A1

Das klingt gut. Ja. Genau, und dann können wir ins Interview einsteigen. Kurze Vorstellung haben wir gerade schon bekommen bezüglich deiner Arbeit. Inwiefern, um jetzt ein bisschen spezifischer zu werden, wird da Machine Learning oder KI benutzt?

I11

Also viel so Richtung Klassifizierungsgeschichten. Jetzt aktuell kommt gerade diese ganze LLM-Thematik ziemlich hoch, weil viele wollen irgendwie Chatbots haben, viele wollen ihre Dokumentation mit so Large Language Modellen anreichern, dass man quasi mit der Dokumentation reden kann. Das ist gerade ein großes Thema. Und ja, also das, was wirklich Machine Learning Basics ist, ist wirklich viel so Klassifizierungsgeschichten oder so Clustering-Geschichten und Anomalie-Erkennungssachen. Solche Themen sind das jetzt oft.

A2

Ich… kannst du noch mal kurz erklären, für mich als nicht Technik Background, was ist XXXXX X? Was ist das Produkt oder was genau machst du oder ihr?

I11

Wir sind im Grunde Dienstleister. Nicht im Grunde, sondern wir sind Dienstleister. Für… Ich schimpfe das immer, Digitalisierung, Data Science, Softwareentwicklung. Also Themen, die sich in diesem ganzen Daten- und Digitalisierungsumfeld bewegen. Wir gehen im Grunde zu Firmen hin oder Firmen fragen bei uns an, wenn die irgendwas mit Daten machen wollen, ganz platt. Es kann wirklich sein, dass die kommen und sagen, wir haben gehört, man braucht jetzt KI, lasst uns irgendwas mit KI machen. Und die haben noch gar nichts. Es kann aber auch sein, dass die schon über Jahre Daten annehmen an ihren Maschinen und wollen jetzt gucken, was man mit diesen Daten machen kann. Kann aber auch sein, dass die schon ganz ausgereifte Sachen haben und da irgendwie noch so eine kleine Idee haben, die mal irgendwer außerhalb so ein bisschen ausprobieren soll. Aber es gibt ganz viele Konstellationen. Aber ein klassischer Fall ist, dass wir irgendeinen Maschinenbauer haben, zum Beispiel eine Firma, die stellt Anlagen her, die Folien herstellen oder Papieranlagen. Das sind Einfamilienhaus-große Anlagen. Anlagen kosten Millionen Euro, Weltmarktführer, sind immer alle Weltmarktführer bei uns. Und die wollen jetzt diese Maschinen so mit Sensorik ausstatten, dass man aus diesen ganzen Maschinedaten schauen kann, ob das Produkt, was am Ende rauskommt, also zum Beispiel die Folie oder das Papier, was da produziert wird, ob das gut ist. Weil das natürlich Prozesse sind. Da steht jetzt nicht jemand oben in Etage 13 und guckt auf die Folie und sagt, ja, ist gut oder schlecht, sondern das muss natürlich alles irgendwie automatisiert laufen. Die ganzen Prozesse mit Daten anzubinden, das ist so ein Fall. Und dann will man am Ende vielleicht, dass der Maschinenbediener eine App hat, wo drin läuft, ja, die letzten 500 Meter waren gut, irgendwie jetzt ist hier gerade schlecht. Das muss alles in einem System sein. Kunden sollen alle ihre Maschinen irgendwie einsehen können. Also solche Thematiken ist so das, was wir machen unter anderem.

A2

Danke, also ich weiß, du hast kurz vorgestellt, aber ich musste jetzt nochmal für mein Verständnis nochmal ein bisschen genauer hören, danke

A1

Genau, was mich in dem Kontext interessieren würde, ist eure Zielgruppe der Firmen, die ihr ansprechen wollt, irgendwie begrenzt oder ist das die komplette Sparte, egal wer ein Problem mit Digitalisierung und automatisierter Datenverarbeitung hat, ihr kümmert euch drum oder ist es schon Fertigung jetzt zum Beispiel, das Beispiel, das du gerade gegeben hast, wäre ja so industrielle Fertigung von einem bestimmten Gut.

I11

Das ist der Kern, der historisch so gewachsen ist. Weil, also wir haben so ein paar wirklich harte Ausschlusskriterien. Das sind irgendwie so Rüstungsgeschichten oder ethisch fragwürdige Sachen. Alles, was mit Personendaten zu tun hat, machen wir nicht. Also wir sind schon eher an Maschinen und Prozessen interessiert. Wir machen jetzt kein Online-Marketing und irgendwie so Zielgruppenfokussierung oder so. Das ist überhaupt nicht unsere Ecke. Finanzdienstleistungen haben wir auch noch gar nicht gemacht. Einfach, weil das nie so mit diesem Physik-Background die Ecke war. Das schließen wir jetzt aber nicht aus ethischen Gründen aus. Aber das ist einfach nicht so unsere Ecke. Wir haben so ein paar Medizintechnik-Geschichten. Ja, manchmal kommen noch so andere Sachen irgendwie so um die Ecke. Aber der Fokus ist im Moment schon wirklich der deutsche Markt und so der Maschinenanlagenbau eher im Mittelstand.

A1

Alles klar, okay cool. Was sind denn so deine arbeitsalltäglichen Prozesse? Und also du hast ja vorhin schon gesagt, dass du auch selber ein bisschen in Softwareentwicklung und so reingeschneit bist. Inwiefern spielt KI da auch eine Rolle in deinen arbeitsalltäglichen Prozessen?

I11

Also es hat so ein bisschen zwei Seiten. Das eine ist das, was ich wirklich tue jeden Tag, was man dann wirklich technisch macht. Das andere ist das, was Kunden haben wollen, was man verkaufen muss. Klingt immer so blöd, aber das gemeine Verständnis von KI und das, was KI jetzt irgendwie tun soll für die deutsche Wirtschaft, ist ja meilenweit entfernt von dem, was man vor dem Rechner sitzt und wirklich tut. Was ich dann wirklich tue, sind, wenn man jetzt aus der Forschung kommt, sind das tatsächlich eher Basics. Wir trainieren keine neuen Modelle jetzt groß. Wir benutzen viel, was schon fertig ist. Für so kleinere Sachen dann einfach irgendwie hier so SKLearn. Also wir sind so ziemlich in der Python-Welt unterwegs, dass man mal so einen kleinen Random Forest trainiert oder so ein bisschen Clustering macht. Also wirklich eher so Basics. In der LLM-Sparte haben wir uns auch so viel mit lokalen Modellen und wie man die laufen lassen kann und wie man damit arbeiten kann befasst. Wir sind jetzt aber nicht so unterwegs, dass wir selber irgendwelche krassen neuronalen Netze bauen und uns da um die Architekturen bemühen und welche Loss-Funktion da jetzt die beste ist. Das ist gar nicht so unser Fokus, sondern wir sind eher wirklich Anwender. Und ja, Anwender. Von dem, was gerade so da ist. Und das ist auch oft so, dass in den Fällen, die wir haben und für das, was man dann wirklich macht in der Industrie, das ist meilenweit entfernt von dem, was man jetzt in der Wissenschaft, auf welchem Level man sich da einfach bewegt, was Leute brauchen, was gerade für den Prozess wichtig ist. Genau, also wirklich technisch sind das eher Basics. Da würde man jetzt, wenn man in Doktorarbeit in Machine Learning schreibt, müde drüber lächeln wahrscheinlich. Die andere Seite ist das, was man, was so an Verständnis da ist bei unseren Kunden, was die haben wollen und was man irgendwie verkaufen soll. Also wenn man dieses Wort KI in den Mund nimmt, dann geht schon eine ganze Spirale los irgendwie. Oh jetzt ist XXXXXXX weg

A2

Ja, weiß ich auch nicht. Aber wir nehmen auf und machen einfach mal weiter.

