

Vereinfachte Minimaltranskription

Grundsätzlich besteht eine Transkriptionsdatei aus einem Transkriptionskopf und der Transkription selbst. Der Kopf enthält Metainformationen über die transkribierte Audiodatei, die Transkription selbst ist in ihrer Feinheit stark vom Forschungsziel abhängig. Für Inhaltsanalysen von Interviews oder Fokusgruppen ist eine Transkription nach dieser Konvention vollkommen ausreichend.

Die folgende Transkriptionskonvention stellt eine vereinfachte Variante der GAT2Minimaltranskription¹ dar. Grundsätzlich wird bei der Transkription alles klein und ohne Satzzeichen geschrieben (auch keine Fragezeichen). Die Rechtschreibung wird eingehalten, Äußerungen wie z.B. kannst werden aber **nicht** schriftsprachlich als kannst du aufgelöst.

Pausen:

- (.) **geschätzte** Mikropause bis ca. 0,2 Sek
- (-) **geschätzte** kurze Pause bis ca 0,5 Sek
- (--) **geschätzte** mittlere Pause bis ca 0,8 Sek
- (---) **geschätzte** Pause bis ca. 1,0 Sek
- (4.0) „gemessene“ Pause von 4,0 Sekunden **mit 0,5 Sekunden Toleranz**

Sonstige segmentale Konvention: und_äh Verschleifungen
innerhalb von Einheiten (Unterstrich) äh öh äm
Verzögerungssignale, sog. „gefüllte Pausen“

Sequenzielle Struktur:

[das war lustig] gleichzeitiges Sprechen von zwei Personen
[total]

Lachen, Weinen, Atmen

haha hihi silbisches Lachen
((steht auf)) Beschreibung einer Handlung
<<ironisch> das war lustig> wie wird etwas gesagt
°h bzw. h° **deutlich hörbares** Ein- bzw. Ausatmen

Unverständliches

(xxx) bzw. (xxx xxx) eine bzw. zwei völlig unverständliche Silben
((unverständlich, ca. 3 Sek)) längere unverständliche Passage
(solche) vermuteter Wortlaut
(solche/welche) zwei vermutliche Alternativen
((...)) Auslassung im Transkript

¹ Selting, M., Auer, P., Barth-Weingarten, D., Bergmann, J. R., Bergmann, P., Birkner, K., ... & Hefter, M. (2008). *Transkriptionskonventionen für die Linguistik*. Berlin: De Gruyter Mouton.

Beispieltranskript:

Datensatz: zoom_0.mp4

Kürzel: TI_02 (hier: VP) , VL

Datum und Ort der Aufnahme: 09/2021 Videokonferenz

Dauer der Aufnahme: 38:27 Minuten

Transkribiert am: 09/2021

Transkribiert von: [ANON]

VL: ((...)) Als erstes würde ich gern wissen wer du bist, was du machst und was ist dein Expertisegebiet ist?

VP: Okay. Also ich heiße [F] Ich bin Doktorand an der [ORG] im Machine Learning Fachbereich. Und ich beschäftige mich hauptsächlich mit Angriffen auf neuronale Netze. Also wie (---) Angreifer theoretisch Trainingsdaten aus den neuronalen Netzen extrahieren können oder anderweitig Informationen über die Trainingsdaten irgendwie aus diesen neuronalen Netzen rausziehen können quasi. (--) Genau.

VL: Genau. Ich erkläre dir noch, dass hab ich vielleicht vergessen, ich erkläre vielleicht noch ganz kurz wer ich bin. Ich bin auch Doktorand am [ORG] hier in Aachen an der [ORG] (---) Hab/ wir machen halt Akzeptanz. Weißt du sicher von [B] also Akzeptanzstudien zu irgendwelchen Technologien. Ich war vorher in der Mobilität, also hatte so Mobilitätsprojekte, autonomes Fahren, Akzeptanz von autonomen Fahren und dann haben wir geguckt was passiert, wenn wir eine KI in das Auto tun, die sich aber nicht ums Fahren kümmert, sondern um die Interaktion mit den Passagieren. (--) Und davon bin ich aber jetzt auch ein bisschen weg und mach auch mit [B] dieses [PRJ] Projekt, wo es wirklich um künstliche Intelligenz geht, aber nicht auf der technischen Seite, sondern auf der User Seite und Stakeholder und was brauchen die verschiedenen Stakeholder, was brauchen die verschiedenen User um künstliche Intelligenz besser zu verstehen, (-) besser zu nutzen und so. Genau. (--) Ja, die nächste Frage ist, beschreib mal die KI, mit der du hauptsächlich arbeitest.

