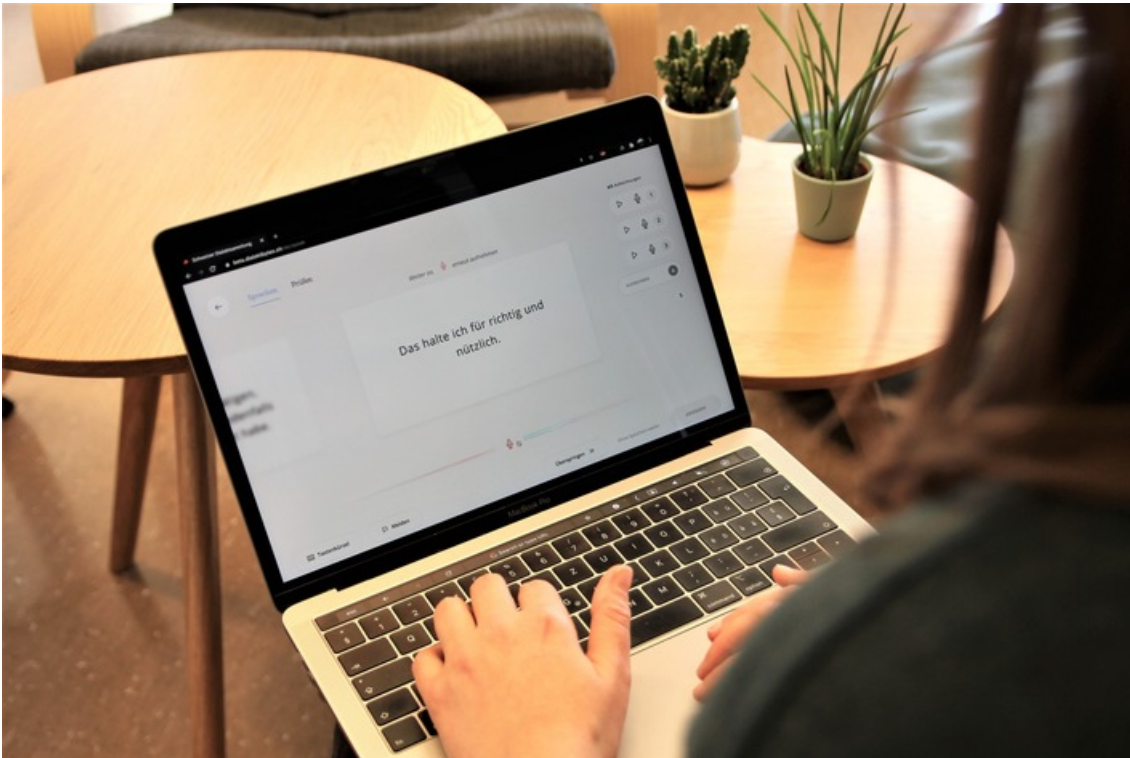


25.05.2021 - 09:09 Uhr

Mithilfe der Bevölkerung: ZHAW und FHNW digitalisieren Schweizer Dialekte



Medienmitteilung vom 25. Mai 2021

ZHAW School of Engineering

Mithilfe der Bevölkerung: ZHAW und FHNW digitalisieren Schweizer Dialekte

Die beiden Hochschulen rufen zu einer Datensammlung von Schweizer Dialekten in der gesamten Deutschschweiz auf. Die Bevölkerung kann ab sofort mithelfen, indem sie mit einer einfachen App Sprachaufnahmen erstellt. Mit den digitalisierten Dialekten können wichtige Computerprogramme trainiert werden.

Chatbots und Sprachassistenten wie Siri oder Alexa verstehen gesprochenes Schweizerdeutsch nur schlecht. Das liegt am Mangel an Audiodateien, die nötig sind, um diese Systeme zu trainieren. Für grosse Technologiefirmen wie Google, Apple oder Amazon ist der Schweizer Markt zu klein, um eine Lösung zu entwickeln, die Schweizerdeutsch versteht. Das soll sich ändern. «Wir wollen Schweizer Dialekte sammeln und digitalisieren», sagt Mark Cieliebak vom ZHAW-Centre for Artificial Intelligence (CAI). «Mindestens 2000 Stunden Aufnahmen von schweizerdeutschen Dialekten sollen zusammenkommen, damit wir eine gute Datenbasis haben». Manfred Vogel von der FHNW ergänzt: «Diese Daten werden wir anschliessend verwenden, um einem Algorithmus basierend auf künstlicher Intelligenz beizubringen, schweizerdeutsche Sprache zu verstehen und automatisch in hochdeutschen Text umzuwandeln».

