

Webinar

Textverstehen mit
Künstlicher Intelligenz

Know-how zu maschineller Textanalyse und Datenerfassung

Grundlagen, praktische Übungen und Tooldemonstrationen



Grundlagen der maschinellen Textanalyse

Die Teilnehmenden lernen im fünfteiligen Webinar, welche Möglichkeiten die maschinelle Analyse von Textdokumenten heute schon bietet. Zudem erfahren sie, welche regelbasierten Verfahren und welche Ansätze des maschinellen Lernens es gibt und für welche Fragestellungen diese optimal geeignet sind. Mittels Praxisbeispielen werden die Chancen der Datenextraktion und der Klassifikation von Texten verständlich gemacht und angewendet.



Herausforderungen und Lösungen für die Datenerfassung

Ebenso werden Herausforderungen und moderne Lösungsansätze für die Datenerfassung, wie Sprach- oder Texterkennung (OCR) und weitere Tools, betrachtet. Die Teilnehmenden vertiefen ihr Wissen in praktischen Übungen und Tooldemonstrationen.

15.09.2020

Bereits **heute** am Morgen auf dem Weg zur Arbeit um **06.42 Uhr** musste ich feststellen, dass mein **Bus der Linie 53 Richtung Feuerbach Bahnhof** (wieder einmal!!) 15 Minuten **zu spät kam**.

...

Unstrukturierte Daten



Anliegen	Beschwerde
Thema	Verspätung
Absender	
Ereignisdatum	15.09.2020
Ereignisuhrzeit	06:42
Verkehrsmittel	53
Haltestelle	
Richtung	Feuerbach Bahnhof
Sentiment	neutral
Erf. Aktion	Antwortschreiben

Strukturierte Daten

Übersicht der Inhalte des Webinars

6 Module für einen umfassenden Einblick ins Thema.

Einführung Textverstehen

- Beispielanwendungen für Textverstehen
- Kernaufgabe: Gewinnung von Fakten aus Texten
- Typische Ein- und Ausgabekanäle
- Klassifikation von Textverstehen-Anwendungen
- Natural Language Processing – Grundlagen und Begriffe
- Pipeline-Modell der Sprachverarbeitung
- Praktische Übung: Satztrennung, Tokenisierung, POS und Syntaxanalyse

OCR und Sprach- erkennung

- Vorverarbeitung, Zeichenerkennung und Layouterkennung
- OCR Ausgabeformate
- Standardablauf in Input-Management-Lösungen
- Spracherkennung – Ablauf und Herausforderungen
- Praktische Übung zu OCR und Ausgabeformaten

Klassifikation und Merkmalsextraktion

- KI vs. regelbasierte Verfahren
- Data Science Aspekte: Machine Learning Algorithmen, Trainingsdaten, Überangepasstheit
- Klassifikation: Dokumentklassen erkennen, Sentiment Detection, Topic Detection
- Extraktion von Merkmalen aus Texten: Verschiedenartige Merkmale und Lösungsansätze für die Extraktion
- Named Entity Recognition
- Erkennen von Zeitangaben
- Auswahl und Konfiguration der passenden Extraktionswerkzeuge
- Reguläre Ausdrücke - Basiskenntnisse
- Erkennung von Kfz-Kennzeichen, Bankverbindungen und Datumsangaben mittels Regex
- Praktische Übungen zu Klassifikation, Extraktion und Regulären Ausdrücken

Maschinelle Lernverfahren

- KI, Maschinelles Lernen und Neuronale Netze
- Aufbereitung von Texten für Maschinelle Lernverfahren: Bag-of-Words, Word Embeddings
- Einsatz von Machine Learning Frameworks für Textverstehen
- Verfügbarkeit von Trainingsdaten und Annotieren von Texten
- Textanalyse mit spaCy
- Praktische Übung zum Training von Machine Learning Modellen mit spaCy

Ergebnisse und Evaluation

- Die Wahrheitsmatrix
- Accuracy, Precision, Recall, F-Score
- Grenzen der Standardmaße
- Besonderheiten bei der Evaluation der Merkmalsextraktion
- Nutzenüberlegungen

Aus der Forschung

- AutoML
- Erzeugung synthetischer Daten
- Foundation Models
- Generative Modelle
- Explainable AI

Eckdaten und Anmeldung

Weitergehende Infos zum Webinar



Dauer und Termine

Das Webinar findet an fünf Terminen vom 22. November bis zum 1. Dezember (22.11.2022 / 24.11.2022 / 29.11.2022 / 30.11.2022 / 01.12.2022) von 09:30 bis 11:30 Uhr statt.



