

<tag-it/> – der Schlüssel zu smarten Inhalten

Metadaten (Tags) spielen in der nächsten Generation von Technischer Dokumentation eine zentrale Rolle. Sie ermöglichen die optimierte Kommunikation zwischen Maschinen, Menschen und Systemen. Die Schlagworte von morgen lauten: personalisierte Inhalte, smarte Dokumentation, intelligente Suchfunktion.

Die Basis bildet Standardisierung: [iiRDS](#) (Intelligent Information Request and Delivery Standard) ermöglicht den strukturierten Austausch von Inhalten. Die Bereitstellung von Inhalten, also Content Delivery, wird plattformunabhängig möglich.

Der entscheidende wirtschaftliche Faktor für Unternehmen liegt in der automatisierten Zuweisung dieser Intelligenz verleihenden Metadaten. Eine KI-basierte Betagging-Lösung beschleunigt den Prozess erheblich, reduziert Kosten und erzielt Ergebnisse, die mit manueller Arbeit nicht zu erreichen wären.

„iiRDS KI-Betagging“ – eine erfolgreiche Kooperation zwischen text-it und dem Campus Hagenberg

Die [text-it Produktdokumentation GmbH](#) und die [Fachhochschule Oberösterreich, Campus Hagenberg](#) haben eine Machbarkeitsstudie für einen KI-basierten Betagger durchgeführt, um die Vergabe von Metadaten für Technische Inhalte zu automatisieren. Die Lösung basiert auf dem iiRDS-Standard der tekomp und nutzt Large Language Models (LLMs) von OpenAI zur automatischen Tag-Generierung.

Im Rahmen des Projekts wurde ein Prototyp entwickelt, der verschiedene Dateiformate (PDF, TXT, DOCX, HTML) unterstützt. Dem Large Language Modell werden extrahierte Textinhalte mit spezifischen Anweisungen übermittelt. Zwei Tagging-Kategorien wurden implementiert:

- Controlled Vocabulary: Tags aus einer vordefinierten (iiRDS-)Liste werden vergeben
- Free-Form: das Large Language Model generiert eigene, kundenspezifische Metadaten

Es wurde ein User Interface (UI) entwickelt, das den Tagging-Prozess visualisiert und es Benutzerinnen und Benutzern ermöglicht, Dokumente hochzuladen, automatisch zu beschlagworten und die Ergebnisse zu prüfen und zu bearbeiten.

The image shows a user interface prototype for document tagging. At the top, there is a toggle switch labeled 'Bearbeiten automatisch befüllter Felder'. Below this, the interface is divided into two main sections: 'Vordefinierte Tags' and 'Generierte Tags'. The 'Vordefinierte Tags' section contains four dropdown menus: 'informationssubject' with tags 'SafetyInstruction', 'Warning', 'Caution', and 'Notice'; 'phaseofproductlifecycle' with tags 'Operation' and 'Transport'; 'topictype' with tags 'GenericTask' and 'GenericTroubles...'; and 'documenttype (Required)' with the value 'Nothing to display'. The 'Generierte Tags' section contains five input fields: 'language (Required)' with 'English', 'title (Required)' with 'General safety instructions', 'productvariant' with 'Write custom tag', 'component', and 'device'.

User Interface Prototyp

Learnings und Potenzial

Die Machbarkeitsstudie zeigte, dass der Einsatz von LLMs für das Betaggen technischer Inhalte äußerst vielversprechend ist und das KI-basierte Betaggen in der technischen Dokumentation großes Potential hat. Ein semi-automatischer Ansatz, bei dem die menschliche Expertise einfließt, scheint derzeit der vielversprechendste Weg zu sein.

Es müssen noch Herausforderungen, insbesondere auch seitens der Standardisierung von iiRDS, gemeistert werden. Es braucht klare Vorgaben für die Beschlagwortung, damit die Vergabe von Tags nach einheitlichen Kriterien erfolgen kann und ein „Unter-, Über- oder Anders-Betaggen“ von Inhalten vermieden wird. Dafür ist ein Leitfaden zum iiRDS-Standard ein wichtiger Schritt.

Die nächsten Schritte

Der Campus Hagenberg und text-it führen das Betagging-Projekt im Rahmen eines Studienprojektes fort. Diese Zusammenarbeit bietet die Gelegenheit, zusätzliche Funktionalitäten zu integrieren und so das volle Potenzial der Technologie zu entfalten und einen zuverlässigen KI-Betagger zu entwickeln.

text-it bietet folgende Beratungsdienstleistungen rund um iIRDS

- Erstellung von kundenspezifischen Metadatenkonzepten
- Modellierung von Wissensgraphen (Wissensnetze)
- Begleitung und Unterstützung bei der Umsetzung