

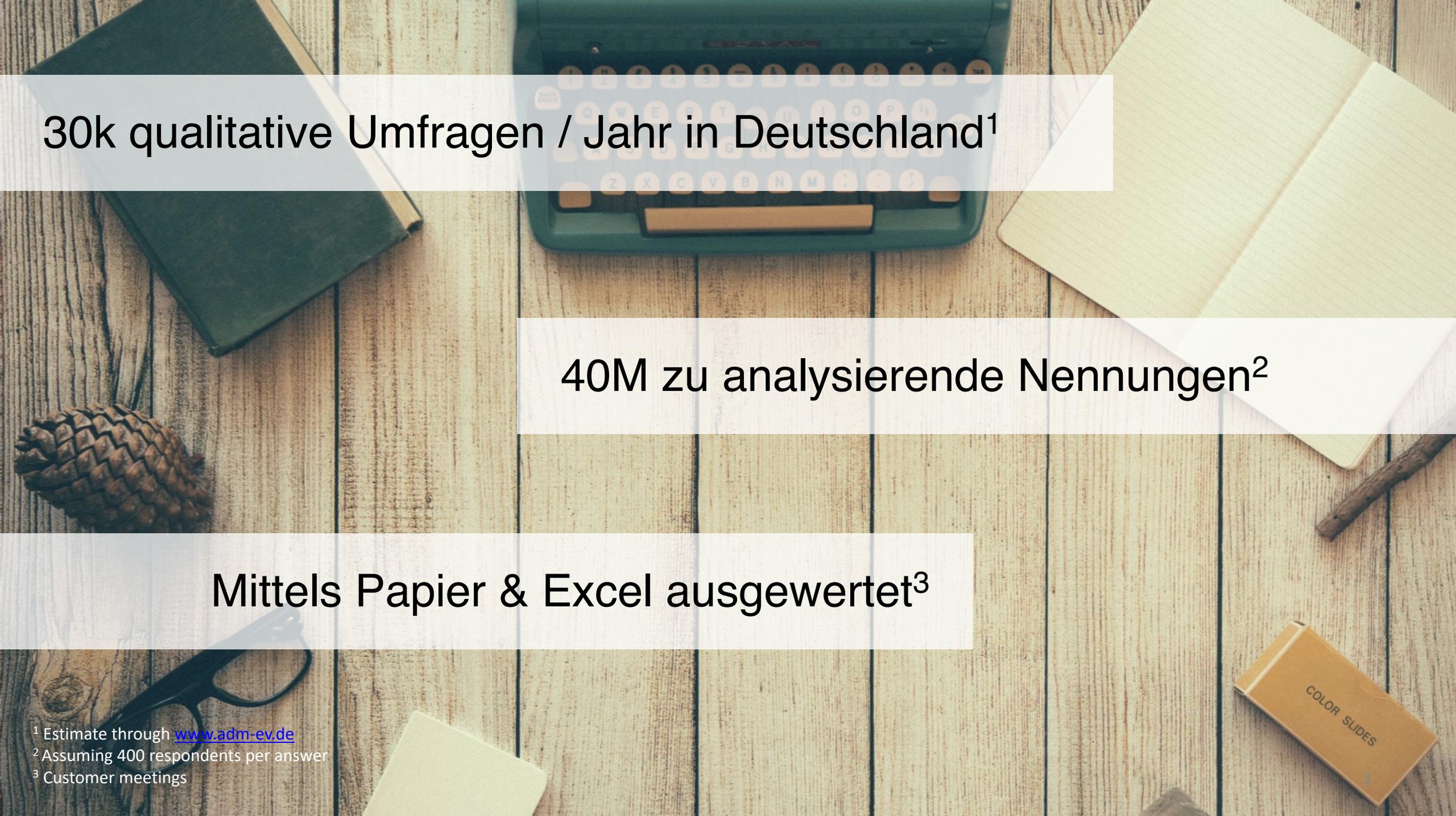
Umfragekodierung mittels künstlicher Intelligenz

Maurice Gonzenbach, Caplena, Zürich, 22.3.2018



Agenda

- 1. Problemstellung: Kodieren**
- 2. Ansätze**
- 3. Technischer Exkurs: KI und Kontext**
- 4. Resultate & Anwendungen**
- 5. Ausblick**



