

## **AI030 AI Data & Information Specialist**

### **Kurzbeschreibung:**

Der Zertifikatslehrgang **AI030 AI Data & Information Specialist** bildet Sie zu einem KI-Experten für Daten- und Informationsverarbeitung aus.

Die Fähigkeit Daten zu visualisieren ist für viele Bereiche der Datenanalyse und des Machine Learning von großer Bedeutung. Oftmals ist es schwierig, Zusammenhänge in Daten zu erkennen, ohne sie visuell darzustellen. Der Lehrgang vermittelt die Grundlagen ansprechender Datenvisualisierung sowie deren Bedeutung für die Datenanalyse.

Darüber hinaus behandelt der Lehrgang Grundlagen des Prompt Engineering explizit für technische Fach- und Entwicklungskräfte. Ziel ist es zu verstehen, wie effektive Prompts aufgebaut und mit Daten angereichert werden können, um verschiedenste Use Cases zu implementieren. Die Teilnehmer erleben das Potenzial der Technologie und lernen Basistechniken zur Integration von LLMs (Large Language Models) kennen.

Um datengetriebene Entscheidungen treffen zu können, müssen Informationen in strukturierte Daten überführt werden. Der Lehrgang vermittelt Techniken zur Datenextraktion und -transformation, wobei sowohl etablierte Verfahren des Natural Language Processing (NLP) als auch neue Möglichkeiten durch Large Language Models (LLMs) wie OpenAI und **Aleph Alpha** vorgestellt werden.

Die Teilnehmer lernen zudem, wie sie mithilfe großer Sprachmodelle mehr relevante Information in weniger Zeit erfassen können. Dabei geht es nicht nur um die Generierung von Zusammenfassungen, sondern auch um die weitere Nutzung der gewonnenen Information z.B. zur automatisierten Erstellung von Dokumenten in unterschiedlichen Ziel-Formaten. Der Kurs behandelt Techniken zur Wissensextraktion, Textkomprimierung und -kompilierung sowie den Export in gewünschte Formate wie PDF und docx.

Der Lehrgang führt Sie zudem ein in die Konzepte der automatisierten Audio-Transkription und zeigt, wie Künstliche Intelligenz dazu verwendet werden kann, natürliche Sprache in strukturierte Informationen umzuwandeln.

### **Zielgruppe:**

- Entwickler
- IT-Fachkräfte

### **Voraussetzungen:**

- AI020 AI & Data Science Practitioner

oder mindestens

- AI200 Einführung in Python für Data Science und KI (alternativ Grundkenntnisse in Python)
- AI220 Einführung in Machine Learning

**Sonstiges:**

**Dauer:** 5 Tage

**Preis:** 2490 Euro plus Mwst.

**Ziele:**

- Grundlagen der Datenvisualisierung erlernen, verschiedene Visualisierungsformen kennenlernen, Visualisierungsbibliotheken in Python verwenden, komplexe Sachverhalte klar darstellen können
- Verstehen von Prompt Engineering, Lernen effektiver Prompting-Techniken, Erkennen von Potenzial und Limitationen, Integration von LLMs
- Informationsextraktion, Named Entity Recognition, Deep Learning, LLMs, Output Parsing, Function Calling, Tool Usage
- Erlernen grundlegender Architektur und Anforderungen, Nutzung großer Sprachmodelle und Python, Daten aus unterschiedlichen Formaten einlesen, Prompting Techniken für die Textkomprimierung und Textkompilierung, Export in gewünschte Formate
- Audiobasierte Sprache in Text umwandeln, Transkripte verarbeiten, Anwendungsfälle erkunden, APIs nutzen können

