

# Zusammenfassung

Florian Fink

CIS, LMU

`finkf@cis.lmu.de`

15. Juli 2021

# Klausur

- ▶ Die Klausur findet am Donnerstag 22.07.2021 statt.
- ▶ Beginn ist 14:00 Uhr s.t.
- ▶ Die Bearbeitungszeit beträgt 90 Minuten.
- ▶ Zusätzlich 10 Minuten Puffer für die finale Abgabe der Klausur.
- ▶ FAQ Eintrag auf der Webseite des Kurses:  
`https://cis-kb21.github.io/faq.html`

# Vorgehen

- ▶ Angabe befindet auf dem Moodle-Kurs: `https://moodle.lmu.de/mod/assign/view.php?id=862529`
- ▶ Anmeldung mit dem Code `ciskb21`
- ▶ Klausur als Jupyter-Notebook `kb21_klausur.ipynb` *oder* als PDF `kb21_klausur.pdf`
- ▶ Abgabe der Lösung auf dem Moodle-Kurs: `https://moodle.lmu.de/mod/assign/view.php?id=862529`

# Abgabe

- ▶ Die Bearbeitungszeit beträgt 90 Minuten.
- ▶ Zusätzlich 10 Minuten zur Abgabe.
- ▶ Eigenständigkeitserklärung bei der Abgabe aktivieren.
- ▶ Formate:
  - ▶ handschriftlich auf Papier als PDF:
    - ▶ Handschriftliche Bearbeitung der Klausur auf separaten Blättern
    - ▶ Einscannen oder Abfotografieren der bearbeiteten Blätter
  - ▶ direkt im Notebook
  - ▶ in einer Textdatei

# Probeklausur

- ▶ Heute Probeklausur
- ▶ Angabe befindet auf dem Moodle-Kurs: `https://moodle.lmu.de/mod/assign/view.php?id=862529`
- ▶ Melden Sie sich mit dem Code: `ciskb21` an.
- ▶ 30 Minuten Bearbeitungszeit
- ▶ Klausur enthält 5 Aufgaben

# Verschiedenes

- ▶ Stabile Internetverbindung
- ▶ Genug Akku / Stromkabel
- ▶ Fragen als private/öffentliche Chatnachricht oder telefonisch:  
+49-89-2180-9712
- ▶ Achten Sie darauf, die richtigen Dateien im Moodle hochzuladen

# Shell-Skripte

- ▶ Verzweigungen (if,else,...)
- ▶ Schleifen (for)
- ▶ Variablen
- ▶ Ausgabe von Variablen
- ▶ Globbs
- ▶ Rückgabewerte (\$?)
- ▶ Quoting

# Unix-Werkzeug

- ▶ Man-Page
- ▶ Laufzeitumgebung
- ▶ logische Operatoren ( | | , & & )
- ▶ Dateipfade
- ▶ Dateiverwaltung



# Ein- und Ausgabe

- ▶ Umleitungsoperatoren (>, <, ...)
- ▶ Pipes (|)
- ▶ Textwerkzeuge

# sed und awk

▶ sed

▶ awk

# sed und awk II

- ▶ sed Kommandos (d, p, q, s)
- ▶ sed Adressen (/laufen/, +1, /foo/, ...)
- ▶ awk BEGIN, END
- ▶ awk Variablen (NR, FS, NF, NR, ...)
- ▶ awk Patterns

# Unicode

- ▶ Unicode-Zeichensatz
- ▶ ASCII, Latin-1, UTF-8, UTF-16, ...
- ▶ Konvertierung (recode, uconv)
- ▶ hexdump, od
- ▶ kombinierende Zeichen
- ▶ Normalisierungsformen (nfc, nfd)

# XML

- ▶ semistrukturierte Datenformate (JSON, YAML, TOML, ...)
- ▶ XML
- ▶ Baumstruktur
- ▶ Attribute
- ▶ Knoten
- ▶ Entitäten
- ▶ DTD's

# Rest

- ▶ Tagsets
- ▶ eigene Tagger  
(`nltk.tag.sequential.SequentialBackoffTagger`)
- ▶ TreeTagger
- ▶ wget, tar
- ▶ Web-API's
- ▶ Crawling

# Fragen?