Veranstaltungsort

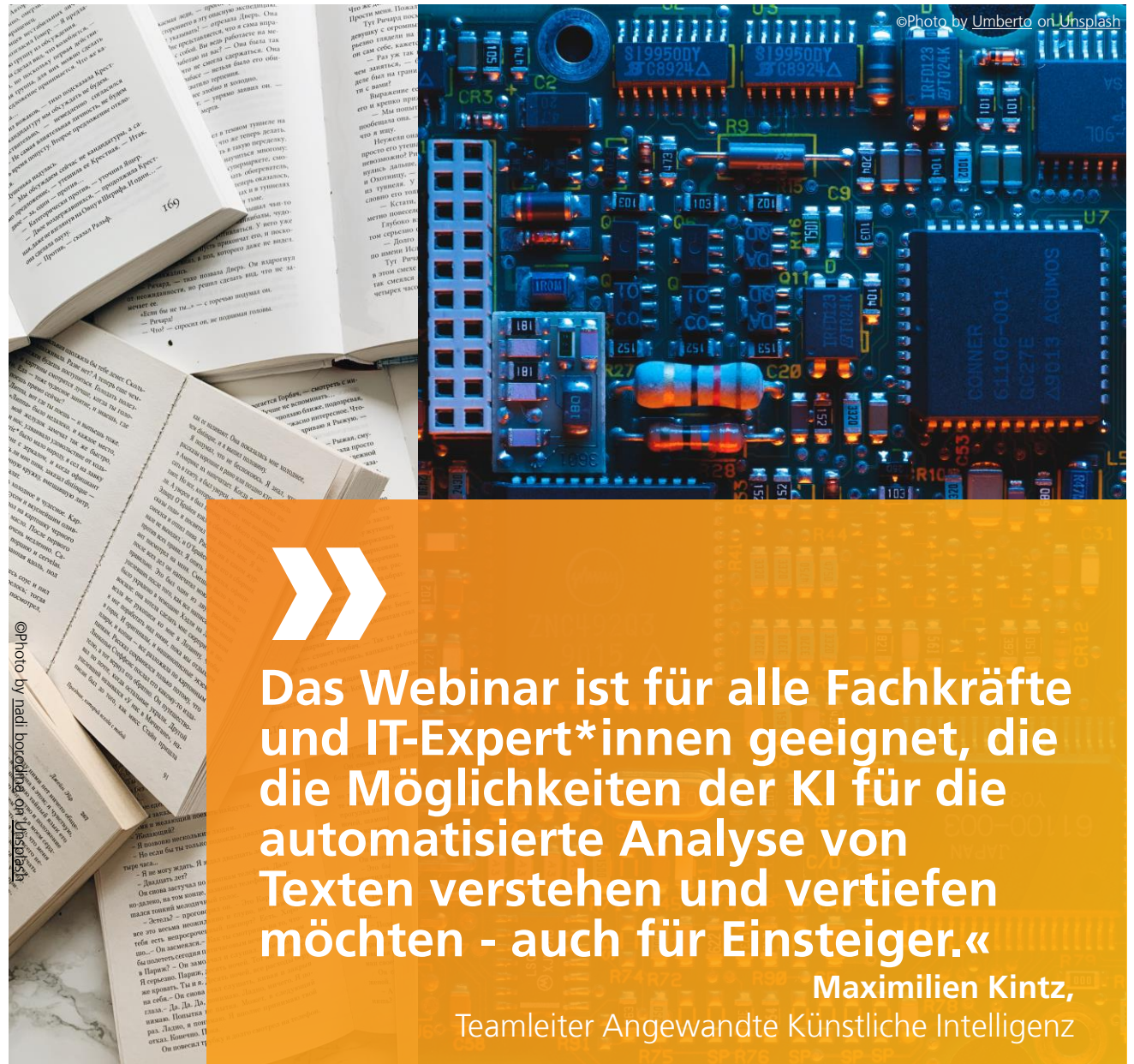
MS Teams



Kosten

Die Teilnahmegebühr für die Veranstaltung beträgt 445€ pro Person.

[Jetzt anmelden!](#)



Das Webinar ist für alle Fachkräfte und IT-Expert*innen geeignet, die die Möglichkeiten der KI für die automatisierte Analyse von Texten verstehen und vertiefen möchten - auch für Einsteiger.«

Maximilien Kintz,
Teamleiter Angewandte Künstliche Intelligenz

Die Referenten

Ihre »Textverstehen mit KI«-Ansprechpartner



Matthias Engelbach

Team

»Angewandte Künstliche Intelligenz«

Schwerpunkte

Maschinelle Sprachverarbeitung,
Big Data, Text-Mining und KI

Telefon

+49 711 970-5211

E-Mail

matthias.engelbach@iao.fraunhofer.de



Jens Drawehn

Team

»Angewandte Künstliche Intelligenz«

Schwerpunkte

Textverstehen
Process Mining

Telefon

+49 711 970-2407

E-Mail

jens.drawehn@iao.fraunhofer.de



Jan Finzen

Team

»Digital Business«

Schwerpunkte

Text- und Sprachverarbeitung,
Softwareentwicklung

Telefon

+49 711 970-2432

E-Mail

jan.finzen@iao.fraunhofer.de



Maximilien Kintz

Team

»Angewandte Künstliche Intelligenz«

Schwerpunkte

Informationsvisualisierung
Textverstehen

Telefon

+49 711 970-2182

E-Mail

maximilien.kintz@iao.fraunhofer.de