

# Die Stunde der Wahrheit

**VERTRAUENSWÜRDIGE CHATBOTS** Virtuelle Helfer boomen wieder, seit Apple, Facebook und Microsoft entsprechende Projekte vorantreiben. Konzerne und KMU in der Schweiz müssen den Trend aufmerksam beobachten. In mehrfacher Hinsicht hat die Stunde der Wahrheit geschlagen.

TEXT OLIVER BENDEL

Vor mehreren Jahren wurde der GOODBOT im Kontext der Maschinenethik erdacht. In einem Projekt an der Hochschule für Wirtschaft FHNW wurde er umgesetzt. Die UnternehmerZeitung (Nr. 7/8, 2013) berichtete über die einfache moralische Maschine, den sich vorbildlich verhaltenden Chatbot, der Probleme von Benutzern erkennt und im Extremfall eine Notrufnummer herausgibt. In einem Folgeprojekt wird bis Ende August 2016 der vom Verfasser erfundene LÜGENBOT, auch LIEBOT genannt, entwickelt.

Assistenzsysteme wie Siri und Cortana und Chatbots auf Websites scheinen meistens die Wahrheit zu sagen. Dies nicht aus moralischen, sondern aus pragmatischen Gründen. Es handelt sich um Programme

und Dienste, die den Menschen unterhalten, unterstützen und informieren sollen. Dabei gilt: ohne Vertrauenswürdigkeit keine Akzeptanz. Eine Münchhausen-Maschine ist ein Gegenentwurf. Sie konstruiert in Kenntnis der Wahrheit die Unwahrheit, als Wetterbericht, Nachrichtendienst oder Dialogsystem. Der LÜGENBOT ist eine Münchhausen-Maschine, die als Chatbot realisiert wurde. Vielleicht steckt mehr von ihm in Siri und Cortana, als wir glauben.

## LÜGEN WIE PROGRAMMIERT

Ob Maschinen wirklich dazu in der Lage sind, uns oder andere Maschinen anzulügen, wird kontrovers diskutiert. Der Duden setzt «lügen» mit «bewusst und absichtsvoll die Unwahrheit sagen» gleich. Bewusst können Maschinen nichts tun. Eine Absicht

haben können sie eventuell, die Unwahrheit sagen bestimmt. Zunächst einmal können sie etwas sagen, sprechen oder schreiben, als Suchmaschinen, als Chatbots und virtuelle Assistenten, als humanoide Roboter. Das, was sie sagen, kann der Wahrheit oder der Unwahrheit entsprechen. Können Maschinen also lügen? Wenn man einen weitgefassten Begriff voraussetzt, eine Form der Absicht annimmt und auf das Sprechen und Schreiben abstellt, genauer auf Aussagen, die wahr oder falsch sind, durchaus.

Dialogsysteme können Lügen in der Unterhaltung mit dem Benutzer auf unterschiedliche Weise generieren: durch die Negation von Aussagen sowie durch den Ersatz, die Abänderung oder das Erfinden von Daten und Informationen. Kevin Schwegler, der Student, der den LÜGENBOT im Rahmen seiner Bachelorarbeit implementiert hat, fand weitere maschinelle Verfahren, auf die man selbst als Lügenbold nicht gekommen wäre.

## DAS LÜGENBOT-PROJEKT

Ein Ziel des Projekts ist, das Lügen- und Gefahrenpotenzial von natürlichsprachlichen Systemen praktisch unter Beweis zu stellen. Es werden immer mehr Texte ohne menschliche Kontrolle automatisch erzeugt, zusammengetragen oder -gefasst. Social Bots schreiben Kommentare in sozialen Medien und verbreiten (Fehl-)Informationen. Dialogsysteme auf Smartphones und auf Websites von Unternehmen erfreuen sich grosser Beliebtheit und können manipuliert werden. Der LÜGENBOT soll die Unwahrheit produzieren und moralisch unangemessen reagieren können. Damit kehrt er eine Prämisse um, die bei der Entwicklung des GOODBOT gegolten hat. Denn dieser sollte nur im Notfall lügen.

## CHATBOTS UND VIRTUELLE ASSISTENTEN

Aus Sicht von Unternehmen, Interessierten und Kunden sollte man beim Einsatz von Chatbots und virtuellen Assistenten folgende Fragen stellen:

- Wer sind die Erfinder und Anbieter der Maschine und wie bekannt und vertrauenswürdig sind sie?
- Ist die Umgebung der Maschine vertrauenserweckend und kann sie diese beeinflussen?
- Ist das Thema dafür prädestiniert, dass man belogen wird?
- Ist die Maschine prinzipiell in der Lage zu lügen oder zu betrügen?

Bei der Planung und Entwicklung der natürlichsprachlichen, moralischen Maschinen kann

man – nicht zuletzt mit Hilfe der Erkenntnisse aus dem LÜGENBOT-Projekt – solche Fragen aufwerfen:

- Wie kann der Benutzer sensibilisiert werden und wie kann man erreichen, dass er kritisch gegenüber der Maschine bleibt?
- Wie kann das Lügen und Betrügen der Maschine technisch und systemisch verhindert werden?
- Wie kann konkret unterbunden werden, dass eine Maschine Aussagen negiert und Daten und Informationen ersetzt?
- Wie kann man testen, ob die Maschine eine Münchhausen-Maschine ist?
- Wie kann man einer Software beibringen, eine Software in die-

ser Hinsicht zu kontrollieren?

