfer Protocols ist die Übertragung von Internetseiten und Daten zwischen einem Webserver und einem Webbrowser.

### **HTTPS**

HTTPS steht für Hypertext Transfer Protocol Secure und ist die verschlüsselte Protokollvariante von http. Mit Hilfe des https können Daten „abhörsicher“ zwischen Client und Server im World Wide Web übertragen werden.

## **Hyperlink (Netzwerk)**

Ein Hyperlink ist ein Querverweis in einem Hypertext, der zu einem anderen elektronischen Dokument verweist. Ein Klick auf einen Hyperlink ruft das Ziel des Hyperlinks direkt auf.

## **Hypermedia**

Hypertexte, die durch weitere Medien (Graphik, Bild, Video, Animation, Ton) ergänzt sind.

I

## **IaaS**

Infrastructure as a Service

## **Implementierung**

Eine Implementierung – auch Implementation genannt – ist das Implementieren oder das Implementiert werden, also die Realisierung oder Umsetzung von festgelegten Strukturen und Prozessabläufen in einem System unter Berücksichtigung von Rahmenbedingungen, Regeln und Zielvorgaben, im Sinne einer Spezifikation.

## **Indexierung**

Methoden und Verfahren der Informations-Erschließung im Information Retrieval. Inhalte von Dokumenten oder anderen Objekten werden durch eine Analyse und die Vergabe von Deskriptoren erschlossen.

## **Information**

Information ist definiert als Wissen in Aktion, d.h. Information ist die Teilmenge von Wissen, die zur Lösung eines Problems benötigt wird.

## **Informationsmanagement**

Mit dem Konstrukt Informationsmanagement wird also das Leitungshandeln (Management) in einer Betriebswirtschaft in Bezug auf Information und Kommunikation bezeichnet, folglich alle Führungsaufgaben, die sich mit Information und Kommunikation in der Betriebswirtschaft befassen.“  
[Heinrich01, S. 8]

## **Inhaltserschließung**

Inhaltserschließung bezeichnet innerhalb der Bibliotheks- und Dokumentationswissenschaft die Erschließung bibliographischer und archivalischer Ressourcen nach inhaltlichen Kriterien. Das bedeutet, dass eine

Ressource intellektuell oder automatisch aufgrund ihres Inhalts beschrieben wird.

### **INSERT (SQL-Befehl)**

Der SQL-Befehl INSERT ermöglicht es ein oder mehrere Datensätze in eine bestehende Tabelle einzufügen.

### **Instanz**

Eine Instanz ist eine konkrete Realisierung eines Modells. Beispiel: In der Programmierung ist ein Objekt eine Instanz einer Klasse.

### **Information Retrieval**

Information Retrieval (IR) ist in erster Linie textuelle Analyse und Repräsentation von Texten. IR umfasst die Erschließung und das Wiederauffinden von Informationsobjekten.

### **IR-Systeme**

Informationssysteme, die Informationen erschließen, organisieren, speichern und wieder auffinden. 'An information retrieval system is an information system, that is, a system used to store items of informations that need to be processed, searched, retrieved, and disseminated to various user populations.' (Salton/McGill 1983, XI)

## **J**

### **Java**

Java ist eine objektorientierte Programmiersprache, die 1995 erschienen ist.

### **Javascript**

JavaScript ist eine Scriptsprache, die Benutzerinteraktion auf Webseiten ermöglicht. JavaScript wird neben HTML und CSS für die Erstellung von Webseiten verwendet.

### **Join (SQL-Befehl)**

Mit JOINS kann man zwei oder mehr Tabellen zusammenfügen, solange es eine Verbindung zwischen den Tabellen gibt. Es gibt verschiedene JOINS: LEFT JOIN, RIGHT JOIN und UNION.