VP: (---) Genau, also das wären ja dann im Prinzip neuronale Netze. (--) Also jetzt die Funktionsweise quasi soll ich beschreiben?

VL: Ja.

VP: Okay. (--) Okay die Funktionsweise ist im Prinzip (---)/ ist gar nicht mal so einfach ((lacht)). Genau, also ein neuronales Netz besteht, wie der Name schon sagt, aus Neuronen. Und im Prinzip ist die Idee, dass man diesem neuronalen Netz ganz viele Beispiele zeigt und dann anhand dieser Beispiele das neuronale Netz abstrahiert im Prinzip Eigenschaften dieser Beispiele lernt. (--) Und dann darauf basierend (---), ja Aufgaben erfüllen kann.

VL: Sind die Beispiele/ also sind das immer die gleiche Art von Beispielen oder hast du da verschiedene?

VP: (--) Ja, das ist das Problem. Also im Idealfall hat man so viele verschiedene wie möglich. In der Realität ist das eben nicht so wirklich machbar, weil man eben immer nur einen Ausschnitt aus der/ also einen Ausschnitt aus der möglichen Menge aller Daten hat.

VL: <<bestätigend> Mhm>. Okay. Was ich eher glaub ich auch meinte war, also sind es Bilder, sind es/was sind die Input-Daten, die du (---)/

VP: Ach so, ja. (-) Sowohl/ Es können Bilder sein, es kann Text sein. (---) Können auch andere Daten sein.

VL: Okay, also ein breiteres Spektrum?

VP: Genau.

VL: Okay super. (---) Eigentlich, ja/ ich wollte dich gerade fragen beschreiben Sie mit eigenen Worten was KI für Sie ist, aber ich glaub, dass

57 hast du schon ganz gut beschrieben in (--) dem Sinne wo du gesagt hast
58 beschreib die KI mit der du arbeitest. Oder gibt es da noch/ würdest du das
59 anders beschreiben? Wenn ich dich fragen würde was ist KI?

60 **VP:** (1.0) Ich würde dann noch hinzufügen/ man hört ja oft, dass die Leute
61 Angst haben vor der KI. Da würde ich noch hinzufügen, dass es eben so ist,
62 dass KI nur das machen kann, was ihr vorher gezeigt wird. (--) Und (--) ja
63 genau, dass das quasi die Beschränkung ist. Wenn ich ihr was nicht zeige,
64 dann kann sie es nicht.

65 **VL:** Okay, okay. Nutzt du KI außerhalb von deinem Beruf? Und wenn ja, wo?

66 **VP:** (--) Also meinst du jetzt in Produkten irgendwie? Oder (---)/

67 **VL:** Ja im privaten Bereich, in Produkten, in deinem Alltag, wenn du gerade
68 nicht in der Uni sitzt.

69 **VP:** Okay. (4.0) Also natürlich ist KI heutzutage fast in jedem Produkt, was
70 man dann halt so verwendet. Und einem ist manchmal gar nicht klar, dass da
71 überhaupt KI drinsteckt, aber so entwickeln in der Freizeit tue ich keine
72 KI, nee.

73 **VL:** Okay. Würdest du sagen, dass du/ dass man/ ist vielleicht ein bisschen
74 (-) provokant oder vielleicht auch so ein bisschen aus den Medien gezogen.
75 Benutzt man KI oder arbeitet man mit KI zusammen?

76 **VP:** (2.0) Ich würde sagen man benutzt KI. Weil das eine Art Werkzeug ist (-
77) quasi.

78 **VL:** Okay. Ja. Was denkst du, sind sinnvolle Einsatzgebiete für KI? Also
79 warum entwickeln wir das?

80 **VP:** Ich denke KI ist da sinnvoll zu verwenden, wo man Dinge automatisieren
81 kann, (---) die aber nicht so klare Regeln zum automatisieren haben. Also
82 ein Beispiel ist zum Beispiel autonomes Fahren. Wenn man jetzt sagt ich
83 stell ein paar Regeln auf, wie man Auto fährt, ist ja nicht so einfach oder
84 fast unmöglich. Deswegen ist da quasi ein gutes Anwendungsgebiet zum
85 Beispiel. Anderes Anwendungsgebiet ist (---) Medizin zum Beispiel. Dass man
86 da die Ärzte mit einer Automation quasi unterstützt. Bilderkennung oder
87 sowas, genau.

88 **VL:** Okay. Also im Groben wo es nicht eine klar definierte Regel gibt, die
89 zum Ergebnis führt, sondern (-) ja (---)/?

