

Beitrag D: Martin Wilden, Antonia van Eek, Matthias Bluhm

Sag „Hallo“ zu deiner Karte

Say “Hello“ to Your Map

Matthias Bluhm, Antonia van Eek, Martin Wilden

con terra GmbH, m.bluhm@conterra.de

Abstract

"Will voice control soon replace typing?" More and more linguistic assistants like Google Assistant, Amazon Alexa, Microsoft Cortana or Apple's Siri have recently entered the market and its usage is already possible in many applications. The market trend shows that many of our interactions with technology take place through conversations today, but does this trend also apply to mapping applications?

Con terra picked up on this trend and carried out its first realizations. It uses natural language understanding (NLU) to control web mapping applications. For example, if one could say "show me the distribution of porpoises in the summer of 2013" instead of searching in a table of content, it would be a great advantage for the users. Also functions like "zoom" or "pan" can be well performed using voice control.

In addition to an introduction to the topic, this article introduces the exemplary implementation of voice control for a mapping application. For the software **map.apps** a voice recognition was implemented and trained. Commands like "zoom", "pan" or the map theme selection have been realized. The results were implemented and tested for applications in biodiversity monitoring.

Zusammenfassung

„Wird Sprachsteuerung das Tippen bald ersetzen?“ In letzter Zeit erreichen mehr und mehr Sprach-Assistenten wie Google Assistant, Amazons Alexa, Microsoft Cortana oder Apples Siri den Markt und der Einsatz ist in vielen Anwendungsfeldern bereits möglich. Der Markttrend zeigt, dass bereits heute viele unserer Interaktionen mit Technologie durch Gespräche stattfinden. Aber gilt dieser Trend auch für Kartenanwendungen?

Die con terra hat diesen Trend aufgegriffen und im Zuge von F&E-Aktivitäten erste Realisierungen durchgeführt. Dabei wird „natural language understanding“ (NLU) genutzt, um Web-Mapping-Anwendungen zu steuern. Wenn man beispielsweise „zeige mir die Verteilung der Schweinswale im Sommer 2013“ sagen könnte, anstatt das Thema in einem Themenbaum zu suchen, wäre das für die Nutzer eine große Erleichterung. Aber auch Funktionen wie „zoom“, „verschieben“ oder den Kartenausschnitt auf ein bestimmtes Gebiet setzen, lassen sich gut über Sprache steuern.

In diesem Beitrag wird neben einer Einführung in die Thematik die exemplarische Umsetzung von Sprachsteuerung für eine Mapping Anwendung vorgestellt. Für die Software **map.apps** wurde eine Sprachsteuerung umgesetzt und trainiert. Befehle wie „zoom“, „verschieben“ oder die Auswahl von Kartenthemen wurden realisiert. Exemplarisch wurden die Ergebnisse für Anwendungen im Biodiversitätsmonitoring umgesetzt und getestet.

1 Zielsetzung und Motivation

Sprachschnittstellen haben den Massenmarkt bereits erobert. 2011 hat Apple mit Siri einen Sprachassistenten auf iPhones gebracht. Smarte Lautsprecher wie Alexa von Amazon oder Google Home drängen derzeit sehr stark auf den Markt. Google Assistant steht auf 400 Millionen Endgeräten zur Verfügung. Endbenutzer sind es zunehmend gewöhnt, über Sprachschnittstellen mit Geräten zu interagieren. Jeder kennt heutzutage Systeme wie Siri, Alexa, Google Assistant oder Cortana. Gartner [Gartner 2015] hatte bereits vor einigen Jahren vorausgesagt, dass bis zum Jahr 2018 30% unserer Interaktionen mit Technologie durch Gespräche stattfinden wird. Besonders auf mobilen Geräten wird besonders oft über Spracheingabe gesucht.

Da die Fehlerrate der Spracherkennung in den letzten 5 Jahren deutlich zurückgegangen ist (z.B. bei Google von ca. 23% auf ca. 5%) [Protalinski, 2017], lässt sich die Spracherkennung jetzt operationell einsetzen. Auch für Geoinformationssysteme sehen wir Anwendungsfälle und beschreiben nachfolgend unsere Umsetzung mit map.apps und die Erfahrungen.

map.apps ist ein Entwicklungstoolkit und eine Betriebsumgebung für Geo-Apps (web / mobile) und Lösungen [Uhlenkücken 2018; Hackmann 2015]. map.apps bietet GIS-Funktionen in ansprechender und leicht nutzbarer Form. Die Anwendungen bieten meist wenige, aber sehr zielgerichtete Funktionen und machen GIS-Funktionalitäten für jedermann überall zugänglich. Als zentrale Betriebsumgebung sorgt map.apps

auch in größeren Organisationen für einen effizienten IT-Betrieb multipler Apps für unterschiedlichste Zwecke, wie z.B. bei der Bundesanstalt für Geowissenschaft und Rohstoffe [Heber & Kruse 2016] oder in der hessischen Umweltverwaltung [Giese & Bluhm 2017]. Weitere Anwendungsbeispiele sind auf der Website⁶ von con terra beschrieben.

