Seminar über Datenbanken, XML und Suchmaschinen

(Sommer 2008)

Prof. Dr. Stefan Brass

Institut für Informatik

Overview

1. Organisatorisches

2. Themen

3. Gute Vorträge

Themen/Lernziele

- Selbständige Einarbeitung in ein Thema anhand wissenschaftlicher Artikel

  Aus Tagungsbänden oder Zeitschriften, auf Englisch.

- Gute und verständliche Vorträge halten

- Implementierung von Datenbanksystemen
- XML und Datenbanken
- Suchmaschinen Spam

Termine

- Vorschlag: Blockseminar.
  28. bis 30. Juli 2008, jeweils ca. 10–13, 14–17

  Das wäre die zweite Woche nach Ende der Vorlesungszeit.
  Eventuell nur 2 oder 2.5 Tage nötig.

- Einführungsveranstaltung: 08.04., 12–14, R. 116.

- Abgabe von Ausarbeitungen und/oder Folien bis 30.06.2008.

  Sie müssen eine Ausarbeitung von mindestens 8 Seiten abgeben, es kann sich dabei aber auch um die Folien mit dem vollständigen, zu sprechenden Text für mindestens 10 Folien handeln. In jedem Fall möchte ich die Folien (ohne Sprechernotizen) spätestens eine Woche vor dem Vortrag sehen. Je früher Sie etwas abgeben, desto eher und sicherer bekommen Sie Feedback.

Ansprechpartner (1)

Dozent: Prof. Dr. Stefan Brass

- Email: `brass@informatik.uni-halle.de`

  Betreff-Zeile sollte mit `[sem08]` beginnen, möglichst aussagefähig.

- Büro: Von-Seckendorff-Platz 1, Raum 313
- Telefon: 0345/55-24740 (Büro)
- Sprechstunde: Freitags, 14^{00}–15^{00}
- Frühere Unis: Braunschweig, Dortmund, Hannover, Hildesheim, Pittsburgh, Gießen, Clausthal.
- Oracle8 Certified Database Administrator.
  IBM Certified Advanced DBA (DB2 UDB 8.1).

Ansprechpartner (2)

Sekretärin: Ramona Vahrenhold

- Büro: Von-Seckendorff-Platz 1, Raum 324
  (Dienstags nicht besetzt)
- Telefon: 0345/55-24750, Fax: 0345/55-27333
- Email: `vahrenho@informatik.uni-halle.de`

WWW:

- `http://www.informatik.uni-halle.de/~brass/sem08/`

Seminarschein (1)

Kriterien:

- Ausarbeitung/Folien
- Guter Vortrag am vereinbarten Termin:
  - ◇ Logischer Aufbau, Vortragsstil, Didaktik
  - ◇ Zeit eingehalten: 40 Minuten (± max. 5 min.)
  - ◇ Vorgegebener Stoff behandelt (Kernpunkte)

    Oft ist es nicht möglich, den Stoff vollständig zu behandeln. Dann sollte man sich auf die grundlegenden Ideen konzentrieren.

  - ◇ Verständnis des Stoffes gezeigt, auch in der anschließenden Diskussion
- Anwesenheit bei allen Vorträgen, aktive Teilnahme

Seminarschein (2)

- Sie müssen mindestens die Kernaussagen des jeweils ersten Artikels zu Ihrem Thema behandeln.

  Es wäre völlig falsch, nur den Titel (das globale Thema) des Vortrages als gegeben anzusehen, und sich dazu selbst Literatur zu besorgen, und den vorgegebenen Artikel zu ignorieren (insbesondere reicht es nicht, nur Lehrbücher zu benutzen). Wenn mehrere Artikel angegeben sind, sind die weiteren Artikel nur zur Ergänzung gedacht.

- Selbstverständlich können Sie sich gern zusätzlich weitere Literatur besorgen.

  Oft müssen Sie sogar Teile der zitierten Literatur beschaffen und lesen. Alle verwendete Literatur muß genannt werden. Falls Sie im Web einen Seminarvortrag zum Thema finden sollten, müssen Sie trotzdem einen eigenen Vortrag ausarbeiten (eigenständige Leistung).