90 **VP:** Genau.

91 **VL:** Wie ist denn deine Erfahrung mit KI im Arbeitsalltag? Also was läuft
92 gut und wo gibt's Schwierigkeiten? Woran scheitert man manchmal? Ja/

93 **VP:** Also meistens läuft es gut, wenn man genug Daten hat. Und auch viele
94 unterschiedliche Daten, die gut die Realität abbilden, in
95 Führungsstrichen jetzt gesagt. Scheitern tut es eben meistens daran, dass
96 man keine Daten hat. Und das dann sehr aufwendig ist die Daten zu bekommen
97 oder man so direkt gar keine Möglichkeit hat die Daten zu bekommen.

98 **VL:** Okay. Hattest du mal besonders positive Erfahrung in deinem Umgang mit
99 KI? Also so ein Aha-Erlebnis, wo du sagst/ wo du gesagt hast ah krass das
100 geht oder so?

101 **VP:** (--) Das war/ Ja das war während meinem Studium. Als ich meine erste
102 Deep Learning Vorlesung gehört hab. Da sollten wir so ein Projekt machen.
103 Wir durften uns aussuchen, was wir machen wollen. Und ich hab mit meinen
104 Kommilitonen/ wir haben zusammen ein neuronales Netz trainiert was Super
105 Mario spielt. Und das hat mich erstaunt, dass man quasi mit, ja mehr oder
106 weniger Mathematik so Super Mario spielen kann ((lacht)).

107 **VL:** ((lacht)). Gibt's denn auch/ also das ist ja jetzt eine positive Sache.
108 Gibt's auch Sachen, wo du mal richtig frustriert bist, wenn du mit KI
109 arbeitest? Also ich stell jetzt diese ganzen Fragen, du kannst/ also das
110 ist in deinem Beruf jetzt. Das kann aber auch abweichen auf deinen Alltag.
111 Das ist so ein bisschen beide Perspektiven. Je nach dem was dir einfällt.

112 **VP:** <<bestätigend> Mhm.> ((nickt)). Okay. (4.0) Ja. Im Alltag bin ich
113 manchmal muss ich ganz ehrlich sagen zum Beispiel von Siri frustriert. Dass
114 die manchmal nicht versteht, was man möchte. Das ist jetzt ein Beispiel aus
115 dem Alltag. Und aus dem Berufsalltag quasi, wäre ein Beispiel, dass (---)

116 man was entwickeln möchte und es ist einfach nicht trainiert und man weiß
117 nicht genau warum. ((lacht)). Dann ist es kompliziert rauszufinden woran es
118 liegt.

119 **VL:** Okay. Ist es, weil du schwierig reinkommen kannst, in dem was da
120 passiert?

121 **VP:** ((nickt)). Das ist ein Punkt, ja.

122 **VL:** Okay. (-) Gut. (---) Also wenn du jetzt/ du hast deine Trainingsdaten/
123 wenn du jetzt den oder/ wenn du jetzt den Prozess erklären müsstest in ein
124 paar Worten, wie KI vom Input bis zum Ergebnis kommt. Also wie ist der
125 Prozess, den KI durchläuft und wie würdest du das einem Laien erklären?

126 **VP:** <<bestätigend> Mhm.> Also die grundlegenden Schritte sind man hat eben
127 seine Trainingsdaten. Man (--) guckt zuallererst wie kann man die
128 Trainingsdaten in ein Format bringen, was man quasi diesem neuronalen Netz
129 oder dem Modell irgendwie füttern kann. Und man überlegt sich zuallererst
130 was für ein Modell man verwendet, was für ein KI Modell. Was am besten
131 passen würde zu diesen Daten. Dann trainiert man das Modell, das KI Modell.
132 Zeigt ihm quasi die Beispiele und was erwartet ist mehr oder weniger. Und
133 das KI Modell lernt dann quasi eine Abstraktion auf diesen Daten zu lernen
134 und dann testet man auf/ testet das KI Modell auf Daten, die man vorher
135 quasi rausgenommen hat und die das Modell noch nicht gesehen hat.

136 **VL:** <<bestätigend> Mhm.> Okay. Warum glaubst du, fällt es, sag ich mal
137 Nutzern so schwierig dieses Konzept zu verstehen? Also warum ist KI für
138 Nutzer so eine Black Box?

