

Dies ist das akzeptierte Manuskript (Author's Accepted Manuscript) des folgenden Beitrages: Kersting, M. (2025). Einfache Mittel, große Wirkung. Wie ChatGPT & Co. die Personalauswahl erleichtern und verbessern. *Forum Strafvollzug*, 2/25, 108-110.

Hinweis: Die Nutzer(innen) dürfen die Inhalte nur zum Zwecke der wissenschaftlichen Forschung ansehen, drucken, kopieren, herunterladen, sowie für Text- und Datamingen verwenden. Die Inhalte dürfen weder ganz noch teilweise wörtlich (wieder)veröffentlicht oder für kommerzielle Zwecke verwendet werden. Die Nutzer(innen) müssen sicherstellen, dass die Urheberpersönlichkeitsrechte des Autors sowie gegebenenfalls bestehende Rechte Dritter an den Inhalten oder Teilen der Inhalte nicht verletzt werden.

Martin Kersting

Einfache Mittel, große Wirkung

Wie ChatGPT & Co. die Personalauswahl erleichtern und verbessern

Der öffentliche Dienst, der Strafvollzug und die Straffälligenhilfe stehen vor einer großen Herausforderung: gute Leute finden und binden. Während immer mehr Menschen aus den geburtenstarken Jahrgängen in den Ruhestand gehen, rücken die deutlich kleineren Generationen Y, Z und Alpha nach. Der Personalauswahlprozess im öffentlichen Sektor läuft häufig noch immer nach bewährten, aber oft umständlichen Mustern ab – mit Stellenanzeigen, die mehr Aktenstaub als Sternenstaub aufwirbeln und mit aufwändigen Auswahlritualen.

Doch es gibt gute Nachrichten: Künstliche Intelligenz (KI) und digitale Lösungen helfen dabei, Stellenanzeigen attraktiver zu formulieren, Bewerbungsprozesse effizienter zu gestalten und Talente gezielter zu identifizieren. Das bedeutet nicht, dass Grundprinzipien wie der Grundsatz der Bestenauslese, Objektivität und Transparenz verloren gehen – im Gegenteil! Durch den klugen Einsatz digitaler Werkzeuge können Prozesse vereinfacht und gleichzeitig optimiert werden.

Im vorliegenden Artikel greife ich beispielhaft die Themen Anforderungsprofil, Stellenanzeigen, Vorstellungsgespräche und Rollenspiele heraus. Es geht bewusst nicht um die neuesten Trends wie Robot Recruiting und die automatische Auswertung sprachlicher und visueller Parameter, Entwicklungen, die allein aus rechtlichen Gründen häufig nicht in Frage kommen.¹ Stattdessen konzentriere ich mich auf die Nutzung gängiger Programme wie Excel sowie vor allem auf den Einsatz von Large Language Models (LLM) wie ChatGPT, Copilot, Gemini, Mistral, die in der Praxis ohne tiefgehende IT-Kenntnisse eingesetzt werden können. Dieser Artikel zeigt praxisnahe Möglichkeiten, wie KI-Modelle, die Jedermann zur Verfügung stehen, im Personalauswahl-Alltag sinnvoll genutzt werden können.

Anforderungsprofil

Die Grundlage der Personalauswahl ist das Anforderungsprofil.² In der Praxis findet aber gerade dieser wichtige erste Schritt häufig zu wenig Beachtung. Die Anforderungsprofile und die Stellenbewertungen sind doch „schon da“ – was soll man sich damit beschäftigen? Wenn man hier auf Sand baut, „wackelt“ allerdings das gesamte Verfahren. Auch wenn man auf ein hinterlegtes Profil zurückgreift, muss man es gedanklich auf die aktuelle Situation der Stelle sowie auf die Situation des Arbeitsmarktes transformieren. Welche der Merkmale sind aktuell besonders relevant? Was bedeutet „Teamfähigkeit“ und „Zuverlässigkeit“ – Begriffe, die in fast allen Anforderungsprofilen stehen – konkret auf dieser Stelle? Ist das Anforderungsprofil noch zeitgemäß? Welche Merkmale fehlen (z.B. Anforderungen an die Digitalkompetenz)? Welches aktuelle und ansprechende Wording beschreibt das Anforderungsprofil treffend?