So funktioniert die Dialektsammlung

Für das Projekt wurde eine Webapplikation entwickelt, mit welcher Freiwillige eigene Audioaufnahmen erstellen können, indem sie hochdeutsche Sätze in natürliche Mundart übersetzen und/oder die Aufnahmen von anderen Teilnehmenden überprüfen. In der App können die Teilnehmenden sehen, aus welchen Kantonen bereits Sprachaufnahmen vorhanden sind – und welche Kantone und Dialekte noch fehlen.

Mögliche Anwendungen

Mit der Datensammlung und den trainierten Sprachmodellen können Sprachschnittstellen zu verschiedenen Anwendungen entwickelt werden. So könnte man beispielsweise mit Sprachassistenten auf Schweizerdeutsch sprechen, Firmen könnten automatisch Kundenfeedback auswerten, etwa Anrufe beim Kundendienst. Auch Untertitel für TV-Sendungen könnten automatisch erstellt werden.

Die Technologien für das Training von Speech-to-Text-Systemen wurden in den letzten Jahren laufend weiterentwickelt und basieren heute meistens auf neuronalen Netzwerken. Für Sprachen wie Englisch und Deutsch liefern diese Methoden bereits sehr

gute Ergebnisse mit Fehlerraten von unter zwei Prozent.

Innovationsstandort Schweiz stärken

«Weil wir den Datensatz für Forschungszwecke veröffentlichen, können Computerprogramme entwickelt werden, die dann in Zusammenarbeit mit lokalen Firmen für verschiedene Zwecke eingesetzt werden können», erklärt Manuela Hürlimann von der Swiss Association for Natural Language Processing (SwissNLP), welche das Projekt koordiniert und leitet. Unterstützt werden die beiden Hochschulen vom Media Technology Center der ETH und der Universität Zürich. Finanziert wird das Projekt mit internen Mitteln der ZHAW und FHNW und mit finanzieller Unterstützung des Schweizer Nationalfonds und der AXA Versicherung.

www.dialektsammlung.ch

Kontakt

Prof. Dr. Mark Cieliebak, Leiter Natural Language Processing Group, Centre for Artificial Intelligence, ZHAW School of Engineering, Tel. 058 934 72 39, E-Mail mark.cieliebak@zhaw.ch

Prof. Dr. Manfred Vogel, Leiter Information Processing, Institut für Data Science, Hochschule für Technik FHNW, Tel. 056 202 77 36, E-Mail manfred.vogel@fhnw.ch

Manuela Hürlimann, Wissenschaftliche Mitarbeiterin, Centre for Artificial Intelligence, ZHAW School of Engineering, Tel. 058 934 45 07, E-Mail manuela.huerlimann@zhaw.ch

Dr. Bettina Mack, Content-/ Media Manager und Science Writer, ZHAW Zürcher Hochschule für Angewandte Wissenschaften, Tel. 058 934 6026, E-Mail bettina.mack@zhaw.ch

ZHAW Zürcher Hochschule für Angewandte Wissenschaften

Corporate Communications

Gertrudstrasse 15

Postfach

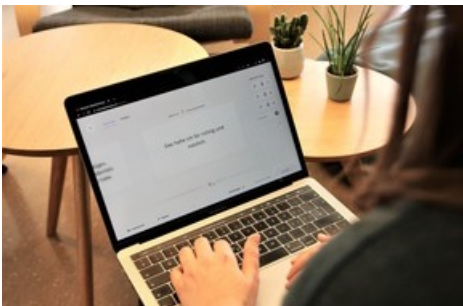
CH-8401 Winterthur

Tel. +41 58 934 75 75

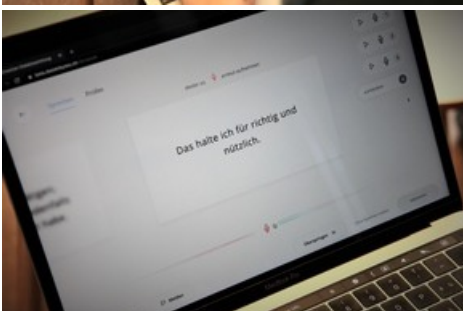
medien@zhaw.ch

www.zhaw.ch/medien

Medieninhalte



Die Bevölkerung kann ab sofort mithelfen, indem sie mit einer einfachen App Sprachaufnahmen erstellt.



In der App können die Teilnehmenden sehen, aus welchen Kantonen bereits Sprachaufnahmen vorhanden sind – und welche Kantone und Dialekte noch fehlen.

Diese Meldung kann unter <https://www.presseportal.ch/de/pm/100018827/100871353> abgerufen werden.