2 Sprachsteuerung für Kartenanwendungen

Ziel der Sprachsteuerung ist es, die natürliche Sprache in strukturelle Daten (hier: Karteninteraktionen) umzuwandeln, die maschinell gelesen werden können.

Eine der Hauptherausforderungen bei der Entwicklung ist es, die völlig unterschiedlichen Ausdrücke und Phrasen, die unterschiedliche Menschen für den gleichen Sachverhalt verwenden, zu modellieren. So sagt beispielsweise ein GIS Experte "Aktiviere den Layer Grundschulen", aber ein GIS-Laie formuliert "Zeige mir Grundschulen".

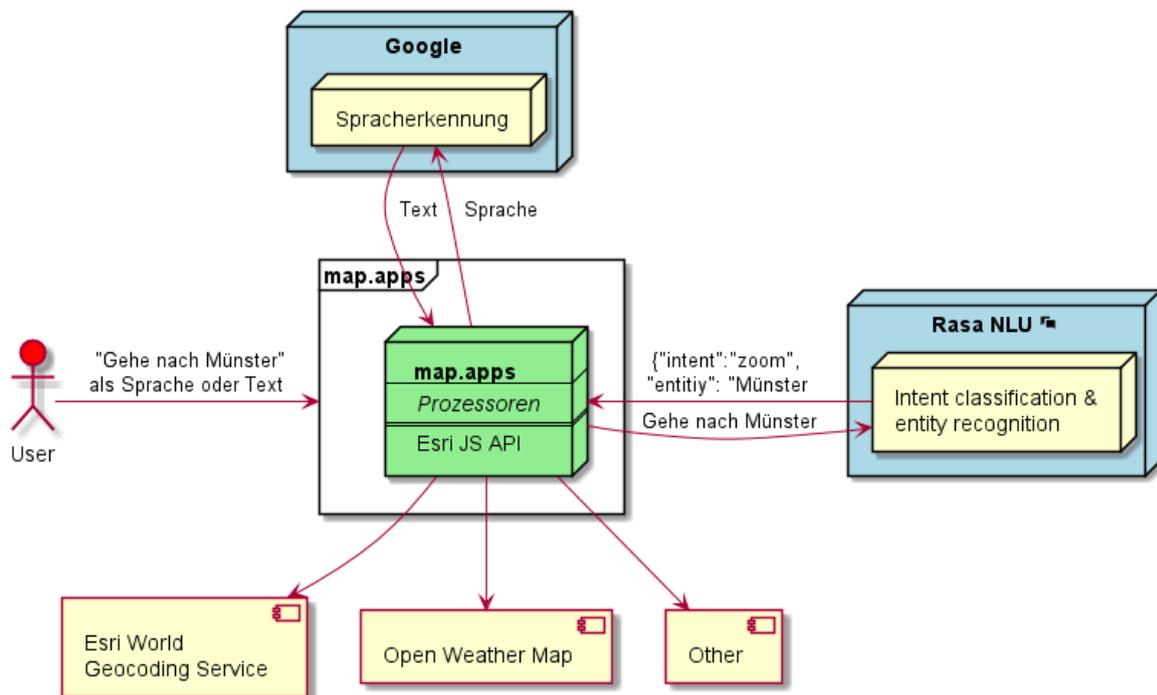


Abbildung 1: Systemskizze

⁶ <https://www.conterra.de/mapapps-demos> zuletzt aufgerufen 30.08.2018

In Abbildung 1 ist die Systemskizze mit einer exemplarischen Anfrage dargestellt. In einem ersten Schritt wird die Anfrage des Anwenders von map.apps an die Spracherkennung geschickt. Die Spracherkennung wandelt die gesprochene Sprache in Text um. In der Pilotimplementierung wird für die Spracherkennung die von Google für den Browser Chrome umgesetzte Implementierung der Web Speech API⁷ verwendet. Diese wäre aber durch eine andere Spracherkennung ersetzbar.

In einem zweiten Schritt muss der Text nun in eine ausführbare Aktion für die Web-Mapping Anwendung umgesetzt werden. Dafür wird die open-source Bibliothek Rasa NLU⁸ genutzt, mit der über eine HTTP Schnittstelle kommuniziert wird. Mit Hilfe von Rasa NLU können aus dem Text *Intents* und *Entities* extrahiert werden. Der Intent besagt, *welche* Funktion ausgeführt werden soll und extrahierte Entities beschreiben genauer, *wie* diese Funktion ausgeführt werden soll. Im Beispiel „Gehe nach Münster“ wäre die auszuführende Funktion „zoom“ mit der zusätzlichen Information, dass nach „Münster“ gezoomt werden soll.

