

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., ... & Heflinger, M. (2008). *Transkriptionskonventionen für die Linguistik*. Berlin: De Gruyter Mouton.

Beispieltranskript:

Datensatz: zoom_0.mp4

Kürzel: TI_01 (hier: VP) , VL

Datum und Ort der Aufnahme: 09/2021 Videokonferenz

Dauer der Aufnahme: 58:30 Minuten

Transkribiert am: 09/2021

Transkribiert von: [ANON]

VL: ((...)) Erstmal würde ich gern wissen wer du bist, was du machst und in welchem/ was ist deine Area of Expertise? Also auf welchem Gebiet bist du sozusagen ein Experte?

VP: Ja okay. Ich kann versuchen, dass zu beantworten. Also mein Name ist [A] Ich bin wissenschaftlicher Mitarbeiter am [ORG] und das (-)/ und wir bearbeiten dort/ oder mein Expertise-Gebiet ist die Anwendung von maschinellern Lernen. Also einer Unterform der künstlichen Intelligenz, um Produktionsprozesse bzw. mechatronische Systeme zu optimieren. Das ist so ein bisschen die übergeordnete/ das übergeordnete Ziel. Also wie wir die Systeme besser machen und das halt mit künstlicher Intelligenz. Und ja (--) dabei hab ich mich mit unterschiedlichen Techniken auseinander gesetzt oder Methoden (--), aber bei uns geht es hauptsächlich darum quasi diese Methoden anzuwenden. Das heißt, wir optimieren die Algorithmen nicht selber, sondern wir wenden sie meistens nur an, (--) vielleicht ist das relevant, vielleicht auch nicht, aber ja (---). Ich persönlich versuche mich in das Thema der generativen Modelle einzuarbeiten. Die beschäftigen sich hauptsächlich damit künstliche Daten zu erzeugen aus Daten, die vorher schon mal gesehen worden sind. Also (---) man sich das ein bisschen so vorstellen, dass/ wir haben ganz viele Bilder von Menschen gesehen und jetzt kann der Algorithmus eigenständig neue Bilder von Menschen generieren.

VL: Okay. (--) Ja. (1.0) Gibt es Funktionen, die du in deinem beruflichen (---)/ in deinem Beruf hast und die du in private Kontexte übernehmen könntest oder anwenden könntest? (--) Also wenn du zum Beispiel sagen wir mal einen Algorithmus hast, könntest du dir auch vorstellen, dass dieser Algorithmus/ du machst jetzt zum Beispiel Generierung von Bildern, könntest du auch vorstellen, dass du diesen Kontext nimmst und einfach einen privaten Kontext nimmst und der da nützlich sein könnte?

VP: (---) Hm, also jetzt bezogen auf die Generierung von Bildern würde mir jetzt spontan einfallen (---) das Thema von Generierung von Kunst. Also das man quasi Bilder erzeugen könnte, die ein künstlerischen Effekt aufweisen oder quasi ein gesamt neues Bilder entstehen zu lassen aus den Erfahrungen, die es davor gemacht hat und anhand der Bilder, die es davor gesehen hat. Das fände ich zum Beispiel cool in der/ in meiner Freizeit das quasi anzuwenden. Was anderes, was mir noch spontan einfallen würde ist/ oder worüber ich mich auch (---)/ viele Gedanken/ nicht viele Gedanken, aber manchmal so drauf komme, ist halt das Thema Spiele. Also irgendwelche Computerspiele oder so und dafür verfolge ich auch immer sehr gerne die Anwendung von Green Enforcement Learning, wie es dort in der/ in den Spielekontext eingesetzt werden kann und wie sie halt Lösungswege erzeugen, die Menschen halt selbst noch nicht herausgefunden haben und das finde ich immer sehr spannend wie sie da auch bei sehr komplexen Spielen quasi zu einer interessanten Lösungsstrategie kommen. (---) Also das sind halt so zwei Bereiche, wo ich künstliche Intelligenz jetzt am ehesten in meiner Freizeit sehen würde.

VL: <<bestätigend> Mhm>. Wenn du jetzt/ wir hatten eben (2.0) im Productioncluster ein Meeting/ hatten wir so eine Sache, da mussten wir kurz definieren, was das/ was ein bestimmter Begriff ist. Könntest du, wenn

58 du jetzt kurz beschreiben müsstest in kurzen zwei Sätzen, was genau KI
59 macht.

60 **VP:** (---) Ok, ah okay. Für mich ist künstliche Intelligenz/ ist für mich
61 das Lernen von Repräsentation. Also (---) was für mich ein KI Algorithmus
62 immer macht, du hast ganz viele Daten und diese Daten sind auf den ersten
63 Blick erstmal unverständlich und was dann der Algorithmus macht ist diese
64 Daten in eine Repräsentation oder in eine andere Form umzu/ zu übersetzen,
65 sodass sie mit dem/ sodass man mit denen bestimmte Regeln ableiten kann. (-
66 -) Das (---) genau, und finde ich sieht man in allen Algorithmen der
67 heutigen Zeit wieder. Also bei der Bildverarbeitung werden aus Pixelwerten
68 abstrakte Features generiert. Also quasi Kanten oder noch höher
69 Gesichtsformen beispielsweise und im Bereich der Visualisierung oder in der
70 Textgenerierung werden zum Beispiel dann semantische Begriffe erzeugt oder
71 in einen semantischen Raum gepackt, sodass sie halt besser voneinander
72 unterscheidbar sind. Also insgesamt ist für mich künstliche Intelligenz das
73 Lernen von sehr niedrigen Informationen zu höherwertigen Informationen.

74 **VL:** Okay, ja verstehe. Kommen wir jetzt zu meiner letzten Frage be/ also
75 zur letzten Frage vom ersten Teil. Benutzt man KI oder arbeitet man mit KI
76 zusammen? Was würdest du sagen?

77 **VP:** Ich würde sagen man benutzt KI. Für mich ist das ein Werkzeug.

78 **VL:** Okay. Also auch in deinem Fall ist es einfach ein Werkzeug. Du bist/ Du
79 hast nicht/ du arbeitest im Team mit 'ner KI zusammen?

80 **VP:** Nee. Für mich ist es ein Werkzeug. Ich versuche dort Probleme zu lösen.
81 Ist wie der Hammer auf einem Nagel oder so.

82 **VL:** Okay. Jetzt in deinem Arbeitsalltag. Du arbeitest ja tagtäglich mit KI
83 zusammen oder benutzt sie in dem Fall. Würdest du/ was würdest du sagen
84 läuft gut und was ist schwierig? Also wo gibt es Probleme, wo gibt es
85 Engstellen? Was bereitet manchmal Schwierigkeiten (--) dir oder anderen
86 Nutzern?

