

PG: Time Machine

Oliver Tüselmann, Fabian Wolf und Gernot A. Fink

Mustererkennung,
Fakultät für Informatik

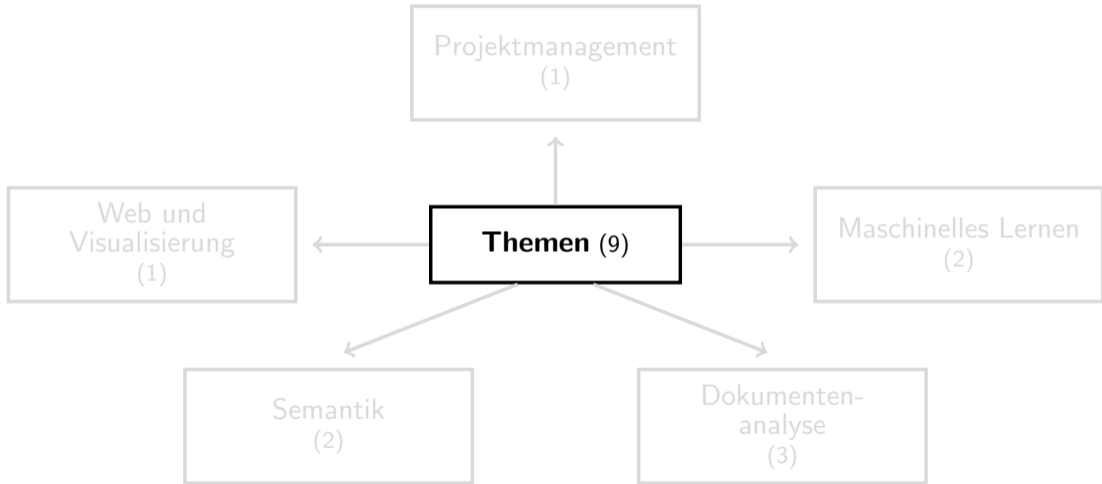
Erstes Treffen: 2. September 2020

- ▶ Seminarphase
- ▶ Themenübersicht
- ▶ Ablauf
- ▶ Social Events
- ▶ Literatur

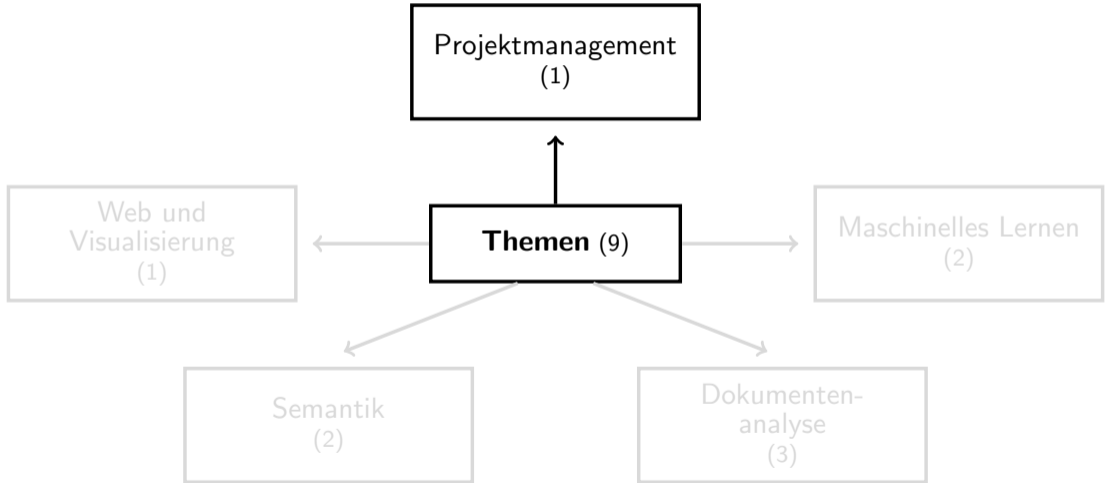
Seminarphase

- ▶ Termine: Woche vor dem Semester 26.10-30.10 (Doodle)
- ▶ Teil A: **Präsentationen**
 - ▶ 9 Themen
 - ▶ 15 Minuten pro Person
 - ▶ Einarbeitung in Themenbereich
 - ▶ Vorstellung der essentieller Methoden und Literatur
- ▶ Teil B: **Workshop**
 - ▶ Erarbeitung der Organisationsstruktur
 - ▶ Bilden der Bearbeitungsgruppen
 - ▶ Festlegen erster Zwischenziele