## **K**

### **Kardinalität**

Kardinalitäten sind Mengenangaben, mit denen in der Datenmodellierung für Entity- Relationship-Diagramme (ER-Diagramme) für jeden Beziehungstyp



festgelegt wird, wie viele Entitäten eines Entitätstyps mit genau einer Entität des anderen am Beziehungstyp beteiligten Entitätstyps (und umgekehrt) in Beziehung stehen können oder müssen.

### **Keywords**

Mit Keywords kannst du Stichwörter für ein Suchprogramm bestimmen.

### **Kommentar**

HTML beinhaltet die Möglichkeit, an fast beliebigen Stellen innerhalb einer HTML-Datei Kommentare einzufügen. Diese werden vom Browser ignoriert, d. h. nicht angezeigt.

### **Kompositazerlegung**

Komposita werden verwendet, um komplexere Sachverhalte in einem einzigen Wort auszudrücken und tragen so wesentlich zur Effizienz der Sprache bei. Die Kompositazerlegung ist deshalb ein wichtiger Bestand der Verarbeitung von natürlichsprachlichen Texten.

### **Konflation**

Unter der Konflation versteht man Kürzen, Kombinieren oder auch Zusammenfügen von Wörtern (Buchstabenfolgen!) und/oder "Wortelementen", um unterschiedliche morphologische Varianten eines Wortes zu matchen. Die Konflation kann auf zweierlei Weisen vorgenommen werden: manuell, indem man die regulären Ausdrücke benutzt, oder automatisch, mit Hilfe von geeigneten Programmen, sogenannten Stemmern.

### **Konzeptionelles Datenmodell**

Das konzeptionelle Datenmodell bildet den zu betrachtenden Realitätsausschnitt implementierungsunabhängig ab, d.h. es wird nicht festgelegt, mit welcher Programmiersprache die Datenbank umgesetzt werden muss bzw. welche Datenbankart aus der Implementierung resultiert. Ein konzeptionelles Datenmodell, das sich als Standard etabliert hat, ist das Entity-Relationship-Modell.

## **L**

### **Leere Antwortmenge**

Das System liefert keine Dokumente als Ergebnis.

### **List**

Dem input-Element wird über das Attribut list ein datalist-Element zugeordnet.

### **Logisches Datenmodell**

Das logische Datenmodelle sind auf einen bestimmten Datenbanktyp, wie z.B.

objektorientiert, hin ausgerichtet. Das daraus resultierende Datenschema ist dementsprechend implementierungsabhängig. Ein Beispiel für ein logisches Datenmodell ist das relationale Datenmodell, mit dessen Hilfe eine relationale Datenbank definiert wird.

## M

### **Manuelle Indexierung**

Die Manuelle Indexierung, Intellektuelle Indexierung oder Verschlagwortung ist ein Verfahren der Sacherschließung von Dokumenten, bei der einem Dokument repräsentative Schlagwörter durch einen Indexierer zugewiesen werden. Da die freie Indexierung mit beliebigen Wörtern sehr ungenau ist, sollten Schlagwörter aus einem kontrolliertem Vokabular (z. B. der Schlagwortnormdatei) oder einer anderen Dokumentationssprache verwendet werden.

### **Method-Attribut**

Das method-Attribut legt die HTTP-Methode get oder post fest, nach der die Daten gesandt werden.

### **Modellierung**

Eine Modellierung ist das Entwickeln eines geeigneten Modells, also einer Abbildung der Realität. Dabei werden nur die Aspekte im Modell abgebildet, die für die Aufgabe relevant sind. Die Realität wird also vereinfacht dargestellt.

### **Morphologie**

In der Morphologie wird die Zusammensetzung von Wörtern analysiert.

### **MySQL**

SQL – Structured Query Language (englisch für strukturierte Abfragesprache) ist eine deklarative Datenbanksprache, um auf relationale Datenbanksysteme zuzugreifen.

## N

### **Navigation**

Eine Navigation ermöglicht das Navigieren (lat. für (An-)Steuern) zwischen mehreren Webseiten oder Kapiteln einer Webseite. Beispiele finden sich in Menüs, Inhaltsverzeichnissen und Linklisten.

