

Name:	Studiengang: <input type="checkbox"/> B.A. <input type="checkbox"/> MA.
Vorname:	
Matrikelnummer:	
Studienfächer:	Fachsemester Informationswissenschaft:

Allgemeine Hinweise:

1. Überprüfen Sie bitte, ob Sie alle Seiten der Klausurangabe vollständig erhalten haben (Gesamtzahl: 5)
2. **Bearbeitungszeit: 90 Minuten**, maximal erreichbare **Punktzahl: 60**. Die jeweils erreichbare Punktzahl ist bei jeder Frage angegeben. Bitte teilen Sie Ihre Arbeitszeit entsprechend ein.
3. Denken Sie daran, die Daten oben einzutragen, **bevor** Sie mit der Bearbeitung beginnen.
4. Treffen Sie bitte die Auswahl Ihrer Antworten bei Multiple-Choice-Fragen **direkt** auf dieser Klausurangabe.
5. Verwenden Sie für die Beantwortung der Freitext-Fragen ebenfalls diese Klausurangabe. Sie können jederzeit auch die Rückseiten beschreiben, falls der Platz auf der Vorderseite nicht ausreichen sollte. Bitte geben Sie in jedem Fall an, auf welche Frage sich die Lösung jeweils bezieht.
6. Benutzen Sie keine Bleistifte, keine rot schreibenden Stifte und kein TippEx, o.ä.
7. Keine Hilfsmittel sind zugelassen d.h. keine Foliensätze oder selbstgeschriebene Notizen.
8. Einen Taschenrechner dürfen Sie gerne benutzen.
9. Mobiltelefone sowie Computer am Arbeitsplatz - auch ausgeschaltet - sind **nicht zugelassen**.
10. Geben Sie keine mehrdeutigen (oder mehrere) Lösungen an. In solchen Fällen wird stets die Lösung mit der geringeren Punktzahl gewertet. Eine richtige und eine falsche Lösung ergeben also null Punkte.
11. Bei Multiple-Choice-Fragen Führen falsche Antworten zu Punktabzug
12. Wenden Sie sich bei Unklarheiten in den Aufgabenstellungen immer an den Aufsichtsführenden. Hinweise und Hilfestellungen werden dann, falls erforderlich, offiziell für alle Teilnehmer durchgegeben.

1) Welche der folgenden Aussagen treffen zu? *Which of the following statements are true?*
(10 Punkte)

- | Wahr | | Falsch |
|--------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------|
| <input type="checkbox"/> | Precision-Recall-Graphen werden typischerweise für einzelne Anfragen analysiert
<i>It is standard to examine Precision-Recall graphs for single queries</i> | <input type="checkbox"/> |
| <input type="checkbox"/> | Precision and Recall sind beide menge-basierte Maße
<i>Precision and Recall are both set-based metrics</i> | <input type="checkbox"/> |
| <input type="checkbox"/> | P@1 ist ein passendes Evaluationsmaß für Navigationswebsuche
<i>P@1 is a suitable metric for navigational web search</i> | <input type="checkbox"/> |
| <input type="checkbox"/> | P@1 ist ein passendes Evaluationsmaß für exploratorische Websuche
<i>P@1 is a suitable metric for exploratory web search</i> | <input type="checkbox"/> |
| <input type="checkbox"/> | Unstimmigkeit in Relevanzbeurteilungen verschiedener Assessors wird mit hoher Wahrscheinlichkeit das optimale Ranking beeinflussen
<i>Disagreement in judgements between relevance assessors is likely to influence the optimal ranking order</i> | <input type="checkbox"/> |

2) Welche der folgenden Aussagen treffen auf Bates Berrypicking-Modell zu? *Which of the following statements are true regarding Bates' berry-picking model:?* (10 Punkte)

