KI-AGENTEN REALISTISCH BEWERTET: ÜBER TRENDS, USE CASES UND GRENZEN

Ausgesprochen digital. Der Podcast für digitale Trends.

Intro

[00:00:00.680] - Steffen Wenzel
Hallo und herzlich willkommen zu einer neuen Folge von Ausgesprochen Digital. Wir beschäftigen uns heute mit dem Thema Agentic AI, also dem sogenannten KI-Agenten. Mein Name ist Steffen Wenzel und zusammen mit Stefanie Liße moderiere ich diesen Podcast. Stefanie, wie geht es dir heute?
[00:00:26.950] - Stefanie Liße
Mir geht es super, Steffen. Wie soll es denn anders sein?
[00:00:29.530] - Steffen Wenzel
Das, glaube ich dir. Auf jeden Fall. Was hast du für einen Zugang zu diesem Thema KI-Agenten? Was denkst du, wenn du das Wort das erste Mal hörst?
[00:00:39.840] - Stefanie Liße
Also Steffen, ich komme natürlich vorbereitet zu unseren Podcast-Folgen. Ich habe mich schon ein bisschen damit auseinandergesetzt und der Zufall, wollte ich, ist, dass wir letzte Woche mit unseren Sales Kollegen zusammensaßen. Ich habe da auch schon mal ein bisschen reingefragt, inwieweit das momentan auch schon bei Kunden adressiert wird, was Sie darüber wissen und freue mich total auf diese Folge heute, weil ich glaube, das war so das einschlägige Feedback, dass da noch ganz viel Verwirrung da ist, ganz viel Fragezeichen im Kopf da sind. Ich selber auch Fragezeichen im Kopf noch habe, aber deswegen freue ich mich ja, dass wir heute einen ganz tollen Gast haben, der ganz viele Fragezeichen in Ausrufezeichen hoffentlich verwandelt.
[00:01:15.960] - Steffen Wenzel
Genau, und das ist Martin Wunderwald. Den kennen wir schon hier, der war nämlich schon mal zu Gast bei dem KI-Avatare-Podcast. Da haben wir uns das mal angeschaut, was KI-Avatare so im Alltag machen. Und jetzt reden wir mit ihm heute über KI-Agenten. Martin ist Principal AI Consultant bei der MMS. Hallo und herzlich willkommen Martin. Hallo. Ja, Martin, hilf uns mal ein bisschen auf die Sprünge. Also was sind KI-Agenten auch im Vergleich zu den ganzen anderen Begrifflichkeiten, die da so herumschwirren?

Definition KI-Agenten

[00:01:47.280] - Martin Wunderwald
Wow, das ist schon eine ziemlich große Frage gleich am Anfang, weil die Definition, die ist wirklich sehr vielfältig und hat viele verschiedene Perspektiven. Aber im Grunde kann man sagen, KI-Agenten erweitern die klassische KI, um die Eigenschaft selbstständig Entscheidungen zu treffen. Also sie denken in Anführungsstrichen selbst, entscheiden, welche Werkzeuge sie nutzen, um ein Problem zu lösen.
[00:02:12.100] - Stefanie Liße
Jetzt sagst du, die können alleine denken? Wo haben sie denn aber das Wissen her, das muss ihm doch jemand beigebracht haben, oder ist da schon der Unterschied?
[00:02:19.780] - Martin Wunderwald
Nein, das ist bei klassischen Sprachmodellen auch jetzt schon der Fall, dass wir Wissen quasi im Modell eintrainiert haben, aber natürlich auch Wissen von extern dazukommen kann. Und so ist es auch bei KI-Agenten. Sie nutzen also diese uns schon bekannten Sprachmodelle, LLMs, um einfach diesen Denkprozess zu vollführen mit dem Wissen. Nur was Neues: Sie haben jetzt die Möglichkeit, Werkzeuge zu benutzen, Schnittstellen zu bedienen et cetera, und das eben selber zu steuern, wie sie das tun, um zum Ziel zu kommen.
[00:02:51.100] - Steffen Wenzel
Kannst du uns mal ein Beispiel bringen, welche Werkzeuge das sein könnten jetzt in einem normalen Prozess?
[00:02:55.620] - Martin Wunderwald