30k qualitative Umfragen / Jahr in Deutschland¹

40M zu analysierende Nennungen²

Mittels Papier & Excel ausgewertet³

¹ Estimate through www.adm-ev.de

² Assuming 400 respondents per answer

³ Customer meetings

Analyse von offenen Fragen

1. Datenerhebung



2. Kodierung

sleek UI, impressive predictions



DESIGN

AI

ML model makes coding efficient



AI

EFFICIENT

amazing speedup and support



EFFICIENT

SUPPORT

3. Auswertung

DESIGN
AI
SUPPORT
EFFICIENT



Ansätze für automatische Textanalyse

“Out of the Box”

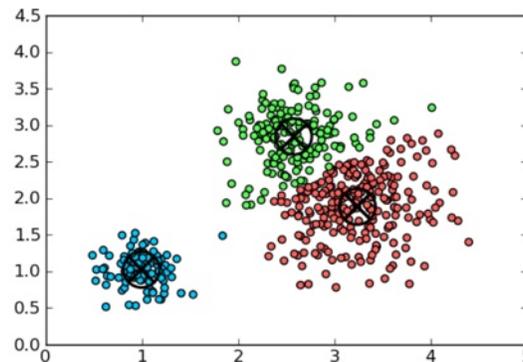
- Einfache APIs
- Vordefinierte Kategorien



Google Cloud Platform

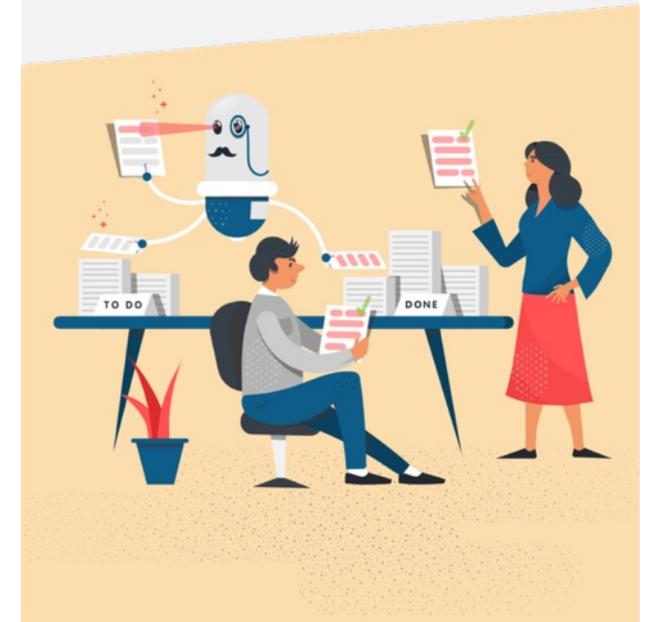
Unüberwacht

- APIs / Open Source
- Grobe Gruppierung



Überwacht

- Open Source / Codit.co
- Lernt von Beispielen

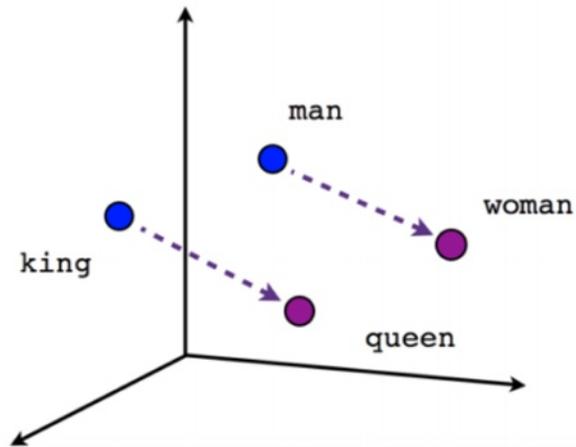


Welches Bild steht auf dem Kopf?

