

Digitale Zeitzeugenschaft – Wenn Algorithmen das digitale Gedächtnis übernehmen

Erfahrungen mit künstlicher Intelligenz im Archiv „Deutsches Gedächtnis“

Almut Leh

Der Mensch im Mittelpunkt: Personengeschichtliche Quellen in Kommunalarchiven

72. Westfälischer Archivtag

17. und 18. März 2021



Douglas Adams: Per Anhalter durch die Galaxis, 1979



Foto: FernUniversität in Hagen

Bericht über das Archiv „Deutsches Gedächtnis“ im
Wissenschaftsmagazin der FernUniversität

„Die Antwort ist 42“ – Wenn Algorithmen das digitale Gedächtnis übernehmen.

Erfahrungen mit künstlicher Intelligenz im Archiv „Deutsches Gedächtnis“

1. „Die Antwort ist 42“
2. Das Archiv „Deutsches Gedächtnis“ als Forschungsdatenrepositorium und digitales Gedächtnis
3. Technologische Projekte zur Sprachtechnologie und zum Wissensmanagement
4. Digitalisierung und biographische Forschung – Fluch und Segen
5. Perspektiven einer digitalen Hermeneutik



Interview mit Klaus Peter Kisker am 15.11.2018
Projekt „Erlebte Geschichte“ zur Geschichte der Freien Universität Berlin

Bestände des Archivs „Deutsches Gedächtnis“

- 3.085 qualitative Interviews (Stand 1.3.2021)
aus über 100 Projekten von 1975 bis heute
- Textdokumente: Autobiographien, Tagebücher ...
- private Fotos
- Breites thematisches Spektrum
 - Faschismus- und Kriegserfahrung, KZ- und Lagerhaft, Verfolgung, Traumatisierung, Flüchtlinge, Migranten
 - DDR (Verfolgung, Bürgerrechtsbewegung, Bildungssystem)
 - Arbeit, Arbeitsschutz, Sucht, Jugendkultur, Religion, Naturschutz

Das Interview im Archiv

- Audio- und/oder Videoaufzeichnung (analog und digital)
- Transkription
- Interviewprotokoll
- Kurzbiographie
- Einverständniserklärung
- Datenbogen
- evtl. Dokumente und Fotos

KA³ - Kölner Zentrum Analyse und Archivierung von AV-Daten

Ziel:

Erforschung und Entwicklung von Werkzeugen zur akustischen Analyse von audiovisuellen Daten

Projektpartner:

- Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme, Sankt Augustin
- Universität zu Köln: Institut für Linguistik, Regionales Rechenzentrum, Data Center for the Humanities
- Fernuniversität Hagen: Institut für Geschichte und Biographie, Archiv „Deutsches Gedächtnis“

Förderung und Laufzeit:

Bundesministerium für Bildung und Forschung
Okt. 2015 bis Sept. 2020

Audio Mining

Herausforderungen von OH-Interviews

- Spontansprache
- keine professionellen Sprecher
- emotionale Sprache
- Dialektfärbung, Slang
- altersbedingt undeutliche Aussprache
- mangelhafte Aufnahmequalität
- Alterungsprozesse der Magnetbänder

Erkennungsergebnisse

Anwendungsbereich Oral History:

- 2015: 55 % Wortfehlerrate
- 2020: 25 % Wortfehlerrate
(breite Streuung: von 15 % bis über 45 %).

Zum Vergleich:

Anwendungsbereich Rundfunk: aktuell 8 % (2010: über 25 %)

Fehlertypen:

- Out-of-Vocabulary: „Grünwald“ statt „Dünnwald“
- Zusammengesetzte Wörter: „Krieger Witwen Söhne“ statt „Kriegerwitwensöhne“



Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS

Audio Mining

- ermöglicht den inhaltlichen Zugriff auf die gesprochenen Daten,
- ist eine Technologie zur Transkription, Erschließung und Analyse von Sprachdaten.

