



## Webinar

Texte verstehen, verarbeiten und erzeugen mit Generativer KI

# 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

Text-  
verstehen

<b>Anliegen</b>	Beschwerde
<b>Thema</b>	Verspätung
<b>Absender</b>	
<b>Ereignisdatum</b>	15.09.2020
<b>Ereignisuhrzeit</b>	06:42
<b>Verkehrsmittel</b>	53
<b>Haltestelle</b>	
<b>Richtung</b>	Feuerbach Bahnhof
<b>Sentiment</b>	neutral
<b>Erf. Aktion</b>	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 Übungen: 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
- Dokumentklassen erkennen, Sentiment & Topic Detection
- Merkmalsextraktion, u.a. Named Entity Recognition und Zeitangaben
- Reguläre Ausdrücke
- Praktische Übungen zu Klassifikation und Extraktion

## 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

## Generative KI und Große Sprachmodelle

- Meilensteine der Entwicklung leistungsfähiger Sprachmodelle
- Begriffsklärung: Generative KI, Transformer-Architektur, Masked Language Modeling, Attention Mechanismus, Encoder und Decoder
- Training und Anpassung von großen Sprachmodellen
- Überblick aktuelle Sprachmodelle
- Bildgenerierung
- Anwendungsfelder Generativer KI
- Demos zum Einsatz Generativer KI: Generierung von Texten, Bildern, Programmcode

## Ergebnisse und Evaluation

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

## Aus der Forschung

- Foundation Models
- Multimodale KI-Modelle
- Video- und 3-D Generierung
- Explainable AI

# Eckdaten und Anmeldung

## Weitergehende Infos zum Webinar



### Dauer und Termine

Das Webinar findet an drei Terminen vom 23. bis zum 25. April 2024 jeweils von 10:00 bis 12:00 Uhr und von 13:00 bis 15:00 Uhr statt.



### Veranstaltungsort

MS Teams



### Kosten

Die Teilnahmegebühr für die Veranstaltung beträgt 745€ pro Person. Sonderkonditionen: Jede weitere teilnehmende Person eines Unternehmens erhält einen Rabatt von 100 €.

Stornierung:

bis 26.03.2024 | 75€

ab 27.03.2024 | volle Teilnahmegebühr

[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