



## UMSETZUNGSPROJEKT KICKBUZZ GMBH

### Künstliche Intelligenz zur Textanalyse mittels Natural Language Processing



#### UNTERNEHMENSPROFIL

Kickbuzz steht für Technologiekonzepte, die helfen Hürden für das Business von Unternehmen abzubauen. Kickbuzz arbeitet enthusiastisch an innovativen IT Lösungen, welche die Unternehmensbedürfnisse von Anfang an berücksichtigen. Ihr Netzwerk hilft Initiativen zur Marktreife zu führen.

**Technologie:** Die Entwickler sind mehr als gute Handwerker. Sie ruhen nicht bis das MVP steht. Business Modelling und Dokumentation jedes Prozessschrittes ist für Kickbuzz selbstverständlich, um Synergien für Geschäftsmodelle zu schaffen.

**Vermarktungskompetenz:** Kickbuzz steht am Anfang einer massiven Umwälzung des Kundenverhaltens und der Möglichkeit der digitalen Vermarktung. Sie setzen schon heute auf die Kundenwünsche von morgen.

**Netzwerk:** Das Unternehmen entwickelt Innovationen im Austausch mit Ihrer B2B Community. Sie profitieren von der Exzellenz vieler einzelner Subject Matter Experts, auf welche Sie zurückgreifen können, um das Chaos auf dem unternehmerischen Schreibtisch zu reduzieren.

**Produkt:** Das erste marktreife Produkt der Kickbuzz GmbH ist DSGVO SCAN. Mithilfe des DSGVO SCAN werden Unternehmen in die Lage versetzt, sowohl das Impressum als auch die Datenschutzerklärung der unternehmenseigenen Website völlig automatisiert auf DSGVO-Konformität zu überprüfen.

Dazu zählen z. B. folgende Fragestellungen: Ist die Rechtsform des Unternehmens im Impressum aufgeführt? Ist die Adresse des Unternehmens im Impressum aufgeführt? Ist die Telefonnummer des Unternehmens im Impressum aufgeführt? Diese und weitere rechtliche Erfordernisse können durch eine Künstliche Intelligenz geprüft werden. Hierfür implementierte das Unternehmen ein sog. Bi-LSTM (Bidirectional Long-Short-Term-Memory), eine besondere Form von Deep Learning, um ein Impressum mittels Natural Language Processing (NLP) auswerten zu können.

Ein weiterer wichtiger Baustein des DSGVO SCAN ist das Abgleichen der in der Datenschutzerklärung vorkommenden Unternehmen (z. B. Facebook, LinkedIn, Instagram, etc.), die Cookies setzen und Daten an ihre Server senden. Die Nutzenden der Website müssen darüber informiert werden, an welche Unternehmen welche Daten gesendet werden. Aufgrund der schier Masse an eingesetzten Werbetrackern ist es häufig nicht einmal für die Website betreibenden Unternehmen ersichtlich, welche Daten der Nutzenden an wen und wohin gesendet werden. Ergo, muss der DSGVO SCAN einen Abgleich zwischen den in der Datenschutzerklärung aufgeführten Unternehmen und den gesendeten Daten der Nutzenden der Werbetracker vornehmen. An diesem Punkt kommt die ausgewiesene Expertise des Mittelstand 4.0 Kompetenzzentrum Cottbus im Bereich Künstliche Intelligenz ins Spiel.



#### HERAUSFORDERUNG

Für die Lösung dieser Problemstellung ergeben sich drei Herausforderungen. Erstens, eine Software muss analysieren, an welche Unternehmen die eingesetzten Werbetracker Daten der Website-Nutzenden senden. Diese Herausforderung wurde von der Kickbuzz GmbH bewältigt. Zweitens, es müssen aus der Datenschutzerklärung diejenigen Unternehmen identifiziert werden, an die Daten gesendet werden. Diese Herausforderung wurde vom Mittelstand 4.0-Kompetenzzentrum Cottbus erfolgreich gemeistert. Die dritte Herausforderung bestand im Abgleich der gesendeten Daten und den in der Datenschutzerklärung identifizierten Unternehmen aus den beiden vorangegangenen Herausforderungen.