Seminarschein (3)

Ganz wichtig zum Bestehen:

- Die Ausarbeitung soll nicht einfach eine deutsche Übersetzung der Literaturstelle sein, sondern von eigenem Verständnis zeugen.
- Sie müssen mit Fragen zum Thema während des Vortrags oder nach dem Vortrag rechnen.

  Der Dozent möchte gerne testen, daß Sie das Thema wirklich verstanden haben, zumindest soweit, wie Sie es in Ihrem Vortrag auch bringen. Sie müssen z.B. Fachworte, die Sie in Ihrem Vortrag benutzen, auch erklären können. Sie müssen ggf. auch ad-hoc Beispiele für etwas machen können. Sie müssen erklären können, was bei leichten Modifikationen Ihrer Beispiele passieren würde, oder warum etwas so und nicht anders ist.

Seminarschein (4)

- Eine Diskussion nach dem Vortrag ist wichtig.
- Ich möchte daher für jede Arbeit außer dem Vortragenden noch einen Sitzungsleiter/Diskussionsleiter bestellen (wählen Sie auch dafür ein Thema):
  - ◇ Hat die Arbeit auch gelesen (zumindest überflogen), eventuell auch den Vortrag vorher gesehen.
  - ◇ Spricht Einleitung (2–3 Sätze), achtet auf Einhaltung der Zeit, leitet Diskussion.
  - ◇ Hat sich mindestens drei Fragen überlegt.

Literatur-Beschaffung

- Von den Rechnern der Universität aus haben Sie kostenlosen Zugriff auf den Volltext (PDF) der von der ACM herausgegebenen Publikationen.

  Die ACM Digital Library [http://www.acm.org/dl/] listet aber auch einige Literatur, die nicht von der ACM herausgegeben ist.
- Schauen Sie auch mal in die Bibliothek!
- Bei der Beschaffung von Texten, die von der IEEE Computer Society herausgegeben sind, kann ich Ihnen eventuell helfen.
- Weitere Werkzeuge: Google, DBLP, CiteSeer.

Overview

Themen (1)

Implementierung von Datenbanken:

- B-Bäume und Caching in CPUs
- Änderung von B-Bäumen
- Online-Reorganisation von B-Baum Indexen
- Diag-Join
- OPT++: Ein erweiterbarer Anfrageoptimierer

Themen (2)

XML und Datenbanken:

- Speicherung von XML-Daten in relationalen Datenbanken (*)

  (*) bedeutet, daß das Thema schon vergeben ist.

- Native XML Indexierungs-Techniken für die Speicherung von XML-Daten in Relationalen Datenbanken (*)
- Datenbankentwurf für hybride Datenbanken (relational und XML) (*)

Themen (3)

Suchmaschinen-Spam:

- Suchmaschinen-Spam: Übersicht und Statistik (*)
- Suchmaschinen-Spam: Pagerank und Link-Farmen (*)
- TrustRank

Overview

1. Organisatorisches
2. Themen
3. Gute Vorträge

Verständlichkeit (1)

- Wichtigstes Kriterium für guten Vortrag:
  Die Hörer sollen hinterher etwas gelernt haben.

  Nicht nur, daß Sie ein schlechter Vortragender sind ...

- Die Hörer investieren eine Stunde ihrer Zeit.
  Wenn sie nichts gelernt haben, ist die Zeit verloren.

- Überlassen Sie es nicht dem Zufall, was die Hörer in Erinnerung behalten: Denken Sie bei der Planung über “Take Home Messages” nach.

  Nützliche Übung: Ganzen Vortrag in 5 Punkten zusammenfassen (à max. 3 Zeilen). Ganze Vorlesung auf einer Seite. Quick Ref. Chart. Verwandt, für längere Veranstaltungen: Lernziele (Klausur entwerfen).

Verständlichkeit (2)

- Der Hörer kann nicht wie bei einem Buch zurückblättern und eine Definition nochmal nachschauen.

- Daher sollten Definitionen möglichst nahe an der Verwendung eingeführt werden oder noch einmal kurz wiederholt werden.

  Zu viele neue Begriffe kann man sich nicht merken.

- Manchmal wird man kurz abgelenkt. Wenn man für den Rest des Vortrags nicht mehr folgen kann, nachdem man ein Detail verpaßt hat, ist es ein schlechter Vortrag. Hörer schalten dann auch ab.