87 **VP:** Was gut läuft denke ich, weil wir beschäftigen uns ja mit der KI, wir
88 arbeiten nicht in dem Sinne mit denen zusammen, dass wir mit denen
89 interagieren, sondern wir machen ja mehr so die Entwicklung dahinter oder
90 wir versuchen da immer ein System hinter zu entwickeln. Und deswegen (--)
91 kann ich aus meiner Perspektive sagen was da gut läuft bei der Entwicklung.
92 Also gut läuft, ist die Community bzw. der Informations/ die
93 Informationsdarstellung. Das heißt, es gibt ein riesiges Angebot an
94 Informationen, die zur Verfügung gestellt werden, die einem dann das
95 Erstellen von künstlichen Intelligenz Systemen vereinfachen. Bspw. gibt es
96 sehr viele vorgeschriebene Algorithmen, die man einfach nutzen kann. Und da
97 ist die Einstiegshürde ziemlich gering, was überraschend ist. Was ich dann
98 nicht so gut finde oder was ich schwer finde ist halt, wenn etwas nicht
99 funktioniert, weiß man nie so genau woran es liegt. Also man kann natürlich
100 lesen, aus Erfahrung kann man sagen hey der hat das so und so gelöst, aber
101 das könnte für dein Anwendungsfeld dann wieder ganz anders sein. Also da
102 quasi eine (---)/ da kommt es halt so ein bisschen auf die Erfahrung des
103 Entwicklers an und da gibt es keine All-fits/One-fits-all-solution.

104 **VL:** Mhm. Gab es mal eine Situation in deiner Arbeit, wo du so einen Aha-
105 Moment hattest? Wo dann eine besonders positive Erfahrung, die du so/ die
106 so herausstechend ist?

107 **VP:** Es ist natürlich immer ein gutes Gefühl, wenn am Ende dann der
108 Algorithmus, nachdem er die Nacht durchtrainiert hat, dann auch eine gute
109 Performance zeigt. Das ist immer ganz nice. Aber so einen richtigen Aha-
110 Moment oder so das Gefühl ok jetzt hab ich was super neues erzeugt oder so
111 jetzt noch nicht.

112 **VL:** Okay. Hast du/ Also wenn du an KI denkst und deine Arbeit mit KI.
113 Kannst du kurz sagen, was sind Vorteile von KI wo du sagst das ist cool,
114 dass wir KI haben und was sind noch Nachteile gerade?

115 **VP:** Vorteile, denke ich, ist hauptsächlich oder meiner Meinung nach die
116 Möglichkeit die ganzen Daten, die jetzt produziert werden in den

117 Unternehmen auch wirklich effizient zu nutzen. Also, dass man da nicht
118 angewiesen ist auf irgendwelche Auswerteprozesse sag ich jetzt mal, die
119 sehr stupide sind. Also ich nehme jetzt mal ein Beispiel von der
120 Qualitätskontrolle, dass dann da wirklich einer sitzen muss und dann sich
121 das alles anschauen muss und dann einzeln jedes Bauteil dann entsprechend
122 aussortieren muss, sondern wirklich das das dann alles automatisiert
123 ablaufen kann. Nachteile von KI ist natürlich das, dass bei der
124 Datengenerierung (---) so ein bisschen (--) ja vielleicht/ also das man
125 manchmal das Gefühl haben könnte, dass einem die Kontrolle quasi entzogen
126 wird. Sobald man natürlich alles automatisiert macht, könnte man befürchten
127 oder könnte ich mir auch vorstellen, dass die Systeme dann plötzlich was
128 machen ohne dass du/ also sich/ ihr Verhalten ändern ohne dass du es
129 gemerkt hast. Das heißt, (---) quasi ein/ man kann es so ein bisschen sagen
130 wie ein Schlupfloch im System, dass man davor nie entdeckt hatte und
131 plötzlich verhält sich dein KI-System total anders, als wie du es vorher
132 gesehen hattest. Das heißt/ Natürlich kann dir das auch in den aktuellen
133 Systemen passieren, aber da ist die Einsicht natürlich viel größer, weil du
134 ja quasi alles per Hand selbst erzeugt hast und nicht auf die Daten
135 angewiesen bist (--) oder auf die Datenmengen.

136 **VL:** Okay, ja. Wenn du jetzt jemandem den Prozess beschreiben müsstest, wie
137 KI vom Input bis zum Ergebnis kommt. Also wie würdest du das einem Laien
138 erklären? Könntest du das mal ganz kurz umreißen?

139 **VP:** Ja (-), aber ich versuchs an einem Beispiel würde ich sagen. Also
140 nehmen wir mal an wir machen eine Klassifikation von Bildern wieder. Hund
141 und Katze. Der Prozess ist eigentlich meistens identisch. Und zwar wie
142 gesagt, es werden Repräsentationen gelernt. Das heißt, man kann sich das so
143 vorstellen in einem Pixel oder ein Datenpunkt ist nichts weiter als ein
144 riesige Zahlenmatrix oder eine Anhäufung von Zahlen. Und diese Zahlen haben
145 erstmal keinen großartigen Sinn für die Maschine oder für den Menschen
146 auch, wenn du sie nur als Zahlen betrachtest. Und was dann halt ein Machine
147 Learning Algorithmus macht ist halt, er lernt quasi von diesem sehr
148 zahlenlastigen System werden also Repräsentationen gelernt, die diesen Raum
149 von Zahlen trennen lassen. Das heißt, was dann da passiert ist, nehmen wir
150 mal an wir haben Hunde und Katzen, die sehen auf einem Pixellevel fast
151 identisch aus, weil du siehst halt nur rot, grün, blau Werte und wenn der
152 Algorithmus gut definiert ist/ was dann passiert ist, dass dann halt die
153 rot, grün, blau Werte übersetzt werden in Repräsentationen, die quasi einen
154 tieferen Sinn oder einen menschenverständlichen Sinn haben. Wie bspw., dass
155 die Katze halt Katzenohren hat und die Hunde halt Hundeohren. Und anhand
156 dieser Unterscheidung in diesem neuen Raum lassen sich dann quasi neue
157 Klassen bzw. lassen sich diese Bilder dann unterscheiden.

158 **VL:** Und warum glaubst du das/ also warum glaubst du das es manchmal Leuten
159 oder auch Laien schwierig fällt sowas zu verstehen?

160 **VP:** Weil die Krux an der ganzen Sache ist, dass man nicht verstehen kann
161 wie dieser Raum aussieht. Also weil am Ende des Tages bleiben es ja Zahlen,
162 die der Algorithmus quasi verarbeitet und diese Zahlen sind in einem Raum,
163 die mit Eigenschaften, mit menschlich bildlichen Eigenschaften
164 korrespondieren und diese Verbindung ist nicht immer sehr trivial. Wie
165 willst du denn einen Vektor/ du kannst ja nicht einfach einen Vektor
166 vergleichen mit einer Pfote beispielsweise. Da/ wieso sollte jetzt dieser
167 Vektor diese Pfote darstellen? Aber genau das passiert halt. Dass dieser
168 Vektor, sagen wir mal in einem vierdimensionalen Raum 1-0-0-0 genau ein
169 Wert für eine Pfote ist. Und genau diese Fähigkeit/ diese Darstellung von
170 komplexen Konzepten, nenn ich es jetzt mal, eines Bildes, wird halt nicht/
171 ist halt nicht trivial oder nicht offensichtlich.