Mit Hilfe dieser Informationen ist es nachfolgend möglich, die in map.apps implementierten Funktionen auszuführen und somit das vom Nutzer gewünschte Ergebnis zu erreichen.

3 Pilotimplementierung Biodiversitätsmonitoring

Als Beispiel aus dem Umweltbereich wurde die Sprachsteuerung prototypisch für eine Anwendung zum Biodiversitätsmonitoring umgesetzt. Die Anwendung informiert über die Verbreitung geschützter Tierarten (Seevögel, Meeressäugetiere) in der deutschen Nord- und Ostsee [Hauswirth & Bluhm 2017]. Die Anwendungen zum Biodiversitätsmonitoring⁹ (ohne Sprachsteuerung) sind beim Bundesamt für Naturschutz mit kurzen Beschreibungen öffentlich nutzbar.

Als Basisfunktionen für die Web-Mapping Anwendung sind beispielsweise „Maßstab setzen“, „Karteninhalt anzeigen“, „Legende anzeigen“, „Kartenausschnitt verschieben“ und „Drucken“ über Sprachbefehle verfügbar.

⁷<https://www.google.com/intl/en/chrome/demos/speech.html> zuletzt aufgerufen 30.08.2018

⁸<https://nlu.rasa.com/> zuletzt aufgerufen 30.08.2018

⁹<https://www.bfn.de/infotehek/karten.html> zuletzt aufgerufen 30.08.2018

Für die Fachanwendung werden weitere Funktionen unterstützt. Dazu gehören die Anzeige von Schutzgebieten („Zeige mir das Naturschutzgebiet Doggerbank“), die Anzeige von Informationen zum Schutzgebiet („Zeige mir Informationen zum Naturschutzgebiet Doggerbank“), sowie die Möglichkeit die Nutzeroberfläche zum Hinzuladen von Themen mit Hilfe von Sprachbefehlen zu steuern. Da letzteres mit Hilfe von Sprachsteuerung teilweise umständlich sein kann, wurden zusätzlich Einsatz-Befehle zum Hinzuladen eines Themas eingeführt, wie folgende Beispiele zeigen:

- Zeige mir Basstölpel im Sommer 2010
- Zeige mir Eiderenten aus dem Sommer 2005
- Lade das Thema Eisente Sommer 2012
- Basstölpel Sommer 2005
- Ich möchte gerne die Daten zur Heringsmöve im Sommer 2007 sehen
- Jetzt zeige mir noch Basstölpel im Herbst 2003

Da es für den Nutzer nicht immer einfach ersichtlich ist, was er überhaupt eingeben kann, sind neben einzelnen Sprachbefehlen zum Steuern von Apps auch Assistenten beziehungsweise Bots sinnvoll. Daher wurde prototypisch für die Anwendung zum Biodiversitätsmonitoring ein einfacher Bot implementiert, der den Nutzer beim Hinzufügen eines neuen Themas unterstützen soll. Für die Umsetzung wurden die selben Technologien genutzt. Ein beispielhafter Gesprächsverlauf könnte wie folgt aussehen:

Nutzer	„Hallo, Bot“
Bot	„Hallo, wie kann ich Ihnen weiterhelfen?“
Nutzer	„Ich möchte gerne ein neues Thema hinzuladen.“
Bot	„Welche Art möchten Sie gerne sehen?“
Nutzer	„Die Art Eiderente.“
Bot	„Für welche Jahreszeit möchten sie gerne Daten sehen?“
Nutzer	„Sommer“
Bot	„Welches Jahr möchten Sie auswählen?“

Nutzer	„2010“
Bot	„Möchten Sie das Thema ‚Eiderenten im Sommer 2010‘ zur Karte hinzuladen oder die Transparenz setzen?“
Nutzer	„Setze die Transparenz auf 60“
Bot	„Möchten Sie das Thema nun zur Karte hinzufügen?“
Nutzer	„Ja“
Bot	„Ich füge das Thema zur Karte hinzu. Auf wiederhören.“

Tabelle 1: Beispiel für einen Gesprächsverlauf zwischen Nutzer und Bot

4 Fazit

Bereits heute werden viele Interaktionen mit Technologie durch Sprache gesteuert. Die exemplarische Umsetzung von Sprachsteuerung für eine Web-Mapping Anwendung hat gezeigt, dass dies auch für Kartenanwendungen realisierbar ist.

Die Spracherkennung funktioniert in der Regel gut, zu Fehlern kommt es insbesondere bei selten gebrauchten Fachbegriffen bzw. Namen. Beispielsweise wird häufig der Begriff „Dreizehenmöwe“ als „13 Möwen“ erkannt.