139 **VP:** Ich könnte mir vorstellen, dass es für Laien schwierig ist den
140 Zusammenhang zwischen (---) der Mathematik, die quasi dahinter steckt zu
141 machen und dem was dann in der Realität quasi passiert. Also wenn man sich
142 das anguckt in einem Produkt oder sowas dann denkt man immer wow, das ist
143 Magie was da passiert. Aber dass da quasi Mathematik und Wissenschaft
144 hinten dran steckt, das ist dann den meisten glaub ich nicht so wirklich
145 klar. Dass das mehr oder weniger mit dieser Mathematik alles erklärt werden
146 kann.

147 **VL:** Also sie sind ja nicht in der Lage das zu erklären. Wenn sie es
148 wüssten, wären sie ja nicht in der Lage es zu erklären. Oder denkst du es
149 reicht aus, zu wissen, dass da Mathematik hinter steckt und keine Magie,
150 damit es besser verständlich ist?

151 **VP:** Ja könnte ich mir vorstellen. (.) Also ich könnte mir vorstellen zum
152 Beispiel einem Laien die ganz, ganz basic Sachen erklärt mit ganz, ganz
153 einfacher Mathe. Dass man dann da schon/ dafür schon eine Intuition hat
154 mehr oder weniger und sagen kann ah ok so ist das im einfachsten Fall. Da
155 passiert was ähnliches, wenn ich da ganz komplizierte Sachen quasi benutze.

156 **VL:** Wir hatten eben mehr oder weniger drüber gesprochen, dass Leute auch
157 Angst vor KI haben. Das wird ja manchmal in den Medien dargestellt. Nicht
158 nur Angst, sondern auch/ also es gibt viele verschiedene Facetten, wie KI
159 in den Medien dargestellt wird. Stört es dich manchmal, wie es dargestellt
160 wird? Und wenn ja, was stört dich?

161 **VP:** Ja, es stört mich manchmal. Und es stört mich vor allem, wenn die Leute
162 oder wenn es/ wenn das Gefühl in den Medien so vermittelt wird KI könnte
163 sich selbstständig machen. (---) Ich mein, ich weiß es nicht. Ich kann
164 nicht in die Zukunft gucken, vielleicht gibt's das irgendwann, aber ich
165 halt das für sehr, sehr, sehr unwahrscheinlich, weil es eben meiner Meinung
166 nach so ist, KI lernt, was man ihr zeigt. Und wenn man ihr was nicht zeigt,
167 dann kann sie es nicht.

168 **VL:** Okay. Welche Eigenschaften/ also wenn du jetzt mit KI arbeitest oder KI
169 nutzt, also sowohl privat als auch beruflich. Welche Eigenschaften muss für
170 dich die KI haben und welche Anforderungen muss erfüllt sein für eine gute
171 Zusammenarbeit oder eine gute Nutzung von KI? Also was sind die wichtigsten
172 Anforderungen und Eigenschaften?

173 **VP:** <<bestätigend> Mhm.> Also das KI Modell sollte in der Mehrheit der
174 Fälle richtig liegen, natürlich ((lacht)). (--) Und ich glaube, dass (3.0)

175 ja bis zu einem gewissen Grad auch eine bestimmte Erklärbarkeit gegeben
176 sein sollte. Also zum Beispiel wenn man jetzt mit Siri redet und die gibt
177 mir eine Antwort steht da oben drüber immer was Siri verstanden hat. Das
178 heißt, das gibt einem dann gewisse Einsicht ok sie hat das und das
179 verstanden, deswegen gibt sie mir die und die Antwort. Das heißt, wenn sie
180 mich falsch verstanden hat, kann ich verstehen, warum sie mir eine komische
181 Antwort gibt, zum Beispiel.

182 **VL:** Okay. Gibt es Ausschlusskriterien? Wo du sagen würdest, okay wenn die
183 KI das nicht hat, dann nutze ich die auch nicht?

184 **VP:** Mhm, ja. Das ist eine schwierige Frage. (7.0) Nee, eigentlich nicht so
185 direkt. Gibt eigentlich kein Ausschlusskriterium. Weil ich glaube, dass man
186 immer nochmal als Mensch checken sollte, ob das schlüssig ist oder Sinn
187 macht, was die KI quasi ausgegeben hat.

188 **VL:** Okay. Meinst du denn, dass hängt vom Kontext ab was KI/ was für
189 Eigenschaften und was für Anforderungen KI haben sollte? (--) Also ist es
190 in manchen/ in anderen Kontexten wichtiger

191 **VP:** Auf jeden Fall, ja. (--) Denke ich. Also zum Beispiel beim autonomen
192 Fahren ist es besonders wichtig, dass die KI nicht fehleranfällig ist, weil
193 eben davon Menschenleben abhängen. Und wenn ich jetzt eine KI frage ist das
194 ein Hund oder eine Katze, wenn sie mir da die falsche Vorhersage gibt, ist
195 das jetzt nicht so schlimm.

