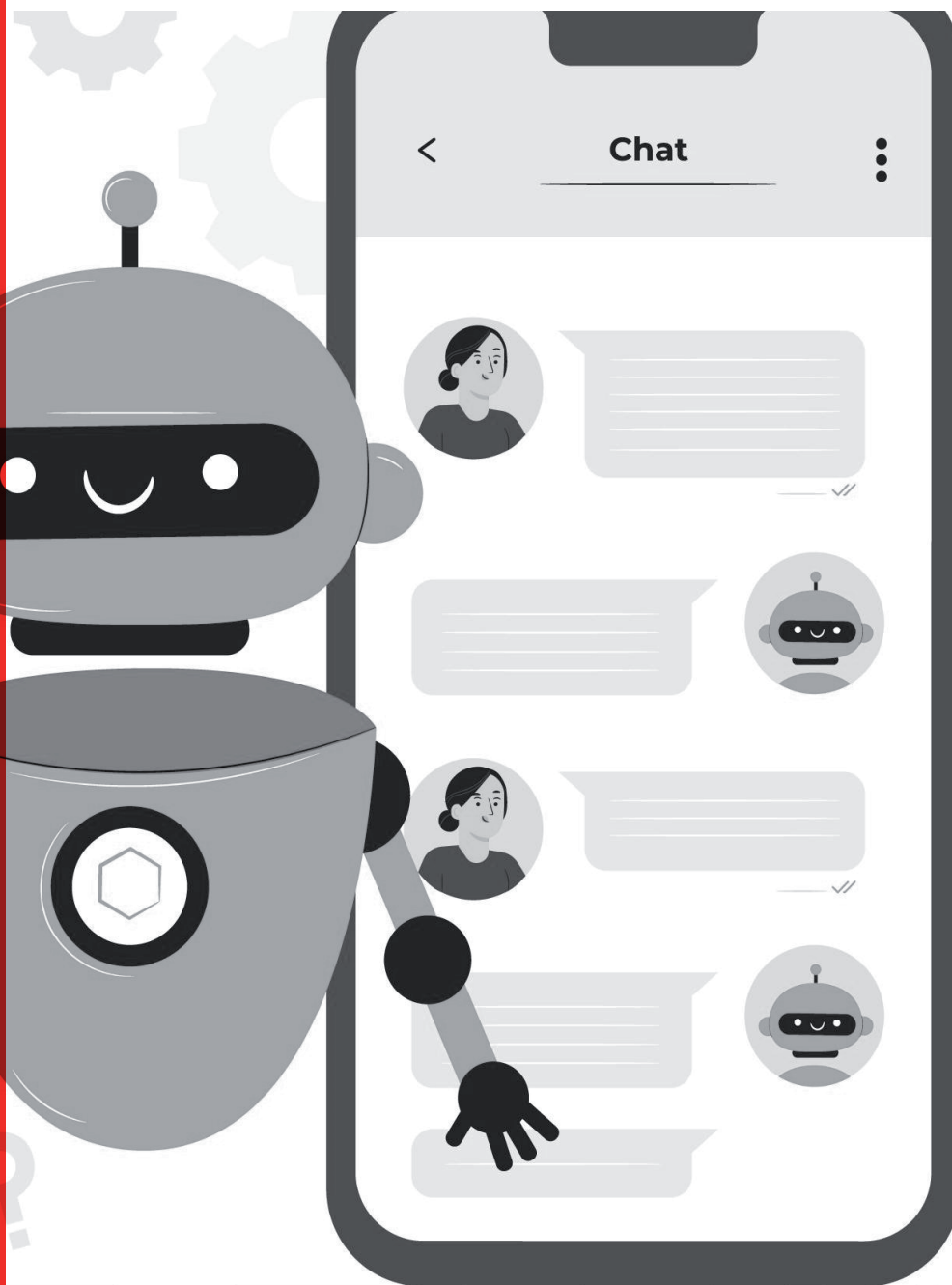


Swiss  
Insights  
News



#04

KI-Chatbots und Ethik –  
Geht das zusammen?