### **N-Gram**

Eine Navigation ermöglicht das Navigieren (lat. für (An-)Steuern) zwischen mehreren Webseiten oder Kapiteln einer Webseite. Beispiele finden sich in Menüs, Inhaltsverzeichnissen und Linklisten.

## O

### **Object-Element**

Das object-Element ermöglicht es externe Dateien zu referenzieren, die dann als Grafik dargestellt oder von externen Plugins verarbeitet werden.

### **Objekt**

Exemplar von Personen, Gegenständen oder nichtmateriellen Dingen, über das Informationen gespeichert werden sollen.

### **Objektorientierte Programmierung**

Die Objektorientierte Programmierung ist ein Programmierparadigma. In der objektorientierten Programmierung werden alle Bestandteile des Quellcodes durch Objekte beschrieben. Dieses Konzept orientiert sich an der Realität, in der Objekte physisch vorstellbare Dinge sind (siehe auch Entität). Wichtige Konzepte in der OOP sind Objekte, Klassen, Vererbung und Polymorphie.

### **Objektyp**

Eine durch einen Objektyp-Namen eindeutig benannte Klasse von Objekten, über die Informationen gespeichert werden und die in prinzipiell gleicher Weise verarbeitet werden.

### **Open-HPI**

OenHPI ist eine webbasierte Plattform für E-Learning, auf der kostenlose und frei zugängliche Massive Open Online Courses zu Themen aus dem Bereich Informationstechnologie angeboten werden. openHPI wird vom Hasso-Plattner-Institut für Digital Engineering in Potsdam entwickelt und betrieben.

## P

### **Performance**

Das Wort Leistung (engl. performance /pə'fɔ:məns/) wird in der Informatik verwendet, um das Vermögen eines Datenverarbeitungssystems (kurz: DV System) zu beschreiben, Aufgaben allgemein (Funktionalität) oder auf bestimmte Weise (schnell, gleichzeitig, ununterbrochen usw.) auszuführen.

### **Petabyte**

Eine Mengeneinheit: 1024 Terabytes sind ein Petabyte (PB, PByte).

### **PHP**

PHP ist eine Skriptsprache mit einer an C bzw. C++ angelehnten Syntax, die hauptsächlich zur Erstellung von dynamischen Webseiten oder Webanwendungen verwendet wird. PHP ist Open-Source-Software.

### **Precision**

Quotient aus allen gefundenen relevanten Dokumenten und allen gefundenen

Dokumenten. Precision beschreibt die Genauigkeit eines Ergebnisses. Das F-Maß ist wichtig für die Berechnung eines kombinierten Maßes aus Recall und Precision.

### **Primärschlüssel**

Ein Primärschlüssel dient der eindeutigen Identifizierung der einzelnen Zeilen in einer Tabelle. Er kann entweder Teil des eigentlichen Datensatzes oder ein künstliches Feld sein (das mit dem Datensatz nichts zu tun hat). Ein Primärschlüssel kann aus einem oder mehreren Feldern einer Tabelle bestehen. Primärschlüssel, die aus mehreren Attributen bestehen, werden als zusammengesetzte Schlüssel bezeichnet.

### **Projektion**

Ein Fremdschlüssel ist ein Attribut oder eine Attributkombination einer Relation, welches auf einen Primärschlüssel (bzw. Schlüsselkandidaten) einer anderen oder der gleichen Relation verweist.

## **Q**

### **Quellcode/Quelltext**

Unter dem Quelltext, auch Quellcode (engl. source code) oder Programmcode, versteht man in der Informatik den für Menschen lesbaren, in einer Programmiersprache geschriebenen, Text eines Computerprogrammes. Abstrakt betrachtet kann man den Quelltext eines Computerprogramms auch als Software-Dokument bezeichnen, welches das Programm so formal exakt und vollständig beschreibt, dass dieses aus ihm vollständig automatisch vom Computer generiert werden kann.

