

# Erfahrungsbericht: Die Bedeutung von Requirements Engineering für KI-Projekte

Prof. Dr. Andrea Herrmann  
AKAD-Hochschule, Heilbronner Straße 86, 70191 Stuttgart, AH\_Science@gmx.de

Requirements Engineering ist die Tätigkeit, mit der Anforderungen ermittelt, dokumentiert, geprüft und verwaltet werden. Anforderungen und Requirements Engineering ist für alle technischen Projekte wichtig und genau genommen auch für nicht-technische Projekte. Darum sollte man erwarten, dass sie auch dann wichtig sind, wenn eine Software des Typs Künstliche Intelligenz (KI) entwickelt wird. Aber ist eventuell irgendetwas anders, überflüssig, zusätzlich nötig? Lassen Sie uns vier meiner bisherigen KI-Projekte betrachten.

## Der Ethicbot

Der Ethicbot [1] ist ein Chatbot, der den Benutzer bei ethischen Konfliktsituationen unterstützt. Hauptziel und Motivation für dessen Entwicklung bestand darin herauszufinden, was die KI kann. Und das nicht anhand einer simplen Klassifikation, sondern einer komplexeren Aufgabe. Die KI sollte hier ausdrücklich selbst keine ethische Wertung vornehmen, sondern den Benutzer durch Fragen bei seinen Überlegungen anleiten. Hinter seinem Gesprächsfluss stecken mehrere Konzepte aus der Literatur der ethischen Entscheidungsfindung. Insbesondere sollte sowohl utilitaristisch als auch deontologisch argumentiert werden. Es wurde eine Nutzwertanalyse durchgeführt, um utilitaristisch die objektiv beste Lösung zu finden, aber auch ermittelt, welche Regeln hier zum Einsatz kommen könnten und wie frühere Benutzer des Ethicbots in einer ähnlichen Situation entschieden hatten. Am kniffligsten erwies sich hier die Spracherkennung, die in einer Freitextgeschichte die handelnden Personen, Gruppen, Maschinen oder Tiere identifizieren sollte, sowie deren Aktionsmöglichkeiten und eine Bewertung wie „Zu sterben ist schlecht. Stimmen Sie mir zu?“

Die Anforderungen an den Ethicbot haben wir in Form von Personas und User Stories spezifiziert. Dabei waren zahlreiche Anwendungsszenarien denkbar, aber: Eine Person, die sich mit ethischen Entscheidungsfindungsmethoden auskennt, benötigt keine Unterstützung eines Chatbots, um eine Nutzwertanalyse zu erstellen, und wer keiner-

lei Vorbildung hat, wird sich eventuell mit dieser systematischen Denkweise schwer tun. Der Ethicbot war und blieb darum ein Forschungsprototyp, dessen Evaluierung mit Testbenutzern zu interessanten Erkenntnissen über Spracherkennung, Ethik, Entscheidungen, Chatbots und Usability von Chatbots führte. Tatsächlich war der Ethicbot aber von der Lösung her gedacht: „Wir entwickeln eine Lösung und suchen anschließend das Problem, das sie löst.“

## Gut oder böse?

Auch in diesem Projekt ging es um Ethik und Sprachanalyse, aber die Forschungsfrage war konkreter: Kann eine KI mit Hilfe von unüberwachtem Lernen zwischen Gut und Böse unterscheiden? [2] Die KI sollte innerhalb eines Märchens die Begriffe clustern, und wir fragten uns, ob sie die Figuren, Symbole, Gegenstände und Handlungen richtig in die Nähe von „Gut“ oder „Böse“ platzieren würde.

Schwierig war die Suche nach Texten, die ethische Fragen behandeln und klar zwischen Gut und Böse unterscheiden. Die Bibel fiel aus, weil sie nur mit viel Hintergrundwissen verständlich ist und die Kapitel sich implizit aufeinander beziehen. Märchen erschienen passender, weil hier die Personen konsequent gut oder böse handeln, und jedes Märchen ein in sich geschlossener Text ist. Es stellte sich dann aber heraus, dass auch Märchen sehr viel implizites Wissen enthalten. Beispielsweise muss man wissen, dass es keine sinnvolle, nutzbringende Tätigkeit ist, zuerst Linsen unter Asche zu mischen, damit jemand anderes sie wieder herausklaubt. Das wertende Wort „Schikane“ muss im Märchen nicht ausdrücklich genannt werden und wird auch nicht genannt. Es wird auch nie ausdrücklich irgendeine Handlung als „gut“ oder „böse“ bezeichnet. Märchen sind also ebenfalls erklärungsbedürftig.