Personalmarketing

Das Anforderungsprofil ist nicht nur für die Personalauswahl, sondern auch für das Recruiting von zentraler Bedeutung. Man muss wissen, wen man sucht, um diese Personen gezielt ansprechen zu können. Wer mit veralteten Anforderungsprofilen und unüberlegten Stellenanzeigen arbeitet riskiert, beim Wettbewerb um die besten Köpfe leer auszugehen.

Unterstützung bei der Erstellung von Anforderungsprofilen

LLM können dabei helfen, Anforderungsprofile für bestimmte Berufe oder Positionen basierend auf Branchenstandards und wissenschaftlichen Erkenntnissen zu erstellen. Man kann die generative KI auffordern, das aktuelle Anforderungsprofil mit Stellenanforderungen ähnlicher Positionen zu vergleichen, um unrealistische, überzogene Erwartungen oder fehlende Merkmale zu identifizieren. Vor allem kann man die LLMs für die unbedingt notwendige „Operationalisierung“ der Eignungsmerkmale nutzen. „Operationalisieren“ bedeutet, die Eignungsmerkmale durch sogenannte „Verhaltensanker“ beobachtbar und messbar werden zu lassen. Unter dem abstrakten Begriff ‚Kommunikationsfähigkeit‘ stellen sich verschiedene Personen Unterschiedliches vor. Verhaltensanker wie „hält Blickkontakt“, „spricht deutlich“ usw. tragen dazu bei, das Verständnis zu vereinheitlichen. LLMs können Entwürfe für die Verhaltensanker erarbeiten und dabei angewiesen werden, auf die Trennschärfe der Eignungsmerkmale zu achten.

Stellenanzeigen

Das Anforderungsprofil muss in die Stellenanzeige übersetzt werden. Hierbei kann die künstliche Intelligenz in vielfältiger Form unterstützen. Im Allgemeinen kann die Maschine analysieren, welche Variante der Ausschreibung in Suchmaschinen gut „gefunden“ wird, bei der Zielgruppe gut ankommt und Talente anspricht, statt abzuschrecken. Im Besonderen kann man generative KI nutzen, um Stellenanzeigen bewusst geschlechtsneutral oder inklusiv zu formulieren und so eine breitere Gruppe anzusprechen. Man kann mit der Stellenanzeige aber auch gezielt Gruppen ansprechen, die aktuell in dem Beruf unterrepräsentiert sind, z. B. Frauen oder Ältere. Stellenanzeigen enthalten nämlich oft stereotype Formulierungen, die bestimmte Gruppen unbewusst bevorzugen oder ausschließen können.³ Dies geschieht durch die Wahl von Worten, die beispielsweise eher als männlich oder weiblich wahrgenommen werden⁴, oder durch eine Sprache, die unbewusst Vorurteile vermittelt.

Es gibt zahlreiche Tools⁵, die Organisationen helfen, ihre Stellenanzeigen inklusiver und attraktiver für diverse

¹ Schepers, 2024

² Hoff und Kersting, 2019

³ Gaucher, Friesen & Kay, 2011

⁴ Bagly et al., 2020

⁵ Transparenzhinweis: Der Autor ist wissenschaftlicher Berater in einem Unternehmen, das solche Tools anbietet