Das können fast alle Werkzeuge, Software-Tools, ganze Computer-Bildschirme können das sein, die gesteuert werden durch Agenten, bis hin zu, wie wir es auch kennen, APIs, die bedient werden, Chat-Programme, E-Mail-Programme, Kalender, je nachdem, was ich automatisieren will oder wo ich letzten Endes den Ablauf in die Hand von einem Agenten legen möchte, bekommt er das Set an Tools, was er braucht oder es braucht, um das Problem zu lösen.
[00:03:26.380] - Stefanie Lißе
Ich würde gerne noch mal einen Schritt zurück mit euch gehen, denn wir kommen ja ein Stück weit aus der Prozessautomatisierung. Ich glaube, das ist allen klar, was wir darunter verstehen. Dort gibt es vordefinierte Workflows, die wir bauen, damit der Prozess automatisiert werden kann. Dann haben wir im Kundengespräch auch schon Themen, wie das sind in letzter Zeit sehr viel Copilot-Anfragen, zum Beispiel, stattfinden und jetzt eine vielleicht sehr vereinfachte Frage. Ich hoffe, du siehst es mir nach. Aber ist denn jetzt zum Beispiel Copilot dann schon so ein KI-Agent oder ist es das noch nicht ganz?
[00:03:58.820] - Martin Wunderwald
Das ist das, wo die Definition so ein bisschen schwimmen. Ja, Copilot ist ja letzten Endes ein Produktname für etwas und bildet in dem Fall eine Assistentenfunktion ab. Und natürlich kann so ein Assistent, wenn er die Aufgabe bekommt, eine bestimmte Aufgabe zu lösen, auch autonom denken. Und traditionell würde er letzten Endes Antworten geben auf Fragen, die ich habe, auf Basis von Daten, auf die er zugreifen kann. Und jetzt kommt eben genau das dazu, dass er eben auch Aufgaben erledigen kann für mich. Also ich kann dann durch dieses Interface eines Assistenten, wie zum
Beispiel Copilot, sprechen oder schreiben und sagen: „Ich brauche einen wiederkehrenden Termin jeden Donnerstag um die und die Uhrzeit. Bitte lade die Leute aus der Gruppe aus meinem Adressbuch dazu ein und erledige das für mich. Und die Schritte, die dafür notwendig sind, die kann ich als Mensch natürlich im Kopf vorher, weiß ich, wie ich durch Programme durchgehe und beispielsweise mein Outlook-oder E-Mail-Programm eben bediene. Und hier ist es jetzt, wo die KI übernehmen kann und selbstständig weiß, wie sie beispielsweise Outlook bedient, um die Information zu haben, auch zu wissen: Welche Leute lade ich ein? Sind das alles Strategien und braucht man sehr viel Wissen, um das zu entscheiden?
[00:05:15.800] - Martin Wunderwald
Und das übernimmt eben der Agent in dem Sinn.
[00:05:18.920] - Steffen Wenzel
Das heißt aber, der Agent braucht auch Zugriff zu meinen sonstigen Tools und Daten, die ich auf meinem Rechner oder in meinem System habe.
[00:05:28.020] - Martin Wunderwald
Das stimmt, genau. Also je nach Kontext in dem sich der Agent bewegen soll, braucht er den entsprechenden Zugriff auf diese, zum Teil Unternehmensdaten, aber auch funktionalen Schnittstellen. Und damit geht natürlich auch einher, dass er diese Rechte hat, bestimmte Dinge zu verändern, zu lesen, zu schreiben. Alles, was wir schon von klassischer Software kennen.

Studie: Vertrauen in KI-Agenten

[00:05:48.340] - Steffen Wenzel
Die MMS hat dazu jetzt auch eine Studie herausgebracht, die heißt „Vertrauen in KI-Agenten". Wenn wir über das Wort Agent reden, denken wir im Deutschen ja oft an ein bisschen auch was Negatives, an ein Spion, und so weiter, während das im Englischen, glaube ich, ein bisschen neutraler ist, eher als Berater gesehen wird. Was hat diese Studie herausgefunden?
[00:06:07.900] - Martin Wunderwald
Ja, das war eine ganz interessante Studie, weil sie, das ist ein bisschen ein Novum, sie hat quasi zum gleichen Zeitpunkt zwei Seiten von Leuten befragt. Einmal potenziell über 1. 000 Nutzer von KI-Agenten, also quasi ein Durchschnitt von 18-Jährigen bis, ich glaube, ein paarundsechzig Jährigen. Und auf der anderen Seite aber auch IT Entscheider, also Leute, die sehr viel Erwartung an diese KIAgenten in der internen Wertschöpfung ihrer Unternehmen sehen oder haben. Und was da spannend war an der Stelle, ist erst mal pauschal sind es, glaube ich, 55% über beide Gruppen hinweg wissen, was KI-Agenten sind. Haben Sie schon mal gehört, können es einordnen. Wo es aber spannend wird, ist, dass also IT-Entscheider ganz andere Erwartungen haben an KI-Agenten. Sie sehen wirklich, ich glaube, mehr als 75% der Entscheider sehen, dass sie damit große Hebel in Bewegung setzen können, große Effizienzen bergen können, automatisieren können und von klassischen IT-Problemen quasi wegkommen durch künstliche Intelligenz, während auf der anderen Seite in der Nutzerbasis die Akzeptanz eher zurückhaltend ist und viele Vorbehalte dem Thema gegenüberstehen. Jetzt haben wir zum einen diese große