172 **VL:** Du bist jetzt sehr tief da drin. Aber es gibt auch nochmal so eine
173 andere Seite wie KI in den Medien dargestellt wird. Man hört manchmal
174 Schlagzeilen so ich sag jetzt mal die ganz krasseste wäre: KI übernimmt die
175 Weltherrschaft. Stören dich solche Schlagzeilen, wenn du sowas hörst?

176 **VP:** Stören würde ich es jetzt nicht bezeichnen, weil/. Mich würde es nur
177 dann stören, wenn meine Freunde mich damit belästigen. Also wenn wirklich
178 Freunde oder Leute in meinem Umfeld diese Schlagzeilen sehr ernst nehmen
179 und wirklich dann mit mir eine Diskussion führen wollen würden, ob/ und
180 darauf beharren, dass diese Aussage korrekt ist. Aber was ich halt auch
181 denke, ich kann auch nachvollziehen, warum diese Angst da ist, weil was man
182 nicht versteht, davor hat man ja auch meistens ein bisschen Angst.

183 **VL:** <<bestätigend> Mhm.> Also du siehst es nicht so (---)

184 **VP:** [überlegt] Nee/ j/ also sagen wir mal ich sehe zumindest es jetzt noch
185 nicht so.

186 **VL:** Okay, aber in/ also das ist eine riesen Frage natürlich, aber du
187 würdest sagen, in näherer Zukunft könnte es vielleicht zu Problemen damit
188 kommen? (--) Oder in weiterer Zukunft?

189 **VP:** Ja, also die Sache ist halt, Weltherrschaft oder die Welt zu übernehmen
190 finde ich ist Definitionssache. Die Frage, die ich mir da stellen würde,
191 ist, warum sollte ein Artificial Intelligence Algorithmus versuchen die
192 Weltherrschaft zu übernehmen? Was hat das da für einen Vorteil? Wenn
193 natürlich dein Algorithmus darauf getrimmt ist, sag ich jetzt mal irgendwas
194 zu optimieren und die Optimierung wäre bspw. die/ das Leid der Menschen (-)
195 als Beispiel, ja, wie auch immer man das messen sollte (-), dann würde der
196 Algorithmus quasi, wenn er natürlich den richtigen Aktionsraum hat, also
197 quasi das/ all die Stra/ all die Macht ausspielen kann oder all die
198 Aktionen durchführen kann, die halt das ermöglichen könnten, nehmen wir mal
199 an Panzer steuern oder so (--) oder Atomraketen steuern oder sowas, dann
200 könnte der Algorithmus auch zu dem Schluss kommen, ja dann müssen wir halt
201 die Menschherrschaft übernehmen übersetzt. Die Sache ist halt, warum sollte
202 man so einen Algorithmus programmieren? Oder warum/ oder wo ist da der
203 Mehrwert? Deswegen (---) kommt es dann doch auf den Menschen an.

204 **VL:** Ja, vielleicht ist das ein ganz guter Übergang zum nächsten Thema. Also
205 das nächste wofür wir uns interessieren ist/ wir haben zwar gesagt du
206 benutzt einen Algorithmus, du arbeitest nicht mit ihm zusammen, aber
207 welche/ wenn du ihn benutzt, welche Eigenschaften muss die KI haben und
208 welche Anforderungen muss sie haben? Also was ist am wichtigsten?

209 **VP:** (5.0) Ich glaub am/ also wenn ich mit einer künstlichen Intelligenz/
210 also ok, ist die Frage so zu verstehen, dass man quasi mit der KI zusammen
211 ein Problem lösen soll oder ist die KI so verstanden, dass die KI einem im
212 Haushalt bspw. unterstützt und dir Sachen abnimmt, die du nicht mehr machen
213 willst? (---) Oder beides?

214 **VL:** Ich würde mal beides ansprechen, damit wir nicht irgendwas
215 ausschließen. Also /

216 **VP:** Ok, wenn die KI quasi meine Arbeit ersetzen soll, dann wäre es mir am
217 wichtigsten, dass das Ergebnis, also die Leistungsfähigkeit des Algorithmus
218 die geringste Fehleranfälligkeit hat. Also, dass sie die Aufgabe quasi
219 perfekt abnimmt. Also nehmen wir mal das Staubsaugen. Dann will ich, dass
220 die das perfekt staubsaugt. Perfekt natürlich/ jetzt kommt man natürlich
221 noch an die Frage was heißt perfekt? Da könnte man natürlich argumentieren,
222 dass sie (---) denn eigenen Präferenzen entspricht und da würde natürlich
223 noch hineinspielen, dass/ ok, was wäre, wenn (--) beim Staubsaugen sie
224 natürlich auch meine Präferenzen wie ich mach hier Yoga auf der Matte oder
225 so, dass sie da mich dann nicht umrollt oder umfährt. Das heißt, das würde
226 dann bedeuten, dass dann natürlich auch das Thema Sicherheit oder im Fall
227 der Fälle die Sicherheit der eigenen Person natürlich gewährleistet ist.
228 Ja, das heißt, das wäre ein, wenn ich zusammenfassen darf, ein
229 Zusammenspiel aus Sicherheit und optimaler Lösungsstrategie.

230 **VL:** Also würdest du sagen, dass die Eigenschaften von KI so ein bisschen
231 vom Kontext abhängen in welcher sie benutzt wird?

232 **VP:** Natürlich ja.

233 **VL:** Okay, das heißt/ Fallen dir noch andere Beispiele ein, wo es vielleicht
234 nochmal andere Punkte geben könnte? Also wir haben ja jetzt zum Beispiel

235 Home Appliances und Problemlösungen in der Industrie sag ich mal. Gibt's da
236 noch andere Sachen wo du sagen würdest: okay da könnte noch was hinzu
237 kommen als Anforderung oder dass du schon mal/ sagen wir, dass du schon mal
238 hattest, dass das passiert ist oder so?

239 **VP:** [überlegt] (2.0) Also ich weiß ja noch, dass (5.0)/ das künstliche
240 Intelligenz auch für Marktwirtschaft und Ads eingesetzt wird, also für
241 Werbung und da wäre es mir schon wichtig, dass ich auch ein paar
242 Werbesachen, die vielleicht/ die ich nicht sehen möchte, dass ich das auch
243 steuern kann. Quasi ihm noch nachträglich sagen kann, was mir wirklich
244 gefällt und nicht, sodass er (---) also das ich da auch ein bisschen
245 Einfluss hab, auf das System.

246 **VL:** Ja okay. (--) Vielleicht auch dann ein bisschen/ nicht vielleicht
247 reinsehen, aber noch Input geben?

248 **VP:** Genau Steuerbarkeit (--)

249 **VL:** Steuerbarkeit ist das, okay. Als nächstes (---) gucken wir uns auch an/
250 also das ist auch einer der anderen Themen hier, einmal KI und dann einmal
251 Transparenz. Und ich würde dich ganz gerne erstmal allgemein fragen was ist
252 Transparenz bei KI-Nutzung und ist dein System transparent? So einfach mal
253 (-)

254 **VP:** Transparenz bedeutet für mich, dass man versteht, wie der Algorithmus
255 zu dem Schluss gekommen ist/ den Output geliefert hat, den er geliefert
256 hat. Das wäre quasi der Weg vom Input bis zum Output. Und dass diese so
257 verständlich dargestellt wird, dass Menschen sie auch verstehen. Weil, wie
258 schon gesagt, im Algorithmus/ den kannst du dir schon anschauen. Das ist
259 das Modell, die Blackbox, aber das sind halt nur Zahlen. Zahlen, die für
260 dich erstmal keinen Sinn ergeben. Das heißt, es muss irgendeine Möglichkeit
261 geben, diese Zahlen zu übersetzen, weil (xxx) Visualisierungen, die dann
262 quasi für den Menschen verständlich sind.