Trotzdem gelang Word2vec die Zuordnung von Personen und Dingen zur richtigen Mannschaft sehr gut. Jedoch manches Ergebnis war nicht leicht nachvollziehbar. Beim Aschenputtelmärchen zeigte der „Schuh“ die größte Nähe zum Adjektiv

„gut“ auf, obwohl nie von „guten Schuhen“ die Rede ist. Besonders nahe an „böse“ waren der Fuß und der Königssohn. Dass Stiefmutter und Stiefschwester ein Team auf der einen Seite und Aschenputtel mit Vater und Mutter auf der anderen Seite bildeten, erkannte die KI jedoch richtig. Aschenputtel sind Herd und Himmel näher als Kleider und Schuhe, die wiederum den Stiefschwester näher liegen. Das macht Sinn! Rapunzel bildet mit diesen Wörtern ein Cluster: schönsten, Schmerz, Gesang, Zöpfe und lieb. Auch das gibt die Essenz des Märchens wieder. Allerdings sind sich Zauberin und Königssohn näher als Rapunzel. Liegt es an deren inniger Feindschaft, die gründlicher thematisiert wird als die Liebe zwischen Königssohn und Rapunzel? Der Königssohn wird fälschlich zusammen mit der Zauberin dem Cluster des Bösen zugeordnet. Die Clusteranalyse könnte hier auf schriftstellerische Defekte des Märchens hinweisen. Rapunzel wird richtig als gut erkannt und durch Turm und Haar charakterisiert.

Diese Studie zeigt gut, was Sprachanalyse und Clustering kann. Die KI kann aber nur abbilden, was im Text ausdrücklich steht. Über implizites Wissen verfügt die KI nicht. Die Stellen, die wir einem Kind erklären müssen, sind auch für die KI nicht selbsterklärend. So gesehen waren auch die Märchen als Datengrundlage nicht ideal.

### **Global Product Classification (GPC)**

Dieses Projekt [3] hatte das Ziel, einen Klassifikationsalgorithmus zu finden, der Produkte anhand ihrer Freitextbeschreibung in eine der Kategorien der GPC-Klassifikation [4] einordnet. Dies war ein klarer Use Case für eine typische KI-Aufgabe. Da diese Klassifikation bisher von Hand durchgeführt wurde, lagen umfangreiche Trainingsdaten für überwachtes Lernen vor. Erreicht wurde schließlich eine sehr gute Klassifikationsqualität.

### **Stellenanzeigenstudie**

In den Jahren 2009, 2012, 2015 und 2018 wurden bisher von Hand Stellenanzeigen im Bereich Requirements Engineering ausgewertet (siehe [5] und die dort genannten früheren Studien). Positionen, Aufgaben und geforderte Kompetenzen wurden identifiziert und klassifiziert. Der Use Case ist klar: Empirische Forschung, die bisher manuell durchgeführt wurde, soll automatisiert werden. Dadurch wird Zeit gespart. Es können mehr Daten ausgewertet werden und die Studie öfter wiederholt. Die Anzeigen und Auswertungen früherer Studien dienen als Trainings- und Testdaten. Momentan ist die Klassifikationsqualität noch schlechter als bei der manuellen Klassifikation, und wir ringen noch mit dem Ausbalancieren von

Präzision versus Recall. Die Qualität verbessert sich mit einer größeren Menge an Testdaten. Außerdem experimentieren wir mit verschiedenen Algorithmen. Die KI tut sich u.a. schwer damit, dass gerade die Terminologie für die Aktivitäten im Requirements Engineering verschieden ist je nach Branche, Standard und Vorgehensmodell.

### **Schlussfolgerungen**

Gute Anforderungen ist für KI-Systeme wichtig, weil...

- die Lösung Nutzen bringen soll,
- technisch vieles machbar ist, aber man den richtigen Algorithmus und die richtigen Daten wählen muss,
- KI immer Fehler macht und man ethische Verantwortung trägt für Ungenauigkeiten und deren Folgen.

Anforderungen an KI-Systeme sollen umfassen:

- funktionale Anforderungen, v.a. Nutzer und Use Cases,
- Anforderungen an Qualität wie Präzision, Recall, Zeitbedarf, Speicherplatz
- Und – neu – Anforderungen an die Trainingsdaten.

Es ist hilfreich, wenn die KI eine Aktivität ausführt, die bisher manuell durchgeführt wurde. Dann steht ein klarer Use Case, oft sogar ein automatisierbarer Algorithmus zur Verfügung sowie zahlreiche Trainings- und Testdaten für überwachtes Lernen.

### **Referenzen**

[1] Alexander Römelt: Ethicbot – ein KI basierter Chatbot zur Modellierung von ethischen Entscheidungssituationen. Masterarbeit, Ruprecht-Karls-Universität Heidelberg, 2019.

[2] Julia Beck: Die Moral der Maschinen – Können neuronale Netze ethische Grundsätze erlernen? Masterarbeit, AKAD University, 2018.

[3] Amritpal Kaur: Maintaining Global Classification Code with Machine Learning. Masterarbeit, Ruprecht-Karls-Universität Heidelberg, 2021.

[4] <https://www.gs1.org>

[5] Andrea Herrmann, Maya Daneva, Chong Wang, Nelly Condori-Fernandez (2020) Requirements Engineering in Job Offers. Requirements Engineering Magazine, 16. September 2020, <https://re-magazine.ireb.org/articles/requirements-engineering-in-job-offers>