196 **VL:** Okay, das heißt, sag ich mal die/ ja wie richtig jetzt die Einschätzung
197 ist, ist verschieden, wenn man jetzt im autonomen Fahren unterwegs ist (-)

198 **VP:** Genau, ja.

199 **VL:** Okay. Kommen wir zum Thema Transparenz. Jetzt mal ganz offen und
200 generell gefragt was ist Transparenz bei KI-Nutzung und ist dein System
201 transparent?

202 **VP:** Okay, Transparenz ist für mich, dass man (---) nachvollziehen kann,
203 warum die KI zu einer bestimmten Entscheidung gekommen ist. (1.0) Und das
204 hängt natürlich vom Modell ab, inwiefern das transparent ist. Die Modelle,
205 mit denen ich arbeite, die sind meistens nicht wirklich transparent. (---)
206 Also das ist zum Beispiel Bildverarbeitung und am Ende kommt eine
207 Klassifikation raus und man weiß nicht genau, warum das Modell diese Klasse
208 vorhergesagt hat.

209 **VL:** Okay. Welche Information würdest du denn brauchen, damit du wissen
210 könntest warum die Klassifikation vorgenommen worden ist?

211 **VP:** Genau, also eine Methode ist zum Beispiel, dass man sich anguckt/ wenn
212 man zum Beispiel ein Bild rein gibt, worauf das neuronale Netz besonders
213 geachtet hat. Also wenn man zum Beispiel ein Bild von einem Hund nimmt und
214 die Klasse Hund rauskommt und diese (---)/ diese Darstellung worauf dieses
215 neuronale Netz geachtet hat bezieht sich nur auf den Hund, dann würde ich
216 sagen ja ok die Klassifikation ist richtig und hat auch aus dem richtigen
217 Grund quasi so stattgefunden. Wenn jetzt stattdessen nur der Himmel
218 angeguckt wurde und die Klasse Hund rauskommt, würde ich sagen okay da ist
219 irgendwas schief gelaufen.

220 **VL:** Ja, okay. Wenn du jetzt dran denkst okay du möchtest natürlich wissen
221 wie kann ich das neuronale Netzwerk verbessern oder wie kann ich meine
222 Vorgehensweise für das Modell verbessern. Wenn du dich jetzt mal in die
223 Schuhe vom Nutzer reinstallst. Was denkst, würde ein Nutzer brauchen (-),
224 wenn er KI nutzt an, an Transparenz? Ist ihm dann wichtig, dass ein
225 bestimmter Teil vom Bild erkannt wird? Oder wenn man mal an andere Kontexte
226 denkt.

227 **VP:** (3.0) Ich denke da kommt es wieder auf den Kontext an. Also zum
228 Beispiel im Kontext Medizin zum Beispiel, denke ich wäre es schon wichtig,
229 weil der Arzt zum Beispiel wissen sollte wo (---)/ wo die Stelle zum
230 Beispiel auf dem Röntgenbild ist, wo die KI einen Tumor oder sowas erkannt
231 hat. (2.0) Bei anderen Kontexten, keine Ahnung Beispiel Siri wieder, ist
232 das glaub ich nicht so wichtig warum Siri jetzt die und die Antwort gegeben
233 hat.