Erwartungshaltung der Industrien, der IT-Leiter, der auch auf C-Level natürlich, an diese neue Technologie, auf der anderen Seite aber eine große Zurückhaltung. Und dieses Gap wird spannend zu füllen in Zukunft.
[00:07:42.160] - Stefanie Liße
Die IT-Entscheider, was du gerade angesprochen hast aus der Studie, da wurde ja jetzt so meines Wissens nach nicht nach Branchen irgendwie selektiert. Kannst du uns dort vielleicht mitnehmen, was du sagen würdest, welche Branche prädestiniert ist, jetzt mit KI-Agenten sofort anzufangen oder wo du einfach denkst, das wird noch einen Moment dauern bei der und der Branche? Kann man da irgendwas schon abschätzen, deiner Meinung nach?
[00:08:05.620] - Martin Wunderwald
Das ist im Prinzip die gleiche Frage, wie zu fragen, wo in meinem Unternehmen, in welchen Wertschöpfungsprozessen könnten Agenten am ehesten helfen. Da hast du recht. Und ähnlich, wie wir das mit klassischen Sprachmodellen beantwortet haben, ist es hier analog dazu, dass natürlich diese ... Weil die Sprachmodelle klassische Hilfsprozesse in Unternehmen, so was wie E-Mails schicken, Kalendereinladungen, solche Paradigmen kennen, werden sie auch in diesen Hilfsprozessen, also in der Buchhaltung, im Sales, in Sachen, die sozusagen horizontal in allen Branchen sind, dort werden sie als erstes greifen und stark automatisieren können. Wo es jetzt ein bisschen diffiziler wird, ist in der Kernwertschöpfung der einzelnen Unternehmen, wenn wir jetzt über produzierendes Gewerbe sprechen. Natürlich wird so schnell kein KI-Agent einen Mitarbeiter an einer großen Maschine ersetzen und Entscheidungen treffen, wie eine Maschine für ein bestimmtes Produkt konfiguriert wird. Also der Weg dahin, da müssen wir uns noch ein bisschen gedulden, bis auch da die ankommen, aber dort ist auch dann wieder ein großes Potenzial.
[00:09:10.520] - Stefanie Liße
Diese Anwendungsfälle, die du gerade skizziert hast, soweit ich das jetzt korrekt verstanden habe, braucht es ja dafür, oder der KI-Agent braucht dafür einen gewissen Freiheitsgrat, sage ich jetzt mal, also eine gewisse Autonomie, damit wir das volle Potenzial sozusagen heben können. Wie schätzt du das ein? Wird das von null auf 100 stattfinden? Gibt es da eine Übergangsphase? Weil ich persönlich stelle mir das schwierig vor, da, ich sage mal, den Schalter herumzulegen und zu sagen: „Und jetzt vertrauen wir mal und das wird schon alles gut gehen." Aber spannend ist natürlich, wie das Unternehmen vielleicht bewerten für ihre Prozesse, die langwierig und manuell noch laufen.
[00:09:44.500] - Martin Wunderwald
Das ist natürlich eine wichtige Frage: Wie viel Autonomie gestattet man der KI und lässt also entsprechend auch Kontrolle los? Das ist ein großes Spannungsfeld eben genau hinsichtlich dieser mit mehr Agency, also mit mehr agentischen Verhalten und Autonomie, erreiche ich letzten Endes natürlich größere Abstraktionen an der Stelle, wo der Agent wirkt. Wenn ich ihn aber ein Stück weit einsperre und Grenzen setze, sind wir wieder stärker bei Workflow, Automatisierung und bei traditionellen Prozessen und haben dann natürlich nicht mehr die großen Potenziale, die wir uns dafür erhoffen. Jedoch ist der Weg dahin, muss eigentlich sein, klein anzufangen und uns entlang dieser uns bekannten Prozesse, die wir vielleicht schon mit so Workflow Automatisierungstools quasi standardisiert haben, wo wir wissen, welche Tools sind da, wir haben eine gewisse Governance, wie dürfen diese Tools benutzt werden, etc. Dort vielleicht Elemente schon rauszunehmen, die jetzt agentisch entschieden werden können. Also bestimmte Entscheidungen jetzt schon, Sprachmodellen zu überlassen und damit einzusteigen letzten Endes. Was das jetzt eben nicht auf einmal klappen wird, zu sagen: „Hier sind alle meine Unternehmensdaten und meine Programme. Du Agent, es gibt jetzt die eine Oberfläche und ich schreibe was rein und er macht alles von alleine. Dafür ist die Datenqualität ungeklärt, die Schnittstellendefinition schwierig. Wir haben auf allen Ebenen, wie wir generell bei KI-Anwendungen jetzt noch Herausforderungen haben, haben sie sie auch in dem agentischen Umfeld dann.
[00:11:26.840] - Stefanie Liße
Das war mein größtes Fragezeichen im Kopf, weil wenn wir überlegen, ein KI-Agent wäre jetzt zum Beispiel die Kollegin, die sonst in der Buchhaltung die Entgeltabrechnungen für unsere Mitarbeiter durchführt. Dann haben wir ja ganz genau diese Themen ja dann auf dem Tisch. Aber ja, danke.

Use Cases: Erste Einsatzfelder

[00:11:43.520] - Steffen Wenzel
Ja, lasst uns vielleicht aber diese konkreten Beispiele mal ein bisschen uns näher anschauen, weil sie es ja dann auch deutlicher machen, wie man diese Schritte in einem Unternehmen gehen kann. Du hast eben