# KI-Chatbots und Ethik – Geht das zusammen?



**Sophie Hundertmark**  
Doktorandin Institut für Finanz-  
dienstleistungen Zug (IFZ),  
selbstständige Chatbot-Beraterin  
Zürich

Was eine KI lernt und ob sie adäquate Antworten auf Fragen der User liefert, müssen ihre Betreiber permanent überprüfen. Auch ein Ethik-Check ist sinnvoll, ganz besonders wenn die KI einen Chatbot steuert und so einen Teil der Kundenkontakte eines Unternehmens pflegt.

Digitale Technologien wie die künstliche Intelligenz (KI) setzen sich zunehmend durch. Es gibt unterschiedliche Ansichten darüber, was KI leisten kann und was nicht. Problematisch ist, dass es keine Anforderungen an die Transparenz der KI gibt, ja nicht einmal klar ist, was genau mit KI gemeint ist.

Das Herzstück der heutigen KI sind die Daten, mit denen sie gefüttert wird. Doch im Gegensatz zu Öl, mit dem sie oft verglichen werden, sind diese Daten lebendig, weil sie auch von uns Menschen produziert werden. Sind die Daten unvollständig, veraltet oder gar falsch, lernt die KI auf falschen Annahmen, was letztlich zu unbefriedigenden Leistungen und Unmut bei den Nutzern führt.

Darüber hinaus können Datensätze bestimmte Verzerrungen, Vorurteile und Voreingenommenheiten enthalten, die zu Diskriminierung führen, insbesondere bei Entscheidungen, die auf personenbezogenen Daten basieren.

Hubert Österle hat die Disziplin des Life Engineering entwickelt, um genau solchen Gefahren oder Problemen entgegenwirken zu können. Er berücksichtigt dabei auch ethische Aspekte der KI-Entwicklung und -Nutzung. Laut Österle H. (Österle H., 2020) besteht die Aufgabe des Life Engineering darin, Regeln für digitale Systeme zu entwickeln, die auf das menschliche Wohlbefinden ausgerichtet sind. Life Engineering ist eine Weiterentwicklung des Business Engineering, jedoch mit einem anderen Fokus: statt auf dem Unternehmen (Gewinnorientierung) liegt der auf der Gesellschaft (Lebensqualität). Mit

dem Fokus auf «Human» wird beim Life Engineering nicht die Gesellschaft als Einheit, sondern der einzelne Mensch als zentraler Punkt der Betrachtung gesehen.

## **Bessere KI-Chatbots dank dem Turing-Test?**

Beeinflusst durch die technologischen und wissenschaftlichen Fortschritte in der KI sowie durch die wachsende Akzeptanz nicht-menschlicher Kommunikationspartner hat in letzter Zeit die Zahl der Unternehmen zugenommen, welche Chatbots oder Conversational Agents (CAs) zur Automatisierung ihrer Kundenkontakte einsetzen. Man kann eigentlich schon sagen, dass KI-basierte CAs wie Amazons Alexa oder Apples Siri zu einer wichtigen Serviceschnittstelle zwischen Anbietern und Nutzern geworden sind.

Zunächst sieht es so aus, als ob diese Digitalen Assistenten dazu entwickelt werden, ihre Nutzer im Alltag als intelligente persönliche Assistenten zu unterstützen. Die Chatbots simulieren dabei die menschliche Kommunikation und können im Vergleich zu anderen Software-Lösungen menschliche Eigenschaften besser annehmen.

Um zu testen, ob ein Conversational Agent so gut ist wie ein menschlicher Gesprächspartner, können Unternehmen den bereits existierenden Turing-Test anwenden. Dieser 1950 von Alan Turing entwickelte Test beschreibt eine Möglichkeit, die Intelligenz von Maschinen zu testen. Beim Turing-Test unterhält sich ein menschlicher Fragesteller mit zwei bis drei anderen Gesprächspartnern. Die Konversation

findet ausschliesslich per Chat statt. Das Interessante an dem Gespräch ist, dass einer der Gesprächspartner eine Maschine ist und die anderen ein oder zwei echte Menschen.

Der Fragesteller weiss nicht, hinter welchem Gesprächspartner sich die Maschine verbirgt. Der Interviewer hat die Aufgabe, mindestens zehn Minuten lang intensiv Fragen zu stellen. Am Ende muss er entscheiden, welcher seiner Gesprächspartner ein Mensch und welcher eine Maschine ist.