Verständlichkeit (3)

- Die Zuhörer wissen oft weniger, als man denkt.

  Man sollte sich natürlich möglichst genau über den Kenntnisstand der Zuhörer informieren.

- Es ist daher in Ordung, auch Stoff zu wiederholen, der eigentlich bekannt sein müßte.

  Auch jemand, der den Stoff wirklich schon kennt, freut sich über eine kurze und prägnante Zusammenfassung.

- Am Ende sollten die Zuhörer aber nicht herausgehen, ohne etwas Neues gelernt zu haben.

  “Die erste Viertelstunde versteht jeder, die zweite Viertelstunde nur die Spezialisten, die letzte Viertelstunde nur der Vortragende.”

Verständlichkeit (4)

- Bei einer Untersuchung wurden einmal
  - ◇ 100 Einheiten neues Material eingeführt,
    das Publikum erinnerte sich an 25 Einheiten,
  - ◇ 50 Einheiten bekanntes und 50 Einheiten neues Material behandelt, das Publikum erinnerte sich an 35 Einheiten des neuen Materials.
- Um Stoff zu lernen, muß man ihn mit Bekanntem verknüpfen (“in Gedankengebäude einbauen”).
- Jeder Lehrer muß wiederholen, wiederholen, ...

Verständlichkeit (5)

- Wählen Sie eine didaktische Anordnung des Stoffes.

  Kommandos in alphabetischer Anordnung zu besprechen ist sicher sinnlos. Man würde zuerst die wichtigsten oder einfachsten besprechen, und dann zu etwas exotischeren/komplizierteren übergehen.

- Oft muß die Darstellung exemplarisch bleiben.

  Es ist z.B. besser, nur die wichtigsten Kommandos zu behandeln, und die so ausführlich, daß die Hörer sie verstehen, als alle Kommandos so schnell, daß gar nichts hängenbleibt. "Coverage is the enemy of understanding" [Howard Gardner]. Wenn Sie unbedingt wollen, können Sie ja auf einer Folie eine Liste aller Kommandos haben, diese aber nur kurz zeigen und nicht besprechen (Eindruck von allen Möglichkeiten).

- Eine lange Liste gleichförmiger Punkte wirkt uninteressant und ermüdend.

Verständlichkeit (6)

- Beispiele sind für das Verständnis häufig nützlicher als allzu formale Definitionen (z.B. Grammatik).

  Die Lernsituation in einem Vortrag mit Folien, der relativ schnell abläuft, ist anders als zu Hause mit einem Buch. Natürlich müssen bei einer Vorlesung auch die formalen Definitionen gelernt werden. Es sollte aber jede Definition mit einem Beispiel illustriert werden.

- Nur Beispiele sind auch unbefriedigend.
  Man möchte auch wissen, was allgemein gilt.

  Wenn man etwas anwenden will, soll man ja nicht nur die gegebenen drei Beispiele behandeln können. Bei einem eher mathematisch orientierten Thema dürfen die formalen Dinge nicht weggelassen werden.

- Bringen Sie Beispiele nicht zu spät im Vortrag.

Verständlichkeit (7)

- Zum Beispiel kann man die Datenbanksprache SQL auf folgende Arten vermitteln:
  - ◇ Formale oder halbformale Definition.

    Formale Syntax (Kontextfreie Grammatik und Syntaxdiagramme) und mehr oder weniger Logik zur Semantikdefinition.

  - ◇ Beschreibung der Ausführung in einem DBMS.

    Zumindest naiver Ausführungsalgorithmus (d.h. ohne Optimierungen). Dies ist eine Art operationale Semantik.

  - ◇ Große Sammlung vieler Beispiele.
- Verschiedene Studierende benötigen unterschiedliche von diesen Dingen.

Verständlichkeit (8)

- "Ein Bild sagt mehr als tausend Worte."
- Bei einem Vortrag gibt es verschiedene Kanäle, auf denen man mit den Zuhörern kommunizieren kann:
  - ◇ Folien oder Tafelbild
  - ◇ Gesprochener Text
  - ◇ Betonungen
  - ◇ Gesten
- Man sollte möglichst viele dieser Kanäle benutzen.