234 **VL:** Braucht es da dann überhaupt Transparenz?
235 **VP:** (2.0) Wie gesagt, kommt auf den Kontext drauf an.
236 **VL:** Wie hilft dir denn Transparenz den/ also den Prozess und die
237 Entscheidung von KI zu verstehen. Und nicht nur verstehen, sondern auch
238 erklären und zu interpretieren?
239 **VP:** (1.0) Also jetzt während dem entwickeln oder (--)/?
240 **VL:** Ja, auch. Aber (--) ja/ wir können auch das Beispiel außerhalb sag ich
241 mal des akademischens nehmen. Aber während des Entwickelns interessiert
242 mich halt auch.
243 **VP:** <<bestätigend> Mhm.> (5.0) Ich denke da kommt es sehr stark auf die
244 Daten an. (-) Also zum Beispiel bei Bilddaten könnte man sich angucken
245 worauf hat die KI geachtet. Bei Textdaten könnte man sich angucken was für
246 Wörter haben jetzt besonders großen Einfluss gehabt auf die Entscheidung.
247 Und (---) ich glaub das ist insofern hilft dann beim Entwickeln zum
248 Beispiel, dass man sagen kann bei einzelnen Fällen, bei denen es kolossal
249 falsch lief (1.0)/ das man da ein Gefühl für bekommt, warum es so falsch
250 lief. Also wenn jetzt das total falsche Wort angeguckt wurde, aber trotzdem
251 die richtige Vorhersage rauskam, muss man vielleicht noch irgendwas ändern.
252 **VL:** <<bestätigend> Mhm.> Okay, okay. So ein bisschen das letzte was wir
253 auch untersuchen, ist Vertrauen in KI. Unter welchen Bedingungen vertraust
254 du KI?
255 **VP:** (5.0) Wenn es/ Also es kommt auch wieder auf den Kontext an. Also wenn
256 es nicht (---) wirklich wichtig ist (---) vertraue ich da eigentlich mehr
257 oder weniger blind drauf. (---) Andererseits wenn ich jetzt zum Beispiel an
258 einen Kontext wie autonomes Fahren denke, wo Leben von abhängen, vertrau
259 ich bis zu einem gewissen Grad, aber ich möchte immer die Möglichkeit haben
260 einzugreifen. Also zum Beispiel beim autonomen Fahren.
261 **VL:** Okay also du braucht eine gewisse Art von Kontrolle, wenn es/
262 **VP:** Genau.
263 **VL:** Okay. Würde denn/ würde es dir helfen, wenn du mehr Einsicht hättest
264 und mehr Transparenz hättest, würde es dir helfen mehr zu vertrauen auch in
265 kritischen Situationen?
266 **VP:** (3.0) Wahrscheinlich nicht, weil (4.0)/ weil ich das Gefühl hätte, dass
267 auch/ selbst wenn ich die Transparenz habe, dass selbst dann Sachen
268 theoretisch schief gehen können.
269 **VL:** Okay, verstehe. Das heißt du kannst dir dann auch, wenn du weißt wie es
270 innendrin so ein bisschen aussieht, kannst du dir nicht sicher sein, dass
271 es wirklich zum gewünschten Output kommt?
272 **VP:** Genau.
273 **VL:** <<bestätigend> Mhm.> Okay. Also wir versuchen immer Nutzer so nah wie
274 möglich an KI heranzubringen, um zu testen wie sie mit KI umgehen, wie sie
275 sich zu KI verhalten, welche Einstellung sie zu KI haben. Klar wir wissen
276 so ein bisschen was die Einsatzgebiete von KI sind, aber wenn man dann sich
277 Nutzertests überlegen muss, ist es eher schwierig natürlich. Und ich wollte
278 dich fragen, hast du eine Idee wie du aus deinem Gebiet, wo du arbeitest,
279 was du gerne mal sag ich mal von Nutzerseite testen möchtest? Also wie geht
280 ein Nutzer damit um, nicht ein Entwickler, der die das aktiv nutzen würde
281 was du da entwickelst. Oder sagen wir mal du überträgst was du gerade
282 entwickelst in Anwendungsbeispiele und einen Use Case. Hast du eine Idee
283 wie man daraus einen Nutzertest basteln könnte?
284 **VP:** (6.0) Das ist jetzt ein bisschen schwierig, weil ich untersuche ja
285 hauptsächlich Attacken mehr oder weniger. (12.0) Nee das würde mir jetzt
286 auf Antrieb nicht so richtig was einfallen.
287 **VL:** Ist natürlich auch schwierig, wenn du sagst/ also Attacken sind ja/ du
288 guckst wie/ was machst du? Du guckst, wie die/ wie man das neuronale
289 Netzwerk austricksen kann, damit man an die Trainingsdaten rankommt?
290 **VP:** Ja genau. Oder Informationen über die Trainingsdaten bekommt, ja.
291 **VL:** Wie läuft das so grob? Also/