Wenn der Fragesteller nicht eindeutig herausfindet, wer eine Maschine ist, hat der Chatbot oder die Maschine den Turing-Test bestanden. Allerdings werden beim Turing Test keinerlei ethische Aspekte berücksichtigt. Es geht nicht darum, ob sich ein Conversational Agent für den User ethisch korrekt verhält, sondern nur, wie nah sein Verhalten dem eines durchschnittlichen menschlichen Gesprächspartners kommt.

### **Der Ethik-Check für Bots**

In einem interdisziplinären Projekt möchten wir nun die Ideen des Life Engineerings und den daraus resultierenden Ideen zur Entwicklung ethischer Systeme mit denen des Turing-Tests kombinieren. Das Ziel ist ein Ethik-Check für Chatbots.

Inspiziert vom Turing-Test, der herausfinden soll, wie menschenähnlich sich eine KI verhält, wollen wir ethisch korrekte Bots identifizieren und denjenigen, die es noch nicht sind, Hinweise zur Optimierung geben. Um ethisch korrekte Bots als solche zu definieren, müssen wir zunächst festlegen, was ethisch korrekt in unserer Sprache und Kultur bedeutet. Und wir müssen berücksichtigen, dass sich diese Werte im Laufe der Zeit ändern können.

Anschliessend müssen wir Verfahren definieren, wie die zuvor entwickelten ethischen Kriterien gemessen werden können. Darauf folgt eine Bewertung, der sich jeder Bot unterziehen kann, um eine Einschätzung des Zustands seiner ethischen Korrektheit einschliesslich eines eventuellen Verbesserungspotenzials zu erhalten. Die Bewertung wird zunächst von Menschen durchgeführt, es ist aber denkbar, dass in Zukunft auch Bots in der Lage sein werden, die Bewertung vorzunehmen.

Der Ethik-Check für Conversational Agents ist in zweifacher Hinsicht neu. Zum einen gibt es weder in der Forschung noch in der Praxis anerkannte und weit verbreitete Richtlinien für Chatbot-Projekte. Zum anderen gibt es, abgesehen vom Turing-Test, keine Benchmarks oder andere Tests, die Conversational Agents bewerten und gleichzeitig Optimierungsmöglichkeiten aufzeigen. Unser Projekt kombiniert beide Aspekte in einer Anwendung.

Im Gegensatz zum Turing-Test, bei dem ein Mensch mit einer Conversational AI chattet, soll unser Ethik-Check am Ende direkt von einem Chatbot durchgeführt werden. Der Chatbot verfügt dann über die definierten Ethikregeln, weiss, welche Fragen er stellen muss und kann die Antworten des zu testenden Chatbots mit den von uns definierten Benchmarks abgleichen und bewerten. Das Ergebnis ist ein vollautomatischer Ethik-Check, der zudem transparent ist, da er genau aufzeigt, welche Kriterien in welchem Umfang in den Entscheidungsprozess eingeflossen sind.

Bevor eine Ethikprüfung stattfinden kann, muss definiert werden, was ethisch korrekt bedeutet. Ethik ist nicht universell. Ethik ist etwas, das sich ständig weiterentwickelt und stark von der Kultur geprägt ist.

Wir definieren daher zuerst ethische Standards für eine spezifische Region. Als erste haben wir den deutschsprachigen Raum ausgewählt. Wir nennen diese Region Smart Region, da wir unsere Standards auf der Grundlage der Ideen von Smart Cities entwickeln wollen. Diese beziehen die Bürgerinnen und Bürger typischerweise in den Prozess der Ideenfindung und Entwicklung ein.

Mit Hilfe der verschiedenen Forschungsmethoden Literaturrecherche, Umfragen, Experteninterviews und Fokusgruppen werden wir ethische Standards für Chat- und Voicebots entwickeln und erste Ansätze präsentieren, wie diese gemessen werden können.

