UNIVERSITÄT MANNHEIM

UB UNIVERSITÄTSBIBLIOTHEK MANNHEIM



An der **Universitätsbibliothek** ist zum nächstmöglichen Zeitpunkt eine Vollzeitstelle als

Data Scientist (m/w/d)

zu besetzen.

Die Stelle ist im Rahmen des DFG-Projektes **FAIR Corpus: Ein kuratierter Datensatz für das Training und die Evaluation von KI-Assistenzsystemen im Forschungsdatenmanagement und in der Open Science** auf 24 Monate befristet.

Das FAIR-Corpus-Projekt hat zum Ziel, den ersten offen lizenzierten, hochwertigen Datensatz für das Training und die Evaluation von KI-Assistenzsystemen zu erstellen, die speziell auf Forschungsdatenmanagement (FDM) und Open Science (OS) ausgerichtet sind. Das Korpus wird Tausende kuratierte Dokumente enthalten, darunter Datenmanagementpläne, Richtlinien, Schulungsmaterialien, Standards und Open-Access-Veröffentlichungen. Es wird mit Metadaten angereichert und in Formate überführt, die für KI-Training, Feintuning, Evaluation und Retrieval-Augmented Generation geeignet sind. Das Korpus soll eine zentrale Grundlage für die Entwicklung vertrauenswürdiger, domänenspezifischer KI-Assistenzsysteme im Bereich FDM und OS bilden – an Universitäten, Forschungsdatenzentren sowie in nationalen und internationalen Forschungsinfrastrukturen.
Die Arbeitssprachen im Projekt sind Deutsch und Englisch.

Ihr Auftrag:
Sie entwickeln im Rahmen des Projekts einen qualitativ hochwertigen, offen nachnutzbaren Datensatz für KI-Anwendungen im FDM und in OS. Im Mittelpunkt stehen die automatisierte Verarbeitung und Qualitätssicherung von Dokumenten sowie die Entwicklung und Evaluation KI-geeigneter Datensätze für Retrieval-Augmented Generation (RAG) und Question Answering (QA).

Ihre Aufgaben:

- Konzeption und technische Umsetzung des FAIR Corpus einschließlich der Auswahl geeigneter Datenquellen, der Entwicklung von Daten- und Metadatenmodellen sowie der Umsetzung von FAIR- und CARE-Prinzipien unter Berücksichtigung rechtlicher, datenschutzbezogener und ethischer Anforderungen
- Entwicklung automatisierter Verfahren zur Akquise, Verarbeitung, Bereinigung und Qualitätssicherung von Dokumenten und Metadaten

Die **Universität Mannheim** ist eine der führenden Hochschulen in Deutschland mit derzeit rund 12.000 Studierenden an fünf Fakultäten. Insbesondere die Wirtschafts- und Sozialwissenschaften gehören national und international zur Spitzengruppe. Ihren über 2.600 Beschäftigten bietet die Uni Mannheim ein spannendes Arbeitsumfeld mit zahlreichen Benefits.

Die Universitätsbibliothek ist die zentrale Informationsinfrastruktureinrichtung der Universität Mannheim und unterstützt Studium, Lehre und Forschung mit einem breiten Serviceangebot.

Eckdaten

Start: nächstmöglicher Zeitpunkt

Befristung: 24 Monate

Eingruppierung: E 13 TV-L BW

Stunden/Woche: 39,5 Stunden, die Stelle ist grundsätzlich teilbar

Arbeitsort:
Mannheim, Innenstadt, Schloss

Bewerbungsfrist:
29.11.2026

Referenznummer: UB-512

- Erstellung eines RAG-Subkorpus einschließlich der Entwicklung und Erprobung geeigneter Segmentierungs- und Retrievalverfahren
- Entwicklung eines QA-Subkorpus mittels LLM-gestützter Generierung und Qualitätssicherung von Frage-Antwort-Paaren
- Entwicklung und Durchführung von Benchmarks zur Evaluation der Retrieval- und QA-Qualität
- Dokumentation und Bereitstellung reproduzierbarer Datenverarbeitungs- und Evaluationsverfahren sowie Präsentation der Projektergebnisse