292 **VP:** Ja also die Hauptidee ist im Prinzip, dass die neuronalen Netze auf den
293 Trainingsdaten sicherer sind, weil sie die Daten schon mal gesehen haben.
294 (-) Und das macht man sich dann quasi zu nutze, um/ wenn man jetzt als
295 Angreifer das Modell hat und irgendeine Menge Daten bekommt, gibt man die
296 quasi dann in das Modell und guckt ok wie verhält sich das Netz und kann
297 ich daraus quasi Unterschiede ableiten zwischen den einzelnen Inputs.
298 **VL:** Ah spannend. Das ist natürlich ein bisschen schwierig wenn man das aus
299 Nutzerseite sieht, weil das sind ja dann/ ja da ist der Use Case nicht so
300 da.
301 **VP:** Ja genau ((lacht)).
302 **VL:** Ja. (-) Wie siehst du eigentlich KI in Zukunft? Also eben hattest du
303 mal gesagt ja weiß ich nicht, so 50 Jahre kann ich nicht sagen. Aber was
304 ist/ was würdest du schätzen?
305 **VP:** Ich würde schätzen, dass KI eigentlich in jedem Bereich in unserem
306 täglichen Leben irgendwann Einzug erfährt. Also autonomes Fahren beginnt
307 jetzt langsam. (--) Aber zum Beispiel im täglichen Leben, Stichwort Smart
308 Home, wird es glaub ich noch viel, viel, viel, viel mehr. Und ich glaube
309 dass es irgendwann so sein wird, dass man quasi von der KI in allen
310 Bereichen unterstützt wird.
311 **VL:** Okay. Aber das/ du glaubst nicht, dass KI irgendwann (---) selbst
312 Sachen lernen kann, die man ihr nicht direkt beigebracht hat? Sagen wir mal
313 du/ man trainiert KI um ein bestimmtes Problem sag ich mal zu optimieren.
314 KI findet jetzt aber eine Möglichkeit das so zu optimieren, wie du es ihr
315 nicht gezeigt hast. Und das glaubst du passiert nicht?
316 **VP:** (--) Doch das kann passieren. Es kommt drauf an. (---) Ein Beispiel ist
317 in diesem Projekt wo wir Super Mario gespielt haben. Da hat die KI
318 irgendwann gelernt einen Bug auszunutzen und das war natürlich so nicht
319 vorgesehen. Aber auch da kommt es natürlich auf den Use Case an. Also wenn
320 der Use Case sehr speziell ist, gibt es glaub ich nur wenige Möglichkeiten
321 wie man quasi/ oder wie KI quasi die Sachen falsch machen kann.
322 **VL:** <<bestätigend> Mhm.> Ja, spannende Sache. Also ich komme da immer, wenn
323 ich diese Interviews mache/ ich komme da immer drauf so, auf/ ist halt
324 nicht das Thema, aber ist trotzdem spannend so ein bisschen weiter in die
325 Zukunft zu gucken. (---) Ja, wir sind eigentlich am Ende angekommen. Was (-
326 -) für uns noch wichtig ist, ist was denkst du welche Aspekte unseres
327 Gespräches waren für dich am wichtigsten oder am interessantesten?
328 **VP:** (4.0) Ich glaub am interessantesten war für mich, dass es/ das die
329 Nutzung von KI sehr stark davon abhängt/ oder ne anders formuliert. Das
330 Vertrauen in die KI sehr stark davon abhängt wo es verwendet wird. (---)
331 Und das war/ darüber hab ich so vorher noch nicht wirklich nachgedacht.
332 Aber ist natürlich klar, wenn Menschenleben davon abhängen ist/ es
333 natürlich wichtiger ist.
334 **VL:** Was natürlich spannend für uns werden könnte, vor allem im Projekt, ist
335 halt, dass die/ das Vertrauen wissen wir schon ungefähr variiert von
336 Kontext zu Kontext, das ist eigentlich ziemlich klar. Dann ist die Frage,
337 was ja interessant ist, ob Transparenz und Vertrauen das besser macht oder
338 es Kontexte gibt, wo es keinen Unterschied macht, ob Transparenz da ist
339 oder nicht. Und das Vertrauen ist unabhängig davon und das ist glaube ich
340 eine Sache die (-) ja/ die spannend sein könnte.
341 **VP:** Ja.
342 **VL:** Ja, wir sind jetzt eigentlich durch. Hast du noch irgendwie Fragen an
343 mich, uns, unser Projekt? (--) Oder an das Interview oder was auch immer?
344 **VP:** (3.0) Mich würde noch interessieren, was bei dem Projekt/ was ist das
345 Ergebnis, was du denkst, was rauskommt? Oder was (---)/ ja doch was ist das
346 Ergebnis, was denkst du denn was am Ende rauskommt? (---) Ist natürliche
347 eine schwierige Frage.
348 **VL:** Also was ist denke oder was hoffe oder was ich/ oder beides?
349 **VP:** Sowohl als auch.