  Verschiedene Typen von Lernern, größere Bandbreite.

Verständlichkeit (9)

- Relative Wichtigkeit verschiedener Kanäle, "to get a message across" (gilt aber wohl kaum für wissenschaftliche Vorträge):
  - ◇ Verbal (was man sagt): 7%
  - ◇ Vocal (wie man es sagt): 38%
  - ◇ Visuell: 55%
- Erinnerung (Einprägen des Stoffes):
  - ◇ Drei Stunden nach dem Vortrag kein großer Unterschied zwischen verbal, visuell, verbal+visuell.
  - ◇ Drei Tage später: verbal+visuell deutlich stärker.

Tafel vs. Folien (1)

- Mit Folien kann man mehr Stoff pro Zeitinheit behandeln.

  Weit über die Aufnahmebandbreite der Zuhörer hinaus ...
- Programmbeispiele anzuschreiben verbraucht (zu) viel Zeit, besonders wenn nur relativ kleine Teile neu sind, aber der ganze Kontext benötigt wird.
- Tafelwischen erfordert Pausen.

  Über die sich manche Zuhörer freuen ...
- Wenn man etwas an die Tafel schreibt, kehrt man dem Publikum den Rücken zu.

Tafel vs. Folien (2)

- Wenn man eine schlechte Schrift hat, oder das Tafelbild nicht übersichtlich organisieren kann, können gut vorbereitete, gedruckte Folien helfen.
- Folien können Bilder enthalten, die man nicht anschreiben kann.
- Die Vorbereitung von Folien kostet viel Zeit.

  Bei einem Tafelvortrag muß man sich aber auch wenigstens den vollständigen Text notieren, der an die Tafel geschrieben werden soll.
- Ausdruck der Folien automatisch eine Art Skript.

  "Kopieren geht über studieren ..."

Tafel vs. Folien (3)

- Mitschreiben bei einem Folienvortrag ist praktisch unmöglich (zu viel Stoff/Zeiteinheit).
- Durch Mitschreiben prägt sich der Stoff besser ein.

  Durch den Zwang zum Mitschreiben können die Hörer auch nicht einschlafen oder an andere Dinge denken.
- Wenn man nicht mitschreiben muß, hat man mehr Zeit zum mitdenken.
- Vermutlich beste Lösung: Mischung, z.B. Folienvortrag, aber mit eingestreuten Aufgaben, die an der Tafel behandelt werden (Aktivität der Hörer).

Folien (1)

- Ausreichend große und lesbare Schrift, z.B. Sans Serif Zeichensatz 22–30 Punkt.

  max. 7–10 Zeilen pro Folie, 42–60 Zeichen pro Zeile.
- Ganz schlimm: Fotokopierte Seite aus einem Buch.
- Experten empfehlen Querformat und Rahmen.

  Fernseher/Computerbildschirm ist den Zuhörern vertraut. Für den Beamer Querformat nötig. Bei Querformat weniger Stoff pro Folie.
- Nicht zuviel Stoff auf einer Folie.

  Wenn man die Folie auflegt, fangen die Hörer an zu lesen, und sind damit für eine kurze Zeit abgelenkt. Komplexe Folien können eventuell stückweise aufgebaut werden.

Folien (2)

- Nicht zu viele Folien!

  Maximal eine pro Minute. Besser: Eine pro zwei Minuten. Eine pro drei Minuten geht auch. Nur die Folien vorzulesen ist schlecht. Man sollte noch mehr zu erzählen haben, als auf den Folien steht. Außerdem sollten auch die Zuhörer genügend Zeit haben die Folie zu lesen und den Inhalt zu verstehen.
- Normalerweise muß man bei einem Vortrag die Zeit genau einhalten: Man braucht dann einige Folien, die man einfügen oder weglassen kann.

  Nachteil eines Beamers: Die Zuhörer merken es.

Folien (3)

- Wenn man sich auf eine Folie später noch einmal beziehen will, sollte man zwei Kopien haben.

  Oder mit dieser Folie sehr vorsichtig sein.
- Telegrammstil (nicht ganze Sätze).

  Meine Folien sind in dieser Hinsicht kein gutes Beispiel.
- Den Zuhörern sollte die Struktur des Vortrags klar sein, sie sollten sich auch während des Vortrags orientieren können und den Fortschritt sehen.
- Farben sollten konsistent verwendet werden.