#### LÖSUNG

Für die Lösung des Problems bestanden zwei größere Schwierigkeiten. Zum einen, wie identifiziert man die Unternehmensnamen in der Datenschutzerklärung. Und zum anderen, wie findet man heraus, an welche dieser Unternehmen laut Datenschutzerklärung tatsächlich Daten gesendet werden. Ein plakatives Beispiel illustriert dieses Problem. So kann in einer Datenschutzerklärung stehen:

“Mittels Google-Suche finden Sie weitere Angebote unseres Unternehmens.“ Ein anderer Satz könnte lauten: “Die eingesetzten Werbetracker senden ihren Standort zu Google.“ In beiden Sätzen erscheint der Unternehmensname „Google“, aber nur im zweiten Beispiel werden tatsächlich Daten gesendet. Zur Lösung dieses Problems wurden Informationen aus 2.000 Datenschutzerklärungen analysiert und anschließend gelabelt, um sie für die Künstliche Intelligenz (KI) nutzbar zu machen. Anschließend wurden die Texte in eine maschinenlesbare Form überführt, um die KI-Verfahren anwenden zu können. Da es verschiedene KI-Verfahren gibt, musste ein Vergleich der Verfahren über eine Monte-Carlo-Simulation durchgeführt werden. Die Parametereinstellungen der einzelnen KI-Verfahren wurden mittels Grid Search identifiziert. Verwendet wurden die folgenden Verfahren: Logistische Regression, Gradient Boosting, Random Forests, Passive Aggressive Classifier, K-nearest Neighbour, Support Vector Machines und Naive Bayes. Unter den sieben Verfahren stellte sich Gradient Boosting, ein state-of-the-art Algorithmus, in diesem Fall den anderen Verfahren gegenüber als überlegen heraus. Durch Gradient Boosting wusste man nun, welche Sätze entscheidend für das Senden von Daten an Unternehmen waren. Im nächsten Schritt wurde mittels Named Entity Recognition (NER) das Unternehmen aus den entscheidenden Sätzen extrahiert.



## ANSPRECHPARTNER

**Dr. Sascha Vökler**

:: Trainer für Künstliche Intelligenz

:: Data Scientist

:: Machine Learning Engineer



## UMSETZUNG

Die Umsetzung fand durch einen notwendigen, regelmäßigen Austausch zwischen den IT-Experten der kickbuzz GmbH und einem der KI-Trainer des Mittelstand 4.0-Kompetenzzentrum Cottbus sowie einer studentischen Hilfskraft der Informatik statt. Als Vorreiter der Digitalisierung bewies das Mittelstand 4.0-Kompetenzzentrum Cottbus, dass es möglich ist, komplexe Fragestellungen zu 100% remote, also nur über Fernkommunikation mittels Video-Calls, zum Erfolg zu führen. Der rege Austausch trug neben den hervorragenden fachlichen Kompetenzen beider Seiten maßgeblich zum überragenden Erfolg dieses Umsetzungsprojekts bei. Die technische Umsetzung mündet in die Implementierung der gefundenen Lösung in das Produktivsystem des Produkts DSGVO SCAN.



### IMPRESSUM

**Herausgeber:** Mittelstand 4.0-Kompetenzzentrum Cottbus, c/o BTU Cottbus – Senftenberg,  
Prof. Dr.-Ing. Ulrich Berger, Siemens-Halske-Ring 14, DE-03046 Cottbus, Telefon: 0355 695171,  
[info@kompetenzzentrum-cottbus.digital](mailto:info@kompetenzzentrum-cottbus.digital), [www.kompetenzzentrum-cottbus.digital](http://www.kompetenzzentrum-cottbus.digital)

**Redaktion & Gestaltung:** Norman Werk, Grit Rehe

Gefördert durch:



Bundesministerium  
für Wirtschaft  
und Klimaschutz

aufgrund eines Beschlusses  
des Deutschen Bundestages