263 **VL:** <<bestätigend> Mhm.> (---) Also der Weg (---) vom Input bis zum Output,
264 würde ich mir jetzt vorstellen/ so drin bin ich nicht, ich weiß nur wie das
265 ungefähr funktioniert, aber es ist ja tendenziell ein langer Weg und es
266 sind viele Daten, es gibt viele verschiedene Sachen, die da einfach
267 mitspielen. Welche Informationen brauchst du denn dann, um/ damit
268 Transparenz gegeben ist?

269 **VP:** (---) Das ist eine gute Frage. (---) Ich glaube für vollständige
270 Transparenz muss sowohl alle Schritte innerhalb dieses Modells abgebildet
271 werden, weil meistens verlaufen, wie gesagt, die schrittweise, die
272 Umwandlung von Input zu Output. Das heißt in jedem Schritt sollte klar
273 sein, wie oder welche Mechanismen da gegriffen haben, anhand welche
274 Informationen die Entscheidung getroffen worden ist und dann auch für den
275 gesamten Datensatz wäre es auch hilfreich. Also das man quasi nicht nur den
276 aktuellen Punkt, also die Begründung [Internetverbindung unterbrochen]/
277 generell greifende Regeln zuerkennen sind, anhand dessen der Algorithmus
278 quasi entscheidet.

279 **VL:** Okay. Du bist ganz kurz zwischendurch abgebrochen, also ich glaub die
280 letzten 20 Sekunden.

281 **VP:** Okay, dann versuche ich es nochmal zu wiederholen. Der zweite Punkt
282 war, dass man nicht nur einen bestimmten Datenpunkt/ die Begründung für
283 einen bestimmten Datenpunkt geliefert bekommt, sondern auch für/ am besten
284 für einen gesamten Datensatz bzw. für alle möglichen Datenpunkte, die
285 aufkommen könnten. Also dass man quasi nicht betrachtete Situationen
286 beurteilen könnte.

287 **VL:** <<bestätigend> Mhm.> Okay, das heißt, du strebst ein bisschen als auch
288 nach/ wenn du Transparenz sagst, dann sagst du auch/ zielst du auf
289 komplette Transparenz ein bisschen ab?

290 **VP:** Genau. Wir nennen es meistens globale Explainability oder Transparenz.

291 **VL:** Und das zieht dann zum Beispiel auch mit ein, den Trainingsdatensatz
292 zum Beispiel? Das ist/

293 **VP:** Genau. Also was ich damit meine ist halt/ oder was ich mir darunter
294 vorstelle zumindest, ist, unter der Annahme du hast einen
295 Trainingsdatensatz für den kannst du dir einzeln die Erklärungen abliefern,
296 aber du willst ja auch wissen was passiert, wenn ich dem jetzt mal was
297 anderes gebe, was er davor noch nicht gesehen hat (-) oder was so weit
298 außerhalb der Norm ist, aber trotzdem irgendwie vorkommen kann. Beispiel
299 ist ja immer dieses schöne automatisierte Fahren. Er sieht ganz viele
300 normale Straßen, aber was, wenn jetzt ein Elefant auf der Straße ist? Ja,
301 dann wollen wir ja auch in solchen Situationen wissen, wie er da
302 entscheiden würde. Natürlich kannst du dem ein Bild vom Elefanten jetzt
303 machen, aber das halt auch diese Situationen, sag ich jetzt mal,
304 durchleuchtet werden, (xxx) oder verstanden und abgedeckt werden. (--) Das
305 wäre natürlich ganz schön.

306 **VL:** Ja, verstehe ich. Also das die Wunschvorstellung, dass er auch diese
307 Situationen (---)/

308 **VP:** Ja

309 **VL:** Ja, verstehe. Meinst du, dass was ist/ Transparenz was wir immer
310 brauchen? Also muss das immer gegeben sein, wenn wir mir Algorithmen/ also
311 mit künstlicher Intelligenz arbeiten oder (---)?

312 **VP:** Ich (2.0), ich glaube es würde/ ich würde es mir wünschen, aber ich
313 glaube persönlich würde ich das aber auch nicht zwangsläufig verlangen.
314 Also man lässt sich meiner Meinung nach schnell dann von dem Service, den
315 dann dieser Algorithmus liefert, ich sags mal, einlullen oder so. Man wird
316 dann bequem, deswegen würde ich es/ deswegen denke ich, dass diese Systeme
317 die auf künstlicher Intelligenz basieren, dass die eine Transparenz
318 natürlich haben sollten, aber das natürlich auch viele Firmen auch am Ende
319 wegkommen nicht diese Transparenz darzulegen. Ich meine das siehst du ja
320 auch bei deinen eigenen/ das sieht man ja dann auch bei den ganzen Social
321 Media Plattformen, die dann natürlich ihre/ sagen ja wir haben viel
322 Transparenz, aber dann wirklich an die Daten ranzukommen und die
323 Algorithmen dann zu durchleuchten machen sie ja dann trotzdem nicht. Und
324 die Nutzer nutzen es ja trotzdem weiter. (---) Also das versuche ich damit
325 auszudrücken.

326 **VL:** Ja, ja voll. Frag ich mal andersrum. Wie schlimm ist es denn, wenn
327 diese Transparenz nicht da ist und vor allen Dingen bei Fehlern? Also ist
328 es, ja (---)

329 **VP:** Der Aufschrei ist dann natürlich groß, also wenn/ das sieht man ja dann
330 meistens/ also das sieht man ja häufig bei dem automatisierten Fahren, wenn
331 dann mal was passiert, sind die Leute natürlich oh wir brauchen
332 Transparenz, Transparenz. Oder nach zwei, drei, vier Jahren sagen sie dann
333 ok wow ist jetzt besser geworden und das Auto fährt jetzt vollautomatisiert
334 dann werden trotzdem alle Leute das System nutzen und automatisiert fahren.

335 **VL:** Ja, ja. Meinst du also nochmal/ wir haben ja eben schon mal drüber
336 gesprochen. Ich hab es glaub ich ein bisschen rausgehört, es kommt auch auf
337 den Kontext an, oder? Also, ob sie wichtig ist oder nicht? Ist doch
338 bestimmt so/ kann mir vorstellen automatisiertes Fahren wichtig, aber wenn
339 ein Unfall passiert, Social Media, Netflix Algorithmus, Recommender geht
340 so? Was würdest du sagen? Würdest du zustimmen oder da hast du da eine
341 andere Meinung?