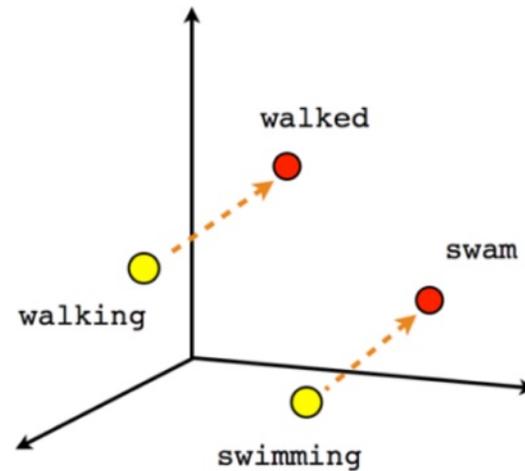


Kontext ist eine Herausforderung für Algorithmen

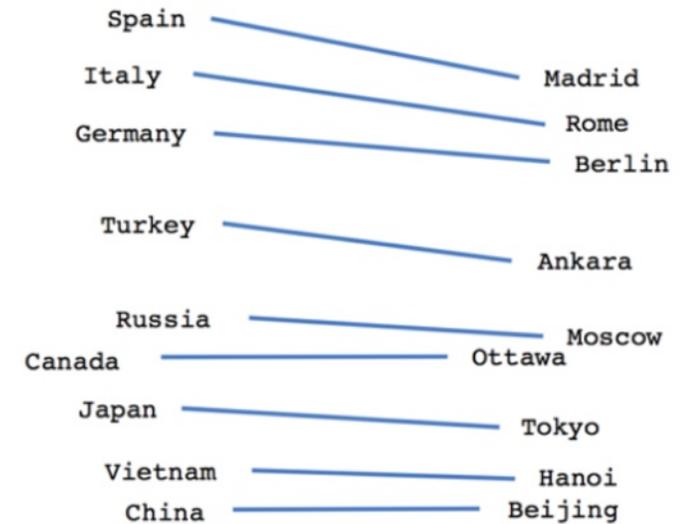
Wie Computer mit Worten rechnen



Male-Female



Verb tense



Country-Capital

Kontext ist entscheidend (1)

DATEN

Filmkritik

- ~ 7.9M Amazon Filmkritiken
- 1.5B Wörter, 600k Einzigartige

Nachrichten

- Verschiedene Nachrichtenportale, 2013-2017
- 4.5B Wörter, 1.2M Einzigartige

ÄHNLICHKEITEN

WATSON
Holmes, Sherlock, Emma, Mycroft...

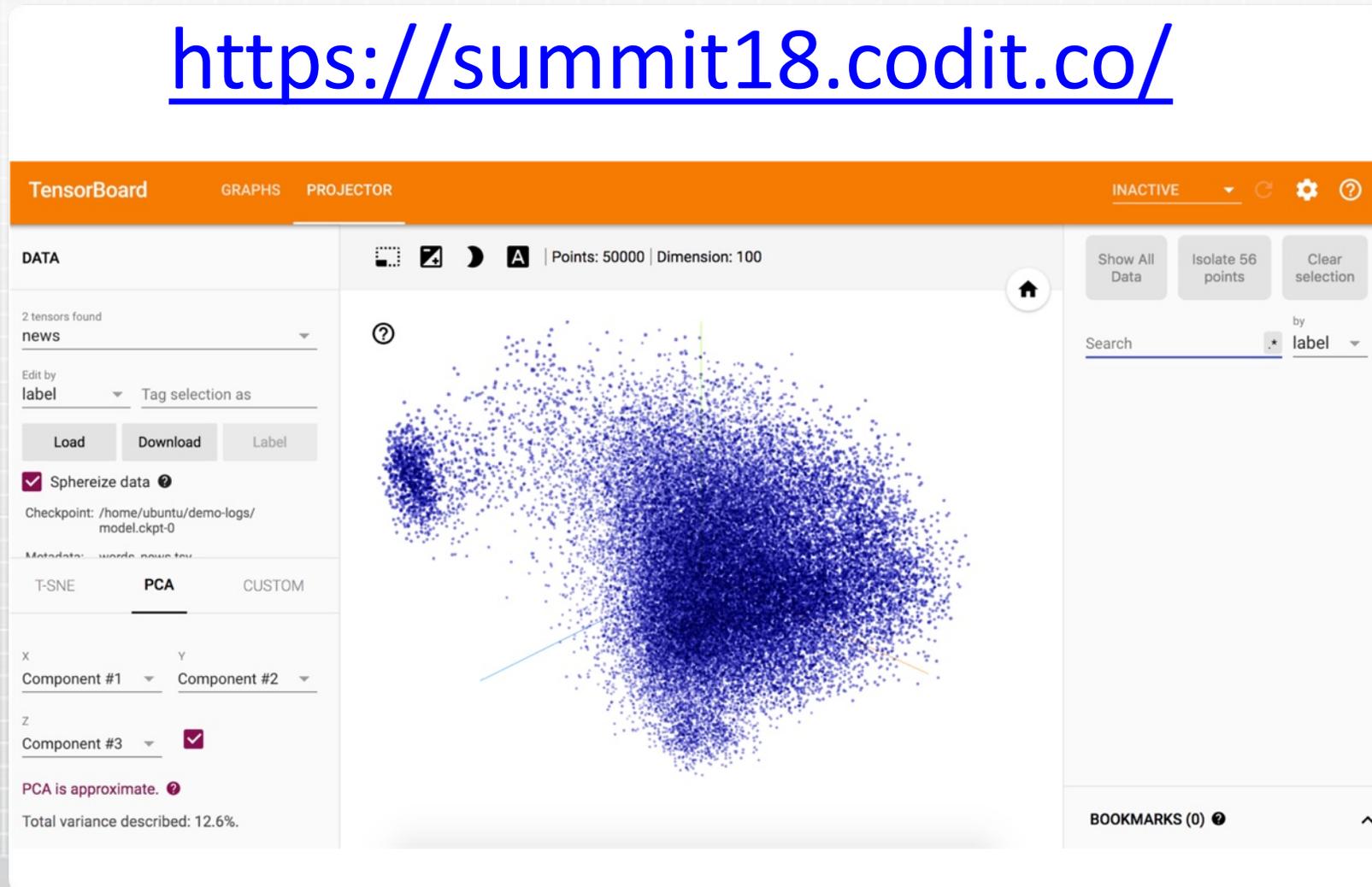
WATSON
Smith, Clarke, Hodge, McGilchrist...