schon mal ein Beispiel genannt: Terminfindung, eigenständig. Zu gucken, wer sind die richtigen Menschen für diesen Termin, die ich in meinem Unternehmen Unternehmen brauche? Wo sind welche Zeitslots? Der Agent kann das vielleicht selbstständig entscheiden. Ist ein sehr simples Beispiel, glaube ich. Es gibt aber auch andere Anwendungsfälle. Kannst du uns da noch mal mitnehmen?
[00:12:10.240] - Martin Wunderwald
Gerne. Wobei ich ein bisschen widerspreche. So simpel ist es nämlich gar nicht, Wer weiß denn, in welchen Abteilungen die richtigen Leute sitzen, ob sie Zeit haben, wie sie erreichbar sind? Also selbst wenn man kleine, scheinbar einfache Entscheidungen nimmt, die man sich selbst aus seinem eigenen Erfahrungsschatz holt, wenn man mal wirklich dahintergeht und sagt: „Welche Informationen brauche ich, um das zu entscheiden? Beispielsweise im Kalender, Termine zu planen. Da gehört jetzt nicht nur: Haben die Kollegen Zeit?, sondern da geht es auch darum: Habe ich vielleicht private Termine, die anders priorisiert sind, oder auch firmeninterne Termine, die man gegen einander abwägen muss. Man kann auch kleinste Probleme sehr komplex machen und das ist auch was, was wir in der Entwicklung dieser Agenten merken, dass der initiale Use Case und die Idee, wie ich so einen Agent baue, sehr schnell da ist. Aber in der Implementierung kommen wir erst an die tatsächlichen Stellschrauben und an die Herausforderungen und die Probleme, dass bestimmte Datenquellen fehlen, dass die Qualität nicht stimmt, dass Governance-Richtlinien nicht da sind, dass es Rollenrechte Systeme gibt, die verschiedenen Nutzerperspektiven und so weiter. Also es wird dann doch wieder ein großes Softwareentwicklungsprojekt, wo man sich mit dem Thema sehr auseinandersetzen muss.
[00:13:26.640] - Martin Wunderwald
Und diese Erwartungshaltung, ich setze jetzt einfach einen Agent davor und der kümmert sich schon, die hat sich jetzt bei uns in der Realität noch nicht bewahrheitet. Aber deine ursprüngliche Frage war ja, was gibt es für Use Cases, die auch komplexer sind? Ich denke, Agenten spielen insbesondere ihre Stärken aus, weil sie, haben wir ja schon gelernt, ermöglichen, mit natürlicher Sprache auch sehr komplexe Sachverhalte darzulegen nach außen. Das heißt, ich kann unten einen Agenten ein sehr kompliziertes und komplexes Problem lösen lassen, vielleicht einfach eine, weiß nicht, Architekturentscheidungen treffen oder eine heterogene Softwarelandschaft bedienen, wie auch immer, ohne dass ich nach vorne hin sehr viel Wissen brauche, weil natürlich der Agent sich in seiner Sprache und in seinem Niveau an das Gegenüber anpassen kann. Insofern denke ich, dass man auch mit einfacher Sprache sehr komplexe Systeme bedienen kann und der Agent durch seine Quasi-Intelligenz und das Zusammenziehen von Wissen dann den fehlenden Teil meiner Erfahrung kompensieren kann.
[00:14:31.430] - Steffen Wenzel
Reden wir da jetzt über Softwareentwicklung?
[00:14:33.400] - Martin Wunderwald
Zum Beispiel, ja. Das kann in allen Dimensionen oder fast in jedem Bereich relevant sein. Wir haben ja jetzt schon gemerkt oder wissen schon von den Sprachmodellen, dass der Kontext, den ich dem Modell mitgebe, also die Informationen, auf Basis der er entscheiden kann oder bestimmte Antworten gibt, sehr relevant ist. Es gibt inzwischen wie Perplexity oder ChatGPT-Suchmaschinen, die also auch schon externe Quellen anzapfen, je nachdem, was ich möchte. Und so muss auch der Inhalt zum einen eben nicht nur für die Entscheidungsfindung gut kuratiert sein, der in diesem Kontext für die Entscheidung stattfindet beim Agenten, sondern auch die Wahl des geeigneten Werkzeugs. Das kommt jetzt einfach dazu, dass er weiß, ein Wetterbericht oder so was, was ganz einfaches wieder haben möchte, wie entscheidet er sich, welchen Wetterdienstleister er anfragt, um eine gute Vorhersage zu haben, welche Informationen zieht er sich einfach. Und insofern müssen auch diese Agentensysteme sehr gut engineered sein. Und je mehr Verantwortung sie übernehmen, desto komplizierter wird es. Und in der Zukunft, da werden wir bestimmt später noch drauf zu sprechen kommen, wird es natürlich irgendwie Expertensysteme geben und Agenten interagieren auch untereinander miteinander.
[00:15:53.720] - Steffen Wenzel
Stefanie hat eben gerade Branchen angesprochen. Es gibt beispielsweise Multi-Agentensysteme in der Logistik. Die werden dort schon eingesetzt. Kannst du dazu uns ein bisschen was erzählen?
[00:16:04.420] - Martin Wunderwald
Ja, da haben wir ein System eingeführt, was beispielsweise eben autonom Frachtbriefe einscannt oder einfach Bilder von Frachtbriefen analysiert und entsprechend validiert, qualitätssichert mit internen Unternehmensdaten und dann auf Basis dessen wieder eine Fracht-oder Routenplanung durchführt, wo

wieder andere natürlich Agentensysteme damit betraut sind, dann zu schauen: Sind alle Bedingungen erfüllt? Habe ich die richtigen Leute an den richtigen Orten? Sind die Transportmittel, also die LKWs, verfügbar? Et cetera. Und somit sozusagen ein sehr komplexer Prozess im Wesentlichen mit vielen kleinen Agenten, die jeweils spezialisiert auf Teilaufgaben sind,
durchorchestriert ist. Aber auch in dem Fall entlang eines doch sequenziellen Workflows der Stelle.
[00:17:00.900] - Steffen Wenzel
Und die entscheiden dann über Routen beispielsweise, selbstständig?
[00:17:04.300] - Martin Wunderwald
Genau. Und insbesondere entscheiden sie: Wo hole ich Informationen? Oder „Welche Informationen muss ich holen, eine Entscheidung zu treffen im nächsten Schritt, beispielsweise.