Vortrag planen (1)

- Nichts ganz Wichtiges am Anfang, manche Zuhörer kommen zu spät und verpassen den Anfang.
- Am Anfang Thema und Struktur des Vortrags kurz erklären.
- Dann Einstieg/Motivation: Interesse wecken.
- Versuchen Sie, eine Geschichte zu erzählen.

  Menschen mögen Geschichten.
- Vermeiden Sie Geheimnisse, die erst später geklärt werden. Sie sollen keinen Kriminalroman erzählen.

  Vorwärtsreferenzen sind normalerweise schlecht.

Vortrag planen (2)

- Machen Sie sich nicht selbst schlecht!

  Sorgen Sie dafür, daß Sie alles verstanden haben (fragen Sie den Dozenten). Wenn Sie etwas gar nicht verstehen, lassen Sie es weg. Vielleicht merkt es niemand. Es ist eine Prüfung, Sie wollen Ihre Zensur doch nicht selbst verschlechtern.
- Kurze Softwaredemonstrationen sind eine nützliche Abwechselung, müssen aber gut geplant sein.
- Am Ende Zusammenfassung, Ausblick.
- Enden Sie auf einer starken Note wie ein klassisches Musikstück und lassen nicht einfach auslaufen.

  Schlecht: "Na ja, ich denke, ich mache jetzt Schluß."

Vortrag planen (3)

- Informieren Sie sich über die technische Ausstattung des Vortragsraums.
- Wird es genug Licht geben, daß Sie Notizen und Uhr lesen können?
- Verlassen Sie sich nicht allein auf elektronische Folien (Beamer). Es gibt öfters technische Schwierigkeiten.

  Sie sollten auch ganz ohne Folien den Vortrag noch irgendwie durchziehen können.

Notizen für den Sprecher (1)

- Gerade Anfänger sollten den kompletten zu sprechenden Text Wort für Wort aufschreiben.

  Bei der Ausarbeitung des Textes merkt man häufig, das die Folien noch modifiziert werden müssen.
- Zumindest sollte man den Anfang genau ausarbeiten: Hilft für guten Start (Nervosität überwinden).

  Das Ende ist häufig auch besonders wichtig, und man steht eventuell unter Zeitdruck.
- Große, leicht lesbare Schrift.
- Viele Überschriften oder andere Markierungen, um leicht die richtige Stelle wiederzufinden.

Notizen für den Sprecher (2)

- Man sollte die Notizen zwischen die Folien legen.

  Man kann nicht zwei verschiedene Stapel parallel verwalten.
- Manche Leute verwenden Karteikarten, die man relativ unabhängig in der Hand halten kann.

  Man sollte sich zwar den vollständigen Text einmal aufschreiben, und für ein mögliches "Blackout" auch dabei haben, aber die planmäßigen Notizen für en Sprecher können auch nur Stichworte sein. Bei kleinen Karteikarten soll das Zittern auch nicht so sichtbar sein ...
- Drei Spalten: Zeit, Text (Stichworte), Bühnenanweisungen (z.B. "Folie 15 auflegen", "Licht an", "Lächeln!").

Vortrag üben

- Man sollte den Vortrag vorher “im stillen Kämmerlein” mehrfach halten (auch laut sprechen).

  Es reicht nicht, sich im Gebiet sicher zu fühlen und die Folien gerade noch rechtzeitig fertig zu haben!

- Man kann zum Üben auch einen Freund bitten, oder es abends im einem wirklichen Vortragssaal halten.
- Man muß auch die Zeit messen.

  Dafür muß der Vortrag wie in echt ablaufen. Man kann dann nicht mehr immer wieder neu ansetzen und eine besser Formulierung versuchen. Meine Erfahrung zeigt, daß der Vortrag später im Ernstfall eher schneller abläuft, weil man wegen der Nervosität schneller spricht oder etwas vergisst, was man eigentlich sagen wollte.

Vortrag halten (1)

- Halten Sie sich an die Zeitvorgabe.

  “Sie können über alles sprechen, aber nicht über eine Stunde.” Informieren Sie sich, ob die Zeit für die Diskussion im Zeitlimit enthalten ist oder nicht.

