

Intelligent navigieren

Das Angebot von APA-DeFacto wird sich in den kommenden Jahren erweitern; eine große Rolle spielen dabei Automatisierung und verbesserte Technologien.

WIEN. „Die Herausforderung der Zukunft wird sein, immer mehr Informationen zeitnäher bereitzustellen, während auf der Kundenseite immer weniger Zeit bleibt, diese Informationen zu verarbeiten“, so Klemens Ganner, Geschäftsführer des Medienbeobachtungsunternehmens APA-DeFacto, über die Challenge sowohl auf Seite der Kunden, aber auch auf der Seite seines Unternehmens, das versucht, genau diese Herausforderungen mit optimalen Lösungen für Kunden zu meistern.

Die Quadratur des Kreises

Was nach der Quadratur des Kreises klingt, will man bei APA-DeFacto mit dem verstärkten Einsatz maschineller Unterstützung schaffen. Damit einher geht auch eine Ausweitung des Angebots.

„Wir verstehen uns als Informationsdienstleister in der Medienbeobachtung. Dabei wollen wir einerseits genau jene Artikel finden, die der Kunde haben will, und andererseits darüber hinausgehende Informationen extrahieren“, beschreibt Ganner das grundsätzliche Aufgabenfeld und ergänzt: „In Zukunft wird es immer wichtiger werden, über den einzelnen Artikel hinaus die bestimmenden Bran-



© APA

die vermehrt von maschinellen Verfahren unterstützt werden soll. Sogenannte Sentiment-Auswertungen, also die Beurteilung von Tonalitäten (positiv, neutral oder negativ) in Texten spielen dabei eine Rolle. Hier zeigen sich auch die derzeitigen Grenzen automatisierter Verfahren: „Auf der Inhaltsebene ist hier die Herausforderung, wie mit Texten umgegangen wird, in denen Firma A positiv, Firma B aber negativ beurteilt wird. Das kann derzeit automatisiert noch nicht zufriedenstellend ausgewertet werden“, so Ganner – die endgültige Entscheidung über eine Zuordnung trifft bei APA-DeFacto jedenfalls immer eine geschulte Fachkraft im Lektorat.

Machine Learning

Einer der wichtigsten Bereiche, in den künftig auch weiter stark investiert werden soll, ist Machine Learning.

Dabei geht es vor allem um lernende Algorithmen, die beispielsweise PDFs mit mehreren Artikeln auswerten sollen, Kategorisierungen treffen können und Relevanz-Bewertungen zur Selektion vornehmen. Bei der sogenannten PDF-Segmentation hat man bereits einige Fortschritte gemacht: „Das Ziel ist, dass auf einer PDF-Seite automatisiert erkannt wird, was ein Artikel ist und welches Bild zu welchem Artikel gehört“, erläutert Ganner. Dabei „lernt“ der Algorithmus von der täglichen intellektuellen Arbeit der Mitarbeiterinnen und Mitarbeiter und erkennt die Systematik hinter verschiedenen Publikationen. Eine manuelle Endkontrolle werde es dabei aber weiter benötigen, schon um die Qualitätssicherung zu garantieren, so Ganner weiter. „Es geht bei der fortschreitenden Automatisierung vor allem darum, unsere Dienstleistungen besser zu machen und unseren

”

Das Ziel am Ende ist es, dass auf einer PDF-Seite automatisiert erkannt wird, was ein Artikel ist und welches Bild zu welchem Artikel gehört.

chenthemen zu erkennen – was also der Kunde, einschließlich seiner Mitarbeiterinnen und Mitarbeiter, wissen sollte.“

Ein weiterer Punkt, der vor allem im Hinblick auf steigende Informationsverdichtung wichtig wird, ist die Beurteilung von Relevanz. Immer mehr Content auf immer mehr Kanälen macht es notwendig, dass Inhalte nicht bloß ungefiltert an den Kunden geschickt werden – vielmehr braucht es eine entsprechende Kontextualisierung,

“

Klemens Ganner

Seit Mitte 2016 führt Klemens Ganner die APA-DeFacto. Davor war er sieben Jahre Geschäftsführer der Bildagentur APA-Picture-Desk.