342 **VP:** Also ja, ich würde da zustimmen. Ich denke, dass/ also ich weiß nicht/
343 also ich bin bei mir selber noch nicht so ganz sicher, aber ich glaube ich
344 würde mich da der Masse, der Bequemlichkeit anschließen, dass bei Systemen,
345 die sagen wir mal nicht lebensbedrohlich sind, dass da die Transparenz eher
346 unwichtiger wird im Vergleich zu Systemen natürlich wo dein Leben davon
347 abhängt. Ja, wenn ein Flugzeug oder so mit künstlicher Intelligenz
348 gesteuert wird, dann hättest du schon gerne irgendwie Sicherheit, dass das
349 System funktioniert. Aber ich denke, die Entwickler haben da ja auch ein
350 Interesse daran, wenn die sich als kritisches System entwickeln, dass da
351 die Transparenz der Algorithmen gegeben ist.

352 **VL:** Ja. Ja ich hatte/ Weil ich hatte die Diskussion sehr lange mit meinen
353 Kollegen und die meinten klar, du/ als erstes denkst du ja Recommender,
354 meistens ist es dir egal, was dir da vorgeschlagen wird, nur dann/
355 irgendwann kam die Frage auf ok, du wachst jeden Morgen auf, du willst
356 jeden Morgen eine Playlist haben, die dir von einer AI vorgeschlagen wird
357 zum Beispiel gehen wir mal ein ganz kleines bisschen in die Zukunft. AI
358 kann deine Stimmung so ein bisschen erklären anhand von Voice Recognition.
359 Über lange Zeit merkt sie halt wie deine Stimmung so ist, und dann schlägt
360 die dir halt morgens immer Musik vor, weil du morgens Musik hörst. Dann
361 irgendwann nervt es dich, dass sie dir die falschen Sachen vorschlägt.
362 Könnte man sich vorstellen, da ist Transparenz vielleicht doch dann
363 wichtig? Wahrscheinlich nicht kritisch, aber schon wichtig, um reingucken
364 zu können?

365 **VP:** (--) Ja, aber ich denke, dass löst sich nicht mit Trans/ Also dieses
366 Problem, finde ich, löst sich nicht mit Transparenz, sondern wie gesagt
367 durch Steuerung. Also ich finde es sollte dann einfach (--)Button geben
368 oder irgendwas wo du dann sagen kannst, ne mach das jetzt weg, weil das hat
369 mich jetzt abgefickt. Und ich vermute, dass dann, wenn die KI/ wenn es
370 diese Buttons gibt, wird es auch Lösungen geben, die das dann auch
371 berücksichtigen und dann quasi dir neue Arten von Liedern vorschlägt oder
372 die Entscheidung anders trifft.

373 **VL:** Ok, ja.

374 **VP:** Also, dass das quasi mitberücksichtigt wird.

375 **VL:** Okay, also Transparenz kann nicht einfach alle Probleme lösen?

376 **VP:** Genau, ich finde das löst dieses Problem bspw. nicht.

377 **VL:** Okay, ja. (---) Könntest du noch erklären, inwiefern Transparenz den
378 Prozess und die Entscheidung von KI/ also wie hilft dir Transparenz den
379 Prozess und die Entscheidung von KI zu verstehen, zu erklären und zu
380 interpretieren?

381 **VP:** Also ich finde Transparenz generell ist ja ein sehr großer Begriff. Und
382 meiner Meinung nach, wenn das System sehr gut transparent ist, dann macht
383 sie alles drei, oder?

384 **VL:** Das ist die Frage. Also wenn Transparenz gegeben ist und eine gute
385 Transparenz gegeben ist, dann sagst du, dass du es verstehen kannst,
386 erklären kannst und interpretieren kannst. Dann müssen alle drei gegeben
387 sein?

388 **VP:** Ne, ich glaube die Interpretation ist ja dann wieder jedem selbst
389 überlassen, meiner Meinung nach. Also Transparenz an sich erzeugt keine
390 Interpretation des Algorithmus, sondern es erklärt nur. So würde ich das
391 unterscheiden.

392 **VL:** <<bestätigend> Mhm.> Okay, aber das heißt jetzt nicht direkt, dass es
393 verstanden wird?

394 **VP:** Genau. Ich finde, nur weil ein Algorithmus transparent ist, heißt es
395 noch nicht, dass ihn alle dann auch verstehen. (--) Also, die Entscheidung.
396 (--) Und ob dann (-) wie nenn ich das jetzt/ ob dann ein transparenter
397 Algorithmus auch das erklären/ also das, sodass das jeder versteht, wie das
398 diese Anforderung erfüllt/ oder eine Erklärung erzeugt, die jeder gleich
399 interpretiert, das finde ich ist dann nicht unter Transparenz zu verstehen
400 (---) oder meiner Meinung nach.

401 **VL:** <<bestätigend> Mhm.> Okay. Okay sagen wir mal wir haben/ also wir sind
402 dann so ein bisschen ok/ Transparenz ist dazu da, dass eine Erklärung/ sie
403 erklärt die KI. Und verstehen und interpretieren sind nochmal so andere
404 Level, die dazu kommen. Wenn jetzt sagen wir mal eine KI wird genutzt. Die
405 KI ist transparent genug, um/ damit sie den Algorithmus erklärt und sie
406 wird auch genutzt von Nutzern. Sagen wir mal es ist ein Laien Nutzer.
407 Welche wichtigen Inhalte, denkst du, also große Themenblöcke müssen
408 transportiert werden oder geschult werden, damit die Leute das dann auch
409 wirklich verstehen, diese Transparenz und diese Erklärbarkeit?

410 **VP:** (3.0) Das ist (1.0) schw/ ich glaube, das wichtigste ist zu verstehen
411 den Zusammenhang zwischen (11.0) so fundamentale Sachen, Elemente der
412 Datenräume vielleicht? Also das/ das Daten eigentlich nichts anderes sind,
413 wie Zahlen und diese halt in einem Raum sich bewegen. Und quasi das dieser
414 Raum eigentlich nur dazu da ist, um Entscheidungen oder Unterschiede
415 festzumachen.

416 **VL:** <<bestätigend> Mhm.> Okay, also (--)

417 **VP:** Aber wenn man/ man könnte es sich auch einfach machen und sagen einfach
418 Machine Learning Fundamentals. Könnte man ja auch versuchen den Leuten
419 beizubringen, als Schulfach oder so. Oder Datenanalyse.

420 **VL:** Könnte man sagen, ja. Also spannend, also entweder man gibt denen ein
421 mentales Modell im Kopf, wo du schon gesagt hast wie mit den Zahlen, Räumen
422 und dann dass aus diesen Zahlen, Räumen eine Repräsentation entsteht. Kann
423 den Leuten helfen. Oder es geht wirklich darum Machine Learning/ mal einen
424 Kurs/ Einführung in Machine Learning zu belegen für verschiedene
425 Schulstufen, also ja okay.

426 **VP:** Also spontan fällt mir da auch nichts ein. Ich bin kein Experte im
427 Bilden von Leuten. ((lacht))

428 **VL:** Ja, ja ((lacht)). (---) Wir kommen zum letzten Block eigentlich und da
429 geht's ein bisschen um Vertrauen in künstliche Intelligenz. Ich stell jetzt
430 mal eine ganz grobe Frage. Unter welchen Bedingungen würdest du KI
431 vertrauen?