Vorgehensweise

[00:17:15.780] - Steffen Wenzel
Du hast eben gesagt, KI-Agenten basieren ja natürlich auch auf Large-Language-Models. Dennoch die Frage: Wenn ich jetzt ein mittelständisches Unternehmen bin, wie soll ich denn anfangen? Ich kann ja nicht einfach mit den LLMs anfangen, die ich besitze, sondern ich brauche „Ich brauche ja mehr Wissen, ich brauche ja wahrscheinlich andere Software noch dazu, Implementation. Was ist da schon auf dem Markt? Wie gehe ich da vor?
[00:17:37.840] - Martin Wunderwald
Genau. Es gibt da eigentlich jetzt aktuell zwei Stoßrichtungen, die man gehen kann. Wenn wir davon ausgehen, wir haben einen Anbieter für Sprachmodelle, egal ob man die jetzt selber in seiner Firma quasi hostet, also souverän, oder ob man sie von einem Cloud-Provider bezieht, wie zum Beispiel OpenAI, hat man die Möglichkeit, zuzugreifen erst mal auf diese Sprachmodelle und sie zu nutzen in entweder Agentischen Workflows, das haben wir gerade schon besprochen, wo ich selber über NoCode-/Low-Code-Plattformen, also ohne selbst zu programmieren, Tools anbinden kann, die ich
verfügbar habe, Schnittstellen meiner Informationen alles einbringen kann, in Form von einem Graphen orchestrieren kann und dort an einzelnen Stellen dann diese Sprachmodelle für so eine Entscheidung hinzuhalten kann. Da entstehen zurzeit hunderttausende Workflows in verschiedensten Firmen. Ein namhafter Anbieter aus Deutschland ist N8N oder manchmal auch Nathan genannt, der eben so eine Plattform anbietet, die dann eben diese Sprachmodelle, die LLMs der Anbieter anbinden kann, aber auch eigene Quellen, wo also auch jemand ohne viel technisches Verständnis so einen Workflow relativ schnell bauen kann. Und die andere Facette, wo man jetzt auch hingeht, ist, Agenten selbst individuell zu implementieren für sehr spezifische Domänen. Wir haben beispielsweise ein Thema, wo wir Professionierung von Applikationen in Cloud Platform, was ein sehr technischer Prozess ist, versuchen über einen Chat-Interface zu abstrahieren und jemandem, der eigentlich keine Ahnung davon hat und nur sehr high-levelig Informationen hier reinstreuen kann, wo er seine App
laufen lassen will, wie viel Redundanz die haben soll und seine Anforderungen: Es kommen so und so viele User und dann dieses Konfigurieren, was dann unten passiert, im Sinne von: So und so viel User kommen, also muss ich die und die Maschine nehmen.
[00:19:36.800] - Martin Wunderwald
Die App hat die und die Eigenschaften, also brauche ich die und die Softwarefragmente und muss das so und so konfigurieren. Diese ganze Komplexität kann dann durch Agenten quasi weggenommen werden oder übergeben werden, diese Komplexität. Und das sind dann individuell entwickelte, spezialisierte Agentensysteme für sehr spitze Nischenprozente Prozesse, sage ich mal.
[00:20:01.360] - Stefanie Liße
Kann ich als Unternehmen, wenn ich jetzt mal angenommen, ich habe wirklich noch kein oder sehr, sehr wenige Prozesse heute automatisiert bei mir laufen, kann ich damit dann schon mit der Plattform anfangen oder ist das wie so ein Vorschritt, den man eigentlich machen muss, damit man dann überlegen kann, an welcher Stelle ist hier sinnvoll, noch einen KI-Agent nutzbar mit reinzubringen?
[00:20:21.800] - Martin Wunderwald
Das ist eine gute Frage, weil natürlich diese Einschätzung, wo ein Sprachmodell oder eine KI helfen kann in einem solchen Prozess, wirklich schwierig ist. Es gibt durchaus Kunden, die schießen mit Kanonen auf Spatzen. Sie tun sehr eigentlich einfache, regelbasierte Prozesse dann einer KI überlassen. Verlieren damit natürlich auch die Kontrolle und da braucht man dann schon Beratung und muss das ein bisschen austangieren. Aber ich würde empfehlen, dass man, zumindest wenn man sich mit Automatisierung beschäftigt, dass diese Tools, die man dafür auswählt, bereit sind, eben auch solche KI-basierten Knoten

in diese Entscheidungsketten zu unterstützen. Und zum Beispiel N8N oder Power Automate von Microsoft, die können das.
[00:21:11.840] - Stefanie Liße
Deswegen frage ich, weil auf dem Markt ganz viel, ich sage mal, sehr viele Anbieter momentan unterwegs sind, die alle irgendwie drauf schreiben: „Unser Tool kann bereits Agentic AI und löst für dich die Aufgabe." Ich hatte vorhin das Beispiel Entgeltabrechnung von mir aus, was ja schon abgefahren ist, wenn man darüber genauer nachdenkt. Und ich glaube, das ist nicht nur für mich, sondern wahrscheinlich auch für unsere Zuhörerschaft auch so ein Thema, wo man ein Fragezeichen bekommt: „Sind wir schon da? Kann ich damit heute schon anfangen? Ist mir geholfen als Unternehmen, wenn ich jetzt zum Beispiel so ein Tool mir zulege und dann sage: Und jetzt geht es aber los? Deswegen hatte ich vorhin auch mit der Autonomie in die Richtung gefragt: Was ist da so deine Einschätzung zum Thema diese ganzen Tools, die auf dem Markt krauchen und flauchen und Agentic AI da versprechen? Sind wir da schon oder hättest du Tipps und Tricks, woran man Dinge erkennen kann?
[00:22:09.140] - Martin Wunderwald
Also wenn man jetzt fragt: Wie kann ich quasi trojanische Pferde erkennen? Also Software, die noch mit Agentic gelabelt ist oder nicht, sollte man die Frage eigentlich noch mal ein Stück zurückstellen und wirklich fragen: Kaufe ich Software, weil KI drauf steht, also technologiebezogen, oder kaufe ich sie nutzenbezogen? Also ich würde die Entscheidung, ob ich sozusagen eine Software einkaufe, immer am wirklichen Wert, den sie dann für mich hat, festmachen. Und wenn ein System verspricht, dass es eben für mich ein bestimmtes Problem, zum Beispiel Customer Relationship Management, löst gut mit KI-Agenten, dann muss man schauen: Tut es das? In welcher Qualität tut es das? Und auch testen, ob insbesondere auch diese agentischen Unterstützungssysteme, die dann drin beworben werden, auch den Anspruch erfüllen, den ich wirklich habe. Ich dränge da sehr darauf, sich wirklich auch insbesondere dieses Chat-Verhalten für die einzelnen Anwendungsszenarien, die man dann mit den verschiedenen Softwaren hat, genau anzuschauen, ob es mir im Alltag auch was nützt, also in Testphasen. Und das lässt jeder zu inzwischen.
[00:23:24.200] - Stefanie Liße
Und würdest du die Frage noch mit reinnehmen, dass die Unternehmen sich auch fragen müssen, ob sie das wollen, weil das geht ja dann auch, wenn ich die Tools anschaffe, so habe ich es verstanden –,muss ich ja auch ein paar Zugänge dann geben.

