

EIN VORSCHLAG FÜR EIN WEBANGEBOT ZUR GESCHICHTE DER AUTOMATISCHEN SPRACHERKENNUNG

Stephan Euler

*Technische Hochschule Mittelhessen, Fachbereich MND
Wilhelm-Leuschner-Straße 13, 61169 Friedberg
Stephan.Euler@mnd.thm.de*

Kurzfassung: Dieser Beitrag beschreibt einen Vorschlag für ein Archiv zur Entwicklung der automatischen Spracherkennung in Deutschland. Als Inhaltskategorien werden Persönlichkeiten, Technologien, Tagungen und Forschungsprojekte vorgeschlagen. Ein Prototyp vermittelt anhand exemplarischer Einträge einen ersten Eindruck einer Web-Anwendung zur Darstellung der Informationen. Für den weiteren Ausbau werden in Hinblick auf die spezifischen Anforderungen eines potentiell langlebigen Systems spezielle Typen von Content-Management-Systemen vorgeschlagen, die eine klare Trennung von Inhalten und Webtechnologien erlauben.

1 Einleitung

Die automatische Spracherkennung ist mittlerweile im Alltagsleben angekommen. Es ist beispielsweise mehr oder weniger selbstverständlich geworden, mittels Sprache mit dem eigenen Smartphone zu kommunizieren. Hinter dieser Technologie steht jedoch eine lange Zeit der Forschung und Entwicklung. Viele Menschen in verschiedenen Arbeitsgruppen in Hochschulen und Firmen arbeiteten intensiv an diesem Thema. Beispielsweise beschreibt Pieraccini [1] die Evolution von spracherkennenden und –verstehenden Systemen aus Sicht eines Entwicklers. Jetzt scheint ein guter Zeitpunkt, Informationen zu dieser Historie gezielt in einem Archiv zu sammeln. Die Inhalte sollen dabei in Form eines Webangebots (Website) allgemein zugänglich sein. In diesem Beitrag sollen als Anstoß und Basis für die weitere Entwicklung grundsätzliche Möglichkeiten aufgezeigt werden.

Bei der Konzeption eines solchen Angebots sind zwei Ebenen zu unterscheiden: Inhalt und Technik. Im folgende wird zunächst ein Vorschlag für die Inhalte vorgestellt. Dabei dient eine prototypische Implementierung zur Veranschaulichung und erster Erprobung. Darauf aufbauend werden anschließend Möglichkeiten der Realisierung der Webanwendung diskutiert.

2 Inhalte

Bezüglich der Inhalte stellt sich zunächst die Frage nach der Abgrenzung. Angesichts des ohnehin großen Themas soll das Archiv auf die Entwicklung der automatischen Spracherkennung in Deutschland beschränkt bleiben. Dazu gehören die damit nahe verwandten Themen wie Sprecherverifikation oder Sprachenidentifikation. Die Bereiche Kodierung und Synthese sowie Sprachdatenbanken werden – zumindest vorerst – nicht berücksichtigt.

Zur Veranschaulichung wurde eine prototypische Website mit einige ersten Inhalten angelegt¹. Dabei sind die Inhalte in einer Baumhierarchie angeordnet. Abbildung 1 zeigt die Startseite. Sie verweist auf die vier vorgeschlagenen Hauptkategorien:

- Personen, die einen wesentlichen Beitrag zu Grundlagen oder dem praktischen Einsatz der automatischen Spracherkennung geleistet habe
- Forschungsprojekte

¹ <https://hosting.iem.thm.de/user/euler/geschichte/>

- Tagungen mit einem wesentlichen Anteil zur Spracherkennung
- Verwendete Technologien wie Dynamic Time Warp (DTW) und Hidden Markov Models (HMM)

Im Sinne des Archiv-Gedankens sollten die Darstellungen zu den Technologien die zeitliche Entwicklung enthalten. Als Beispiel kann die *Timeline of speech and voice recognition*² dienen. Die Kategorie-Seiten können auch direkt über die Navigationsleiste ausgewählt werden. Jede dieser Seiten besteht wiederum aus einem Inhaltsbereich sowie einer Auflistung der weiteren Unterseiten.