I11

Wir nehmen auf, dann kann ich weiterreden. Genau, also diese Vorstellung von KI, wo man da irgendwie mit was man da zu tun hat, da ist es eben oft so, dass so kleine Sachen, wie so eine kleine Klassifizierung, und ich habe hier so einen Datensatz und ich will mal klassifizieren, ob mein Produkt, was ich da produziere, gut oder schlecht war so ungefähr. Und das kann man dann oft, was weiß ich, mit einem kleinen Random Forest machen. Dann ist das für die Kunden schon die riesengroße KI. Hallo, da bist du wieder. Noch bist du gemutet.

A1

Okay, sorry, ja, keine Ahnung, hier war anscheinend gerade kurz ein Uni-Aussetzer. Das Meeting lief weiter?

A2

Ja, wir haben ganz souverän weiter…

A1

Dann haben wir hier schon mal einen guten Fail-Check einmal gehabt, den wir bisher noch nicht hatten. Ja, sorry. Ja.

I11

Genau, ich hatte gerade erzählt, dass für unsere Kunden, wenn man da irgendwas mit KI macht und wenn man da irgendwie so ein kleines, so einen kleinen random Forest trainiert, dass das oft schon für die irgendwie, ja, große Magie ist und großen Impact hat. Und ja, man da manchmal so ein bisschen mehr erzählen muss und für die diese Welt so ein bisschen blumiger, magischer ist, als man das jetzt aus der Forschung gewohnt ist.

A2

das heißt wenn ich das jetzt richtig verstanden haben du bzw. ihr baut eure eigenen Systeme nutzt aber auch irgendwie andere Systeme oder baut auf anderen Sachen auf integrierte bei euch

I11

Genau. Also jetzt speziell in dieser, ich sag mal, in dieser LLM-Thematik, das sind halt so Modelle und neuronale Netze, grob gesagt. Das ist halt so groß und man braucht so viele Daten, um das zu trainieren. Das kann man nicht mal eben nachbauen, sondern das ist irgendwie vortrainiert. Das sind große Systeme, die mit ganz viel Hardware auch irgendwie unterstützt sind. Da kann man sich quasi dran stöpseln und kann das benutzen und baut das nicht, weiß nicht irgendwie selber nach. Und da nutzen wir fertige Sachen. Ist auch immer ein Riesenthema mit Datenschutz und so und nochmal eine andere Sache. Genau. Und für so kleinere Sachen, wo man mal so ein bisschen rumprobieren soll, gibt es aber auch viele Sachen, die man wirklich bei sich selber lokal baut, die man from scratch im Grunde bei sich selber machen kann, wo man sich nicht an irgendwelche externen Systeme anstöpseln kann.

A1

Vor allem, weil ja viele Firmen auch gar nicht solche Datenmengen überhaupt haben, dass es sich lohnen würde. Ich bin nebenbei auch noch in einem Industrieprojekt drin, wo wir auch KI versuchen, in die industrielle Fertigung irgendwie reinzubringen. Und ja, da sind die Daten ja auch nicht so, wie man das, wir als Machine-Learner kennen, ein perfekt kuratierter Datensatz, der vielleicht nicht perfekt kuratiert, aber schon sehr strukturiert und so. Da ist ja viel erstmal Data und Business Understanding, was man erstmal irgendwie leisten muss, damit man dahin kommt, dass man überhaupt irgendwas trainieren kann.

I11

Es ist auch oft 80 Prozent der Arbeit, die wir machen, tatsächlich. Gerade so in Anfangsprojekten, wo wir den Kunden neu kennenlernen, die Prozesse neu kennenlernen. Das ist nie so, dass man da hinkommt und dann liegen da 80 Gigabyte hervorragend aufbereitete Daten, die man gut nutzen kann. Das ist, da ist ganz, ganz viel Vorarbeit. Das ist die Hauptarbeit. Also wenn ich mal dahin komme, dass ich einen Punkt fit irgendwo ausfüllen kann, dann habe ich schon sehr, sehr viele Stunden dafür gearbeitet, so ungefähr.

A1

Ja, auf jeden Fall. Aber super cool, dass ihr praktisch auch beides macht, weil ich hatte jetzt auch gerade, wir hatten das jetzt auch schon in den ersten Interviews so ein bisschen, es gibt halt auch viele Leute, für die ist jetzt KI und Machine Learning ist jetzt nur noch LLMs so und also es gibt halt ganze Unternehmen, die KI machen, die aber eigentlich nur LLMs und Chatbots machen so, was auch fair enough ist, weil das ist halt gerade auch irgendwie das, was am meisten...

I11

Braucht überall, also alle kommen da dran.

A1

Und ich glaube es ist auch kaum möglich als KI Data Science, Machine Learner Start Up Team, was auch immer, arbeit zu machen ohne das Thema komplett auszublenden und zu sagen: Wir machen in die Richtung gar nichts. Weil dann schreibt man viel Markt ab sozusagen

I11

Auf jeden Fall. Also da kann man sich nicht, das ist nicht das, was wir hauptsächlich machen. Und wir haben auch so ein bisschen für uns den Anspruch, nicht KI so als Magic Black Box zu verkaufen, so dass wir alle eure Probleme lösen, sondern wir gehen da wirklich sehr hemdsärmelig erstmal dran versuchen, die Prozesse zu verstehen und benutzen KI oder Machine Learning so in unserem Sprech, dann, wenn das Sinn macht. Wir gehen nicht hin und sagen, wir verkaufen euch eine KI-Lösung, die auf jeden Fall KI enthält und dann ist es am Ende eine lineare Regression oder so, sondern wir gucken erstmal, wo das Sinn macht. Da haben wir Machine Learning-Thematiken drin, das ist so ein bisschen unser Ansatz.

A1

Ja, sehr cool. Cool, dass ihr beides ein bisschen macht und dass ihr da auch in der Tiefe unterwegs seid. Was sind denn so die größten Herausforderungen bei euch, wo du siehst, das macht diese Bereitstellung von KI-Lösungen oder ML-Lösungen, Terminologie an der Stelle jetzt mal austauschbar benutzt. Was sind so die größten Herausforderungen bei der Bereitstellung oder Entwicklung von solchen Systemen?

I11

Also, wenn wir das Menschliche mal rausnehmen, sprich Kommunikation mit dem Kunden, Erwartungsmanagement, so, das ist eine Sache. Technisch ist es ganz klar Datenqualität, würde ich sagen, Datenqualität und eben vorhandenes Labeling, weil natürlich irgendwie Prozesse in der Industrie nicht von Anfang an so aufgebaut wurden, dass wir an jeder guten Produktion, es ist eine gute Produktion dran stehen haben und an jedem Ausfall oder schlechter Produktion haben wir genau dran stehen, ist schlecht und warum das schlecht ist, ja, dann hat man oft, so, wenn wir so über Klassifizierung sprechen, hat man, wir hatten mal einen Fall, da ging es um die Kündigung von Verträgen und dann gab es einen ganzen Datensatz mit 100.000 Verträgen, irgendwie wann die abgeschlossen wurden und mit welchen Randparametern ohne persönliche Daten und dann sollten wir irgendwie gucken, wann so ein Vertrag endet und dann hatten wir genau ein Event irgendwie, wo mal so ein Vertrag geendet ist. Also solche Sachen, wo überhaupt gar nicht die Datenbasis da ist, dass man da irgendwas machen kann, also solche Themen.