Kontext ist entscheidend (2)

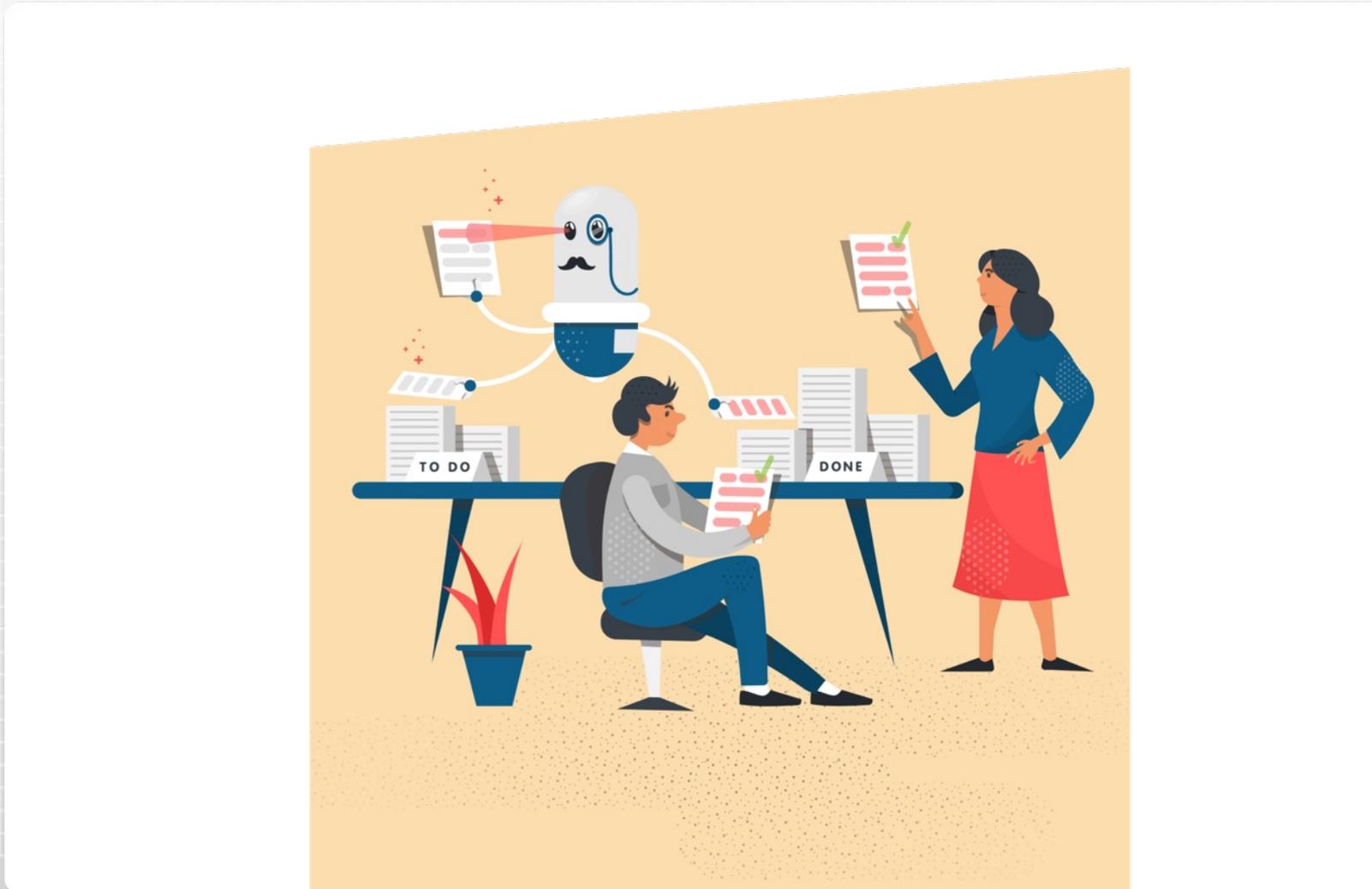
DATEN	Filmkritik	Nachrichten
ÄHNLICHKEITEN	ABSORBING engrossing, compelling, riveting...	ABSORBING reabsorbing, high-concentration
ANALOGIEN	LUKE to HERO like DARTH to VILLAIN	US to FOOTBALL like ENGLAND to RUGBY