432 **VP:** (2.0) Vertrauen. Ja vertrauen ist immer so ein/ so ein schwerer/ ein
433 komplexer Begriff. Ich meine, ich glaube ich würde das gleiche/ oder den
434 gleichen Ansatz wie beim Menschen gehen, würde ich jetzt spontan sagen.
435 Also ich würde mich dann fragen ok wann vertraue ich einer bestimmten
436 Person? Die ist ja auch nicht immer transparent. Aber, obwohl sie ja nicht
437 transparent ist, vertraue ich ihr ja dann trotzdem. Das heißt, ich/
438 deswegen vermute ich, dass das Vertrauen durch Konsistenz erzeugt wird.
439 Also das der Algorithmus quasi immer gute Arbeit leistet, gute Ergebnisse
440 liefert und mich quasi noch nie im Stich gelassen hat. Deswegen glaube ich,
441 würde ich anhand dessen mein Vertrauen aussprechen.

442 **VL:** Spannend. Also es geht nicht nur um Transparenz. Also man könnte jetzt
443 nicht sagen, wenn der Algorithmus komplett transparent ist, würde ich im
444 mehr vertrauen?

445 **VP:** Das hab ich nicht gesagt. Ich habe gesagt, es würde vielleicht schon
446 reichen, wenn der Algorithmus quasi nur eine gute Leistungsfähigkeit
447 aufweist, dass ich ihm dann vertraue.

448 **VL:** Okay.

449 **VP:** Aber natürlich, wenn er vollständig transparent ist und quasi alles
450 offenbart und jede mögliche Fehl/ jedes mögliche Fehlverhalten/ dass jedes
451 mögliche Fehlverhalten ausschließt, dann würde ich ihm natürlich mehr
452 vertrauen.

453 **VL:** <<bestätigend> Mhm.> Okay. Und (---) also einmal haben wir Vertrauen/
454 also Transparenz könnte sein, dass es zu mehr Vertrauen führt bei dir. Was
455 ist, wenn du den Algorithmus verstehst? Führt das auch zu mehr Vertrauen
456 dann?

457 **VP:** <<überlegend> Mhm. (3.0)> Ich denke schon, ja.

458 **VL:** Ja, okay.

459 **VP:** Also würde ich erstmal bejahen.

460 **VL:** <<bestätigend> Mhm.> Ok. (3.0) Jetzt kommen wir nochmal zurück zum
461 Kontext. Meinst du, dass dein Vertrauen also global ist in KI ist oder dass
462 es auch wieder nur Kontextsache ist?

463 **VP:** (3.0) Ich denke, die Anforderungen/ ja die Anforderungen/ also die
464 Notwendigkeit des Vertrauens ist natürlich höher bei
465 sicherheitskritischeren Fragestellungen. Das denke ich ist schon der Fall,
466 deswegen würde ich sagen, ja das ist auch eine Kontextfrage. wie immer. Ich
467 finde das ist doch immer bei solchen soziotechnischen Themen, dass es

468 immer auf den Kontext ankommt, weil halt solchen generalen Aussagen nicht
469 funktionieren, weil du nicht alles abdecken kannst.

470 **VL:** <<bestätigend> Mhm.> Ja ok. Ja also, ich komme jetzt so ein bisschen
471 zum Schluss. Wir suchen immer mal wieder/ wollen irgendwann mal so
472 Nutzertests machen, wo wir KI/ also Nutzer mit einer KI konfrontieren und
473 gucken wie sie reagieren, wie sie damit umgehen, sozusagen. Fallen dir gute
474 Beispiele ein für/ also KI-Beispiele für/ die man mal bei Nutzern oder mit
475 Laien nutzen könnte? Also vielleicht auch aus deinen eigenen Erfahrungen
476 oder wo du denkst das wäre mal lustig da jemanden vor zu stellen und zu
477 sehen vertraut der der? Was macht der mit der? Braucht der Transparenz usw.

478 **VP:** <<bestätigend> Mhm.> (---) Ich/ jetzt konkrete/ also es gibt ja viele
479 Beispiele aus der Industrie. Du hast es ja schon genannt dieses
480 Recommendation System ist ja ein Themenfeld, dass jetzt schon, ich würde
481 mal sagen State of the Art ist. Und wenn du quasi in der Industrie
482 angekommen bist, dann gibt es, vielleicht das Thema des/ also
483 sicherheitskritische Themen wie das automatisierte Fahren. Da kenne ich ein
484 Beispiel glaub ich, das hat mal ein Freund von mir gezeigt. Da ging es
485 darum dem Agenten/ oder da steht der Agent quasi vor unmöglichen
486 Entscheidungen, also er muss sich entweder entscheiden einen Hund
487 umzufahren oder einen Menschen oder so. Und dann soll der Mensch
488 entscheiden, wie der Agent zu reagieren hat und daraus soll dann abgeleitet
489 werden, wie der Agent sich zu entscheiden hat. Sowas finde ich natürlich
490 spannend, aber das ist jetzt auch nicht/ aber da findet man glaub ich nicht
491 heraus, ob der Mensch dem Agenten vertraut.

492 **VL:** Doch. In eine ethische Fragestellung. Moralische, ne?

493 **VP:** Ja das ist auch so ein Thema. Verhalten mit KI. Was könnte man noch
494 machen? Vielleicht/ ah ich wollte noch sagen das Thema Sprache könnte man
495 auch untersuchen. Also es gibt ja ganz viele Modelle, die in der Lage sind
496 Text zu generieren und da könnte man Leute hinsetzen und die sollen mit
497 dem/ mit dieser Person chatten und sie wissen vielleicht gar nicht ob das
498 eine KI ist oder nicht, oder die müssen herausfinden, irgendwas kaufen und/
499 ich glaub, dass da, ohne dass die wissen, dass es eine KI ist, wäre es
500 vielleicht nochmal ganz spannend. Aber natürlich setzt voraus, dass die
501 Modelle schon gut genug sind und da gibt es auch viele technische
502 Fragestellungen. Ja, aber da könnte man vielleicht einen ganz coolen Use
503 Case draus basteln. Sprache, Bilder (---)

504 **VL:** Siehst du denn/ ((...)) Siehst du aus der Bildgenerierung einen Use Case?
505 Also nur, wenn es dir spontan einfällt. Also, sagen wir mal (---)

506 **VP:** Um die Interaktion mit Menschen zu testen?

507 **VL:** <<bestätigend> Mhm.>

508 **VP:** Ich sehe in der Bildgenerierung an sich noch keine
509 sicherheitskritischen/ oder noch keinen Anwendungs Use Case für den Otto
510 Normalbürger. Für mich ist diese Bildgenerierung eigentlich eine
511 Prozessmodellierung. Also wie eine Simulation der Wirklichkeit, nur dass
512 sie datengetrieben ist. Und deswegen übernimmt sie ja für dich nicht/ in
513 erster Linie nicht irgendwelche Aufgaben, sondern wir nutzen sie, um den
514 Prozess besser zu verstehen. Deswegen fällt mir dazu jetzt erstmal konkret
515 nichts ein. Man könnte natürlich/ es gibt natürlich Arbeiten, die wie
516 gesagt diese Kunst erzeugen und dann könnte man die Leute fragen, ob das
517 interessant wäre oder ob das sie anspricht. Ob die Kunst quasi bei denen
518 den selben Effekt hat, wie andere künstlerische Werke und (---) ob dann
519 Transparenz dafür, für Künstler interessant sein könnte. Aber das ist alles
520 sehr, sehr weit ausgeholt und sehr, sehr spezifisch.