Transparenz, Regulierung und Entscheidungswege

[00:23:38.880] - Martin Wunderwald
Das stimmt. Das bringt uns ja zu dem größten Thema, was quasi auch in der Studie gezeigt ist, also diese Befürchtung, dass die Qualität, die dann der Agent hat in der Ausführung und in der Entscheidung, die er betrifft, dass die minderwertig ist, dass die nicht kontrollierbar ist, dass die Entscheidungswege nicht transparent sind, dass es in der Grauzone von: Was darf der eigentlich? Da sind wir, glaube ich, noch am weitesten weg. Es gibt noch nicht das Transparency Governance Framework für KI-Agenten, wo ich sage, das installiert man, das nutzt man und ich kann jetzt in jede Agentenentscheidung nachvollziehbar, retrospektiv reinschauen. Ja, das fehlt, weil mit der Autonomie übergebe ich ja Verantwortung und die ist aktuell Blackbox.
[00:24:22.880] - Steffen Wenzel
Weißt du denn, was der EU AI Act dazu sagt? Weil die fordern ja dort genau diese Transparenz.
[00:24:29.040] - Martin Wunderwald
Also spezifisch auf die KI-Agenten weiß ich es nicht, aber wenn wir ihn allgemein anschauen, gelten alle Anforderungen, die an generelle KI-Systeme gelten, auch dafür. Und das heißt also, KI-Agenten fallen natürlich auch unter die ganzen Auflagen, was Governance, Datenschutz, Security angeht.
[00:24:47.600] - Steffen Wenzel
Also da sind wir noch nicht so weit, was die Transparenz von Entscheidungsprozessen insbesondere bedeutet und natürlich auch die Steuerbarkeit und Regulierbarkeit. Das heißt, es könnte ja sein, dass der Agent aus seinem Wissen heraus eine Entscheidung trifft, wo vielleicht eine kleine, minimale Entscheidung fehlt, was überhaupt nicht seine Schuld ist, weil er zum Beispiel einen gewissen Zugang zu manchen Daten nicht hat. Aber dann wäre es ja gut zu sagen: „Okay, ich verstehe, warum er diese Entscheidung getroffen hat und ich korrigiere die noch mal. Ist das denn möglich?
[00:25:19.800] - Martin Wunderwald

Super spannende Themen, die du da ansprichst. Zum einen diese Nachvollziehbarkeit und dieses unter die Haube zu schauen, wie so ein Agent funktioniert, ist nicht so trivial, aber genauso schwierig ist es, selber überhaupt den KI-Agenten so zu entwickeln, dass er eben gute Entscheidungen treffen kann. Man muss genau entscheiden, zu welchem Zeitpunkt kann er auf welche Informationen zugreifen, welche braucht er, eine gute Entscheidung zu treffen. Zu viel Information ist auch falsch. Die richtige Information zum richtigen Zeitpunkt in der Verarbeitungskette ist entscheidend. Deswegen gibt es jetzt auch neue Berufe. Es ist nicht mehr der Prompt Engineer, sondern der Context Engineer, ist jetzt gefragt als nächstes. Also es entwickeln sich ganz neue wieder Paradigmen und
Schwerpunkte innerhalb dieses, ich sage mal, KI-lifecycles durch diese Einführung von Agenten, weil eben dieser Kontext das Wichtigste ist. Wir haben ein beschränktes, wie bei Menschen eigentlich, wir haben eine beschränkte Kapazität der Aufnahme von Informationen und die müssen wir mit den richtigen und wichtigsten Daten quasi füllen, um die Entscheidung zu treffen. Und genauso verhalten sich diese Agenten auch. Sie müssen zum richtigen Zeitpunkt alles haben. Und da kommen wir genau dahin: Wie kann man das beobachten? Wie kann man das über die Zeit auch verbessern? Da sind wir noch in Lernphasen. Es gibt nicht den KI-Agent von der Stange.
[00:26:39.340] - Steffen Wenzel
Ist das auch eine Weiterentwicklung, wie wir Large Language Models bislang kennengelernt haben, nämlich so viel Information wie möglich reinzugeben? Es wurde ja immer gesagt, dieses LLM ist besonders groß und dementsprechend besonders mächtig. Und jetzt haben wir so eine Bewegung, eher bei KI-Agenten hin zu sagen, es kommt gar nicht auf die Masse an Informationen sondern auf die besondere Richtigkeit, die besondere Qualität von Informationen.
[00:27:06.280] - Martin Wunderwald
Wir dürfen zwei Dinge nicht vermischen. Einmal das Sprachmodell selber, was ja quasi in sich selbst die Information schon mitbringt, also quasi einfach ein Modell mit dem Redaktionsschluss, was wir schon besprochen haben, und dann das, was man quasi dem Modell übergibt, während man die Frage stellt. Und das ist ja der sogenannte Kontext. Und diesen Kontext reichert man ja üblicherweise auch mit Teilen der Lösung an, also mit dem relevanten Wissen, um eine Frage zu beantworten. Und diese Passung, das ist das Zentrale. Das haben wir aber auch bei klassischen Sprachmodellen und in
klassischen Chats. Also wenn ich tausende Dokumente habe in meiner Dokumentenbasis sollen ja die richtigen kommen, meine Frage zu beantworten. Also wird sich erst meine Frage angeguckt, dann guckt, welches Dokument passt am meisten und das wird dann mit dazu genommen. Und so ist auch bei den Agenten.
[00:28:01.200] - Steffen Wenzel
Ich würde noch mal ganz gerne wissen, weil du da jetzt ja auch eben gesagt hast, das ist noch nicht so ganz deutlich mit der Transparenz und der Nachvollziehbarkeit von Entscheidungen. Was sagst du denn in einem Unternehmen, was jetzt auf dich zukommt und sagt: „Ich würde das gerne ausprobieren. Ich bin bereit, weil ich sehe, wie mir das helfen kann, aber ich habe diese Befürchtung?"
[00:28:20.620] - Martin Wunderwald
Dem sage ich, man muss sich ein Stück weit diesem Risiko beugen, weil das ist auch ein Punkt, den du vorhin schon adressiert hattest. KI ist ja in sich, habe ich auch schon mal gesagt, ist nicht deterministisch. Also die Entscheidung kann man schwer nachvollziehen. Sie ist nicht entlang eines Pfades zu treffen mit klaren Regeln, wie wenn ich starr eine Software entwickle, sondern sobald ich ein Sprachmodell oder KI reinnehme, haben wir immer einen Ansatz von Varianz drin, von nicht Vorhersagbarem, wie der Mensch selber ja auch ist. Nicht jeder Mensch wird dieselbe Antwort geben auf die gleiche Frage. So ist es auch bei der KI. Sie wird sich immer anders verhalten, je nachdem, in welchem Kontext sie eben agiert. Und so ist es, dass also auch ein Unternehmer hier das Risiko eingehen muss, dass er zum einen entweder sozusagen immer andere Antworten auf die gleiche Frage bekommt, die sich leicht unterscheiden, aber vielleicht im Grunde das Gleiche was wir meinen und auch auslösen, oder – und das ist eben, wo wir noch nicht so weit sind – sehr umfangreiche Testbatterien, also sehr große Benchmarks, aufzubauen, die dann solche Systeme gut testbar machen. Also, wenn man sich jetzt vorstellt, wir nehmen mal so einen agentischen Servicebot, der bei
einer Fluggesellschaft beispielsweise ist, der kann ja, wenn ich dort anrufe, kann Auskunft sein, es kann irgendwie: „Wo ist mein Gepäckstück?"
[00:29:42.580] - Martin Wunderwald
Es gibt ganz viele Szenarien, die da auflaufen. Und jetzt so ein Agent, der entscheidet, was er macht und dann etwas ausführt. Das heißt, ich habe eigentlich exponentiell viele Szenarien, die ich testen muss. Und