Entdecken Sie es selbst

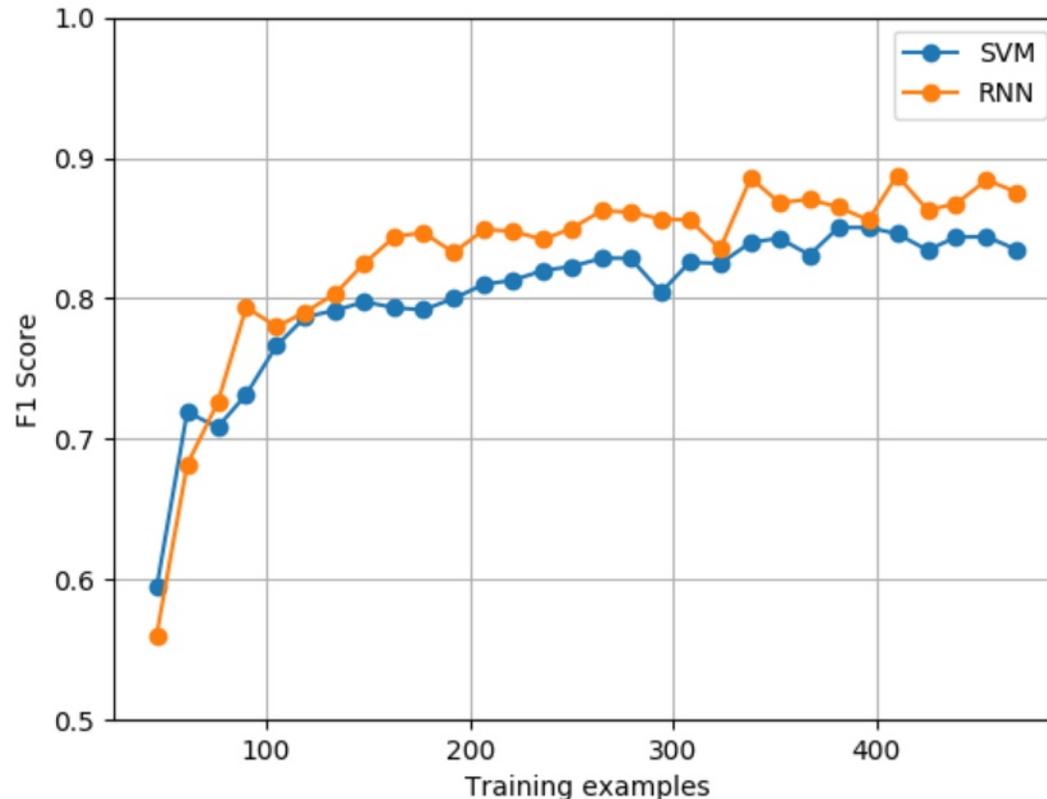
<https://summit18.codit.co/>



Menschen als Lehrer & Supervisor



Auf Umfragen spezialisierte KI