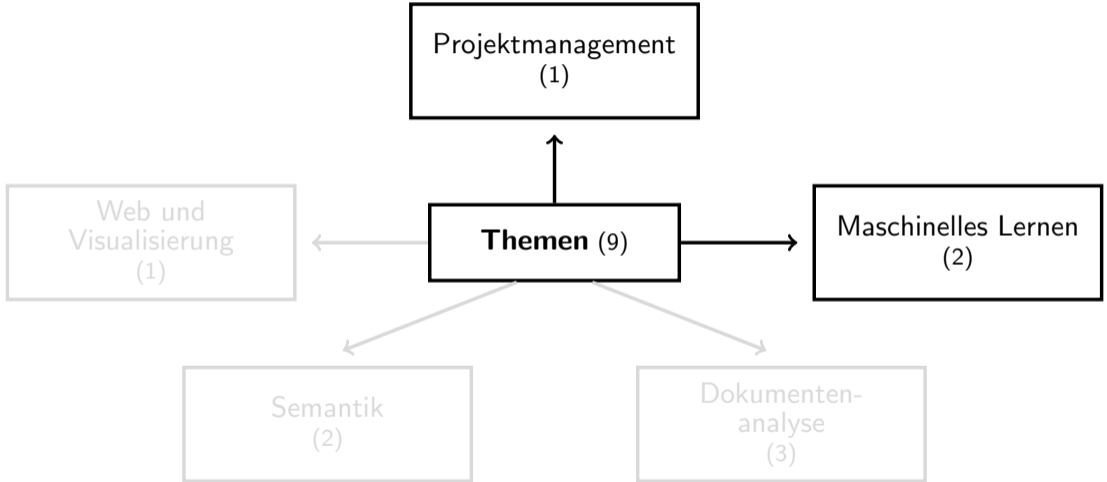
Themenübersicht



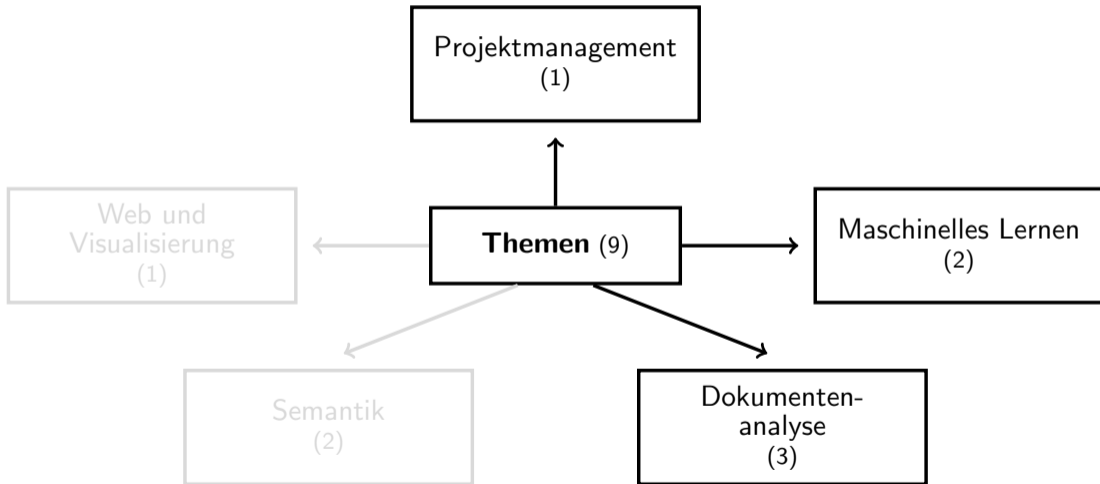
Themenübersicht



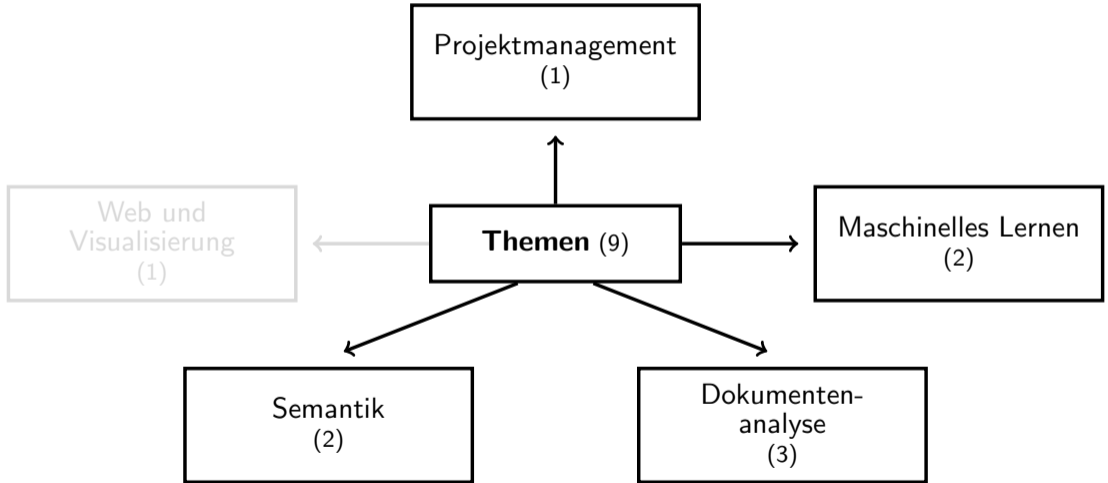
Themenübersicht



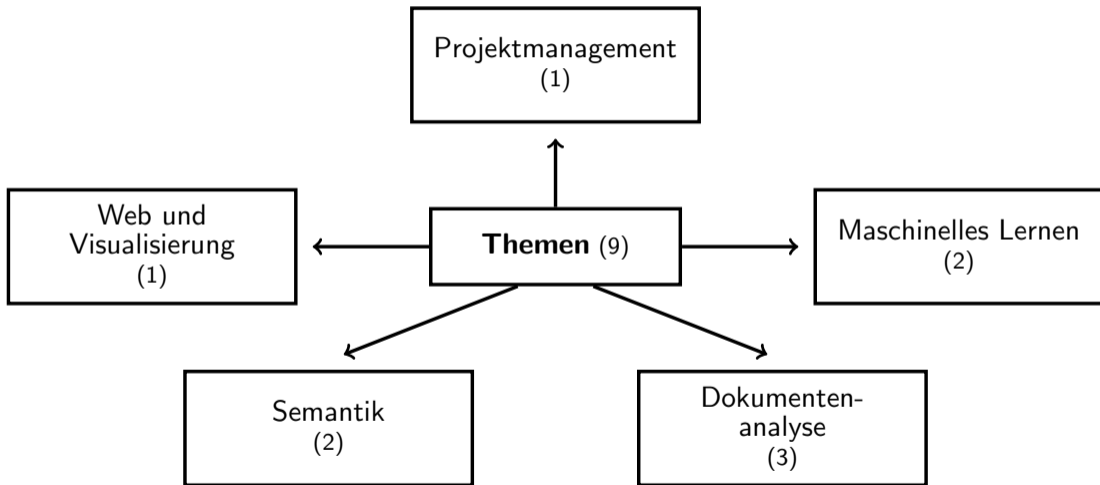
Themenübersicht



Themenübersicht



Themenübersicht



Projektmanagement

1950 1960 1970 1980 1990 2000 2010

2020 2030 2040 2050 2060 2070 2080

1. Projektmanagement



Leitfragen:

- ▶ Wie managen wir die Arbeit in unserer Projektgruppe?
- ▶ Welche Tools sind hilfreich: Slack, Trello, Zoom, GitLab ...?
- ▶ Umfrage unter den Teilnehmern: Erfahrung mit X
- ▶ Spezielle Herausforderung: Arbeiten im Home Office

Maschinelles Lernen

1950 1960 1970 1980 1990 2000 2010

2020 2030 2040 2050 2060 2070 2080 2090

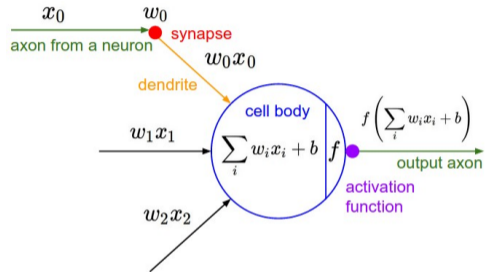
2. Datensätze



Leitfragen:

- ▶ Welche wissenschaftliche Benchmarks existieren?
 - ▶ historische Dokumente
 - ▶ Tasks und Protokolle: Word Spotting [GSGN17], Schrifterkennung [PC16], Layout. . .
 - ▶ Competitions (ICFHR, ICDAR, DAS, CVPR)
- ▶ Postkartendatensatz: Vorbereitung und Vortstellung des aktuellen Stands

3. Deep Learning and Neural Networks



Leitfragen:

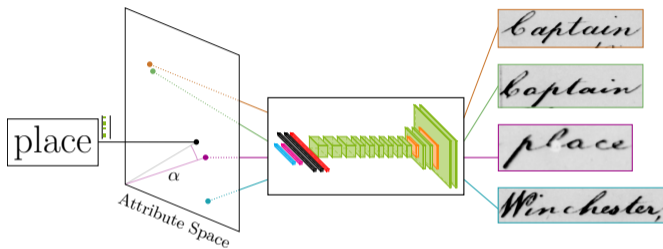
- ▶ Was sind neuronale Netze, insbesondere Faltungsnetze? [GBC17]
- ▶ Welche Schichten werden verwendet? [FFJK17]
- ▶ Training und Optimierung [Nie15]

Dokumentenanalyse

1950 1960 1970 1980 1990 2000 2010

2020 2030 2040 2050 2060 2070 2080 2090

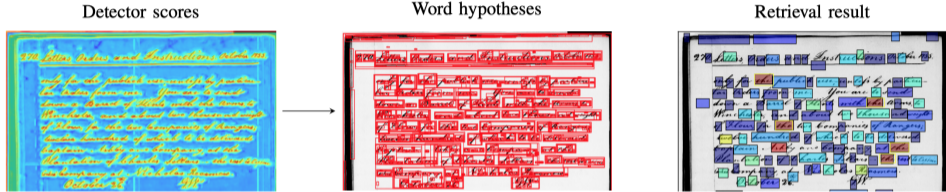
4. Word Spotting



Leitfragen:

- ▶ Was ist Word Spotting und wie funktioniert es? [GSGN17]
- ▶ Taxonomie: Query by String/Example, segmentation-based/free. . .
- ▶ Klassische Methoden: Deskriptorbasiert, Spatial Pyramid
- ▶ Deep Learning: PHOCNet [SF16, SF18]

5. Segmentierung und Layout



Leitfragen:

- ▶ Segmentierungsprobleme: Zeilenextraktion, Wordsegmentierung (Hypothesen) [PSKC10, RSR⁺17, GSGN17]
- ▶ Welche Layoutprobleme aus der Literatur könnten relevant sein?
- ▶ Graphical Object Detection zur Extraktion von Poststemepeln oder Briefmarken?

6. Handwriting Recognition / OCR

Optical Character Recognition
is designed to convert your
handwriting into text.

Optical Character Recognition
is designed to convert your
handwriting into text.

Leitfragen:

- ▶ Überblick OCR Methoden für Handschrift
- ▶ Aktueller Stand der Forschung [Pui17, PC16]
- ▶ Einschätzung zur Anwendbarkeit auf Postkarten Datensatz

Semantik

1950 1960 1970 1980 1990 2000 2010

2020 2030 2040 2050 2060 2070 2080 2090

7. Chatbots



Leitfragen:

- ▶ Welche Systeme können genutzt werden, Vor- und Nachteile? [DCLT19, NH18]
- ▶ Entwicklung eines eigenen Chatbots?
- ▶ Vorgehen für Entwicklung beschreiben

8. Question Answering

Passage Sentence

In meteorology, precipitation is any product of the condensation of atmospheric water vapor that falls under gravity.

Question

What causes precipitation to fall?

Answer Candidate

gravity

Leitfragen:

- ▶ Übersicht der Ansätze
- ▶ Aktueller Stand der Forschung [LCG⁺20, RZLL16]
- ▶ Deutsch als Eingabesprache problematisch?
- ▶ Einschätzung zur Anwendbarkeit auf Postkarten Datensatz



Web und Visualisierung

1950 1960 1970 1980 1990 2000 2010

2020 2030 2040 2050 2060 2070 2080 2090

9. Web und Visualisierung



Leitfragen:

- ▶ Was gibt es für Frameworks, Vor und Nachteile?
- ▶ Datenspeicherung: Wie und wo werden die extrahierten Daten gespeichert?
- ▶ Welche Hardware wird benötigt?
- ▶ Empfehlungen: Python und Django