### **Query**

Eine Anfrage (zum Beispiel an eine Datenbank), die häufig einen formalen Ausdruck erwartet, siehe Datenbanksprache, Abfragesprache.

## **R**

### **Ranking**

Abstufung nach Rang oder Bedeutung von Elementen.

### **Ranking-Systeme**

Ranking-Systeme ordnen die von ihnen selektierten Dokumente nach der Relevanz. Voraussetzung dafür ist eine gewichtete Indexierung.

### **Recall**

Quotient aus allen relevanten gefundenen Dokumenten und allen relevanten Dokumenten. Recall beschreibt die Vollständigkeit eines Ergebnisses. Das F-Maß ist wichtig für die Berechnung eines kombinierten Maßes aus Recall und

Precision.

### **Recall-Precision-Graph**

Graphische Darstellung der Beziehung von recall und precision bei ranking-Systemen. Dabei wird die Liste der nachgewiesenen Dokumente abgearbeitet, und nach jedem Dokument werden r und p auf der Basis aller bislang betrachteten Dokumente berechnet. Die Wertepaare werden in ein Koordinatensystem eingetragen.

### **Redundanz**

Die Redundanz (latein. redundare "im Überfluss vorhanden sein") bezeichnet allgemein das mehrfache Vorhandensein funktionsgleicher, inhaltsgleicher oder wesensgleicher Objekte.

### **Relation**

Eine Relation (Beziehung) verbindet zwei oder mehrere Entitäten wechselseitig miteinander.

### **Relationale Datenbank**

Eine relationale Datenbank verwendet Relationen oder zweidimensionale Tabellen zum Speichern von Informationen.

### **Relationenmodell**

In einem Relationalen Datenmodell werden die Daten in Tabellenform gespeichert, wobei jede Spalte der Tabelle ein Datenelement bzw. ein Attribut repräsentiert. Jedem Dateielement wird ein einfacher Datentyp zugeordnet. In jeder Tabellenzeile wird ein Datensatz gespeichert, der durch einen Primärschlüssel eindeutig identifizierbar ist. Beziehungen zwischen mehreren Tabellen werden über Fremdschlüssel hergestellt.

### **Relevanz**

Relevanz bedeutet, dass ein Dokument dazu in der Lage ist, zur Lösung des Informationsproblems eines Informationssuchenden beizutragen.

## **S**

### **Schlagwort**

Als Schlagwort wird ein Ausdruck oder Spruch bezeichnet, durch den besondere Beachtung erzeugt werden soll. Davon zu unterscheiden ist das Schlagwort in der Bedeutung „Stichwort, Deskriptor“, wie es der inhaltlichen Erschließung von Dokumenten mittels Verschlagwortung gebraucht wird, zum Beispiel in Bibliotheken.

### **Schlüssel**

Ein Schlüssel dient in einer relationalen Datenbank dazu, die Tupel (Datensätze) einer Relation (Tabelle) eindeutig zu identifizieren, sie zu nummern. Ein

Schlüssel ist dann eine Gruppe von Spalten, die so ausgewählt wird, dass jede Tabellenzeile über den Werten dieser Spaltengruppe eine einmalige Wertekombination hat.

### **Schnittstelle**

Die Schnittstelle, auch Interface (englisch "Grenzfläche", Begriff ursprünglich aus der Naturwissenschaft), ist ein Teil eines Systems und dient der Kommunikation.

### **SELECT (SQL-Befehl)**

Der SQL SELECT Befehl ist der Grundstein für zahlreiche SQL-Abfragen, die auf eine Datenbank ausgeführt werden können. Mithilfe des SQL Select Befehls ist es möglich Daten aus einer oder mehreren Tabellen über ein JOIN (Verbindung) abzufragen. Zusätzlich sollte der SELECT Befehl niemals mit einem Stern, also SELECT \* FROM Tabelle ausgeführt werden, da dann alle Spalten der SQL SELECT Abfrage zurückgeliefert werden.