da der Agent jetzt natürlich autonom agiert, ist jede kleine Stellschraube, wird das Verhalten von dem Agenten verändern und das muss eben entsprechend auch dann testseitig abgedeckt sein. Was ist, wenn ich das Sprachmodell einfach austausche? Dafür brauche ich, wenn es wirklich kritischer, businesskritischer Prozess ist, wo ich so was einführen will, brauche ich die entsprechende Testabdeckung, um sicher zu sein, dass wenn ich das quasi veröffentliche, dass es funktioniert. Und was mich da immer wieder wundert, ist, dass diese Ansprüche an solche KI-Systeme massiv hoch sind, was die ausgebrachte Qualität angeht und das wird auch für Agenten so sein, was die Entscheidungsqualität angeht oder die Informationsgehalt et cetera. Wenn man in einem Callcenter anruft, wo die Menschen sitzen, da wird die Qualität nur über eine Metrik "bist du zufrieden oder nicht" geprüft. Aber wenn wir KI-Agenten an solchen Mensch-Computer-Schnittstellen einsetzen, da wird die Richtigkeit der Information, die Schnelligkeit, die Qualität, die Tiefe, ganz viele Dimensionen, will man ganz genau wissen. Ist doch verrückt, oder?
[00:31:03.480] - Steffen Wenzel
Ja, da können wir, glaube ich, jetzt tiefer reingehen. Ich glaube, das hat sehr viel damit zu tun, dass man Menschen Fehler eher verzeiht als einer Technologie. Richtig. Und ich finde das aber auch richtig so, dass man da einen anderen Maßstab anlegt.

Zukunftsausblick

[00:31:16.400] - Stefanie Liße
Martin, bei den ganzen Ausführungen, habe ich mir die ganze Zeit also, kam mir so ein Gedanke: Wo geht denn die Reise noch hin? Also wenn wir jetzt in drei Jahren mal gucken: Habe ich dann persönlich nur noch KI-Agenten um mich herum? Gibt es noch wirkliche Kollegen, mit denen ich interagiere? Also was ich dort wirklich auch von dir wissen will, wie deine persönliche Meinung vielleicht dazu auch ist: Was denkst du, inwieweit die Agentic AI oder auch die Nutzung von KI-Agenten Auswirkungen auf das Zusammenarbeiten zwischen uns Menschen vielleicht hat?
[00:31:46.860] - Martin Wunderwald
Also ich bin ja ziemlich tief drin in den Entwicklungen und auch an den Forschungsthemen, die sich schon um diese Szenarien drehen und ich muss sagen, wir müssen uns da wirklich warm anziehen. Also jetzt ist schon die, sage ich mal, kommerzielle Forschung, also in Unternehmen, wie die großen Tech-Konzerne, die ist schon viel schneller als unsere akademische Forschung, was das angeht. Das heißt also, hier ist eine massive Beschleunigung durch dieses ganze Geld, was in dieses Thema reingepumpt wird und jeder will der Erste sein, wird sich das massiv beschleunigen und massiv in den Markt drücken. Und die Frage ist, kommt man mit Regulierungen und Governance diesem technologischen Vorsprung überhaupt hinterher? Also jetzt wird schon geforscht wie sich Agentensysteme mit hunderten und tausenden Agenten, die miteinander kommunizieren, selbst orchestrieren und wie die sich miteinander absprechen. Was entstehen da für Dynamiken? Ähnlich wie bei Menschengruppen. All das ist schon Forschungsgegenstand heute und das sind aber
Szenarien, die wir jetzt natürlich noch nicht implementiert sehen. Aber es kann sein, dass ich also von diesen monolithischen prozessualen Gedanken, die wir jetzt haben, wenn wir über IT-Systeme oder Prozesse nachdenken, das sehr schnell auf intelligente, agentische Expertensysteme, die miteinander in Netzwerken zusammen kooperieren, verändert.
[00:33:09.480] - Steffen Wenzel
Glaubst du, dass die Regulierung da hinterherkommt?
[00:33:15.980] - Martin Wunderwald
Also aktuell wird sie in bestimmten Teilen der Welt gar nicht zugelassen, dahinterher zu kommen. Ich denke, dass das insbesondere für Deutschland und Europa ein großer Vorsprung ist, dass wir diese Regulierung können und das auch gut können. Aber persönlich ist meine Meinung, dass das natürlich nicht ausreicht. Und ich denke, es ist ähnlich wie mit – ich mache mal ein Beispiel, so früher mit dem iPhone. Wenn man einen Nutzen davon hat, dass man seine Daten teilt, ist man viel bereiter, die zu teilen. Und wenn ich jetzt, mein Beispiel, ich habe jetzt so ein agentisches Supertelefon wie das T-Phone und ich sage: „Du darfst mir den ganzen Tag zuhören und auch wissen, wo ich bin und auch meine Kamera oder meine Brille nutzen, die ich aufhabe. Wenn du mir dafür ganz viel Arbeit abnimmst und mich freihältst von sonstigen Sachen oder schwere Aufgaben für mich löst, dann ist es total okay für mich, dass du diese Daten nutzt. Das sieht man schon. Da kommt man in ganz komische Sphären, die jeder für sich selber beantworten muss, ob er später vielleicht mit auch Chips– das kennen wir auch, das klingt jetzt alles