A2

Und wenn du, was mich jetzt interessieren würde, wenn du bzw. wenn ihr Systeme praktisch euch dazu holt, also die ihr nicht selber entwickelt habt, wie findet also der Entscheidungsprozess statt und was sind da Herausforderungen, die du wahrnimmst?

I11

Also der Entscheidungsprozess ist da sehr stark von dem getrieben, was die Kunden haben wollen. Also jetzt speziell nochmal auf diese LLM-Thematik, weil das ist eben das, wo man gerade im Moment eben oft darüber spricht, dass man da irgendwie externe Services nutzt. Da muss man so ein bisschen gucken, wo die schon unterwegs sind. Also Microsoft Azure ist irgendwie so, das ist so die Umgebung, wo eigentlich alle oder sehr, sehr viele mittlerweile unterwegs sind. Microsoft hat mit OpenAI so ein Deal geschlossen, dass man die ChatGPT oder die GPT-Instanz selber lokal quasi in seiner Umgebung nutzen kann. Wie das technisch funktioniert, da bin ich auch noch nicht so hundertprozentig hintergestiegen. Aber es ist so, dass die Kunden oft Microsoft schon vertrauen, weil die in dieser ganzen Office-Welt leben und dann sagen, ja okay, dann können wir auch das GPT-Modell da benutzen. Und ist das so ein bisschen von denen vorgegeben. Und ansonsten, ja, gucken wir uns an, was geht. Wir versuchen so in unserem physikal Grundspirit immer irgendwie Open Source zu arbeiten, wo das geht. Wenn man sich natürlich in solchen Thematiken wie jetzt so LLM und Chatbots und so, da stößt man mit so kleinen lokalen Open Source-Sachen irgendwann an seine Grenzen. Wenn das was ist, was produktiv bei einem Kunden laufen soll, so, da muss man sich irgendwie entscheiden. Und ja, da ist... Also ich kann gar nicht sagen, wie das dann so zustande kommt, aber das ist oft in Absprache mit dem Kunden, mit dem, was die schon haben und wo es geht, irgendwie Open Source und umsonst Lösungen, die Sinn machen.

A1

Okay. Fallen dir konkret Sachen ein, wo du sagst, das wäre ein riesiger Enabler, das würde einige der Herausforderungen oder alltäglichen Probleme krass verringern?

I11

Für die Entwicklung von so machine learning getriebenen Anwendungen meint ihr jetzt?

A1

Ja. Also die Entwicklung oder auch die Bereitstellung. Ich würde das so unterteilen. Also entweder entwickelt ihr eine eigene Software sozusagen, um konkreten Datensatz, also konkretes Problem im Kunden zu lösen. Oder es ist halt wirklich so, dass ihr euch Lösungen von OpenAI oder so reinholt und die bereitstellt. Das sind ja so ein bisschen die zwei Wege, die ihr, glaube ich, im Moment habt.

I11

Ja. Also klar, in den Prozessen, also jetzt fantasiert, wenn man natürlich schöne Daten hätte, wäre das gut. Klarer Fall. Ist aber in dem Sinne oft nicht zu machen, einfach weil das historisch nicht so gewachsen ist. Also wenn so grundsätzlich ein Mitdenken von solchen Thematiken, also so grundsätzlich so, wenn man so Maschinen aufsetzt und wenn man anfängt Daten zu nehmen, wenn man da grundsätzlich schon mal mitdenken würde, dass man vielleicht gelabelte Daten gebrauchen kann, dass man das vielleicht irgendwie in Strukturen, also in Dateninfrastruktur gießt, wo man einfach auch vielleicht als externer Mensch, der nicht komplett in der IT-Welt des Kunden lebt, irgendwas rauskriegt. So solche Sachen sind dann oft einfach Hürden, aber da gibt es jetzt glaube ich nicht ein Produkt, was das irgendwie lösen will. Also das ist wieder so derjenige, der das Produkt entwirft, was es lösen will, der entwirft wieder das 101. Produkt, was es lösen soll, aber es dann doch nicht tut. Ja, solche Sachen und der Rest ist einfach sehr viel Recherche dann oft, weil man sich das zusammenbasteln muss, von der Hardware bis zu dem Modell, was man dann braucht, was man haben will, wenn es ja jetzt das ideale Recherchetool gäbe, was mir genau das rausbastelt, aber da ist auch die Welt so schnelllebig. Wenn ich mir in drei Monaten irgendwas angucke, dann gibt es wieder allein wie diese LLM-Thematik jetzt, also da haben wir wirklich angefangen vor einem Jahr oder was, dass man wirklich in der Konsole sich so ein Modell runtergeladen hat und das irgendwie neu gebaut hat und die Architektur davon eingehen musste und jetzt hat man irgendwie drei Zeilen Python-Code und kann das alles so, also das ist halt super schnelllebig.

A1

Ja, auf jeden Fall.

I11

Von daher, ja, also ich glaube, da das Recherchieren und das Zusammenbasteln und das auf dem Laufenden bleiben, kann einem da auch keiner abnehmen. Also das sind so die Punkte, würde ich sagen.

A1

Ja, leuchtet total ein. Wir haben ja bei unserer Studie auch eine konkrete Möglichkeit, so ein bisschen im Sinn, wo man vielleicht helfen könnte. Da geht es halt Richtung Modellauswahl oder Reporting auch von Machine Learning Eigenschaften. Weil das ist zumindest was, was jetzt auch mir in den XXXXX XXXXXX XXXXXX XXXXXXXXX halt irgendwie immer wieder auch im Kontakt mit Firmen aufgefallen ist, dass so diese Kommunikation über Machine Learning halt ein Riesenproblem ist. Und die ML-Wissenschaftler, die wollen halt irgendwie Source Code sehen oder zumindest das Paper und die ganzen Formeln so. Aber das ist wieder was, das kann man ja keinem zeigen, der nicht diese Domänen-Expertise hat. Und wir versuchen da so ein bisschen den Gap zu bridgen. Die Lücke zu schließen zwischen Entwicklern und Experten und Nicht-Experten und Anwendern. Oder halt auch Firmen, die halt irgendwie zwar KI-Lösungen entwickeln oder jetzt versuchen, KI in ihre Lösung einzuarbeiten, aber eigentlich noch nicht dieses umfangreiche Domänenwissen haben. Ich teile dazu einmal kurz meinen Bildschirm. Und dann reden wir da ein bisschen drüber, was ich dann so zeigen will. Ich muss mal einmal kurz die Präsentation. Noch starten. Soweit sich PowerPoint mal bequemt. Hello. Okay, sehr gut. Das war das. Und jetzt zeige ich noch meinen passenden Bildschirm dazu. Das wäre dann jenes. So, seht ihr meinen Bildschirm?

I11

Noch nicht.

A1

ich kann eventuell

I11

Ah jetzt. Sogar in Vollbild, Wahnsinn.

A1

Sehr gut, so war es gehofft. Perfekt, genau. Das ist jetzt einfach mal so ganz direkt das Beispiel, was wir zeigen wollen. Das wäre jetzt so ein KI-Label, wie wir es irgendwie so gedraftet haben, sage ich mal. Und bevor ich hier irgendwas im Detail erkläre, würde ich gerne einmal so die spontane Reaktion abfragen da drauf. Was siehst du, was geht dir durch den Kopf?

I11

Ich sehe, es erinnert an den Nutri-Score. Ich sehe jetzt erst auf den zweiten Blick, dass es gar nicht so ein schon bekanntes Label ist. Also ja, okay. Also soll es ein, ich sehe daraus jetzt ein Label, wenn ich ein Machine Learning Modell habe, was mir sagt, wie viel, was sehen wir denn, was das powermäßig kostet, wie schnell das ist, wie akkurat das ist und wie robust das ist. Ja, das macht Sinn. Sehr schön.