### **Selektion**

Teilmengenbildung, d.h. Auswahl bestimmter Tupel in einer Relation (Zeilen einer Tabelle), die eine vorgegebene Bedingung erfüllen.

### **SERP**

Search Engine Result Page, Suchergebnisseite.

### **Stemming**

Als Stemming bezeichnet man im Information Retrieval sowie in der linguistischen Informatik ein Verfahren, mit dem verschiedene morphologische Varianten eines Wortes auf ihren gemeinsamen Wortstamm zurückgeführt werden, z. B. die Deklination von Wortes oder Wörter zu Wort und Konjugation von gesehen oder sah zu seh.

### **Structured Query Language (SQL)**

Die Abkürzung SQL steht für den Begriff Structured Query Language und bezeichnet eine Sprache für die Kommunikation mit relationalen Datenbanken. Mit SQL-Befehlen lassen sich Daten relativ einfach einfügen, verändern oder löschen.

### **Suchanfrage**

Der Ausdruck Suchanfrage bezeichnet den Vorgang, bei dem ein Nutzer in eine Suchmaschine einen beliebigen Inhalt (Schlagwort, Stichwort) eingibt, der mit einer bestimmten Suchintention verbunden ist. Dabei kann es sich je nach Suchintention um einzelne Wörter, (Teil-)Sätze oder konkrete Fragen handeln. Am Ende des Suchvorgangs werden dem Nutzer erzielte Ergebnisse zu seinem Suchwort angezeigt.

### **Stichwort**

Ein Stichwort ist in der Dokumentation ein Wort, das der inhaltlichen

Erschließung eines Dokumentes dient. Vom Standpunkt des Suchenden aus ist es der Suchbegriff.

### **Synonymie**

Die Synonymie ist eine semantische Beziehung zwischen Wörtern oder Ausdrücken einer Sprache. Wenn zwei Wörter synonym sind, haben sie die gleiche Bedeutung. Aber sie haben nicht die gleiche Form, d.h. sie sind phonetisch unterschiedlich. Erstens nehmen wir ein einfaches Beispiel: Das Verb „sprechen“ ist ein Synonym des Verbes „diskutieren“.

## **T**

### **Tabelle**

Eine Tabelle ist eine Zusammenstellung von Daten, die in Zeilen und Spalten angeordnet sind. So können wichtige Daten in einem strukturierten Format in einer Datenbank gespeichert werden. Zusätzlich können die Elemente in der Tabelle bearbeitet oder auch in Relation gesetzt werden.

### **TCP IP**

Transmission Control Protocol / Internet Protocol (TCP/IP) ist eine Familie von Netzwerkprotokollen und wird wegen ihrer großen Bedeutung für das Internet auch als Internetprotokollfamilie bezeichnet.

### **Templates**

Templates (der englische Begriff für Schablonen), sind Vorlagen, die mit Inhalt gefüllt werden können.

### **Traffic**

Der Daten- bzw. Besucherstrom auf einer Webseite. Der Traffic der eigenen Seite kann anhand der Logfiles bzw. eines Analysetools eingesehen werden. Für fremde Webseiten bietet Alexa einen Anhaltspunkt über den Traffic anhand des Alexa Traffic Ranks.

### **Tupel**

Der Begriff Tupel wird in der Informatik für geordnete Wertesammlungen (eindimensionale Arrays) und – insbesondere in der relationalen Algebra – als Synonym für Datensatz verwendet. Seine Werte werden als Attribute (Datenfeld) bezeichnet.

## U

### **Usability-Test**

Bei diesen Tests lösen potentielle Benutzer vorgegebene Aufgaben und stehen dabei unter Beobachtung von Entwicklern und z.B. Psychologen. Ihre Interaktionen mit der Software werden mit Log-Dateien, Video- und Tonaufnahmen über mehrere Testreihen hinweg dokumentiert und anschließend analysiert. Ziel ist es durch Identifikation von ergonomischen Problemen bei der Interaktion mit einem Softwaresystem gezielt Verbesserungen zur Steigerung der Gebrauchstauglichkeit einer Software zu erreichen.