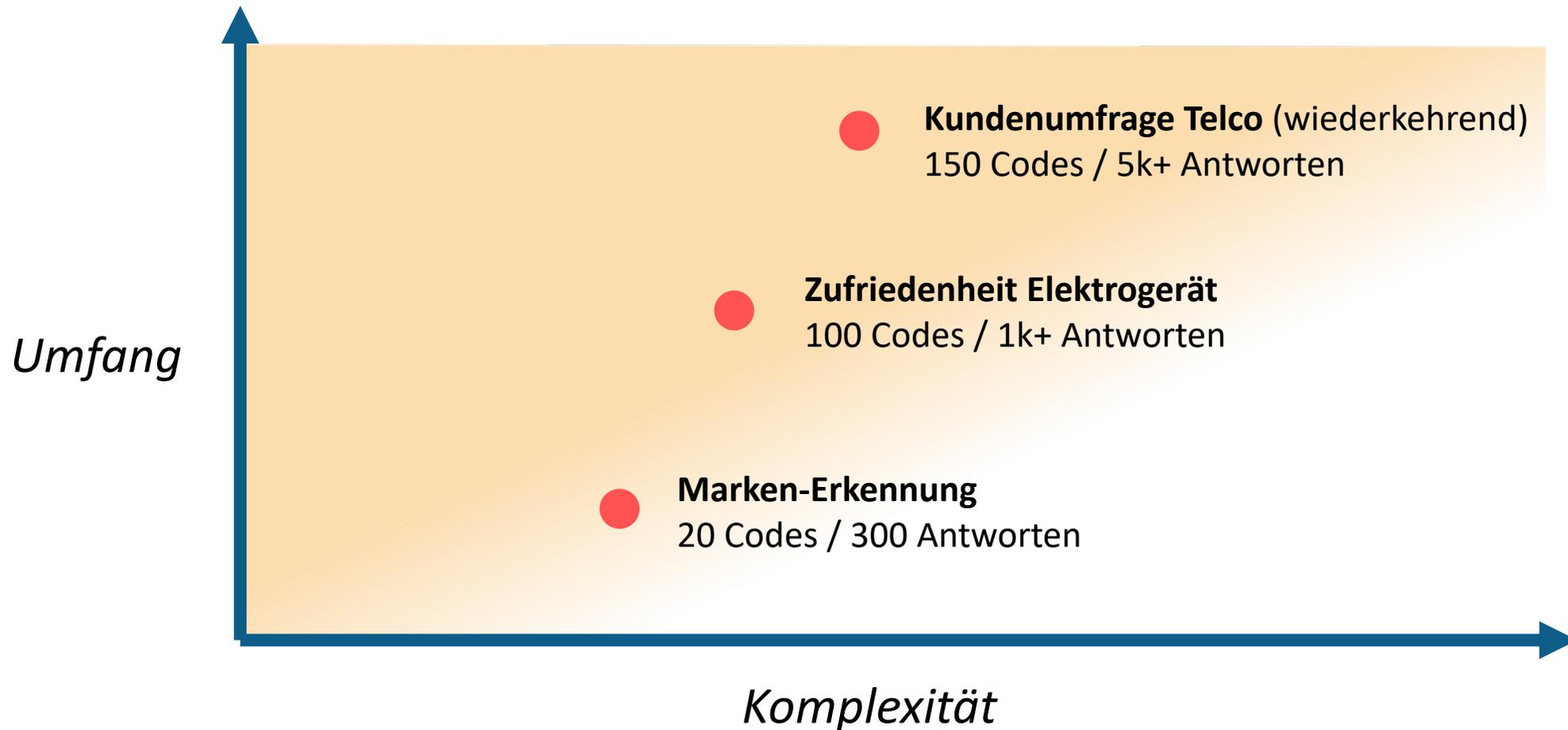


Umfrage über Marken-Erkennung mit 15 Codes

Lernt schnell durch

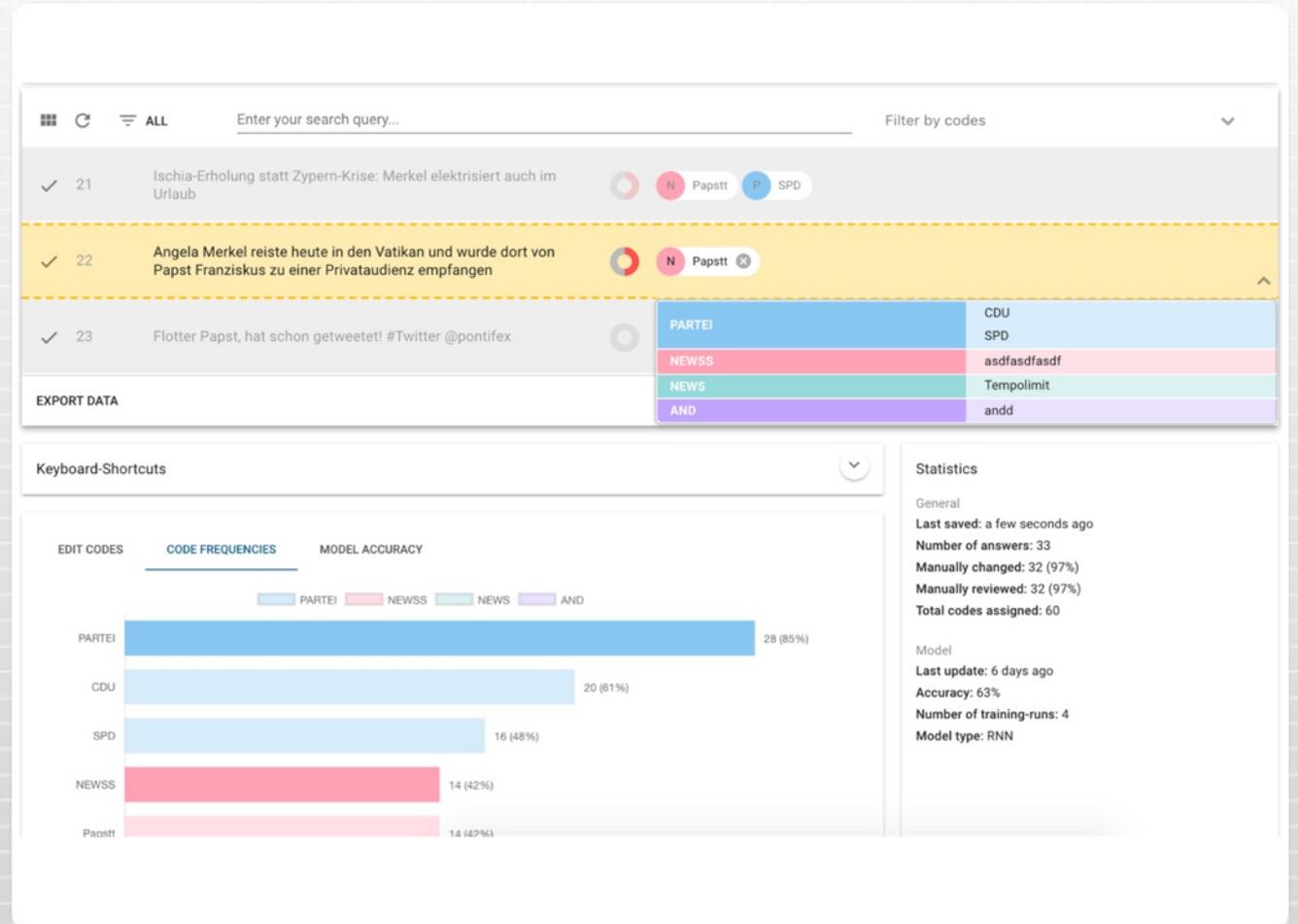
- Aufbau von Sprachverständnis mittels Millionen von Dokumenten
- **Transfer-Lernen** auf ähnlichen Umfragen

Wo bringt maschinelle Kodierung am Meisten?