A1

Sehr schön. Das ist gut. Das ist eine schöne, das können wir so ins Paper schreiben. Ne, genau, also vollkommen richtig. Wir haben unten so ein paar Detaileigenschaften, konkrete Messwerte und halt so ein Gesamtscoring, was natürlich jetzt optisch so ein bisschen an den Nutri-Score angelehnt ist. Es ist tatsächlich gar nicht so einfach, ein neues Label zu schaffen, ohne dass Leute sagen, das sieht so aus wie.

I11

Ja, das verstehe ich. Und es simpel zu halten und

A1

Ja, so der Nutri-Score ist halt so ziemlich das Simpelste, was man machen kann mit so einer Fünf-Stufen-Skala. Aber dadurch sieht es halt auch direkt so aus. Ja, genau. Wenn man das jetzt so sieht, siehst du da vielleicht schon irgendwie eine Brücke zu euren Themen oder Problemen? Hat das irgendwie eine praktische Relevanz bei euch? potenziell?

I11

Also vielleicht irgendwann, wenn man da hinkommt, dass man auf der Basis überhaupt irgendwas optimieren will. Also in dem, was wir so täglich machen jetzt, wenn man das Ganze mal nimmt, ist halt der Schritt am Ende, das Modell richtig zu wählen und dann da vielleicht noch zu gucken, was da ein bisschen weniger kostet und ein bisschen schneller ist. Aber so weit kommen wir oft gar nicht. Also wir sind sehr viel erstmal so im Prototyping unterwegs und man guckt erstmal, ob überhaupt irgendwas funktioniert. Von daher ist das bestimmt gut, wenn man wirklich an den Punkt kommt, dass man auf dem Level Dinge optimieren will. Ehrlicherweise in unserer Arbeit ist das gar nicht so oft der Fall. Es ist nicht jetzt so, dass wir in jedem Projekt irgendwann an den Punkt kommen, wo wir uns über das Machine Learning Modell und die Performance Gedanken machen.

A1

Sondern so dieses Ganze aus den Daten überhaupt was Sinnvolles rauszukriegen und so ein Proof of Concept zu liefern ist tatsächlich die größere Hürde und das danach noch zu optimieren wäre dann eher so ein nachgelagerter Schritt. Wie ist das bei LLMs? Also ich meine, da seid ihr ja immer in der Frage, welches OpenAI-Modell nehmen wir denn jetzt für die Task?

I11

Ja, da sind wir tatsächlich, also bei dem, was ich bis jetzt so mitgemacht habe, ist es tatsächlich so gewesen, dass die Leute mal gesagt haben, auf keinen Fall OpenAI und da können wir noch gar nichts mit, weil Datenschutz und so. Deswegen habe ich mich jetzt eher viel damit beschäftigt, wie man das mit lokalen Modellen so abdecken kann. Da ist natürlich so eine Accuracy irgendwie zu fest zu nageln noch ein bisschen schwieriger, so in der Performance.

A1

vor allem gegeben der Kundensituation also man kann das vielleicht für irgendeinen Benchmark machen aber ob das dann bei den Kundendaten auch ähnlich ist das weiß man halt nicht

I11

Ja. Und das ist, wenn ich gar nicht drauf aus bin, irgendwie einen sprachlich schönen, grammatikalisch richtigen Satz zu haben und keinen Prosa-Text haben will, sondern irgendwas anderes, also das ist ein bisschen schwierig für also die Auswahl, welche externen großen Modelle, da haben wir uns ehrlich gesagt, also da ist dann immer irgendwas da, wenn dann da was da ist, weil die schon so ein GPT in ihrer Azure-Instanz haben, dann wird das erstmal benutzt. Also das ist relativ pragmatisch. Weil irgendwie, weil die, also die Modelle, die man einkaufen kann, die sind so gut, dass für das, was wir da machen, ist es fast ja, fast egal schon, was man da hat, weil die oft einfach für den sehr speziellen Task, den wir haben, eh gut genug sind.

A1

Ja, alles klar. Wir haben noch ein zweites Label tatsächlich mitgebracht, also genau die Idee so ein bisschen statt schwer verständlichen Papern oder Implementierungsdetails kann man irgendwie praktische Eigenschaften auch irgendwie auf einfach verständliche Art und Weise kommunizieren. Hier jetzt als Beispiel noch ein zweites Label. Können wir da vielleicht ein bisschen Vergleiche anstellen oder kannst du das machen?

I11

Ich sehe, ihr habt die Farben schön angepasst von Rot nach Grün. Also wenn ich jetzt eine grünere Accuracy, ein grüneres Icon sehe, dann sehe ich sofort, das hat eine bessere Accuracy. Und ich nehme an, oben ist dann irgendwie so ein gebildeter Durchschnitt der verschiedenen Eigenschaften. Ja, würde ich aber auch sagen, also macht Sinn und ich kann jetzt auf den ersten Blick, also ich würde auf den ersten Blick sehen, wenn ich jetzt weiß, ich brauche eben diese Accuracy oder ich weiß, ich habe hier irgendwie, keine Ahnung, Time-Probleme, weil irgendwas schnell laufen muss, dann würde ich da, ich würde sehen, was es tut.

A1

Ja, genau, also das ist so ein bisschen auch die Idee dafür, das so abstrakt darzustellen, weil ich meine, klar, wenn man die Ergebnisse in einem Pandas Data Frame oder so hat, dann kann man auch einfach basierend auf Spalten sortieren, aber auch gerade so tabellarische Darstellungen, wo man dann hinsichtlich eines Kriteriums irgendwie sortiert und sagt, wir wollen aber best accuracy, das ist ja immer ein Trade. Also du bist ja aus der Domäne auch, ich brauche dir das nicht sagen, dass ich irgendwie mehr Modellgenauigkeit in der Regel negativ tradet mit Modellgröße und am Ende muss man immer abwägen, was ist so die perfekte Lösung für unser Problem, gegeben auch der Priorisierung in dem Projekt sozusagen oder der Firma, wo man den Wert drauf setzt, auf welche Eigenschaften. Wie ist das für dich vom Verständnis? Gibt es hier Sachen auf den Labels, die für dich völlig unklar sind oder wo du jetzt sagst, das verstehe ich gar nicht, was das bedeutet oder wie das zustande kommt?

I11

Ähm Nö, eigentlich also klar kann man jetzt fragen, wie ihr die Accuracy bestimmt habt, was da für ein Algorithmus dahinter steckt, dass ihr auf die Zahlen kommt, so klar erstmal vom bei Robustness wäre jetzt die Frage, was ist 0% und was ist 100% so ansonsten macht das alles Sinn

A1

Ja, es ist wahrscheinlich bist du ein bisschen zu sehr auch schon in der methodischen Tiefe irgendwie unterwegs und vielleicht denkst du dir sogar, ich würde jetzt lieber die Implementierung sehen, wie habt ihr denn die Accuracy berechnet das würde mir noch mehr Aufschluss geben Ziel ist natürlich wie gesagt auch diese Brücke zu schlagen zu Leuten, die nicht so die Expertise haben, wenn du dich versuchst in die Rolle so ein bisschen zu versetzen hast du das Gefühl das ist was, was auch bei eurer Kommunikation mit dem Kunden helfen könnte

I11

Oh, das bestimmt. Die stehen total auf Bilder und Farben und wenn das... Also mit so Kategorien Rot, Gelb, Grün kann man Leute immer abholen, meiner Erfahrung nach. Also das bestimmt. Es ist natürlich so, dass sich unsere Kunden, die meisten davon, wir haben ja auch manchmal mit Menschen zu tun, die sehr, sehr auf unserer Augenhöhe sind, mit denen kann man sowas bestimmen oder drüber sprechen. Bei einigen ist es natürlich so, dass die noch nicht mal das interessiert, so ungefähr. Die wollen wissen, was da rauskommt und was das in Euro Mehrwert bringt. Aber klar, mit Menschen auf der anderen Seite, die vielleicht ein bisschen in diesen Datenthemen drin sind, aber da nicht so tief drin, da hilft das bestimmt, wenn man eben an dem Punkt ist, dass man an der Stelle optimieren muss und dass man sich da Gedanken machen muss, was man da benutzt.

