

Proseminar Programmieren in natürlicher Sprache

Auftaktveranstaltung, 16. Oktober 2018

Alexander Wachtel

IPD Tichy, Fakultät für Informatik



Inhalt des heutigen Treffens

- Allgemeines & Organisation
- Sinn und Zweck des Proseminars

- Motivation
 - Programmieren in natürlichen Sprache

- How To:
 - wissenschaftliches Arbeiten
 - Ausarbeitung
 - Abschlusspräsentation

- Vorstellung der Themen & Themenwahl

Allgemeines

- Semesterveranstaltung am IPD Tichy
- Betreuer: Alexander Wachtel, Alexander.Wachtel@kit.edu
- 11 Themen zu vergeben
- Termine:

Beschreibung	Datum
Gemeinsame Auftaktveranstaltung	16. Oktober 2018
Persönliches Treffen (Thema & Inhalt)	10 bis 14. Dezember 2018
Abgabe der Ausarbeitung & Gemeinsame Abschlusspräsentation	25. Januar 2018, 13.00 Uhr

Organisation

■ Anmeldung im Portal:

- campus.studium.kit.edu → Prüfungen → Prüfungsanmeldung
→ Proseminar-Modul
→ Es erscheint eine Liste mit den angebotenen Proseminare
→ **7500238 Programmieren in natürlicher Sprache**
→ Anmeldung (auf der Warteliste)

Anmeldebeginn: 08.11.2018 12 Uhr

Anmeldeschluss: 29.11.2018 12 Uhr

Wer es nicht im Portal sehen kann, meldet sich hier an:

- Für das Proseminar (PRN 201) für das Wintersemester 2018/2019

Sinn und Ziel des Proseminars

- Im Proseminar geht es in erster Linie darum, den Studenten das wissenschaftliche Arbeiten beizubringen
- Kein Programmieren!
- Ein Teilnehmer bekommt mehrere Forschungsartikeln zu je einem Thema zum selbstständigen Einarbeiten und schreibt den Inhalt in einer Ausarbeitung zusammen
- Zuletzt findet ein Abschlussvertrag zu jedem Thema statt
- Proseminar bereitet Euch auf die Abschlussarbeit vor!

Programmieren in natürlicher Sprache

- Siri von Apple
- Google Voice auf Android Smartphones
- Windows Speech Recognition von Microsoft
- Erkennung und Verarbeitung natürlich geschriebener und gesprochener Sprache
- Programmieren anhand Anweisungen und Beschreibungen der Benutzer



Wissenschaftliches Arbeiten: How To

- URLs in Fußnoten oder ins Literaturverzeichnis, nicht in den Fließtext
- Zitate korrekt und vollständig
 - Zitate nicht als Fußnote, sondern “amerikanisch”
 - Abschreiben kann schwere Folgen haben
- Originalliteratur verwenden
- Wir verwenden Bibtex
 - Sehr effizienter Aufbau des Verzeichnisses
 - Pflege des Verzeichnisses über eine Software, nicht von Hand
- Wikipedia ist nicht verboten, aber Achtung Ungenauigkeit!
 - Aber: Explizit nicht empfohlen als (einzige) Quelle!

Schriftliche Ausarbeitung: How To

- **Umfang:** ca. 15 Seiten
- **Vorlage:** Vorlage vom Lehrstuhl

- **Einleitung**

- **Euer Thema**
 - Entwurf...
 - Syntax...

- **Zusammenfassung**

- **Literatur**

Abschlusspräsentation: How To

- **Umfang:** ca. 15 Folien
- **Zeit:** 10 Minuten Vortrag
- **Vorlage:** Vorlage vom Lehrstuhl

- **Einleitung**

- **Euer Thema**
 - Entwurf...
 - Syntax...

- **Zusammenfassung**

- **Literatur**

Thema 1

Thema

- Aufbau von Intelligenten Systemen

Literatur

- Bruce Ballard and Alan Biermann. *Programming in natural language: NLC as a prototype*. Association for Computing Machinery. 1979
- Mitchell Waldrop, *Natural Language Understanding*, 2003
- Artikeln des Lehrstuhls zum Thema

Aufgabe

- Darstellung erster Ansätze:
 - Ideen, Einschränkungen, Ergebnisse & Bewertung
- Vergleich mit modernen Ansätzen, zeitliche Entwicklung

