



CASE STUDY

SEMANTISCHE TECHNOLOGIE VON RETRESCO MACHT DIE VIELFALT DER MULTIMEDIALEN STASI-MEDIATHEK SICHTBAR

DIE HERAUSFORDERUNG

1992 hat es sich das Stasi-Unterlagen-Archiv zur Aufgabe gemacht, die Hinterlassenschaften der Stasi zu sammeln und zu dokumentieren; 2012 kamen erste Überlegungen auf, Teile dieses Archivs der breiten Öffentlichkeit in einer multimedialen Stasi-Mediathek zugänglich zu machen. Vom Fachpublikum bis zum interessierten Laien – jeder kann sich seitdem über die Machenschaften der DDR-Behörden und der Stasi informieren.

Angesichts einer nach wie vor steigenden Menge an Material bestand eine Herausforderung darin, trotz unterschiedlichem Vorwissen und uneinheitlichen Nutzungsgewohnheiten der Anwender, den Archiv-Bestand für jede Zielgruppe optimal zugänglich zu machen. Als zentrales Element für eine zielgerichtete Nutzerführung kristallisierte sich die Notwendigkeit einer semantischen Suche heraus.

DIE LÖSUNG

Um die Vielfalt des Archivs sichtbar zu machen, pflegen die Archivare seit dem initialen Upload einer großen Menge von Dokumentseiten, Einzelbildern, Tonaufzeichnungen bzw. Filmmaterial fortlaufend weitere Zeitzeugnisse in die interaktive Datenbank ein.

Im Hintergrund analysiert die Retresco-Semantik den Bestand automatisch und identifiziert aus jedem einzelnen Inhalt entscheidende Begriffe - eine wichtige Vorstufe für die passgenaue Zusammenstellung von Rechercheergebnissen.

Im Ergebnis profitieren die Nutzer von einer intuitiven Suche. Das Suchfeld auf <https://www.stasi-mediathek.de/> ist einer der wichtigsten Einstiegspunkte der Recherche, beginnt ein User nach Begriffen zu suchen, sorgt die zusätzliche Funktion „Autovervollständigung“ für die dynamische Unterbreitung weiterer relevanter Schlagworte und Entitäten.

DIE TECHNOLOGIE

Unabhängig von der Art der Inhalte analysiert die Retresco-Semantik Überschriften, Teaser, Texte und Metadaten sämtlicher Dateien. Ein Algorithmus durchsucht dabei den Content auf bestimmte Schlüsselworte und identifiziert sogenannte Entitäten, das sind relevante Personen, Orte, Organisationen, Produkte, Veranstaltungen sowie allgemeine Schlagworte. Zusätzlich wird die Relevanz der einzelnen Entitäten für jedes Dokument berechnet.

Im konkreten Anwendungsfall der Stasi-Mediathek fließen alle Ergebnisse der semantischen Analyse in den Suchindex ein und ermöglichen es dem User, immer die relevantesten Ergebnisse zu einer Suchanfrage zuerst zu erhalten.

Die User haben zusätzlich die Wahl, Inhalte über einen Zeitstrahl zu finden oder sich alle Dateien nach den Formaten Audio, Video, Foto oder Text ausgeben zu lassen. Retresco semantische Suche filtert also sowohl inhaltlich und zeitlich-historisch, als auch nach Beschaffenheit der Unterlagen. Das Resultat ist ein optimaler Zugang zu den Dokumenten der Datenbank.

WIE WIR IHNEN HELFEN KÖNNEN

Sie möchten wissen, wie Lösungen von Retresco auch Ihrem Unternehmen bei der Digitalisierung von Geschäftsprozessen helfen? Erfahren Sie mehr auf unserer Website oder vereinbaren Sie eine persönliche Beratung: welcome@retresco.de