AudioMining
powered by Fraunhofer IAIS

Transkript

Transkript

Analysedatum

Suchen

Leeren

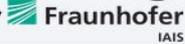
00:02:31

Titel: Battis_Ulrich.mp4
Mittlere ASR Konfidenz: 84
Sendedatum: 21.7.2017
Dauer: 00:00:13

Keywords:
[Jahr](#) [Hagen](#) [Rektor](#) [Rektorat](#) [Münster](#) [Frau](#) [Öl](#) [Professor](#) [Uni](#)
[Berlin](#)

Überblick | **Transkript** | Upload

knapp fünf Jahren nach Münster gezogen zu meinen Großeltern weil deren an einziger Sohn Sohn war wichtig weil der Geschäft übernehmen sollte weil wir hatten sie Todesnachricht bekommen der war in Russlands in Russland geblieben äh also **opfere** als Kriegsgefangene und für ja wieder so war dann sollte der älteste Enkel um den Sohn Ersatz spielen und für möglichst auch das Geschäft übernehmen das Folge Folge war dass sich sehr früh **alte** komplette Uniform hatte als er konnte Tor mit hoher Mütze und Schwarz **Weiß** und zum Ärger meines Vaters mich aber auch nur Ulrich Kerkow nannte wie man **wusste** nach meinem Großeltern und für auch um so ein geschult worden **bin** unter diesem Namen der bis mein Vater irgendwann habe ich dann gesagt nein **der** werden wollte lieber **sich Battiston Schabbat** Gottes Sohn dämlicher nur haben ja und das forderten dass äh der Tropfen der das Fass zum Überlaufen gebracht da nicht zu meinen Eltern die inzwischen nachgezogen waren nach **Museum** an **unser** Großvater hatte dort an ausgebaut für die **junge Frau** die zwischen drei Kinder hatte so dann habe ich für ganz normale Grundschule **lagen** in Münster offen Pollino was sicherlich prägend ist weil da kriegt man sehr früh eigene Leute nicht geplaudert ein vermittelt dass das angeblich das älteste Gymnasium Deutschlands ist das Carolinum inne äh Osnabrück ist noch zwei Jahre älter aber den fehlen bei Jahrhundert im Mittelalter und dann moma klar sagen dass wir die ältesten sind das war damals noch so die **alt** gefragt **es** Gymnasium so wo die **ja** die besseren Schüler hingehen und wo man auch hohe also sehr geschützte Umgebung für das heißt nun **Nichte** dass es nur sogenannte bessere Leute dorthin ging es gegen wirklich nach Leistung als war ziemlich

AudioMining
powered by 

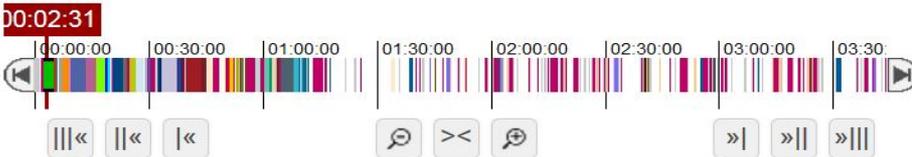
Transkript

Transkript

Analysedatum

Suchen

Leeren



Titel: Battis_Ulrich.mp4
Mittlere ASR Konfidenz: 84
Sendedatum: 21.7.2017
Dauer: 00:00:13

Keywords:
[Jahr](#) [Hagen](#) [Rektor](#) [Rektorat](#) [Münster](#) [Frau](#) [Öl](#) [Professor](#) [Uni](#)
[Berlin](#)