Zielgruppen zu gestalten. Man kann aber, sehr pragmatisch, auch LLMs nutzen, um Stellenanzeigen zu analysieren und zu optimieren. So kann man die Maschinen prompten, Ausschreibungen mit den aktuellen Erwartungen am Arbeitsmarkt abzugleichen, die Präferenzen der Zielgruppe zu beachten, Texte auf stereotype Formulierungen zu analysieren und alternative Formulierungen vorzuschlagen, die entweder inklusiver ausfallen oder gezielt unterrepräsentierte Gruppen ansprechen. Die Betonung der sozialen Wirkung der Arbeit und der Bedeutung von Kooperation spricht tendenziell eher Frauen an. Will man lebensältere Bewerberinnen und Bewerber gewinnen, stellt man die Bedeutung von Berufserfahrung heraus und betont die Jobsicherheit, die Vereinbarkeit von Beruf und Familie sowie das altersgemischte Team. Die generative KI erstellt gezielt optimierte Ausschreibungsvorschläge.

Management(systeme)

Es gibt zahlreiche Programme, die es Organisationen ermöglichen, den Bewerbungsprozess effizient zu gestalten. Hierzu zählt u.a. die Veröffentlichung von Stellenanzeigen auf mehreren Plattformen, die zentrale Erfassung und automatische „Sichtung“ aller Bewerbungen inklusive des Screenings der Lebensläufe, automatisierte Vorschläge für die Vorauswahl, die Kommunikation mit Bewerberinnen und Bewerbern (ggf. mittels Chatbots) einschließlich des Terminmanagement. Die Programme liefern „nebenbei“ wichtige Kennzahlen über die Quantität und Qualität von Bewerbungen und Kanälen, Daten zur Bearbeitungs- und Besetzungsgeschwindigkeit, Klickzahlen von Stellenausschreibungen usw. Die Analysen werden zumeist in Form von Dashboards mit Visualisierungen zur Verfügung gestellt. Wenn man keine „große Lösung“ eingekauft hat, können LLMs auch hier hilfreich sein, um repetitive Aufgaben zu automatisieren, die Daten strukturiert zu erfassen, Vorschläge für Texte (z.B. Einladungsschreiben, Absagen) zu generieren usw. Ich möchte mich im Folgenden aber auf die Eignungsdiagnostik im engeren Sinne konzentrieren.

Eignungsdiagnostik

Sobald das Anforderungsprofil definiert ist, startet die Planung des Eignungsbeurteilungsprozesses. Dabei wird festgelegt, welche Methoden zur Eignungsbeurteilung zum Einsatz kommen und in welcher Reihenfolge sie durchgeführt werden.⁵ Die zentrale Herausforderung besteht darin, Verfahren zu wählen, die einerseits gezielt Anreize schaffen, um das interessierende Verhalten hervorzurufen. Andererseits ist es wichtig, dass die Bewerberinnen und Bewerber die Situation als offen genug empfinden, um frei zu agieren. Im öffentlichen Dienst führt ein falsches Verständnis von „Standardisierung“ häufig dazu, dass die Personalauswahlverfahren so restriktiv gestaltet werden, dass alle Teilnehmenden mehr oder minder gleich agieren, anstatt ihr individuelles Verhalten zu zeigen.

Der Prozess kann vielfältig digital unterstützt werden. Auf der Basis des Anforderungsprofils sowie Informationen zu den Rahmenbedingungen (z.B. Anzahl der Bewerberinnen und Bewerber) kann die KI Vorschläge für die Prozessstufen und die Ausgestaltung einzelner Verfahren generieren.