521 **VL:** Aber nee spannend. Also doch man könnte fragen was sind die/ warum
522 spricht die Kunst einen nicht an und was wären sonst Merkmale, die sonst
523 ansprechend für die Personen sind. Und dann gucken was muss der/ die KI da
524 können, um ansprechende Kunst generieren zu können, die auch bestimmte
525 Nutzergruppen zum Beispiel anspricht. Ja ne doch, schon spannend. (3.0) Wie
526 siehst du KI in der Zukunft?

527 **VP:** Ich sehe, für mich ist es wie gesagt, es muss meiner Meinung nach, oder
528 es ist einfach ein Werkzeug der Automatisierung. Für mich ist es einfach
529 eine effiziente Nutzung der Daten, die wir haben, um Probleme der
530 Menschheit zu lösen bzw. uns die Arbeit abzunehmen, damit wir uns mit
531 Sachen beschäftigen wollen, wo wir mehr Lust drauf haben [(und (---))]
532 **VL:** [Solange]/ Solange wir eingreifen können und kontrollieren können
533 zwischendurch, ne? Hattest du auch gesagt.
534 **VP:** Ja genau, natürlich. Weil die Interessen/ weil mir die Arbeit
535 abzunehmen ist zwar gut, aber ich will trotzdem auch wenn ich bspw. doch
536 die Arbeit machen, will ich doch schon eher/ dann sollte die KI trotzdem
537 eher zweitrangig werden und mich die Arbeit machen lassen bspw.
538 **VL:** Ok. Ok, verstehe. (---) Zuletzt welche Aspekte von unserem Gespräch
539 fandest du am interessantesten und welche fandest du am wichtigsten zu
540 besprechen?
541 **VP:** (5.0) Ich fand am interessantesten/ oder am komplexesten die Frage wie
542 man alle Menschen dazu bringt KI richtig zu verstehen, weil wir selbst oder
543 wir, die dran arbeiten, haben es ja selbst noch nicht zu 100% begriffen,
544 wie/ oder diese komplexen Konzepte beizubringen wird glaub ich nicht
545 trivial. Man kann das/ man sollte es auch einfach nicht reduzieren zu ja
546 das macht jetzt einfach irgendwas und nur an Beispielen erklären, wie die
547 KI funktioniert. Weil genau dann erzeugt/ das erzeugt halt genau diese
548 Gefahr, dass die Leute dann halt denken, die KI könnte außer Kontrolle
549 geraten und die Weltherrschaft übernehmen oder so. Das ist/ sobald/ also so
550 lange man diese Systeme immer als Black Box Modelle betrachtet, wird es
551 immer Menschen geben, die das nicht/ die dem nicht vertrauen werden. Das
552 ist da. (-) Das fand ich am interessantesten, weil darüber habe ich mir
553 noch gar keine Gedanken gemacht. (---) Und am wichtigsten denke ich, (2.0)
554 ist dass man sich mit dem Thema Transparenz auseinander setzt. Also das
555 das/ das da auch die Forschung und quasi die Fragestellung und eure/ quasi
556 das Gespräch an sich ist das wichtig, dass quasi dieses Gespräch geführt
557 wird zwischen den Leuten, die es entwickeln, die es nutzen usw. Das
558 ist/also das es quasi die Metaebene, (--) wenn das Sinn macht. (---) Also
559 das dieser Diskurs überhaupt entsteht.
560 **VL:** Ja ne, auf jeden Fall. Sehe ich genauso auf jeden Fall. Ja also wir
561 sind am Ende angekommen. Hast du noch Fragen an mich (-) generell?
562 **VP:** Ich bin nur ein bisschen neugierig. Die [C] ist da noch mit dabei
563 und das ist ein neues Projekt, dass ihr quasi erzeugt habt. Oder wo/ also
564 die Rahmenbedingungen des Projektes hab ich noch nicht vollständig
565 verstanden, das würde mich noch interessieren.
566 **VL:** Ja, inhaltliche Rahmenbedingungen oder administrative
567 Rahmenbedingungen?
568 **VP:** Sowohl als auch. Inhaltlich wohl eher.
569 **VL:** Also kurz das Projekt: [B] und [C] kennen sich und/ der eben
570 kurz da war/ und [C] meinte da ist so ein Seedfund open call sollen wir
571 da mal was beantragen? Und dann haben wir das gerade mal, weil das gerade
572 so in unsere Themen passte, die uns einfach interessieren, haben wir das
573 mal gemacht und auch relativ schnell sag ich mal und auch nicht mehr damit
574 gerechnet, dass da irgendwas zurück kam. Aber dann kam irgendwann die
575 Bewilligung und genau jetzt sitzen von unserer Seite/ sitzen [B]
576 [D] der macht auch viel mit KI und ich da drin. Und bei [C] bei
577 euch, [C] und [E] ne? Und ja, ist ganz spannend und deswegen bin
578 ich auch/ deswegen bin ich dann auch weiter ins Productioncluster
579 reingerutscht, weil Dissertation bei mir geht es um KI. Und im Moment sitz
580 ich an einem Projekt da geht es um innovative Mobilität im ländlichen
581 Bereich, nicht wirklich was, was mit KI zu tun hat. Also klar autonomes
582 Fahren ist KI, aber damit hab ich halt nichts zu tun. Ja deswegen/ (---)
583 Genau und inhaltlich geht es darum/ eigentlich genau sowas was du gesagt
584 hast. Die Leute, die über Transparenz sprechen zusammenzubringen.
585 Verschiedene Stakeholder anzusprechen und zu wissen was ist Transparenz für