noch so abwegig, aber so neuronale Chips als Input-Modalität, dass meine Gedanken, also dass quasi ein Sprachmodell schon antizipiert, was ich möchte
zu einem bestimmten Zeitpunkt am Tag in einer bestimmten Umgebung und für mich antizipiert, was es als Nächstes tut.
[00:34:46.560] - Martin Wunderwald
Und wenn ich nach Hause komme, ist der Ofen schon vorgeheizt, weil es gemerkt hat, ich habe Hunger auf ein bestimmtes Thema. Jetzt wird es ganz abgefahren. Ja, aber das ist diese komplette Vernetzung. Das sind die Themen, mit denen sich jetzt gerade schon Leute beschäftigen, wo es nicht mehr den Laptop gibt, sondern wo wir an irgendeine Display-Wand vorbeilaufen und das dann unsere Eingabe für ein ganz privates Thema sein kann. Aber jetzt schweife ich schon wieder ab. Ich bin in meinen Visionen gefangen.
[00:35:10.770] - Steffen Wenzel
Nein, du schweifst es gar nicht ab.
[00:35:12.190] - Martin Wunderwald
Ich finde es sehr spannend.
[00:35:12.810] - Steffen Wenzel
Du gibst uns, glaube ich, noch mal einen Ausblick in all das, über was wir hier noch reden wollen und werden mit dir. Martin, vielen Dank, dass du heute hier zu Gast warst mit diesem Thema und ich bin mir sicher, du wirst nochmals wiederkommen und wir werden über all das noch diskutieren können.
Vielen Dank, Martin.
[00:35:28.060] - Martin Wunderwald
Ich danke euch.
[00:35:28.820] - Stefanie Liße
Danke dir.

Check-out

[00:35:29.700] - Steffen Wenzel
Hast du jetzt deine Fragezeichen, die du am Anfang hattest, beantwortet bekommen von Martin heute oder hast du ein bisschen mehr Befürchtung? Wir haben ja auch ein bisschen über die ganzen Probleme gesprochen, Herausforderungen, die damit verbunden sind mit dem Thema. Wie geht es dir jetzt damit? Am Anfang ging es dir gut.
[00:35:45.100] - Stefanie Liße
Mir geht es auch immer noch gut, Steffen. Das ist gut. Doch, es sind einige Fragezeichen in Ausrufezeichen umgewandelt worden. Vor allen Dingen der Unterschied war noch mal gut, wie er es erklärt hat. Das hilft, glaube ich, auch. Der Ausblick ist Wahnsinn, wo die Reise hingeht. Du hattest den Punkt angebracht mit KI-Agent und das Wording: "Agent". Das habe ich dann auch so für mich noch mal überlegt. Ich sagte, ich freue mich auf KI-Kollegen, die Dinge für mich tun oder mir abnehmen. Aber KI-Agent, das Finde ich schon ... für mich persönlich ist das ein Begriff, da bin ich wieder bei „I,Robot" im Kopf.
[00:36:20.040] - Steffen Wenzel
Ja, ich glaube, das ist wirklich ein Riesenthema, das Wording, weil „Agent, hatte ich am Anfang auch gesagt, das assoziiert immer so ein bisschen was Negatives, Spion oder so was. Ich glaube, da brauchen wir noch einen besseren Begriff. Da müssen wir noch mal daran arbeiten, so KI-Buttler oder keine Ahnung was. Vielleicht finden wir da noch was Neues, aber das wird unsere Aufgabe sein.
[00:36:38.590] - Stefanie Liße
Ja, bloß gut, dass wir kreativ sind.
[00:36:41.200] - Steffen Wenzel
Und wie immer finden Sie in den shownotes natürlich noch weitere Informationen zu diesem Thema, insbesondere die verlinkte Studie, über die wir heute hier gesprochen haben und natürlich auch noch mal einen Link zu dem Podcast KI-Avatare mit Martin Wunderwald. Und ansonsten wünschen wir Ihnen einen wunderschönen Tag und danke, dass Sie uns zugehört haben.
[00:37:04.240] - Stefanie Liße
Tschüss.
[00:37:05.120] - Martin Wunderwald
Tschüss