Planung für Seminarphase

- ▶ Email mit Prioritäten (1-9)
- ▶ Eine Woche vor dem Seminar: Folienentwurf (Grobgliederung)
- ▶ Seminarphase 26.10-30.10 zwei Termine a zwei Stunden

PG-Arbeitsaufwand

Auszug Modulhandbuch:
Modul INF-MSc-101: Projektgruppe

- ▶ Projektgruppe: 25 Credits
- ▶ 30 Stunden/Credit = 750 Stunden
 - ▶ Bei 50 Wochen im Jahr = 15 Stunden/Woche
 - ▶ Bei 50 Wochen im Jahr minus 13 Wochen Semesterferien im Jahr = 20 Stunden/Woche
- ▶ 4 Stunden pro Tag oder 2,5 ganze Tage (a 8 Stunden) pro Woche

Ablauf generell

- ▶ Alle zwei Wochen ein Treffen
- ▶ Zwischenbericht / Endbericht
- ▶ Fachgespräch für Zwischen- und Endbericht
- ▶ Arbeitsort: Homeoffice
- ▶ Infrastruktur: Zugang zu GPU Maschine

Social Events

**Houseparty
is a
face to face
social
network.**

[DOWNLOAD](#)



References I

- [DCLT19] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova.
BERT: Pre-training of deep bidirectional transformers for language understanding.
In *Proc. Conf. of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, volume 1, Minneapolis, Minnesota, jun 2019. Association for Computational Linguistics.
- [FFJK17] Li Fei-Fei, Justin Johnson, and Andrej Karpathy.
Cs231n: Convolutional neural networks for visual recognition., 2017.
- [GBC17] Ian Goodfellow, Yoshua Bengio, and Aaron Courville.
Deep Learning.
The MIT Press, 2017.
- [GSGN17] Angelos P. Giotis, Giorgos Sfikas, Basilis Gatos, and Christophoros Nikou.
A survey of document image word spotting techniques.
Pattern Recognition, 68:310–332, 2017.

References II

- [LCG⁺20] Zhenzhong Lan, Mingda Chen, Sebastian Goodman, Kevin Gimpel, Piyush Sharma, and Radu Soricut.
Albert: A lite bert for self-supervised learning of language representations.
ArXiv, abs/1909.11942, 2020.
- [NH18] Mohammad Nuruzzaman and Omar Hussain.
A survey on chatbot implementation in customer service industry through deep neural networks.
In *International Conference on e-Business Engineering*, pages 54–61, 10 2018.
- [Nie15] Michael Nielsen.
Neural Networks and Deep Learning.
Determination Press, 2015.
- [PC16] Ayush Purohit and Shardul Chauhan.
A literature survey on handwritten character recognition.
International Journal of Computer Science and Information Technology, 7:1–5, 02 2016.

References III

- [PSKC10] Vassilis Papavassiliou, Themis Stafylakis, Vassilis Katsouros, and George Carayannis. Handwritten document image segmentation into text lines and words. *Pattern Recognition*, 43(1):369–377, 2010.
- [Pui17] J. Puigcerver. Are multidimensional recurrent layers really necessary for handwritten text recognition? In *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, volume 01, pages 67–72, 2017.
- [RSR⁺17] Leonard Rothacker, Sebastian Sudholt, Eugen Rusakov, Matthias Kasperidus, and Gernot A. Fink. Word hypotheses for segmentation-free word spotting in historic document images. In *Proc. Int. Conf. on Document Analysis and Recognition, Kyoto, Japan, 2017*.
- [RZLL16] Pranav Rajpurkar, Jian Zhang, Konstantin Lopyrev, and Percy Liang. SQuAD: 100,000+ questions for machine comprehension of text. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 2383–2392, Austin, Texas, November 2016. Association for Computational Linguistics.

References IV

- [SF16] Sebastian Sudholt and Gernot A. Fink.
PHOCNet: A deep convolutional neural network for word spotting in handwritten documents.
In *Proc. Int. Conf. on Frontiers in Handwriting Recognition*, pages 277–282, Shenzhen, China, 2016.
- [SF18] Sebastian Sudholt and Gernot A. Fink.
Attribute CNNs for word spotting in handwritten documents.
Int. Journal on Document Analysis and Recognition, 21(3):199–218, 2018.