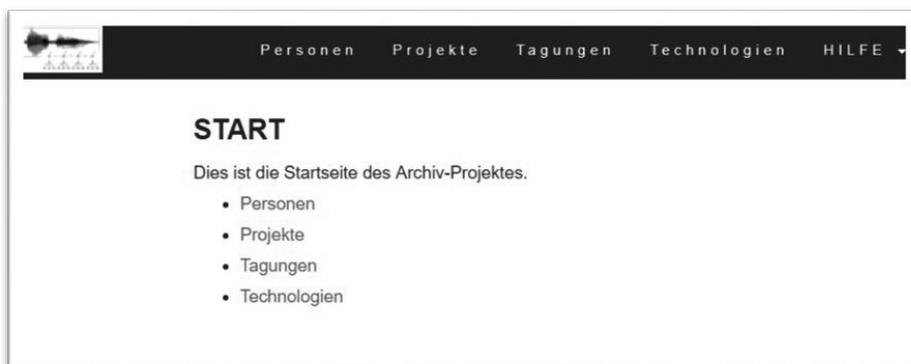


Abbildung 1 – Startseite

So zeigt Abbildung 2 die Unterseite zu Tagungen. Aufgeführt sind Tagungsreihen, bei denen zumindest einige Veranstaltungen thematisch passen. Nato ASI ist ein Beispiel für eine internationale Tagungsreihe. Hierzu ist der Vorschlag, Tagungen daraus aufzunehmen, die in Deutschland stattgefunden haben (in diesem Fall Bad Windsheim 1987). Der erste Verweis in der Aufzählung führt zur – noch unvollständigen – Liste der DFG-Kolloquien *Sprachverarbeitung* (Abbildung 3). Als Beispiel für zusätzliche Inhalte ist ein Auszug aus dem Tagungsprogramm 1977 eingefügt.



Abbildung 2 Unterseite zu Tagungen

² https://en.wikipedia.org/wiki/Timeline_of_speech_and_voice_recognition

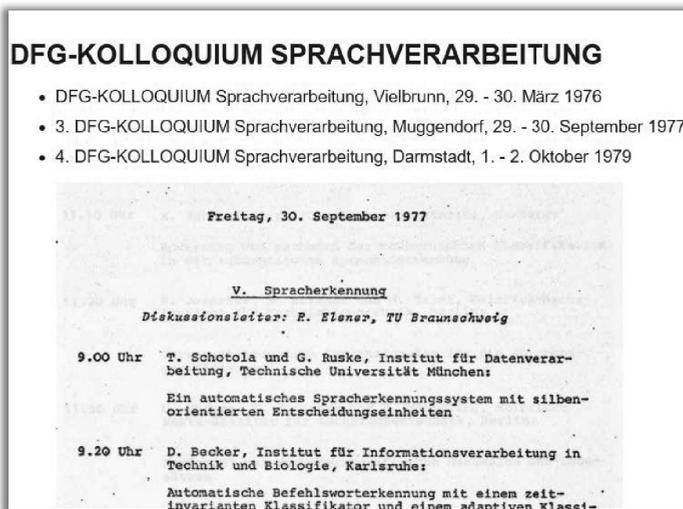


Abbildung 3 Seite zu DFG-Kolloquium Sprachverarbeitung

3 Technologie

Zur technischen Realisierung einer derartigen Website steht eine Vielzahl von Technologien zur Auswahl. Eine grundsätzliche Alternative besteht zwischen der Integration in ein vorgegebenes Content-Management-System (CMS) wie Joomla oder Typo3 oder die Realisierung einer eigenständigen Webanwendung. Für einer eigenen Entwicklung wiederum kommen zahlreiche Frameworks mit unterschiedlichen Programmiersprachen (Java, PHP, JavaScript, Ruby, ...) in Frage.

Im vorliegenden Fall ist die Anwendung als langlebiges Archiv konzipiert. Kern sind die Dokumente – im wesentlichen Texte und Bilder – im eigentlichen Archiv. Die Webanwendung bietet eine Sicht auf diese Dokumente und erlaubt gegebenenfalls die Bearbeitung. Dabei besteht ein wesentlicher Unterschied: Die Dokumente selbst sind in der Regel langlebig während die Webtechnologien einem schnellen Wandel unterworfen sind. Mit einiger Wahrscheinlichkeit werden im Laufe der Zeit zu mindestens Versionswechsel in der Technologie notwendig. Eventuell könnte auch ein Wechsel zu einer anderen Technologie erforderlich werden. Daher sollte auf eine möglichst weitgehende Trennung zwischen Dokumenten und Webtechnologie geachtet werden.

In der gegebenen Anwendung ist weiterhin anzunehmen, dass die Anzahl der Dokumente nur langsam wächst. Auch sind nur selten Änderungen an den Dokumenten zu erwarten. Insgesamt erfordert die Anwendung nur einen geringen Grad an Benutzerinteraktion. Aus diesen Überlegungen ergeben sich folgende Anforderungen:

- Möglichst lose Kopplung zwischen Dokumenten und Web-Sicht
- Darstellung von überwiegend statische Inhalte mit Navigation zwischen den einzelnen Bereichen
- Zwei Benutzerrollen: Besucher und Administrator

Insbesondere der erste Punkt spricht gegen komplexe CMS oder Eigenentwicklungen mit umfangreichen Frameworks. Da – anders als etwa in einem Shop-System – keine dynamischen

Daten zu einzelnen Benutzern oder Inhalten abgelegt werden, kann die Anwendung ohne Datenbank realisiert werden. Für derartige Fälle bieten sich zwei spezielle Varianten von Content Management Systemen an: Flat-File-CMS und Static Website Generatoren. Im Folgenden

Sogenannte Flat-File-CMS verzichten auf eine Datenbank. Inhalte und Konfigurationen werden ausschließlich in Dateien gespeichert. Die Pflege der Inhalte kann dann durch direktes Arbeiten mit den Verzeichnissen und Dateien erfolgen. Die Verzeichnisstruktur wird häufig genutzt, um daraus den Seitenbaum zu generieren.