Für die Fokusgruppen werden Experten aus verschiedenen Disziplinen ausgewählt. Im Mittelpunkt stehen Psychologie, Datenschutz, Informatik, Digitalisierung, Bildung, Data Science, Marketing, Wirtschaft. Um sich ändernden ethischen Werten gerecht zu werden, integrieren

wir einen Mechanismus zur kontinuierlichen Anpassung an die aktuell geltenden Normen.

In weiteren Iterationen können konkrete Methoden oder Fragen entwickelt werden, um herauszufinden, wie ethisch korrekt sich ein Bot verhält. Anbieter, meist Unternehmen, die Bots für ihre Kunden und Mitarbeiter einsetzen, erhalten zudem Anregungen, wie sie ihre Bots ethisch korrekter reagieren lassen können.

An diesem Punkt haben wir wahrscheinlich eine weitere Herausforderung, nämlich dass viele KI-Projekte auf spezifische Anwendungsfälle ausgerichtet sind. Wir müssen es also schaffen, die ethische Korrektheit trotz dieser Einschränkungen zu messen, und möglicherweise Regeln dafür definieren.

In der ersten Phase wird die ethische Prüfung von Menschen durchgeführt. Die Menschen chatten mit dem Bot, stellen die relevanten Fragen, notieren die Antworten und bewerten diese anhand eines zuvor festgelegten Bewertungsrasters. In weiteren Phasen soll das Chatting und die Ethikprüfung von einem Chatbot übernommen werden, sodass langfristig die gesamte Ethikprüfung voll automatisiert werden kann.

Sobald eine KI einen anderen Bot bewerten muss, werden wir auf die Ansätze der Fuzzy-Logik und des Computing With Words (CWW) zurückgreifen. Fuzzy-Systeme können mit unscharfen Daten umgehen und sind daher sehr gut geeignet, wenn es darum geht, die Äusserungen von Menschen oder Bots zu charakterisieren oder ihre ethische Reife zu testen.

Forscher, die Fuzziness anwenden, ordnen Wörter nicht nur in Kategorien ein, sondern berücksichtigen auch ihre Position innerhalb der Kategorie.

Darüber hinaus ist Computing With Words ein auf der Fuzzy-Logik basierendes Rechensystem, in dem die Objekte der Berechnung vor allem Wörter, Sätze und Propositionen aus einer natürlichen Sprache sind, wie wir sie in unseren Chat-Gesprächen verwenden.

### **Entwicklungsstand des Ethik-Checks**

Auf Basis einer ersten Literaturrecherche haben wir erste Design-Richtlinien für ethische

Prinzipien definiert. Nun geht es darum, das Feedback unserer definierten Smart Region zu bekommen und dazu wurde unser so genannter Evaluation Bot entwickelt. Dieser Bot stellt alle von uns gefundenen Design-Prinzipien vor und fragt die Nutzer nach ihrer Einschätzung zur ethischen Wichtigkeit der jeweiligen Prinzipien.

Da die Evaluation immer noch läuft, freuen wir uns natürlich über alle User, die sich fünf bis zehn Minuten Zeit nehmen und ihr Feedback unserem Evaluation Bot mitteilen. Der Bot kann hier besucht werden: <https://eggheads.ai/chat/2738>

### **Die Autorin**

Sophie Hundertmark gehört zu den ersten Master-Studentinnen in der Schweiz, die zu Chatbots geforscht haben. Seitdem arbeitet sie als selbständige Chatbot Beraterin und ist zudem wissenschaftliche Mitarbeiterin an der Hochschule Luzern. Anfang 2020 hat sie ihren eignen Chatbot Podcast gestartet, unter welchem alle zwei Wochen eine neue Folge herauskommt. Ende 2020 ist ihr erstes Buch erschienen «Digitale Freunde», ein praxisnahes Sachbuch über Chatbots. Seit 2021 ist Sophie Doktorandin an der Universität Fribourg und forscht und publiziert regelmässig zu Conversational Agents und Ethics.

### **Kontakt**

Sophie Hundertmark  
+41 78 900 5346, [sophie@hundertmark.ch](mailto:sophie@hundertmark.ch)



# Swiss Insights

Swiss Insights ist der Verband und die Interessensvertretung aller Unternehmen, die Daten und prädiktive Modelle im Rahmen von Marketing, Innovationsprozessen, Kundenservice, Angebotsgestaltung, Kommunikation und Zielgruppendefinitionen erheben, analysieren, einsetzen und daraus Handlungsempfehlungen ableiten.