A1

Ja das ist eigentlich ganz schön. Bei Management bräuchte man eigentlich unten nur ein rotes Dollar für wie teuer ist das, das Modell laufen zu lassen und ein grünes Dollar für wie viel Geld können wir damit verdienen? Das ist der Trade, an dem die interessiert sind, nachvollziehbarer Weise. Cost and Benefits, so. Ja, genau, also vielleicht kurz als Hintergrund noch, also die Label tatsächlich, jetzt nicht 1 zu 1 die Label, aber sehr ähnliche Label, ich habe da eine Software für geschrieben, um die vollautomatisch aus einem Pandas DataFrame zu generieren. Das heißt, falls du das Gefühl hast, ihr könnt das bei euch irgendwie sinnvoll einsetzen, dann kannst du mir gerne auch nochmal eine E-Mail schreiben, vielleicht kann ich euch da in dem Github irgendwie was in der Art und Weise bereitstellen.

I11

Dann mache ich das.

A1

Das nur so als Side-Note. Hast du irgendwelche Ideen, was man an dem Label vielleicht sinnvoll anders gestalten könnte oder sollte? Wir machen jetzt hier keine User Experience Study im Fokus, aber vielleicht sind da trotzdem Learnings, die man irgendwie so ein bisschen diskutieren kann.

I11

Also wenn man, man könnte jetzt vielleicht noch hingehen, also man könnte jetzt vielleicht noch hingehen, diese Werte oder diese Prozentzahlen irgendwie noch so in so Balken, dass man irgendwie die Zahl noch ein bisschen anders sieht als jetzt nur in der Farbcodierung, dann ist aber glaube ich wieder mein Blick als Mensch, der sich jeden Tag mit sowas befasst. Das ist schwierig, also für Leute, die da, mir hilft das, ob ich sehe, ob ich da 60 oder 40 Prozent habe, wenn ich irgendwie einen Balken oder irgendwas habe, wo keine Ahnung, 0 bis 100 und ich so sehe, weil ich das, das ist meine Art zu arbeiten, aber Menschen, die da nicht so viel draus ziehen können, denen hilft natürlich einfach diese Farbcodierung. Sehr viel mehr und daher das das Icon für Robustness finde ich sehr schön. Da muss ich einen kurzen Moment überlegen, aber ich wüsste auch nicht, wie man das anders gut machen kann. Ansonsten finde ich das sehr gut. Nö, ansonsten.

A1

Ja, also... Ja, XXXXXXXXX.

A2

Hast du denn das Gefühl, dass jetzt für dich, wenn du damit jetzt arbeiten müsstest, sagen wir, du musst jetzt mit diesem KI-Label, anhand dieses KI-Labels eine Entscheidung treffen, welches System du auswählst oder so. Welche Informationen wären jetzt hier für dich wichtig, die vielleicht fehlen oder würdest du sagen, das ist jetzt das, was ich brauche? Hängt natürlich auch von der Aufgabenstellung, vom Kunden ab, das sehe ich auch, aber vielleicht kommt dir da was in den Sinn.

I11

Also ja Zeit und so Accuracy sie Geschichten ist irgendwie so das was ich wo ich irgendwie dann immer unterwegs bin weil das ist mal das erste wenn etwas lange dauert ist es doof [undeutlich] erstmal an. Die powerdraw Geschichte ist natürlich auch sehr interessant wenn man das ganze irgendwie größer skaliert was fehlt also was dann fehlt in der in der Arbeit wenn ich dann wirklich davor sitze sind halt so dinge die man fast nur durch ausprobieren dann irgendwie meistens rausfinden kann also wie gut das dann wirklich funktioniert auch den Problemen den daten die ich da genau habe wie einfach das zugänglich ist aus welchen Paketen das kommt wie gut diese Pakete dann gepflegt werden und so also dass das spielt ja alles damit rein wenn ich irgendwie auf GitHub gehe und da irgendwelche Projekte sehe die super Modelle trainiert haben irgendwie oder da so Architektur noch geschrieben haben wo ich aber sehe dass da vor vier Jahren das letzte Mal irgendwer mitgemacht hat so ungefähr. Solche Sachen gehen dann natürlich auch mit rein die man natürlich in so einem Label glaube ich schwer integrieren kann aber das sind noch so Zusatzinformationen die ich dann mit da einbeziehe

A1

Ja, das ist tatsächlich ein interessanter Punkt. Ich meine, wir haben ja das Datum oben stehen, wann das Label ausgestellt wurde. So wie das jetzt gerade funktioniert und implementiert ist, wird praktisch das Modell dann halt auch geladen und einmal durchgemessen. Das heißt, man hat zumindest da so ein bisschen eine Zeitinformation, wie aktuell ist das noch, ähnlich wie letzter Commit. Also das heißt natürlich nicht, dass das eins zu eins garantiert reproduzierbar ist. Das kann auch heute schon wieder deprecated sein. Aber ich kenne das ja bei mir auch. Ich gucke mir das Repo an und wenn ich sehe, okay, das hat irgendwie ein Star und ist irgendwie seit fünf Jahren nicht geupdatet worden, dann kann ich es eigentlich nicht mehr benutzen. Weil das einfach die Abwärtskompatibilität da kaum gegeben ist und die Technologie sich so schnell entwickelt. Das war tatsächlich auch mal ein Feedback, was ich in einem Paper Review bekommen habe, dass man auf jeden Fall so eine grobe Zeitinformation auf dem Label braucht.

I11

Ja, das macht auch Sinn. Also das ist auch was, wo ich dann stark nach gucke.

A1

Wenn du so nachguckst, wo Info, also wenn, es scheint ja jetzt dann irgendwie daran, nicht unbedingt im Alltag die ganze Zeit zu hapern, dass man jetzt die richtige Lösung entwickelt für das Problem, sondern es geht eher darum, überhaupt dieses Problem richtig zu verstehen und zu modellieren. Aber wenn du in der Situation wärst oder wenn du manchmal in der Situation bist, wo du solche Entscheidungen, was benutze ich jetzt hier am besten hast, wo informierst du dich da? Welche Ressourcen nutzt du da?