- Sprechen Sie ausreichend laut und deutlich.

  Man selbst kann das häufig schlecht einschätzen. Man muß etwas lauter sprechen als zu Hause oder wenn man dem Gesprächspartner direkt gegenüber steht. Es ist schlecht, wenn die Stimme am Satzende immer abfällt (oder wenn Sie sich unsicher sind).

- Nicht zuviel ablesen. Schauen Sie die Zuschauer an.

  “To make oneself understood to people, one must first speak to their eyes.” [Napoleon] Besonders schlecht: Ständig zur Leinwand schauen, dabei den Zuhörern den Rücken zukehren.

Vortrag halten (2)

- Vermeiden Sie zuviel “Hm”, “Äh”, “ja?”.
- Sprechen Sie nicht langweilig (Betonungen!).
- Strahlen Sie eigene Begeisterung für das Thema aus.

  Wenn Sie das Thema schon nicht interessiert, wie sollte es dann die Zuhörer interessieren? Sie können ein paar kritische Bemerkungen einstreuen (vielleicht eher gegen Ende), das zeugt von eigener Beschäftigung mit dem Stoff. Aber der Grundtenor sollte positiv sein.

- Sprechen Sie nicht zu schnell.

  Z.B. zwei Worte pro Sekunde. Die Zuhörer können Ihre Gedanken nicht so schnell verstehen wie Sie sie sprechen können. “Everything worth doing is worth doing slowly.”

Vortrag halten (3)

- Benehmen Sie sich normal. Vermeiden Sie ablenkendes Verhalten, oder Ihre Nervosität zu zeigen.

  Wringen Sie nicht Ihre Hände, spielen Sie nicht mit dem Zeigestock. Verstecken Sie sich nicht hinter dem Podium.

- Sie sollten sich während des Vortrags etwas bewegen, aber zuviel hin und her ist ablenkend.

  Zum Beispiel ist es schlecht, absolut angewurzelt und auf den Zeigestock gestützt zu stehen. Sie müssen ja auch manchmal auf der Folie etwas zeigen. Etwas Bewegung ist gut, damit die Zuhörer nicht einschlafen.

Vortrag halten (4)

- Seien Sie rechtzeitig da, probieren Sie den Projektor aus, informieren Sie sich über die Lichtschalter.

  Nutzen Sie die letzte halbe Stunde vor dem Vortrag, die Folien noch einmal durchzugehen (bringen Sie sie aber nicht durcheinander). Gehen Sie noch einmal auf Toilette.

- Kurze Verständnisfragen können während des Vortrags geklärt werden (wenn nicht zu viele).
- Längere Fragen/Diskussionen bitten Sie, erst am Ende des Vortrags beantworten zu können.

Vortrag halten (5)

- Ggf. müssen Sie ein Glas Wasser mitbringen.

  Manchmal braucht man das, um besser sprechen zu können. Normalerweise sollte sich der Organisator darum kümmern, aber das geschieht nicht immer.

- Vermeiden Sie zu viele Metakommentare über den Vortrag, warum Sie etwas so gemacht haben.

  Natürlich müssen Sie die Struktur des Vortrags klar machen. Ansonsten sollte man sich aber auf den eigentlichen Inhalt konzentrieren.

- Wirken Sie nicht schulmeisterlich.

  Es ist klar, daß Sie über dieses Thema besser Bescheid wissen als die Zuhörer. Aber ansonsten ist das Verhältnis gleichberechtigt. Vermeiden Sie alles, was nach geistiger Bevormundung aussieht.

Nervosität

- Etwas Nervosität verbessert die Leistung.
- Normalerweise sind die Zuhörer freundlich gesinnt.

  Die Zuhörer sind in der gleichen Situation. Sie haben für den Vortrag nichts bezahlt, der Dozent wird sogar dafür bezahlt, daß er zuhört.

- Vorher gut vorbereiten!

  Alle Verständnisfragen klären, möglichst genau über Publikum informieren, üben unter realen Bedingungen.

- Ignorieren Sie Zweifel, die Ihnen während des Vortrags kommen.
- Man lernt aus Fehlern häufig mehr als aus Erfolgen.

  Die Schmerzen und das Gefühl der Blamage vergehen.