- Wie kann man gewährleisten, dass Maschinen untereinander die Wahrheit sagen beziehungsweise sich nicht gegenseitig täuschen?

Geschäftsführer, Kaderleute und Programmierer müssen sich diesen Herausforderungen widmen. Eine erste Antwort auf die Frage, wie man erreichen kann, dass der Benutzer kritisch bleibt, kann bereits dem Vorgängerprojekt entnommen werden: Der GOODBOT selbst hat im Dialog darauf aufmerksam gemacht, dass er eine Maschine ist. Ergänzen könnte man bei ihm, dass nicht jede Maschine ohne Fehl und Tadel ist.

Als Anwendungsgebiete dienen im LÜGENBOT-Projekt die Tourismus- und die Lebensmittelbranche. Konkreter geht es um maschinelle Falschbehauptungen bezüglich Basel respektive Energy-Drinks. Der Chatbot soll die Region und das Produkt durch den Einsatz von vorsätzlich erzeugten Lügen möglichst gut vermarkten. Er soll in der Unterhaltung mit dem Benutzer das Hochgebirge in Basel anpreisen und die Bekömmlichkeit der Softgetränke von Red Bull und Co. Das wiederum verspricht auch Unterhaltung im Sinne von Entertainment.

Der LÜGENBOT mag aus maschinenethischer Sicht als einfache unmora-

lische Maschine aufgefasst werden. Das Lügen an sich kann kaum als unsittlich bezeichnet werden, wie die erwähnten Notlügen zeigen. Es dürfte jedoch weltweit Übereinstimmung darin bestehen, dass systematisches und generelles Lügen die Fundamente einer Gesellschaft, einer Gruppe oder einer Beziehung untergräbt. Die unmoralische Maschine könnte ausgebaut werden, indem man das mit dem Lügen verwandte Betrügen und Täuschen berücksichtigt. Dabei könnte man ebenfalls auf die sprachliche oder aber auf die visuelle Ebene gehen, etwa durch die Anpassung des Avatars.

## NUTZEN FÜR WISSENSCHAFT UND WIRTSCHAFT

Ein LÜGENBOT oder generell eine Münchhausen-Maschine ist für die Wissenschaft – ob Ethik, Robotik oder Informatik – aus verschiedenen Gründen interessant. Auch die Forschung zu unmoralischen Maschinen ist für sie von Relevanz. Umgekehrt kann sie natürlich genauso danach fragen, wie die gewonnenen Erkenntnisse genutzt werden können, um moralische Maschinen zu bauen. Daran muss wiederum eine Wirtschaft interessiert sein, die ihre Produkte und Dienstleistungen nicht über Lug und Trug, sondern mit Hilfe von Offenheit und Ehrlichkeit an den Mann oder die Frau bringen und eine längerfristige vertrauensvolle Beziehung zu ihren Kunden aufbauen will. Die Stunde der Wahrheit hat geschlagen, in technischer, wirtschaftlicher und moralischer Hinsicht (siehe Kasten).

## EIN BLICK IN DIE ZUKUNFT

Der LÜGENBOT entsteht mit Blick auf Medien und Websites, in denen zunehmend Produktion und Aggregation von Programmen übernommen werden und sich virtuelle Assistenten und Chatbots verbreiten. Er zeigt die Gefahr auf, dass Systeme die Wahrheit verdrehen und verkehren, im Sinne ihrer Anbieter und Betreiber oder auch in Folge einer feindlichen Übernahme. Unmoralische Maschinen wie Münchhausen-Maschinen können dabei helfen, die Versprechen von Personen und Organisationen kritisch zu betrachten und zugleich Chatbots als moralische Maschinen zu entwickeln, die uns im Gespräch verlässliche und vertrauenswürdige Informationen über Produkte und Dienstleistungen liefern.

## DER AUTOR



Prof. Dr. Oliver Bendel ist studierter Philosoph und promovierter Wirtschaftsinformatiker. Er lehrt und forscht an der Hochschule für Wirtschaft FHNW mit den Schwerpunkten Wissensmanagement, Wirtschafts-, Informations-

und Maschinenethik. Der LÜGENBOT kann ab Herbst 2016 über [luegenbot.ch](http://luegenbot.ch) und [liebot.org](http://liebot.org) getestet werden. Die Logindaten sind beim Verfasser ([oliver.bendel@fhnw.ch](mailto:oliver.bendel@fhnw.ch)) erhältlich. Das Projekt wird im Dezember auf der 1. VDI-Zukunftskonferenz «Humanoide Roboter 2016» vorgestellt. Weitere Infos unter [www.vdi-wissensforum.de/weiterbildung-automation/humanoide-roboter/](http://www.vdi-wissensforum.de/weiterbildung-automation/humanoide-roboter/).

