

QUALITÄTSSICHERUNG FÜR DIE SEMANTISCHE ANALYSE ARCHIVIERTER GESUNDHEITSAKTEN

Krebs, Lena, DMI GmbH & Co. KG
Müller, Annett, DMI GmbH & Co. KG

Die fortschreitende Digitalisierung von Gesundheitsdaten hat zu einer stetig wachsenden Menge an medizinischen Daten in digitaler Form geführt. Allerdings werden viele relevante Daten nach wie vor in unstrukturierten Freitexten erfasst, was ihre Auswertung erschwert. Um Qualitätsberichte oder Krankenhausstatistiken zu erstellen und verschiedenen Forschungsfragen nachzugehen, ist die Kodierung dieser Informationen mithilfe von Klassifikations- und Ontologiesystemen wie ICD-10, OPS und SNOMED CT von entscheidender Bedeutung.

Das Dienstleistungsportfolio Archivar 4.0 von DMI umfasst die Anwendung von Natural Language Processing (NLP), um medizinische Informationen aus archivierten Akten zu extrahieren. Diese werden mit aktuellen Terminologien klassifiziert und interoperabel nutzbar gemacht. Zur Sicherstellung der Qualität dieser semantischen Analyse haben wir ein umfassendes Qualitätskonzept entwickelt, welches kontinuierlich die extrahierten Daten überprüft und die schnelle Erkennung fehlerhafter Kodierungen ermöglicht. Dafür haben wir Qualitätsindikatoren entwickelt, die als Basis für Prüflisten dienen, welche anschließend manuell überprüft werden.

Durch statistische Analysen identifizieren wir Ausreißer in der Häufigkeit bestimmter Kodierungen, die mögliche Fehler aufzeigen. Das Feedback aus den manuellen Nachbearbeitungen fließt in eine kontinuierliche Verbesserung unserer NLP-Pipeline ein. Dieser Ansatz sorgt für eine klare und präzise semantische Analyse, die eine solide Basis für die Nutzung von Daten in der Gesundheitsforschung und -versorgung bietet.