Systematische Nutzung der Daten

Bei der Personalauswahl fallen zahlreiche Daten an. Die Daten aus den Bewerbungsunterlagen, die Ergebnisse aus Tests, die ggf. vorab (mit oder ohne Aufsicht) bearbeitet wurden, die Eindrücke aus dem (Video-)Vorstellungsgespräch usw. Häufig sind die einzelnen Bausteine der Personalauswahl hochwertig – die daraus abgeleiteten Informationen werden aber nicht regelgeleitet zusammengefasst. Mit gängigen Programmen wie Excel lassen sich Daten strukturiert erfassen, filtern und analysieren. Der Einsatz solcher Programme fördert eine systematische Betrachtung der Daten, um deren volles Potenzial auszuschöpfen. Die datenbasierte Abbildung des Prozesses macht bisher vernachlässigte Fragen sichtbar. Soll der im ersten Schritt durchgeführte Test nur zur Vorauswahl dienen und im Vorstellungsgespräch gilt „neues Spiel, neues Glück“? Oder wird das Testergebnis systematisch in die Gesamtbewertung einbezogen? Verlässt man sich auf seinen subjektiven Eindruck, dass die Beobachterinnen und Beobachter in ihren Urteilen recht gut übereinstimmen, oder wird diese Übereinstimmung tatsächlich berechnet, um mögliche Abweichungen in der Beurteilung einzelner Eignungsmerkmale zu erkennen? Ein Scoring-Modell für die Eignungsbeurteilung festzulegen und mit geeigneten Analysetools umzusetzen, ermöglicht nicht nur eine effizientere Arbeitsweise, sondern trägt auch wesentlich zur Verbesserung des Eignungsurteils bei.

Ausgestaltung von Einstellungsgesprächen und Rollenspielen

Das Interview⁷ ist Bestandteil fast jeden Auswahlverfahrens. LLMs können die Personalverantwortlichen sowohl bei der Gestaltung der Fragen als auch bei der Bewertung der Antworten unterstützen. Zunächst fordert man das LLM auf, einen Interviewleitfaden zu entwickeln. Das Modell soll für jedes Eignungsmerkmal mehrere Fragen generieren, und zwar sowohl biographische als auch situative Fragen. Darüber hinaus kann das LLM sinnvolle Nachfragen vorschlagen, die in den Interviewleitfaden aufgenommen werden, um die Antworten der Bewerbenden weiter zu vertiefen.

Ein häufig sinnvolles Element im Auswahlverfahren ist ein Rollenspiel, das es ermöglicht, die Bewerbenden in simulierten beruflichen Situationen zu beobachten und ihre Kompetenzen praxisnah zu bewerten. Die Herausforderung besteht in der Erstellung eines realistischen, anforderungsbezogenen Szenarios mit klaren Rollen, Aufgaben und Bewertungskriterien. Auch diese Aufgabe kann man an ein LLM delegieren. Das LLM kann u. a. die Instruktionen verfassen, verschiedene mögliche Gesprächsverläufe skizzieren, einschließlich kritischer Wendepunkte oder Eskalationsstufen.

Unterstützung bei der Einstufung

Eignungsdiagnostische Interviews sind dann erfolgreich, wenn man sich im Vorhinein nicht nur Gedanken über mögliche Fragen, sondern auch über die erwünschten Antworten gemacht hat. Welche Antwort erwartet man von einer guten Bewerberin/ einem guten Bewerber? Welche Antwort bewertet man als unzureichend? Rollenspiele sind erfolgreich, wenn nicht nur die Szenarien sorgfältig durchdacht, sondern

auch klare Erwartungshorizonte für das gewünschte Verhalten der Bewerbenden definiert wurden. Solche Bewertungsraster kann man in Form sogenannter „Behavior-Anchored Rating Scales“ erstellen. Hier werden beispielhafte Antworten auf die Interviewfragen und Reaktionen im Rollenspiel mit einer Punkteskala verknüpft. Dadurch erhalten die Urteilenden eine klare Orientierung und die Einschätzungen werden vergleichbar. Auch diese Skalen kann man sich von einem LLM erstellen lassen.



Prof. Dr. Martin Kersting
Lehrstuhl für Psychologische Diagnostik an der Justus-Liebig-Universität Gießen (JLU)
www.kersting-internet.de

Zudem kann das LLM auch die Erstellung aller Materialien, vom Anforderungsprofil mit Verhaltensankern bis zu den Beobachtungsbogen und der Matrix zur Integration aller Werte zu einem Gesamturteil sowie den Feedbackbogen übernehmen.