### **UPDATE (SQL-Befehl)**

Mit dem SQL-Befehl UPDATE, kannst du Werte aktualisieren. Wenn man bei der UPDATE-Operation keine Eingrenzung mit WHERE setzt, werden alle Werte der Spalte für die jeweilige Tabelle auf den neuen Wert geändert. In der Regel möchte man aber nur den Wert einer bestimmten Zeile ändern, daher sollte man unbedingt darauf achten, die WHERE-Bedingung zu setzen.

### **URL**

Uniform Resource Locator; Ein eindeutiger Identifikator bzw. Adresse eines HTML-Dokuments im Internet. Die URL besteht aus einer Domäne und der Angabe des Ortes des Dokuments auf dem Server.

## V

### **Vektor**

Ein Vektor ist eine mathematische Repräsentation eines Objektes durch seine Eigenschaften.

### **VPN-Client**

Mit VPN oder Virtual Private Network können User über das Internet eine sichere Verbindung zu einem anderen öffentlichen Netzwerk herzustellen. Ein VPN-Client ist eine softwarebasierte Technologie, die diese sichere Verbindung zwischen dem Benutzer und einem VPN-Server herstellt

## W

### **Wissen**

Wissen definiert gespeicherte Erkenntnisse, die durch Erfahrungen oder Lernen erworben werden.



## **Wissensrepräsentation**

Als Repräsentation werden mentale Organisationsformen beschrieben, die das individuelle Wissen, Prozesse der Veränderung dieses Wissens und die Ableitung von neuem Wissen durch bewusste, unbewusste Schlussfolgerungsprozesse und die Generierung von Handlungsplänen umfasst.

### **Weaver-Memorandum**

Das Fachgebiet der MÜ entstand Anfang der 50er Jahre. Wegweisend war das sogenannte Weaver-Memorandum (1949), das die Möglichkeit der Übersetzung von Sprache durch den Computer propagierte, indem es den Übersetzungsprozess mit einer Art „Decodieren“ gleichsetzt.

### **WHERE (SQL-Befehl)**

Mithilfe des SQL WHERE-Befehls werden in SQL Abfragen nur bestimmten Datensätze ausgewählt. Der SQL WHERE-Befehl funktioniert im Prinzip wie ein Filter, der es ermöglicht, nur Datensätze anzuzeigen, die bestimmte Kriterien erfüllen. Soll ein SQL Statement eine bestimmte Bedingung erfüllen, muss eine WHERE-Bedingung eingebaut und erfüllt werden, damit die Abfrage eine Ergebnismenge liefern kann.

## **X**

### **XML**

XML (Extensible Markup Language) ist ein text-basiertes Format für den Austausch strukturierter Information. Dies können Dokumente, Konfigurationen, Bücher, Rechnungen und vieles mehr sein und ist aus der Not entstanden, dass HTML an seine Grenzen gestoßen ist.

## **Y**

### **Yale-Web-Styleguide**

Der Yale UI Style Guide wurde erstellt, um Entwicklern dabei zu helfen, Schnittstellen zu erstellen, die der visuellen Identität/ Vision von Yale entsprechen.

## Z

### **Zeichenkette**

Eine Zeichenkette ist eine endliche Folge von Zeichen (z. B. Buchstaben, Ziffern, Sonderzeichen und Steuerzeichen) aus einem definierten Zeichensatz. Zeichen können sich in einer Zeichenkette wiederholen, die Reihenfolge der Zeichen ist definiert. Eine Zeichenkette kann auch leer sein, also kein Zeichen enthalten und die Länge 0 haben.

### **ZIP-Dateiformat**

ZIP ist ein Format für komprimierte Dateien, das einerseits den Platzbedarf reduziert und andererseits als Containerdatei fungiert, in der mehrere zusammengehörige Dateien oder auch ganze Verzeichnisbäume zusammengefasst werden können.