Swiss Insights pflegt einen aktiven Dialog mit politisch und gesellschaftlich wichtigen Akteuren und fördert den Austausch mit anderen nationalen und internationalen Fachorganisationen.

Eine der Hauptaufgaben des Verbands ist die Förderung der Markt-, Meinungs- und Sozialforschung im Allgemeinen und der Wissenschaftlichkeit im Besonderen. Er entwickelt, definiert und unterhält strenge Leitlinien zur Qualitätssicherung und grenzt sich im Bereich

der Markt- und Sozialforschung klar von Werbung und Direktmarketing ab. Hierzu führt der Verband das Qualitätslabel «Market & Social Research by Swiss Insights».

Darüber hinaus engagiert sich Swiss Insights dafür, dass die Nutzung von Daten und die Anwendung von datengetriebenen Modellen transparent, nachvollziehbar und in diesem Sinne fair gestaltet wird. Hierfür wurde das Label «Data Fairness by Swiss Insights» geschaffen.



## Herausgeber und Kontakt

Swiss Insights, Swiss Data Insights Association, Gruebengasse 10, 6055 Alpnach, Switzerland  
+41 44 3501960, [info@swiss-insights.ch](mailto:info@swiss-insights.ch), [www.swiss-insights.ch](http://www.swiss-insights.ch)

# Unsere Mitglieder

## SWISS INSIGHTS Institute Member

Im Verband sind alle relevanten **Markt- und Sozialforschungsinstitute** unter einem Dach organisiert. Alle Mitgliedsinstitute unterliegen einem strengen Regelwerk von schweizerischen und internationalen Normen.

Mitgliedsinstitute dürfen das Label **Market & Social Research by SWISS INSIGHTS** und je nach Tätigkeitsgebiet das Label **Data Fairness by SWISS INSIGHTS** tragen.

Die Mitgliederliste finden Sie auf der nachfolgenden Seite.

## SWISS INSIGHTS Corporate Member

Die Corporate Mitglieder sind **Unternehmen**, die sich für den fairen Umgang mit Auskunftspersonen und Auftraggeber sowie für den Schutz der Privatsphäre engagieren.

Corporate Member, die das Label **Data Fairness by SWISS INSIGHTS** tragen, stehen für den wissenschaftlichen, seriösen und respektvollen Umgang mit Daten ein.

Die Mitgliederliste finden Sie auf der nachfolgenden Seite.



# SWISS INSIGHTS

## Institute Member

### amPuls Market Research

Hirschengraben 49, 6000 Luzern 7  
+41 41 612 14 14 / info@ampuls.ch  
www.ampuls.ch

### amrein+heller MarktforschungsTreuhand AG

Südweid 7, 6274 Eschenbach  
+41 748 63 70 / contact@ah-feedback.ch  
www.ah-feedback.ch

### Bilendi Schweiz AG

Reinhardstrasse 19, 8008 Zürich  
+41 79 801 88 80 / contact.ch@bilendi.com  
www.bilendi.ch

### Boomerang Ideas AG

Sihlquai 131, 8005 Zürich  
+41 44 500 88 60 / raphael@boomerangideas.com  
www.boomerangideas.com

### Constant Dialog

Alte Steinhäuserstrasse 33, 6330 Cham  
+41 41 310 05 40 / info@constant-dialog.ch  
www.constant-dialog.ch

### DemoSCOPE Data + Research

Klusenstrasse 17, 6043 Adligenswil  
+41 41 375 40 00 / demoscope@demoscope.ch  
www.demoscope.ch

### dr-ouwerkerk ag – just-medical!

Blegistrasse 5, 6340 Baar  
+41 41 766 11 55 / info@just-medical.com  
www.pharmaagentur.ch

### Gallup AG

Reinhardstrasse 19, 8008 Zürich,  
+41 78 891 31 15 / office@gallup.swiss  
www.gallup.swiss

### gff Swiss Research Services

Baarerstrasse 25, 6300 Zug  
+41 41 560 01 60 / gut@gff.ag,  
www.gff.ag

### GfK Switzerland AG

Suurstoffi 18A, 6343 Rotkreuz  
+41 41 632 91 11 / info.ch@gfk.com  
www.gfk.ch / www.gfk.com