Erstellung der Materialien

Zudem kann das LLM auch die Erstellung aller Materialien, vom Anforderungsprofil mit Verhaltensankern bis zu den Beobachtungsbogen und der Matrix zur Integration aller Werte zu einem Gesamturteil sowie den Feedbackbogen übernehmen.

Zudem kann das LLM auch die Erstellung aller Materialien, vom Anforderungsprofil mit Verhaltensankern bis zu den Beobachtungsbogen und der Matrix zur Integration aller Werte zu einem Gesamturteil sowie den Feedbackbogen übernehmen.

Chancen nutzen, Risiken kennen

Wer LLMs in der Personalauswahl klug einsetzt, kann Zeit sparen und Qualität hinzugewinnen – sollte aber auch wissen, wo die Grenzen der Technologie liegen. Künstliche Intelligenz ist kein Ersatz für menschliche Expertise. Die Qualität der generierten Inhalte hängt stark davon ab, wie sie genutzt wird. Daher ist es wichtig, einige Grundregeln im Umgang mit LLMs zu beachten. Damit ein LLM sinnvolle Unterstützung leisten kann, braucht es präzise Anweisungen – sogenannte „Prompts“. Es lohnt sich in jedem Fall, sich damit zu beschäftigen, welche Art von Eingaben zu den besten Ergebnissen führen.⁸ Auch das Ausprobieren unterschiedlicher LLMs lohnt sich. Die Vorschläge der LLMs müssen in jedem Fall kritisch geprüft werden, voraussichtlich sind einige gut, viele in Ordnung und einige nicht zu gebrauchen. Besondere Aufmerksamkeit sollte man der Datenqualität und dem Thema Diskriminierung widmen. Der Ethikbeirat HR Tech schreibt: „Personalverantwortliche sollen sich bewusst sein, dass datenbasierte und/oder automatisierte Entscheidungsprozesse nicht automatisch objektiv und fair sind. Historische Daten, die genutzt werden, um KI-Lösungen zu entwickeln und zu trainieren, sind von normativen Vorgaben zu unterscheiden.“⁹ Welche Art von Mitarbeiterinnen und Mitarbeitern man sich wünscht, muss der Mensch entscheiden, nicht die Maschine.

Datenschutz

Selbstverständlich ist der Datenschutz einzuhalten. Viele Modelle sind cloudbasiert, eingegebene Informationen können zur Weiterverarbeitung gespeichert werden. Daher sollten keinerlei Daten eingegeben werden, die sich auf eine identifizierte oder identifizierbare natürliche Person beziehen.

Das ist für alle hier im Artikel dargestellten Anwendungen aber auch nicht notwendig.

Hase und Igel

Nicht nur die Personalverantwortlichen, auch die Bewerbenden rüsten digital auf. Bewerberinnen und Bewerber können die KI für die Jobsuche nutzen, und Lebensläufe und Anschreiben automatisch generieren lassen. Und Bewerbende, die heutzutage keine KI nutzen, sind in vielen Fällen eher als rückständig denn als qualifiziert einzustufen. Kritisch wird es jedoch, wenn KI zur Täuschung eingesetzt wird – etwa indem sie unbeaufsichtigte Eignungstests bearbeitet oder in Echtzeit Antwortvorschläge für Videointerviews liefert. Personalverantwortliche müssen diese Entwicklungen berücksichtigen und sorgfältig abwägen, welche Informationen sie als verlässlich einstufen.

Digitale Lösungen lohnen sich

Auch ohne teure HR-Software können Personalverantwortliche mit Excel und Co. sowie vor allem mit KI-gestützten Sprachmodellen die Arbeit erheblich erleichtern und zugleich Effizienz und Qualität steigern. Warum also nicht einfach ausprobieren?