Thema 2

Thema

- Natural Language Processing

Literatur

- Christopher D. Manning. *The Stanford CoreNLP Natural Language Processing Toolkit*. 2014.
- Hugo Liu. ConceptNet — a practical commonsense reasoning toolkit. 2004
- Computerlinguistik und Sprachtechnologie – eine Einführung, 2010

Aufgabe

- Vorstellung, Techniken & Methodik
- NLU, DM, NLG

Thema 3

Thema

- Übersicht der NLP Tools

Literatur

- AddStructure
- Angel.ai
- API.ai
- StanfornNLP, SharpNLP, OpenNLP
- ParZu
- TensorFlow

Aufgabe

- Vorstellung, Techniken & Methodik

Thema 4

Thema

- Aufbau einer Wissensbasis

Literatur

- Didier Guzzoni, Charles Baur and Adam Cheyer. *Active, a Platform for building intelligent software*. 2006
- Johan Schalkwyk. *Google Search by Voice: A case study*. 2010
- QnAMaker

Aufgabe

- Vorstellung, Bewertung, Vergleich, Stand der Technik
- Patterns, Ontologien, Active Ontologien, SQL, ...

Thema 5

Thema

- End User Programming

Literatur

- Brad A. Myers, Andrew J. Ko, and Margaret M. Burnett, *Invited research overview: end-user programming*. 2006
- Henry Liebermann. *End-User Development: An Emerging Paradigm*. 2006

Aufgabe

- Was ist EUP? Was ist unter EUD zu verstehen?
- Darstellung des Forschungsgebiets
- Abgrenzung und Vergleich der Ansätze

Thema 6

Thema

- Künstliche Intelligenz bei natürlicher Sprachverarbeitung

Literatur

- Ray Kurzweil, *How to make a mind*. 2013
- George Ferguson, *TRIPS: an integrated intelligent problem-solving assistant*. 1998.
- Goh Ong Sing, *Towards a more natural and intelligent interface with embodied conversation agent*. 2006

Aufgabe

- Darstellung des Forschungsgebiets
- Abgrenzung und Vergleich der Ansätze

Thema 7

Thema

- Maschinelles Lernen: Entscheidungsbäume, Klassifikatoren, Neuronale Netze

Literatur

- Miroslav Kubat, *An Introduction to Machine Learning*. 2015
- Rosenblatt, F.: The Perceptron: A probabilistic model for information storage and organization in the brain, 1958, S. 386-408
- Pitts, W. ; McCulloch, W. S.: A Logical Calculus of Ideas Immanent in Nervous Activity, 1943, S. 115-133

Aufgabe

- Darstellung, Abgrenzung, Typischer Einsatz, Bewertung

Thema 8

Thema

- Dialogsysteme fürs Programmieren

Literatur

- Sumit Gulwani, NLyze: interactive programming by natural language for spreadsheet data analysis and manipulation, 2014
- Alexander Wachtel, NLUI, 2018

Aufgabe

- Darstellung, Vergleich, Bewertung

Thema 9

Thema

- Bot Frameworks

Literatur

- Microsoft Language Understanding Intelligent Service (LUIS)
- Watson
- Pandorabots

Aufgabe

- Was ist ein Bot Framework?
- Darstellung, Vergleich, Bewertung

Thema 10

Thema

- Übersicht ChatBot Tools

Literatur

- CleverBot
- Mitsuku
- Tutor
- Colombina
- Uberbot

Aufgabe

- Was ist ein Bot?
- Darstellung, Vergleich, Bewertung

Thema 11

Thema

- Question Answering Systeme

Literatur

- David Ferucci, *Building Watson: an overview of the DeepQA Project. 2010*

Aufgabe

- Darstellung, Abgrenzung, Bewertung
- Antwortgenerierung
- Methodik

Inhaltlicher Überblick und Teilnehmer

Themen	Bearbeiter
1. Aufbau von Intelligenten Systemen	Cristian Albu
2. Natural Language Processing	Luise Häuser
3. Übersicht der NLP Tools	Oliver Wirth
4. Aufbau einer Wissensbasis	Daniel Jungkind
5. End User Programming	Theo Pötzl
6. Künstliche Intelligenz bei natürlicher Sprachverarbeitung	Thomas Bartel
7. Maschinelles Lernen	Philipp Weinmann
8. Dialogsysteme fürs Programmieren	Niklas Betten
9. Bot Frameworks	Marius Kilian
10. Übersicht der ChatBot Tools	Jennifer Frankenfeld
11. Question Answering Systeme	Alexander Heck