Zusammenarbeit zwischen Mensch & Algorithmus

- Simple Benutzeroberfläche
- KI schlägt Codes für Antworten vor
- Mit jeder manuell verkodeten Antwort wird die KI schlauer
- Anwender bleibt Entscheidungsträger



The screenshot displays a web application for manual coding of survey responses. It features a search bar at the top, a list of survey items with suggested codes, a table for exporting data, keyboard shortcuts, and a statistics section.

Survey Items and Suggested Codes:

ID	Text	Suggested Codes
21	Ischia-Erholung statt Zypern-Krise: Merkel elektrisiert auch im Urlaub	Papstt, SPD
22	Angela Merkel reiste heute in den Vatikan und wurde dort von Papst Franziskus zu einer Privataudienz empfangen	Papstt
23	Flotter Papst, hat schon getweetet! #Twitter @pontifex	

EXPORT DATA Table:

PARTEI	CDU
NEWSS	asdfasdfsdf
NEWS	Tempolimit
AND	andd

Keyboard-Shortcuts:

Statistics:

- General
- Last saved: a few seconds ago
- Number of answers: 33
- Manually changed: 32 (97%)
- Manually reviewed: 32 (97%)
- Total codes assigned: 60

Model:

- Last update: 6 days ago
- Accuracy: 63%
- Number of training-runs: 4
- Model type: RNN

Code Frequencies Bar Chart:

Code	Count	Percentage
PARTEI	28	85%
CDU	20	61%
SPD	16	48%
NEWSS	14	42%
Papstt	14	42%

Wie unsere Kunden Cudit.co benutzen

Marktforschung

- Verarbeitung diverser Umfragen während der Mensch in Kontrolle bleibt
- Effizienzgewinne, insbesondere bei grösseren oder wiederholten Umfragen

Schweizer Marktforschungsunternehmen

B2C

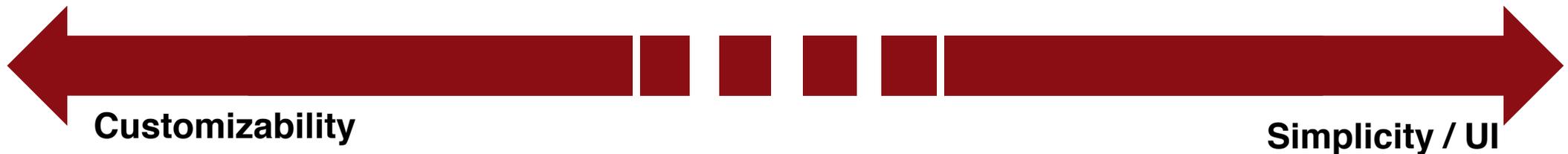
- Analyse von Kunden-Feedback
- Schnellere Entscheidungswege durch automatische Analyse von neuen Antworten

Deutsche Telco

Wir wollen durch Fokus & Einfachheit überzeugen



MACHINE BOX



Unser Weg bis anhin



Stats

80k + Antworten analysiert über unsere Plattform



Lancierung

Februar 2018



Unsere Kunden

„cudit.co macht Kodieren extrem leicht. Schneller Einstieg, simple Bedienung und die hilfreiche KI machen das Tool momentan zur besten Option für mich.“

Marc Haupt, freiberuflicher Marktforscher

Branchentrends & Weiterentwicklung

Marktforschung

- Klassische Umfragen werden bleiben
- Differenzierung durch Kombination von Technologie & Qualität

Codit.co

- Dezentralisierung der Kodierleistung
- Aufbau von «Modellbibliothek»
- Marktplatz zum Austausch von Modellen



Photo: Bekir Dönmez via Unsplash

?

Probieren Sie es aus! <https://codit.co>
Interaktive Visualisierung der Technologie:
<https://summit18.codit.co>

CAPLENA

Kontakt



Maurice Gonzenbach

maurice@caplena.com

Tel. +41 79 455 52 17

www.caplena.com



Pascal de Buren

pascal@caplena.com

Tel. +41 79 518 91 75

www.caplena.com