Bei der Zuordnung von Aktionen ist für das richtige Erkennen von „Intent“ und „Entity“ eine gewisse Anzahl an Texten zum Lernen notwendig. Dies ist abhängig von der Anzahl der Intents und Entities, die für eine Anwendung definiert wurden und demnach von der Komplexität der Anwendung. Prägnante Texte mit Schlüsselwörtern erleichtern dabei das Erkennen.

Bezüglich der Aktionen in der Karte kann festgehalten werden, dass folgende Aktionen sehr gut umsetzbar sind: Auswahl von Themen, Vergrößern, Verkleinern, Verschieben, Objekt über Namen auswählen und Info zu Objekt abfragen.

Nicht so gut funktionieren verständlicherweise Aktionen wie „in Karte selektieren“ oder in der Karte messen. Hier kann ein multimodaler Ansatz helfen, bei dem unterschiedliche Eingabemöglichkeiten zur Verfügung gestellt werden. Der Anwender kann dann die für ihn beste Möglichkeit verwenden und somit den Bedienkomfort erhöhen. Die dargestellte map.apps Anwendung bietet dies. Diese Möglichkeit ist allerdings nicht gegeben, wenn Sprachsteuerung in Situationen genutzt werden soll,

in denen die Bedienung mit Tastatur und Maus nicht möglich ist, z.B. wenn die Hände anderweitig benötigt werden oder im Sinne der Barrierefreiheit für Nutzer mit Einschränkungen.

5 Literaturverzeichnis

Bundesamt für Naturschutz (2017): Schweinswale und Seevögel in Nord- und Ostsee, Pressemitteilung Online-Karten:

(https://www.bfn.de/presse/pressearchiv/2017/detailseite.html?tx_ttnews%5Btt_news%5D=6232&cHash=2254765636fc68d9a87acd026a400f80); zuletzt abgerufen: 26.07.2018

Gartner (2015): Market Trends: Voice as a UI on Consumer Devices — What Do Users Want? (<https://www.gartner.com/doc/3021226/market-trends-voice-ui-consumer>), (zuletzt abgerufen: 26.07.2018)

Giese, M., Bluhm, M. (2017): GruSchu Hessen - Die neue Informationsplattform zum Thema Grund- und Trinkwasserschutz; In: *Tagungsband des 24. Workshops "Umweltinformationssysteme 2017 - Vernetzte Umweltdaten (UIS 2017)" des Arbeitskreises "Umweltinformationssysteme" der Fachgruppe "Informatik im Umweltschutz" der Gesellschaft für Informatik (GI)*; <http://ceur-ws.org/Vol-1919/paper3.pdf>. (zuletzt aufgerufen 26.07.2018); Brandenburg an der Havel, Deutschland, 18.-19. Mai 2017. Herausgegeben von Ulrike Freitag, Frank Fuchs-Kittowski, Friedhelm Hosenfeld, Andreas Abecker, Dietmar Wikarski

Hackmann, Ralf (2015): Building and managing next generation geo apps. In: *Geospatial World Forum 2015*, Lisbon. <https://geospatialworldforum.org/speaker/SpeakersImages/%20Ralf%20Hackmann.pdf> (zuletzt aufgerufen 26.07.2018)

Hauswirth, M., Bluhm, M. (2017): Biodiversitätsmonitoring in der Ausschließlichen Wirtschaftszone (AWZ); In: *Tagungsband des 24. Workshops "Umweltinformationssysteme 2017 - Vernetzte Umweltdaten (UIS 2017)" des Arbeitskreises "Umweltinformationssysteme" der Fachgruppe "Informatik im Umweltschutz" der Gesellschaft für Informatik (GI)*; Brandenburg an der Havel, Deutschland, 18.-19. Mai 2017. Herausgegeben von Ulrike Freitag, Frank Fuchs-Kittowski, Friedhelm Hosenfeld, Andreas Abecker, Dietmar Wikarski. <http://ceur-ws.org/Vol-1919/paper4.pdf> (zuletzt aufgerufen 26.07.2018)

Heber M., Kruse K. (2016): map.apps Kartenanwendung - „Bodenatlas Deutschland“ digital. In: ESRI-Anwendertreffen Norddeutschland, LGLN, Hannover, 10. Mai 2016. http://www.pointsgeoinformatik.de/fileadmin/dateien/BGR_Kruse_Bodenatlas.pdf (zuletzt abgerufen 27.07.2018)

Protalinski, Emil (2017): Google's speech recognition technology now has a 4.9% word error rate. <https://venturebeat.com/2017/05/17/googles-speech-recognition-technology-now-has-a-4-9-word-error-rate/> zuletzt abgerufen: 26.07.2018

Uhlenkükén, Christoph (2018): „map.apps - Innovative Apps für Web und Mobile“. <https://www.conterra.de/mapapps> (zuletzt aufgerufen 26.07.2018)