350 **VL:** Also was ich denke was rauskommen wird, sind viele verschiedene Sichten
351 von Transparenz und viele verschiedene Auffassungen was Transparenz ist,
352 abhängig vom Kontext, weil das ganz verschieden sein kann. Und was glaub
353 ich in einem ersten Schritt rauskommen wird, ist/ wird glaub ich eine große
354 Map sein, welche Transparenz/ also was ich mir dann wünsche was dabei
355 rauskommt ist, dass man sozusagen einen Design Space hat. Und wir haben
356 Medizin, wir haben Cyber Security, wir haben Law, wir haben Mobility. Und
357 da dran noch/ so ein mehrdimensionaler Raum, wo wir halt dann noch
358 verschiedene Kritikalitäten haben und sagen das ist jetzt eine kritische
359 Situation. Und dann darein schieben können was/ ist da Transparenz wichtig,
360 ist Vertrauen abhängig von der Transparenz, Erklärbarkeit wichtig,
361 Verstehen wichtig? Brauch ich das überhaupt, um zu verstehen? Genau, das
362 ist glaub ich was in erster Linie so rauskommen wird, sind so was sind die
363 verschiedenen Arten von Transparenzen, die es pro Kontext braucht? Das ist
364 glaub ich so der erste Schritt und dann am Ende muss es halt gemessen
365 werden, ne? Also in den verschiedenen Kontexten. Da müssen wir halt
366 quantitativ mit sehr vielen Leuten gucken, ist das denn wirklich so oder
367 ist das jetzt die Meinung von einer einzelnen Person? Und dann kann man
368 gucken, ja/ Was ich mir wünsche, was bei rauskommt ist, dass Leute oder sag
369 ich mal Researcher später da rein gucken können und sagen können ah guck ma
370 da in diesem Design Space da bewege ich mich gerade drin, da haben Leute
371 schon mal was zu gemacht, da gibt es die und die Arten von Transparenz, die
372 und die Vertrauensrelationen. Lass mal gucken, ob das stimmt oder lass mal
373 noch was weiter drauf setzen (--) und dieses Framework sozusagen erweitern.
374 **VP:** <<bestätigend> Mhm.>
375 **VL:** Das ist, was wir eigentlich als Ziel definiert haben. Ich würde aber
376 sagen es ist ein relativ kleines Projekt. Mal sehen wie weit wir kommen.
377 Also wir sind/ also das ist ein Jahr, also das ist, ne? (--) Das ist
378 relativ kurz, aber ja (--) mal sehen.
379 **VP:** Hört sich aber interessant an, weil das auch für/ ich sag jetzt mal für
380 die Anwendung eine Auswirkung hat, weil man auch eben darauf gucken kann
381 und sagen kann ok, wenn ich jetzt zum Beispiel autonomes Fahren mache,
382 bräuchte ich vielleicht die und die Visualisierungen, damit sich der Nutzer
383 sicher fühlt oder sowas.
384 **VL:** Ja, ja. (-) Ja und genau da/ genau da kommt es bei uns so ein bisschen
385 drauf an, weil wir machen ja/ was wir ja sehr oft machen ist, wir
386 untersuchen ja Technologien, die sich gerade entwickeln oder die noch nicht
387 auf dem Markt sind. Also wir machen sehr selten/ nehmen wir eine
388 Technologie, lassen den Nutzer die nutzen und sagen dann okay was läuft
389 gut, was läuft schlecht, weil es oft dann zu spät einfach ist. Weil man
390 dann halt/ dann ist man nicht mehr drin so wirklich. Du kannst nicht mehr
391 so viel verändern und wenn du was verändern willst, kennst du das ja. In
392 Produktionsketten ist es sehr kostspielig. Deswegen gucken wir schon
393 vorher. Und das ist halt das, was wir erreichen wollen, auch für
394 Entwickler, dass die da drauf gucken können und sagen so ja ok das müssen
395 wir vielleicht noch machen, damit es Erfolg haben kann und damit es auch
396 genutzt wird. Also das braucht der Nutzer, wo der Entwickler sagen würde
397 das ist doch sowas on schein egal, aber der Nutzer sagt ja aber wenn ich
398 das nicht hab, dann nutz ich das nicht. (---) Und das ist halt der Punkt.
399 Und deswegen halt machen wir die Interviews und dann wird es auch
400 irgendwann mit sag ich mal Laien dann irgendwann nochmal untersucht.
401 **VP:** Okay. (-) Ja cool, hört sich echt spannend an ja.
402 **VL:** Oder ich kann dir gerne mal, sobald wir/ ich schreib das hier immer
403 auf, wenn du dran interessiert bist, wenn irgendwann mal sag ich mal
404 Arbeitspaket 1 abgearbeitet ist, kann ich dir gerne mal so ein bisschen was
405 davon zukommen lassen und was dabei rausgekommen ist.
406 **VP:** Ja gerne, ja.
407 ((...))