- | Wahr | | Falsch |
|--------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------|
| <input type="checkbox"/> | Das Modell ist für die Indexierung der Dokumente sehr nützlich
<i>The model offers insight with respect to how documents should be indexed</i> | <input type="checkbox"/> |
| <input type="checkbox"/> | Das Model unterstützt die Relevanzunabhängigkeitsannahmen des Cranfield-Paradigma
<i>The model supports the independent relevance assumptions of Cranfield</i> | <input type="checkbox"/> |
| <input type="checkbox"/> | Das Model unterstützt die „1-Query pro Aufgabe“-Annahme des Cranfield-Paradigma
<i>The model supports the one-shot retrieval assumptions of Cranfield</i> | <input type="checkbox"/> |
| <input type="checkbox"/> | Das Modell zeigt wie Emotionen den Suchprozess beeinflussen
<i>The model shows how emotions influence the search process</i> | <input type="checkbox"/> |
| <input type="checkbox"/> | Das Modell stellt den Suchprozess im Zeitverlauf dar.
<i>The model provides insight into how searches progress over time</i> | <input type="checkbox"/> |

3) Wenden Sie das LM-Modell an, um für die Suchanfrage "information class" ein Ranking für die unten aufgeführten Dokumente auszurechnen. Es gibt nur diese 5 Dokumente in der Sammlung. *Using a LM with smoothing as discussed in class ($\lambda = 0.5$) provide a ranking for the following documents given the query "information class" (we assume that these are the only documents in the collection) (10 Punkte)*

D1: "information retrieval is the best class I have taken in Regensburg"

D2: "I learned in information retrieval class that long documents like this one have an advantage and are always ranked higher than they should be"

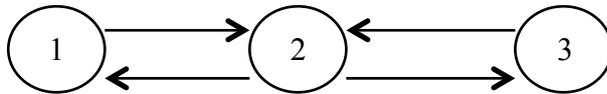
D3: "Robotics and information retrieval make use of the same technologies"

D4: "information retrieval existed long before google"

D5: "After this exam I am going for a beer. This class was tough and so was the exam".

4a) Nennen Sie 4 Unterschiede zwischen Websuche und klassischem Information Retrieval
Name 4 differences between web search and classic information retrieval (4 Punkte)

b) Was sind die Übergangswahrscheinlichkeitsmatrizen für den unten gezeigten Webgraphen, wenn die Teleportwahrscheinlichkeit (i) $\alpha = 0$; (ii) $\alpha = 0.5$ and (iii) $\alpha = 1$ ist? (3 Punkte)
Given the web graph below write down the transition probability matrices for the surfer's walk with teleporting, for the following three values of the teleport probability: (i) $\alpha = 0$; (ii) $\alpha = 0.5$ and (iii) $\alpha = 1$.



c) Die stationären Zustände der Matrizen informieren über die Qualität / Beliebtheit einer Seite. Erklären Sie wie wir diese Informationen in einem Language-Modell (siehe Frage 3) benutzen können. 1 (3 Punkte)
the steady state transition calculated from these tells you about a page's quality / popularity. Describe how this can be implemented in a language model, such as in 3)

5) Nennen Sie 2 Suchoberflächenkomponenten für jede Klasse in Max Wilsons Taxonomie (i.e. input, control, informational, personalised).

Name 2 SUI-elements from each class in Max Wilson's taxonomy e.g. (input, control, informational, personalised) (10 Punkte)

6) Welche der folgenden Konzepte können benutzt werden um Rocchios Relevanzfeedbackalgorithmus zu implementieren? (10 Punkte)

Which of the following can be used to implement relevance feedback using Rocchio's method in VSM?

Nützlich

Nicht nützlich

- | | | |
|--------------------------|---------------------------------------------------------------------------------------|--------------------------|
| <input type="checkbox"/> | Explizite Relevanzbeurteilungen
<i>explicit relevance judgements</i> | <input type="checkbox"/> |
| <input type="checkbox"/> | Das F1-Maß
<i>The F1-Score</i> | <input type="checkbox"/> |
| <input type="checkbox"/> | Eye-tracking und Maus-Daten
<i>Eye tracking and mouse interaction data</i> | <input type="checkbox"/> |
| <input type="checkbox"/> | Die ursprüngliche Anfrage des Nutzers
<i>the user's originally submitted query</i> | <input type="checkbox"/> |
| <input type="checkbox"/> | boolische Logik
<i>boolean syntax</i> | <input type="checkbox"/> |