I11

Also erstmal frage ich weil die Menschen bei uns in der Firma, das sind doch alles Physiker oder viele davon das ist natürlich so die erste und dann wird ja, also wenn ich irgendwas habe, wo ich also letztens hatte ich zum Beispiel so einen Fall da wollte jemand so ein Time Series Forecast haben und da habe ich 2018 das letztes Mal mitgearbeitet so da sind ein paar Jahre vergangen, da ist natürlich irgendwie viel passiert da haben wir dann irgendwie erstmal von Grund auf wieder angefangen und gegoogelt, Hacker News ein bisschen gucken, was so ja, was so Sache ist und dann hat man aber dann auch nicht so viel Zeit, man kann nicht drei Tage recherchieren mit so Kundenprojekten, weil die wollen morgen das Ergebnis haben also muss man sich da so ein bisschen auf dem Bauchgefühl verlassen ehrlicherweise und dann guckt man ja, was dann irgendwie erstmal Sinn macht, dann probiert man das aus und wenn es beim ersten Ausprobieren noch sinnvoll aussieht dann nimmt man das erstmal und dann wartet man, bis es kaputt geht und wer sich beschwert und dann macht man weiter ist dann so ein bisschen der Flow

A1

Genau, leuchtet total ein. Wir haben tatsächlich mal so ein bisschen gesammelt, wer es für andere Informationsquellen man irgendwie nutzen könnte. Also klar, wissenschaftliche Publikationen, zum Beispiel auf Archive, aber auch irgendwo sonst. Jetzt gerade bei Hugging Face, LLM und so, da gibt es ja jetzt auch oft, dass die Implementierung und die Schnittstelle auch direkt so eine Art Modelcard hat. Das geht ursprünglich zurück auf Google. Dann gibt es so Benchmarks und Leaderboards, zum Beispiel auf Papers with Code, wo man so ein bisschen ranken kann, welches Modell ist auf welchem Datensatz wie gut. Dann gibt es natürlich irgendwie Blogs, Hacker News hast du ja gerade irgendwie selber schon genannt, wo man sich so ein bisschen so einen Überblick verschaffen kann, Library-Dokumentation. Und von IBM gibt es auch das Konzept der Factsheets, wobei das nur für IBM-interne Modelle auch so richtig genutzt wird. Und das hat jetzt keine Public Software. Hier jetzt Beispiele für MobileNet V3, wofür wir vorhin das Label gesehen hatten, für die Publikation, die dem zugrunde liegt, eine Modelcard, die Übersichtsseite auf Papers with Code, Blogposts zu dem Thema, die Dokumentation von Keras und ein Factsheet von IBM, auch wenn es jetzt nicht das passende für MobileNet ist. Wenn man jetzt die ganzen Dokumentationsformen irgendwie sieht in der Gegenüberstellung, wo siehst du da Vor- und Nachteile, auch in Bezug auf die Labels, die wir so ein bisschen uns zusammen gedraftet haben?

I11

naja, es ist immer der klassische Spread zwischen was einfach und simpel und auf einen Blick ersichtlich und muss ich mich tiefer eingraben. Also klar, wenn ich ein Label habe, wo ich mir, also das ich mir angucke und eine Entscheidung treffen kann, ist das natürlich super gut. Trotzdem würde ich wahrscheinlich immer noch irgendwie tiefer recherchieren, ob das jetzt das Problem so löst. Zumal, also das sind ja jetzt irgendwie viele so neuronale Netzsachen, vielleicht ist aber das neuronale Netz gar nicht so das, was ich brauche, um mein Problem da gerade zu lösen. Also da steckt ja irgendwie so ein bisschen mehr hinter und ob ich wirklich alle Modelle, die ich so benutzen kann, alle Ansätze, die ich so habe, mit den gleichen Parametern in das gleiche Label quetschen kann. Das würde ich jetzt, wenn wir jetzt, dann ist schon wieder ein Schritt weiter gedacht, wenn wir in so Richtung Vertrauen und wie interpretiere ich das Ganze gehen, würde ich das fast schon so ein bisschen in Frage stellen und am Ende dann doch wieder so ein bisschen recherchieren. Aber für den ersten Überblick, wenn ich weiß, dass dieses Label für eine Klasse an Lösungen, sage ich jetzt mal, irgendwie gebaut wurde und in dieser Klasse an Lösungen, nämlich, weiß ich nicht, neuronale Netze, die man so und so irgendwie evaluieren kann und dann da einen Stempel drauf drücken kann, wenn es dafür gut ist und ich weiß, dass ich so eine Lösung brauche, dann kann ich das angucken und dann sehe ich auf den ersten Blick, ob das was taugt so. Muss aber dann im Grunde nach dem Schritt sein, wo ich mich schon wirklich informiert habe und schon geguckt habe, was ich hier überhaupt verwenden will. Aber wenn man an dem Schritt ist, dann macht das auf jeden Fall Sinn und dann muss ich nicht noch mal alles lesen, weil ehrlicherweise man liest ja dann doch nicht noch mal alle Paper im Detail durch. Kann man gar nicht…

A1

Man guckt sich dann den einen Table an, wo die relevanten Informationen sind oder im Code-Repo guckt man sich das einmal das Getting Started an und gräbt jetzt aber nicht in der Tiefe nach, ob das alles wirklich das tut, was es angeblich tut. Das ist generell so ein Problem, was wir im Moment im Feld haben. Und was du vorhin gesagt hattest, ich glaube, das muss man auch ganz klar sagen, also diese Labels, die sind ja eigentlich gedacht für spezifische Modelle, die praktisch fest zusammengestellt wurden, vielleicht pre-trained sind. Wenn wir jetzt sagen, wir bauen für einen tabularen Datensatz mit 10.000 Samples und 10 Features ein komplett neues System auf, klar, da muss man im Endeffekt dann entweder über AutoML oder händische Überparameteroptimierung erstmal 20 Pipelines, vielleicht 200 Pipelines ausprobieren, dann macht man im Endeffekt ein Ranking und dann trifft man seine Entscheidung. Und wenn man sich 2 Modelle oder 200 verschiedene Modelle anschaut, dann kann man sich auch nicht die Labels anschauen, dann ist es nicht übersichtlich. Da hilft dann die Tabelle tatsächlich mehr. Aber ja, es ist ja auch keine Repräsentationsform, die jetzt die anderen Möglichkeiten ablösen soll, sondern eher ergänzend hinzukommt. Weil wir so ein bisschen das Gefühl haben, diese ganz abstrakte Darstellung, das fehlt einfach irgendwie. So ein bisschen. Gerade wenn wir jetzt gucken, dass halt irgendwie immer mehr Leute auch KI einsetzen wollen, die eben nicht diese Domänen-Expertise haben und vielleicht selber gar nicht im Detail sich mit Libraries auskennen und trotzdem schnell eine Entscheidung treffen wollen, welches Modell passt hier vielleicht. Für die ist das vielleicht eine hilfreiche Brücke.

I11

Ja, das auf jeden Fall. Also wenn man so gar kein Gefühl dafür hat, was vielleicht auch so gerade so bei diesen klassischen neuronalen Netzen, wenn man da schon mal selber mit rumgespielt hat und weiß, wenn ich da noch ein Layer und noch ein Layer und noch ein Layer, dann wird das irgendwie, dann dauert das länger. Wenn man diese Erfahrung vielleicht noch nie gemacht hat, dann ist das sicherlich, dann hilft das auf jeden Fall.

A1

Okay, sehr cool. Jetzt kam ja vorhin schon das Thema Vertrauenswürdigkeit auf und wenn ich das richtig verstanden habe, korrigiere mich, wenn ich falsch liege, hast du gesagt, wenn man jetzt nur diese abstrakte Darstellungsform hat, dann könnte man auch zu Recht irgendwie kritisch anmerken, woher weiß ich denn, dass das tatsächlich auch Ground Truth ist, also dass das tatsächlich auch stimmt, was auf dem Label zu sehen ist.

I11

Ja, auf mein Problem bezogen. Also ich würde jetzt gar nicht mal das irgendwie implizieren, dass da jemand böswillig was Falsches in das Label schreibt, um Gottes Willen. Aber so dieses, dass das, was da ja gelabelt ist, wirklich auch zu meinem Problem passt. Ja gut, das ist dann ein bisschen der...