Literatur

- Eagly, A. H., Nater, C., Miller, D. I., Kaufmann, M. & Sczesny, S. (2020). Gender stereotypes have changed: A cross-temporal meta-analysis of U.S. public opinion polls from 1946 to 2018. *American Psychologist*, 75(3), 301-315. <https://doi.org/10.1037/amp0000494>
- Hewing, M. & Leinhos, V. (2024). The prompt canvas: A literature-based practitioner guide for creating effective prompts in Large Language Models. arXiv:2412.05127, <https://doi.org/10.48550/arXiv.2412.05127>
- Höft, S. & Kersting, M. (2018). Die Anforderungsanalyse als Grundlage der beruflichen Eignungsdiagnostik. In: Diagnostik- und Testkuratorium (Hrsg.), *Personalauswahl kompetent gestalten: Grundlagen und Praxis der Eignungsdiagnostik nach DIN 33430* (S. 97-112). Berlin: Springer. <https://doi.org/10.1007/978-3-662-53772-5>
- Gaucher, D., Friesen, J. & Kay, A. C. (2011). Evidence that gendered wording in job advertisements exists and sustains gender inequality. *Journal of Personality and Social Psychology*, 101(1), 109-128. <https://doi.org/10.1037/a0022530>
- Kersting, M. (2018). Verfahren der Eignungsbeurteilung sowie ihre Möglichkeiten und Grenzen. In: Diagnostik- und Testkuratorium (Hrsg.), *Personalauswahl kompetent gestalten: Grundlagen und Praxis der Eignungsdiagnostik nach DIN 33430* (S. 112-139). Berlin: Springer. <https://doi.org/10.1007/978-3-662-53772-5>
- Schuler, H. (2018). *Das Einstellungsinterview*. 2. Überarbeitete Auflage. Göttingen: Hogrefe. <http://doi.org/10.1026/02871-000>
- Schepers, J. (2024). *Digitalisierte Personalauswahl. Die Zulässigkeit algorithmischer Verfahren der Eignungsdiagnostik*. Wiesbaden: Springer Fachmedien. <https://doi.org/10.1007/978-3-658-44466-2>

⁸ Hewing & Leinhos, 2024.

⁹ <https://www.ethikbeirat-hrtecti.de>

FS

Forum Strafvollzug

Zeitschrift für Strafvollzug
und Straffälligenhilfe

KI im Justizvollzug – Algorithmen, maschinelles Lernen oder einfach eine Krasse Idee?

Einführung in den Schwerpunkt

*Heidi Drescher, Jochen Goerdeler, Marcel Guéridon,
Daniel Wolter*

Künstliche Intelligenz – was ist das?

Marcel Guéridon

Künstliche Intelligenz im Justizvollzug

Marcel Guéridon

Multimodale Verhaltensanalyse im Strafvollzug

Martin Kempel, Thomas Heitzinger, Robert Rothmann

Sprachbarrieren mit KI überwinden

Tabea Mehić

KI@JVA – Die Zukunft der Gefängnissicherheit?

Oliver Werthwein, Dominik Beyer, Wilhelm Stork

Einsatz Künstlicher Intelligenz in der Rechtspflege

Bettina Mielke

KI in der Justiz – Potenziale & Herausforderungen

Christian Hermann

Einfache Mittel, große Wirkung: ChatGPT und Personalauswahl

Martin Kersting

Forschung & Entwicklung

Diagnose, Behandlung und Prognose von jungen Brandstiftern

Jürgen Ptucha

Neue Autorität im Offenen Vollzug

Uta Czynnick-Leber, Jan Goofsen

Recht & Reform

Neue Ansätze der Medienarbeit in den Justizvollzugs- anstalten

Michael Kubink

Herausgeber

Gesellschaft für Fortbildung der
Strafvollzugsbediensteten e.V.



Redaktion

Frank Arloth
Heidi Drescher
Susanne Gerlach
Jochen Goerdeler
Gesa Lürßen
Stephanie Pfalzer
Stefanie Roos
Karin Roth
Stefan Suhling
Daniel Wolter

2
25