Überblick | **Transkript** | Upload

knapp fünf Jahren
 nach Münster gezogen zu meinen Großeltern weil deren an einziger Sohn
 Sohn war wichtig weil der Geschäft übernehmen sollte weil wir hatten sie
 Todesnachricht bekommen der war in Russlands in Russland geblieben
 äh also
opfere als Kriegsgefangene
 und für ja wieder so war dann sollte der ältester Enkel um
 den Sohn Ersatz spielen und für möglichst auch das Geschäft übernehmen
 das Folge Folge war
 dass sich sehr früh **alte** komplette Uniform hatte als er konnte Tor mit hoher
 Mütze und Schwarz **Weiß** und
 zum Ärger meines Vaters mich aber auch nur Ulrich Kerkow nannte wie man
wusste nach meinem Großeltern
 und für auch um so ein geschult worden **bin** unter diesem Namen der bis
 mein Vater irgendwann habe ich dann gesagt nein **der** werden wollte lieber
sich Battiston Schabbat
Gottes Sohn dämlicher nur haben ja und das forderten dass äh der Tropfen
 der das Fass zum Überlaufen gebracht da nicht zu meinen Eltern die
 inzwischen nachgezogen waren nach **Museum** an **unser** Großvater hatte dort
 an ausgebaut für die **junge Frau** die zwischen drei Kinder hatte
 so dann habe ich **für** ganz normale Grundschule **lagen** in Münster offen
 Pollino was sicherlich prägend ist weil da kriegt man sehr früh eigene Leute
 nicht geplaudert ein vermittelt
 dass das angeblich das älteste Gymnasium Deutschlands ist
 das Carolinum inne
 äh
 Osnabrück ist noch zwei Jahre älter aber den fehlen bei Jahrhundert im
 Mittelalter und dann moma klar sagen dass wir die ältesten sind
 das war damals noch so die **alt** gefragt **es** Gymnasium so wo die **ja** die
 besseren Schüler hingehen und wo man auch hohe also sehr geschützte
 Umgebung für das heißt nun **Nichte** dass es nur sogenannte bessere Leute
 dorthin ging es gegen wirklich nach Leistung als war ziemlich

Audio Mining

- Transkription, Keywords
- Recherche direkt im Audiosignal
- synchrone Darstellung von Audio/Video und Text
- qualitative Analysen:
sprachliche und nicht-sprachliche Aspekte der Kommunikation
(Analyse der Interaktion)
Veränderung der Sprechweise (Hinweis auf Emotionen)
- quantitative Analysen: Auswertung größerer Fallzahlen



Interview-Sammlungen als Forschungsdaten

PROJEKT PARTNER AKTUELLES

Zur Übersichtsseite Partner

FernUniversität in Hagen (Institut für Geschichte und Biographie)

Forschungsstelle für Zeitgeschichte in Hamburg (Werkstatt der Erinnerung)

Freie Universität Berlin (UB/CeDIS)

Friedrich-Alexander-Universität Erlangen-Nürnberg (Neuere und Neueste Geschichte)

Ludwig-Maximilians-Universität München (BAS)

Otto-Friedrich-Universität Bamberg (Lehrstuhl Medieninformatik)

Infrastruktur für die Annotation von Interviews

entfaltet eine digitale Informationsinfrastruktur für die Annotation von markierten narrativen Interviews, v.a. für die Unterstützung von Archiven, Museen und Bibliotheken bei der Erfassung, Speicherung und Bereitstellung sowie der Sammlungs- und Datenverwaltung.

Vorteile für Forscher*innen

<https://www.oral-history.digital>



Interview mit Eva Strommenger am 31.5.2018
Projekt „Erlebte Geschichte“ zur Geschichte der Freien Universität Berlin

„Kein Algorithmus ist klug genug, um zu verstehen, dass bestimmte Entscheidungen für bestimmte Bevölkerungsgruppen diskriminierend sind. Und kein Algorithmus ist klug genug, um zu verstehen, dass wir uns eine Zukunft wünschen, die sich von der Vergangenheit unterscheidet.“

Manuele Lenzen: Künstliche Intelligenz. Fakten, Chancen, Risiken (2020)



Vielen Dank