A1

Ja gut, das ist dann ein bisschen das Problem von Transfer Learning. Ich meine, klar, die Accuracy hier ist halt die Accuracy auf den Testdaten und ob das dann auf dem konkreten Datensatz der Firma auch hält, das muss man ausprobieren, indem man es durchmisst. Kann man ja aber in der Regel dann auch, es sei denn, man ist so ganz im unsupervised und hat gar keine Label, dann wird es schwierig. Fällt dir trotzdem im Kontext Vertrauenswürdigkeit irgendwie was ein, wo du das Gefühl hast, das wird das Label vertrauensvoller machen, gerade auch hinsichtlich, wer so ein Label vielleicht ausstellt? Also wo kommt das Label her?

I11

Also für mich persönlich, wenn ich irgendwie weiß, wer und was dahinter steckt, also ja, man würde immer wissen wollen, wer und was dahintersteckt, wenn da jetzt natürlich so ein Uni-Stempel draufsteht oder so, dann wirkt das natürlich immer direkt viel vertrauenswürdiger.

A1

Glaubst du, es würde einer Firma auch so gehen? Dass ein Unistempel sehr vertrauenswürdig wirkt? Also ich meine, du kommst von der Uni, du weißt, wie das hier läuft. Das ist alles Rapid Prototyping. Ja, das ist…

I11

Ja, das ist mir klar. Zweifelsfalle, also ja, natürlich weiß ich, wie das läuft und ich weiß auch, wie naja, egal. Ich weiß gar nicht mal, ob so eine Firma das jetzt, also wenn da so ein, ich sag jetzt mal, wenn da ein XX XXXXXXXX-Stempel drauf ist, ich weiß nicht, ob das jetzt eine Firma in Leipzig juckt, ob da ein XX XXXXXXXX-Stempel drauf ist, so ungefähr. Ja, keine Ahnung. Schwierig.

A1

Also wir fragen uns halt auch so, wer wäre denn so eine Instanz, die sinnvoll so ein Label bereitstellen könnte, weil wir erleben ja gerade eh auch so ein Misstrauen irgendwie, sei es jetzt in Politik oder sei es jetzt in Firmen, Thema Impfskepsis und und, das sind natürlich alles riesige Probleme, die wir als Machine Learner gar nicht lösen können. Das geht dann eher wieder auch stark Richtung Social Sciences und so, die da irgendwie, die da wahrscheinlich ein bisschen mehr Fachexpertise einfach haben, um sich mit so einem Thema zu beschäftigen. Aber wir fragen uns halt wirklich so, wer könnte so ein Label ausstellen, sodass Leute sagen, okay, jetzt vertraue ich der KI, weil oft vertrauen Leute KI erstmal nicht so.

I11

Ja. Das sind ja nur zwei Fragestellungen. Also ein Vertrauen in KI generell und ein Vertrauen in ein Label, was die Performance des Modells angeht. Das sind zwei verschiedene Sachen oder zwei verschiedene Vertrauensprobleme, sag ich jetzt mal. Ja, technisch, keine Ahnung. Weil wenn jetzt natürlich irgendwie so ein, also für den Techie jetzt so ein Label irgendeiner großen Firma, I don't know, Microsoft, Google, irgendwie sowas, wenn das da draufsteht. Aber dann wird jeder, der so komplett im Open Source unterwegs ist, der denkt dann schon wieder, wenn das so Microsoft ist, dann kann das nicht gut sein. Da eine Instanz zu finden, die alle total glücklich macht, wird man wahrscheinlich auch nicht schaffen. Vielleicht muss es sich auch als Marke, als Eigenmarke behaupten. Keine Ahnung

A1

ja also wir haben da ja auch keine Antwort darauf das ist deswegen holen wir ja nur Meinung ein was so relevant ist.

A2

Du hast es tatsächlich auch gerade ganz schön beschrieben. Also das sehe ich so bei mir. Das ist eine eigene Vertrauensfrage, Verschiebung. Also erst wird gefragt, kann ich der KI vertrauen? Und wenn man das nicht richtig beantworten kann, dann fragt man, kann ich dem Label vertrauen? Also das nennen wir dann so eine Second-Wave-Trust-Assessment. Und das dann verschiebt man eigentlich nur. Also es geht wieder, man steht irgendwie genauso ahnungslos da und muss dann halt nochmal neu schauen, ja, kann ich jetzt hier vertrauen? Also das ist ganz gut zusammengefasst.

A1

aber an sich hast du das Gefühl dass so Labels irgendwie eine Zertifizierung irgendwie also hilfreich sein können um so KI Produkte zu zertifizieren gerade auch hinsichtlich AI-Act und so da passiert ja gerade auch viel dass wir da auch eine Haftbarkeit zukünftig haben werden für KI Lösungen und Firmen ihre Prozesse zumindest offen legen müssen wie sie so KI Produkte zusammengestellt und haben und einsetzen wollen

I11

Ja, ich weiß nicht, also gerade jetzt in diesem, ich stecke da nicht im Detail drin, aber also in diesen Thematiken, wenn eine Firma offenlegen muss, wie so eine KI zu einer Entscheidung gekommen ist, wie das aufgebaut ist, fürchte ich, dass es so komplex ist, dass jetzt so ein Label, was Aussagen über Accuracy und, und Power Draw aussagt, dass das gar nicht ausreicht irgendwie, um solche Fragen dann irgendwie zu beantworten. Also wenn ich sage, ich vertraue der KI nicht oder ich vertraue der Antwort nicht, die da rauskommt, dann interessiert mich nicht, wie lange, also wie langsam das ist oder wie robust das ist so ungefähr wahrscheinlich. Also ich glaube, das ist so ein bisschen anders. Ich weiß nicht, ob das Label...

A2

sondern was würde dich dann interessieren?

I11

Ja, also mich als technischer Mensch wird da natürlich irgendwie sehr viel mehr interessieren, wie das aufgebaut ist, was da auch noch dran gebaut ist. Also das, was wir ja auch noch haben, so als zusätzlichen Komplexitäts-Layer, ist ja nicht, dass die Antwort, die jetzt so ein KI-Modell in irgendeiner Weise gibt, so pure so genommen wird, sondern da wird ja oft noch irgendwas drumherum gebaut und dann laufen da mehrere Schleifen und irgendwas wird quasi da reingebaut, dass gewisse Antworten in irgendeiner Weise verhindert werden oder verstärkt werden oder so solche Sachen, die da so drumherum gebaut sind. Das sind alles so Sachen, die da irgendwie mit reinspielen, die man, glaube ich, in so einem Label gar nicht unbedingt abdecken kann. Und ob dann die harten Performance-Parameter helfen jetzt bei so einer Vertrauensfrage und hast du nicht gesehen, würde ich ein bisschen anzweifeln, wenn ich ehrlich bin.

A1

Ja, also es ist definitiv so, dass auch diese Form von Label, so wie wir sie jetzt hier haben, diese Zwecke natürlich nicht komplett erfüllt. Das ist halt die Frage.

I11

Kann auch gar nicht.

A1

Genau, so kann es gar nicht. Dafür ist auch das, also auf der einen Seite muss nachgewiesen werden nach dem AI-Act, dass der Schaffensprozess des Modells irgendwie nachvollziehbar, fair, transparent und alles ist. Da ist man dann aber auch ganz oft schon bei der Datenakquise von vornherein. Da geht es dann gar nicht um Modellierung, da geht es nur darum, wie sind die Daten zustande gekommen, wie wurde das assembled, wurde darauf geachtet, wo die Daten herkommen und so. Stand jetzt ist, die vier machen das halt einfach und zeigen dann hinterher ein Produkt, was selber Ergebnisse hat und dann ist allen egal, wo die Daten hergekommen sind. Und das geht jetzt so langsam los, dass sich da auch irgendwie Rechtsanwälte hinterklemmen, Leute verklagen, weil irgendwo die Trainingsdaten rekonstruierbar sind und solche Sachen, Privacy Issues und Co. Ja, XXXXX.