## Inhalte/Agenda:

- **◆ Modul 1: Einführung in die Datenvisualisierung**
  - ◆ Bedeutung der Visualisierung für die Datenanalyse
  - ◆ Visualisierungsbibliotheken in Python: matplotlib, seaborn, plotly
  - ◆ Verschiedene Visualisierungsformen: Scatter Plots, Bar Charts, Violin Plots, etc.
  - ◆ Grundlagen der Visualisierung: Dos und Don'ts
  - ◆ Erstellen ansprechender Visualisierungen
  - ◆ Übungen und Beispiele zur Anwendung der gelernten Konzepte
  - ◆ Einführung in interaktive Visualisierungen
  - ◆ Best Practices und Tipps zur Auswahl der richtigen Visualisierungsform
  - ◆ Einführung in die Datenvisualisierung
  - ◆ Bedeutung der Visualisierung für die Datenanalyse
  - ◆ Visualisierungsbibliotheken in Python: matplotlib, seaborn, plotly
  - ◆ Verschiedene Visualisierungsformen: Scatter Plots, Bar Charts, Violin Plots, etc.
  - ◆ Grundlagen der Visualisierung: Dos und Don'ts
  - ◆ Erstellen ansprechender Visualisierungen
  - ◆ Übungen und Beispiele zur Anwendung der gelernten Konzepte
  - ◆ Einführung in interaktive Visualisierungen
  - ◆ Best Practices und Tipps zur Auswahl der richtigen Visualisierungsform
- **◆ Modul 2: Einführung Prompt Engineering für Entwickler**
  - ◆ Einführung in Prompt Engineering und Begrifflichkeiten: LLMs, Prompts, Kontext, etc.
  - ◆ Grundlagen verschiedener Prompting-Techniken: Best Practices, Beispiele und Übungen
  - ◆ Anwendungsszenarien: Diskussion verschiedener Use Cases und deren Umsetzung
  - ◆ Möglichkeiten und Limitationen von Prompts: Erkennen von Chancen und Grenzen anhand von Beispielen
  - ◆ Integration von LLMs: Basistechniken zur Einbindung von LLMs in eigene Projekte
  - ◆ Abschlussdiskussion: Erfahrungen, Herausforderungen und zukünftige Entwicklungen
- **◆ Modul 3: Extraktion und Strukturierung von Daten mit KI**
  - ◆ Einführung in die Problemstellung: Notwendigkeit einer Schnittstelle zwischen unstrukturierten Daten und strukturierten Datenbanken und Prozessen
  - ◆ Einführung in Tools und Bibliotheken: Python, NLP-Bibliotheken, LLMs (Open AI / **Aleph Alpha**)
  - ◆ Techniken zur Informationsextraktion: Named Entity Recognition, Deep Learning, LLMs
  - ◆ Output Parsing, Function Calling und Tool Usage: Anwendung der erlernten Techniken zum Extrahieren und Integrieren von Informationen
  - ◆ Viele Praktische Übungen: Anwendung der erlernten Techniken auf realistische Beispiele und Daten
  - ◆ Abschlussdiskussion: Erfahrungen und Herausforderungen / Wie lässt sich das Gelernte auf den eigenen Kontext übertragen
- **◆ Modul 4: Extraktion und Transformation von Informationen mit KI**
  - ◆ Einführung und grundlegende Architektur
  - ◆ Nutzung von großen Sprachmodellen wie OpenAI und **Aleph Alpha** API (Completion, Summarization etc.) und Python
  - ◆ Einlesen von Daten aus unterschiedlichen Formaten und Chunking-Strategien
  - ◆ Textkomprimierungstechniken
  - ◆ Textkompilierungstechniken: Informationen in ein Format wie einen Bericht transformieren
  - ◆ Export in gewünschtes Format wie PDF oder docx
  - ◆ Diskussion über Anwendungsfälle und Strategien für den effektiven Einsatz von KI-basierten Informationstransformation in der Praxis
- **◆ Modul 5: Von automatischer Transkription zu strukturierter Information mit KI**
  - ◆ Einführung in die automatische Audio-Transkription und die zugrunde liegenden Konzepte
  - ◆ Verwendung von API-Lösungen zur Transkription von Audio-Dateien, z.B. Whisper von OpenAI
  - ◆ Verarbeitung der Transkripte und Exploration verschiedener Anwendungsfälle, wie etwa
    - ◆ · Klassifizierung von Kundenfeedback
    - ◆ · Datenextraktion aus transkribiertem Text
    - ◆ · Übersetzung von Text in andere Sprachen
  - ◆ Praktische Übungen zur Implementierung der erlernten Konzepte in eigenen Projekten
  - ◆ Diskussion über mögliche zukünftige Entwicklungen und Herausforderungen im Bereich der automatischen Transkription
- **◆ Zertifikatsprüfung**