### gfs.bern. Menschen. Meinungen. Märkte.

Effingerstrasse 14, Postfach, 3001 Bern  
+41 31 311 08 06 / info@gfsbern.ch  
www.gfsbern.ch

### gfs-zürich, Markt- & Sozialforschung

Riedtlistrasse 9, 8006 Zürich  
+41 44 360 40 20 / gfs@gfs-zh.ch  
www.gfs-zh.ch

### gfs-befragungsdienst

Schaffhauserstrasse 491, 8052 Zürich  
+41 44 360 26 40 / info@gfs-bd.ch  
www.gfs-bd.ch

### GIM Suisse AG

General-Wille-Strasse 10, 8002 Zürich  
+41 44 283 18 18 / info@g-i-m.ch  
www.g-i-m.ch

### Happy Thinking People AG

Staufacherstrasse 101, 8048 Zürich  
+41 44 204 16 26 / contact-zurich@happythinkingpeople.com  
www.happythinkingpeople.com

### INNOFACT (Schweiz) AG Research & Consulting

Flurstrasse 50, 8048 Zürich  
+41 43 931 77 82, Info@innofact.ch  
www.innofact.ch

### Instight Institute AG

Bergstrasse 138, 8032 Zürich  
+41 44 387 90 90 / info@insightinstitute.ch  
www.insightinstitute.ch

### intervista

Optingenstrasse 5, 3013 Bern  
+41 31 511 39 00 / anfragen@intervista.ch  
www.intervista.ch

### IPSOS Suisse SA

11, Chemin du Château-Bloch, 1219 Le Lignon  
+41 22 591 06 00 / Contact\_Switzerland@ipsos.com  
www.ipsos.com/de-ch

### Kantar Media Switzerland AG

Bahnhofstrasse 4, 3073 Gümligen  
+41 31 537 79 00 / ch.panel@kantarmedia.com  
www.kantarmedia.com

# SWISS INSIGHTS

## Institute Member

### LINK

Baslerstrasse 60, 8048 Zürich  
+41 41 367 73 73 / zurich@link.ch  
www.link.ch

### Marketagent.com Schweiz AG

Seefeldstrasse 19, 8008 Zürich  
+41 43 555 06 50, schweiz@marketagent.com  
www.marketagent.com

### M.I.S. Trend SA

Pont Bessières 3, 1005 Lausanne  
+41 21 320 95 03 / info@mistrend.ch  
www.mistrend.ch

### onlineumfragen.com GmbH

Kernserstrasse 15, 6056 Kägiswil  
+41 44 500 50 54 / info@onlineumfragen.com  
www.onlineumfragen.com

### POLYQUEST AG

Flurstrasse 26, 3014 Bern  
+41 31 335 64 00 / info@polyquest.ch  
www.polyquest.ch

### Publicom AG

Alte Landstrasse 55, 8802 Kilchberg  
+41 44 716 55 11 / publicom@publicom.ch  
www.publicom.ch

### Qualitest AG, Institut für Marketing- und Sozialforschung

Rosenberghöhe 3, 6004 Luzern  
+41 41 712 12 21 / qualitest@qualitestag.ch  
www.qualitestag.ch

### SensoPLUS

Industriestrasse 16, 6300 Zug  
+41 41 710 71 61 / info@sensoplus.ch  
www.sensoplus.ch

### NielsenIQ (Switzerland) GmbH

Park 4, 6039 Root D4  
+41 41 445 64 64 / nielsen-ch@nielsen.com  
www.nielsen.com

### TransferPlus AG Market Research

Haldenstrasse 11, 6006 Luzern  
+41 41 618 33 11 / transfer@transferplus.ch  
www.transferplus.ch

# SWISS INSIGHTS

## Corporate Member

### TX Group AG

Werdstrasse 21, 8021 Zürich  
+41 44 248 41 11  
www.tx.group

### BSI Business Systems Integration AG

Täferweg 1, 5405 Baden  
+41 58 255 90 00, info@bsi-software.com  
www.bsi-software.com