A2

Ich muss einmal noch eine Verständnisnachfrage. Also wir haben ja auch mit dir, oder wir haben gefragt, was für Informationen brauchst du damit, dass, und da kamen wir gar nicht so in das Gespräch darüber, dass du eigentlich noch diese Informationen auch bräuchtest, um Vertrauen zu nehmen oder das zu nutzen. Aber da haben wir auch nicht explizit das Wort Vertrauen benutzt. Da haben wir über, glaube ich, Benutzung oder Entscheidungsprozesse gesprochen. Ich bin jetzt einfach, ich bin es sehr, ich bin ein bisschen überrascht, dass es jetzt um diese Informationen geht. Und ich frage mich, liegt es daran, dass du es nicht vorgeschlagen hast, weil du weißt, dass es technisch einfach so unfassbar schwer ist, das umzusetzen? Oder hast du es in dem Kontext nicht auf dem Schirm, weil wir da nicht über Vertrauen gesprochen haben?

I11

Meinst du jetzt, als ich mich gefragt habe, was ich brauche, um die Entscheidung zu treffen,

A2

Ja, genau, ja.

I11

Da hatte ich jetzt gerade einen ganz anderen Prozess im Kopf. Ob ich jetzt wirklich als Entwickler davor stehe und in einem ganzen Konstrukt aus Prozessen, Kundenwünschen etc. mich da technisch für ein Modell entscheide, dann sind diese Dinge, die da gelabelt sind, wichtig. Wenn ich als Nutzer ein Produkt benutze, wo irgendwie KI draufsteht, dann ist ja wichtiger, was der Entwickler, der das entwickelt hat oder die ganzen Menschen und die Daten, mit denen das trainiert worden sind, was da drumherum gebaut ist, wie das zustande gekommen ist. Das waren gerade in meinem Kopf zwei völlig getrennte Prozesse. Deswegen habe ich jetzt mit Vertrauen in dem Sinne gar kein Problem, wenn ich selber der Mensch bin, der entscheidet, was für ein Modell ich da jetzt benutze. Zumal ich ja, also wenn ich jetzt darüber rede, dass ich mir selber mit TensorFlow ein Modell zusammenbastele, was mir, ich weiß nicht, Zeitreihenvorhersagen macht oder Bilder klassifiziert, dann habe ich da kein Vertrauensproblem. Wenn ich ein Chatbot schreibe und ein fertiges Modell benutze und mich entscheiden muss zwischen einer offenen Sache oder OpenAI oder ich weiß nicht, was gerade bei Amazon läuft oder Google Modelle oder so, da muss ich dann natürlich sehr stark auch gucken, in welchem Kontext ich das benutze, ob... also wie sehr ich den Antworten vertrauen kann, die da rauskommen und was ich da dann noch drauf baue und dann vertraue ich dem Ganzen natürlich erstmal nicht. Ich würde jetzt niemals den Antworten, die ChatGPT gibt, zu 100% vertrauen. Also das sind so verschiedene Aspekte irgendwie von dem Ganzen.

A2

Ich versuche das mal gerade zu paraphrasieren. Also es ist ein bisschen kontextabhängig, welche Informationen wichtig sind. Also wenn es irgendwie so anwendungsspezifisch ist, nee, nicht anwendungsspezifisch, weil Chat-GPT ist ja auch eine sehr anwendungsspezifische Aufgabe. Also kontextabhängig.

I11

Ja, das eine ist auch vielleicht so ein bisschen in der Rolle, in der ich mir das gerade angucke. Also bin ich da die Entwicklerin, die da vorsitzt und ich muss nur Daten klassifizieren und ich habe den Prozess sehr gut in der Hand, weil weiß ich nicht… ich sage mal, ich baue einen Random Forest, das ist ja auch ein Modell, der mir sagt, ob da auf dem Bild ein Apfel oder eine Birne ist oder so, keine Ahnung. Dann habe ich da kein Vertrauensproblem, weil dann kann ich das Modell selber bauen, dann habe ich da keinen Stress mit. Wenn ich ein Chatbot schreiben soll, der für eine Firma irgendwie die Betriebsanleitung einer riesengroßen Maschine kennt und ich will ein Chatbot schreiben, wo sich der Operator mit dem Ding unterhalten kann und mir sagt, hier, ich habe das und das Problem, welche Schraube muss ich anziehen? Und ChatGPT sagt mir hier Schraube 43 und Schraube 43 bringt aber die komplette Maschine zum Abstürzen so ungefähr. Dann ist mein Vertrauen in das Modell natürlich viel, viel geringer, da muss ich noch viel mehr drumherum bauen, um diese Antwort irgendwie zu verifizieren und zu gucken, das ist eine andere technische Frage irgendwie. Ich glaube tatsächlich…

A1

Dürfte ich nochmal… Entschuldigung, sprich erstmal weiter.

I11

Ne, als drittes wollte ich noch sagen, also diese Geschichte, wenn ich wirklich als reiner Nutzer von so einem KI-Ding, dann habe ich natürlich am allerwenigsten Vertrauen, weil ich ja gar nicht weiß, ich kenne das Grundmodell nicht und ich weiß nicht, was noch drumherum gebaut wurde und so. Das sind so die drei Rollen, aus denen ich da auch das Problem gucke.

A1

Ich glaube, das ist auch so ein bisschen der Punkt, klar, als Entwickler oder Entwicklerin hat man natürlich die Möglichkeit, dann auch bei jeder Entscheidung das Modell zu debuggen, rauszufinden, wo ist das Problem und so. Aber es geht ja gerade darum, auch das ist am Ende ja ein Modell, was Leuten bereitgestellt wird in Form einer Software, die die nutzen und vielleicht ihr misstrauen. Und da ist ja die Frage, ob so ein Label nicht vielleicht da die Brücke schlagen könnte zu hier, ich habe das entwickelt, ich habe die und die wichtigen Eigenschaften rausgefunden, die stelle ich euch jetzt bereit, damit ihr ein bisschen Transparenz habt, wie sich dieses Modell verhält und ihr auch seht, okay, ich habe hier irgendwie mir verschiedene Aspekte auch angeschaut und so, um halt dieses Vertrauen dann bei der Zielgruppe, die hinter diesem Entwicklungsprozess irgendwie oder die in diesem Entwicklungsprozess irgendwie im Fokus steht, um die irgendwie abzuholen. Und zu informieren. Reporting, ja. Cool. XXXXXXXXX, hast du noch Rückfragen dazu?

A2

Nee, ich habe keine Rückfragen, sorry.

A1

Perfekt, dann sind wir nämlich tatsächlich auch am Ende.

A2

Aber XXXX, wolltest du jetzt noch was sagen?

I11

Nee, alles gut erstmal soweit.

A2

Erstmal.

I11

Dazu muss ich mir erstmal sehr viele Gedanken zu machen. Weil ich finde das Grundsätzlich sehr sehr spannend. Das ist auf jeden Fall ein cooles Thema und auch ein sehr, sehr wichtiges Thema. So grundsätzlich.

A1

Ja, vielen lieben Dank erst mal für die Zeit und auch für das Feedback. Wir sind tatsächlich jetzt mit dem inhaltlichen Teil der Studie durch, können die Aufnahme entsprechend auch stoppen.