586 die. Alles was wir besprochen haben, wie wichtig ist Transparenz, also (--)
587 ja (-) zu gucken, wann braucht es Transparenz. Ist es kontextabhängig,
588 führt es zu Vertrauen? Welche Arten von Transpa/ das fand ich ganz
589 besonders spannend: welche Arten von Transparenz gibt es und welche Leute
590 brauchen welche Arten von Transparenz. Also das ist für mich das Aspekt,
591 der da irgendwie/ Vertrauen ist eine coole Sache, ist eine spannende Sache
592 zu untersuchen, aber was sehr spannend ist, ist/ wenn ich jetzt in einem
593 bestimmten Szenario bin, was brauche/ was muss die KI jetzt mir sagen. Weil
594 es kann schon helfen, dass sie mir sagt: ja die Entscheidung der KI ist
595 jetzt so ausgefallen, weil die Trainingsdaten aus nur New York, Brooklyn
596 kommen. Und das kannst du nicht auf New Jersey übertragen. Oder bei manchen
597 Sachen hilft es: ach ja stimmt, da ist ja dieses/ da steckt ja dieses
598 Modell dahinter, deswegen kommen wir dahin. So das sind die Sachen, die ich
599 irgendwie spannend finde und auch mal ein bisschen weiter aus der Box
600 denken, zu sagen vielleicht ist es für manche Leute schon wichtig zu sagen,
601 dass eine KI überhaupt da ist, weil das wissen viele auch nicht. Oder zu
602 sagen, dass die KI von der Person entwickelt worden ist. Also das ist auch
603 eine Sache, die Transparenz ist. Zwar jetzt nicht auf Algorithmusebene,
604 aber auf der noch Ebene die danach oder daneben liegt. Auch diese ganzen
605 Sachen. Deswegen ist es halt so wichtig einmal mit Experten zu sprechen und
606 dann aber auch einfach mit Laien zu sprechen, mit Nutzern, mit Entwicklern,
607 aber auch mit Policymakern, um auch diesen ganzen Ethics Bereich abzudecken
608 und so. Und genau darum geht's. Lange Rede, kurzer Sinn: es geht um so eine
609 Transparency Map, so einen Framework zu bauen und zu sehen, was sind die
610 verschiedenen Arten. Also so ein Designspace, was sind die verschiedenen
611 Arten von Transparenz, was sind die verschiedenen Stakeholder, was sind die
612 verschiedenen Faktoren, die das beeinflussen und wer braucht was,
613 sozusagen. Und das ist das Projekt eigentlich.

614 **VP:** Okay, ich hab es ein bisschen verstanden. Ich finde es nur/ ich fands
615 persönlich relativ schwer auf eure Fragen eine zufriedenstellende Antwort
616 zu geben, weil ich/ ich finde, weil/ wenn du jetzt besonders Laien
617 ansprichst, wirst du/ vermutlich wirst du viele bunte Antworten bekommen
618 oder sehr oberflächliche Antworten. Und obwohl ich damit arbeite, finde ich
619 es super schwer da eine Antwort zu geben was/ wie man Transparenz zu
620 definieren hat, welche Arten von Transparenz gibt es. Deswegen befürchte
621 ich, dass diese Map sehr, sehr bunt wird. Das ist nur/ so würde meine
622 Intuition jetzt sagen. Und quasi alles das ein bisschen einzudampfen, aber
623 auch nicht die Kernaussagen zu verlieren oder die/ quasi eine Direction
624 abzuleiten könnte ich mir als eine große Herausforderung vorstellen. Aber
625 in der Literatur gibt es dazu auch schon Ansätze, die wahrscheinlich das
626 auch schon mal versucht haben. Da kann man bestimmt auch noch ein paar
627 Insights bekommen.

628 **VL:** Also ich würde sagen für mich ist es so/ ich hoffe, dass die Map so
629 bunt wird, weil das ist halt der erste Schritt, jetzt ist so
630 Bestandsaufnahme, was sind die ganzen Meinungen dazu oder Einschätzungen.
631 Und um dann im nächsten Schritt/ wir haben halt auch zum Beispiel, das ist
632 jetzt halt ein Experteninterview würde ich mal so sagen. Ich mach ein
633 Interview mit dir, wenn ich dir eine Frage stelle die um Interpretability
634 geht, weißt du was ich meine, so. Wenn wir natürlich jetzt Laien fragen,
635 haben wir natürlich andere Methoden dafür. Dann machen wir natürlich nicht
636 ein Interview. Das ist dann natürlich viel schwieriger für die, darauf zu
637 antworten, würd ich mal sagen. Wenn du schon sagst es war schwierig für
638 dich auf die Fragen eine zufriedenstellende Antwort zu geben, ist halt auch
639 die Sache. Vielleicht mag die Antwort für dich nicht zufriedenstellend
640 sein, aber für uns ist sie sehr, sehr wichtig sozusagen, wenn wir deine
641 Sicht daraus sehen, weil da ganz viele Aspekte drin sind, die wir zum
642 Beispiel gar nicht auf dem Schirm haben. Und da ist dann egal, ob die
643 Antwort irgendwie sagen wir mal zufriedenstellend ist, weil darum geht's
644 halt nicht. Es geht halt um den Aspekt, den wir noch nicht aufgedeckt

645 haben, der jetzt plötzlich da ist, so. Deswegen sind diese Interviews halt
646 voll wichtig sozusagen.

647 **VP:** Okay. Okay, dann ja. Dann finde ich es gut, weil/ oder dann finde ich
648 es/ oder dann kann ich es besser nachvollziehen, weil ich würde mich jetzt
649 nicht als einen Transparenz oder Interpretability Experten bezeichnen. Aber
650 das war wahrscheinlich gar nicht der Sinn dahinter quasi die richtige
651 Antwort darauf zu bekommen, sondern eher eine Antwort und dann zu schauen,
652 ob es da Konsens gibt, ob es da unterschiedliche Meinungen zu gibt und (--
653)/ okay dann hab ich das jetzt auch besser verstanden.

654 **VL:** Genau, also was man sich immer vor Augen halten muss, ist das, wenn wir
655 zum Beispiel ein Interview machen oder einen Fragebogen, es gibt keine
656 falschen Antworten oder richtigen oder komplette Antworten. Es gibt halt
657 nur, sag ich mal, wichtige Antworten. Weil es ist alles wichtig, damit wir
658 daraus/ man kann daraus voll viel ziehen, weil wir als/ wir sind/ klar wir
659 sind ein bisschen im Feld drin, aber nicht so tief wie du drin bist und
660 dann sind Sachen, die wir dann weiter mittragen können, die wir noch so gar
661 nicht auf dem Schirm haben. So und das ist dann eine neue Kategorie, die
662 für uns/ die wir aufmachen und die wir dann weitertragen können und
663 deswegen ist es halt so wichtig so Interviews wie heute zu führen
664 sozusagen.

665 **VP:** Okay. Okay cool.

666 **VL:** Okay. Ja vielen, vielen Dank auf jeden Fall, dass du mitgemacht. Ja
667 super. Punktlandung fast. Ja, cool, dass wir uns vielleicht im Cluster
668 sehen.

669 **VP:** Ja. Keep me updatet bzgl. eurer Fortschritte, das würde mich auch
670 interessieren, was ihr da für Ergebnisse erzielt. Wahrscheinlich/ ich werde
671 es wahrscheinlich von [C] so ein bisschen mitkriegen in den
672 Teamsitzungen, aber mich interessiert das Thema auch.

673 **VL:** Okay, ja. Du wirst von [C] mitkriegen, aber ich hab es mir
674 aufgeschrieben, weil sonst vergesse ich sowas. Ich werde es dir auf jeden
675 Fall dann, ich würde mal sagen nach unserem ersten Milestone, kann ich mal
676 so ein paar Sachen rumschicken und dann kannst du ja mal reingucken.

677 **VP:** Cool.

678 **VL:** Ok. Dann wünsch ich dir noch einen schönen Tag und vielen Dank nochmal.

679 **VP:** Gerne, hat Spaß gemacht.

680 **VL:** Cool, freut mich. Ciao.

681 **VP:** Ciao.