Ihr Profil:

Vorausgesetzt werden:

- Abgeschlossenes wissenschaftliches Hochschulstudium (Master oder vergleichbar) in Informatik, Data Science, Computerlinguistik oder einer verwandten Fachrichtung
- Gute Programmierkenntnisse in Python sowie Erfahrung in der Verarbeitung von Text- und Metadaten
- Kenntnisse in Natural Language Processing (NLP) und im Umgang mit Large Language Models (LLMs)
- Interesse an der Entwicklung und Evaluation von KI-basierten Verfahren sowie an Forschungsdatenmanagement und Open Science
- Selbstständige, strukturierte und teamorientierte Arbeitsweise
- Sehr gute Deutsch- und Englischkenntnisse

Wünschenswert sind:

- Erfahrung mit Retrieval-Augmented Generation (RAG), Information Retrieval oder Question Answering (QA)
- Kenntnisse in der Entwicklung automatisierter Datenpipelines, Datenvalidierung und Qualitätssicherung
- Erfahrung mit LLM-gestützter Datengenerierung, Prompt Engineering oder der Evaluation von KI-Systemen
- Kenntnisse im Forschungsdatenmanagement, in Open Science sowie in FAIR- und CARE-Prinzipien
- Erfahrung mit Git, reproduzierbaren Workflows und Softwareentwicklung
- Erfahrung in der wissenschaftlichen Kommunikation und Präsentation von Forschungsergebnissen

Unser Angebot:

- eine anspruchsvolle und abwechslungsreiche Tätigkeit in einem universitären Umfeld mit nationalen und internationalen Partnern
- einen Arbeitsplatz im repräsentativen Barockschloss inmitten des Universitätscampus; Innenstadtlage, unmittelbare Anbindung an den ÖPNV, 5 Minuten zum Hauptbahnhof
- ein flexibles Arbeitszeitmodell, Möglichkeit zur Telearbeit (bis zu 50%), 30 Tage Jahresurlaub
- Möglichkeit zum bezuschussten Erwerb des Deutschlandticket Job
- ein Jahresgehalt nach TV-L E13 und die üblichen Sozialleistungen und Vorteile des öffentlichen Dienstes, wie eine betriebliche Altersvorsorge (VBL)

Datenschutz

Die Informationen zur Erhebung von personenbezogenen Daten bei der betroffenen Person nach Artikel 13 DSGVO können der Homepage der Universität entnommen werden: www.uni-mannheim.de/datenschutz-bei-bewerbungen.

Die Rücksendung der eingereichten Unterlagen erfolgt nur bei gleichzeitiger Übersendung eines ausreichend frankierten Rückumschlags. Andernfalls werden sie nach Abschluss des Bewerbungsverfahrens nach den Vorgaben des Datenschutzrechts vernichtet. Elektronische Bewerbungen werden entsprechend gelöscht.

Bitte beachten Sie, dass eine Gefährdung der Vertraulichkeit und der unbefugte Zugriff Dritter bei einer Kommunikation per unverschlüsselter E-Mail nicht ausgeschlossen werden können.



Innovatives, universitäres Arbeitsumfeld



Flexible Arbeitszeit- und Arbeitsortmodelle (Gleitzeit, Telearbeit)



Familienfreundliche Hochschule



Internationalität



Deutschlandticket



Sportprogramm für alle Beschäftigten

Bewerbungen von Schwerbehinderten werden bei entsprechender Eignung bevorzugt berücksichtigt. Die Universität Mannheim strebt in Bereichen, in denen Frauen unterrepräsentiert sind, die Erhöhung des Frauenanteils an und fordert daher entsprechend qualifizierte Frauen ausdrücklich zur Bewerbung auf.

Wir freuen uns auf Ihre Bewerbung über das Bewerbungsportal der Universität Mannheim bis zum 29.11.2026.

Für fachliche Fragen steht Ihnen Dr. Irene Schumm (Leiterin der Forschungsdatenzentrums) unter

irene.schumm@uni-mannheim.de

gerne zur Verfügung.

Zertifikat seit 2006 audit familiengerechte hochschule

FAMILIE IN DER HOCHSCHULE