Der oben beschriebene Prototyp arbeitet nach diesem Prinzip. Der allgemeine Aufbau der Seiten mit Kopf- und Fußzeile sowie Inhaltsbereich ist in einer Vorlage definiert. Zu jeder einzelnen Webseite gehört ein Verzeichnis. Beim Aufruf einer Seite werden die im entsprechenden Verzeichnis gefundenen Dateien in den Inhaltsbereich der Vorlage eingefügt. Die Unterverzeichnisse werden in einer Aufzählung als Verweise auf Unterseiten dargestellt. Der Zusammenbau einer Seite erfolgt dynamisch beim Aufruf mittels PHP-Skripten. Bei festen Inhalten ist diese Dynamik nicht erforderlich. Dieser Ansatz ist in so genannten Static Website Generatoren umgesetzt. Dabei übernimmt ein Generator die Aufgabe, aus den Dokumenten alle Webseiten zu erzeugen (Erstellungsprozess). Diese Arbeit kann lokal durchgeführt werden. Dann müssen nur noch die fertigen Seiten auf einen Webserver kopiert werden. Abbildung 4 zeigt den prinzipiellen Aufbau.

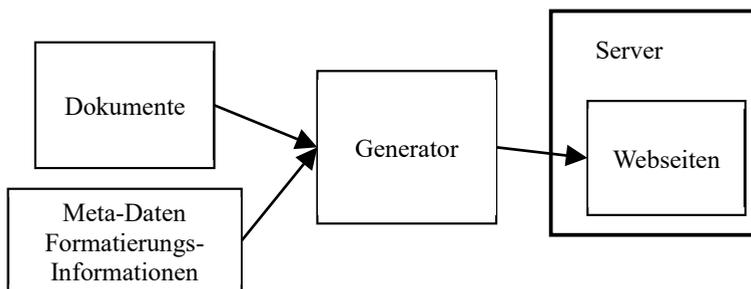


Abbildung 4 Static Website Generator

Neben der Entkopplung von Inhalten und Web-Darstellung sind die geringen Anforderungen an den Webserver ein praktischer Vorteil. Da nur vorgefertigte Seiten ausgegeben werden, ist die Antwortzeit sehr kurz. Weiterhin bestehen kaum Sicherheitsprobleme, da auf dem Server weder Skripte ausgeführt werden noch eine Datenbank verwendet wird. Erkauft werden diese Vorteile durch einen langwierigeren Prozess bei Änderungen. Bearbeitungen werden erst nach Ablauf des Erstellungsprozesses sichtbar.

Camden et al [3] geben folgende Kriterien für Anwendungsfälle an, die sich besonders für Static Website Generatoren eignen:

- *Content focused*
- *Infrequently updated*
- *Do not require a high degree of user interaction or personalization*

Dies bestätigt die gute Eignung dieses Ansatzes für die beschriebene Webanwendung. Zusammenfassend erscheinen sowohl Flat-File-CMS als auch Static Website Generatoren gute Optionen für die Archiv-Anwendung. Zu beiden Konzepten steht eine ganze Reihe von Systemen zur Auswahl. Vor einer konkreten Entscheidung für eines dieser Systeme sollen zunächst wei-

tere Inhalte gesammelt und in den Prototypen integriert werden. Die dabei gewonnenen Erfahrungen sollen dann in eine verfeinerte Anforderungsanalyse als Grundlage für die Systemauswahl einfließen.

Zusammenfassung

In diesem Beitrag wurde ein Konzept für ein Archiv zur geschichtlichen Entwicklung der automatischen Spracherkennung in Deutschland vorgestellt. Basierend auf einem Vorschlag für Inhalte wurde ein erster Prototyp einer Web-Anwendung zur Darstellung der Inhalte implementiert. Für die weitere Entwicklung werden zwei Typen von Content-Management-Systemen vorgeschlagen, die beide eine starke Entkopplung zwischen Inhalten und Webtechnologien gewährleisten. Als nächster Schritt ist vorgesehen, zunächst deutlich mehr Inhalte einzutragen um Erfahrungen zu sammeln und Rückmeldungen einzuholen.

Literatur

- [1] PIERACCINI, R.: *The Voice in the Machine: Building Computers That Understand Speech*. MIT Press, 2012.
- [2] SPILLER, J.: *Decoder Ring: The CMS is Flat*. *Against the Grain*, 26(2), 55, 2017.
- [3] CAMDEN, R. AND RINALDI, B.: *Working with Static Sites: Bringing the Power of Simple to Modern Websites*. O'Reilly, 2017.