

Kompass by ZeugnisPilot: Ein evidenzbasierter Standard für moderne Schweizer Arbeitszeugnisse

Fabrice Bircher, Felix Anderegg, Carim Chenna und Nils Tosoni
ZeugnisPilot / rhyno solutions ag, Schaffhausen und Zürich, Schweiz

Zusammenfassung

Das Schweizer Arbeitszeugnis ist ein institutionell etabliertes Vertrauensdokument, dessen normative Bedeutung weit über eine blossе Austrittsformalität hinausgeht. Gleichzeitig steht die heutige Zeugnispraxis unter einem erheblichen Spannungsverhältnis: Das Dokument muss das berufliche Fortkommen der Arbeitnehmerin oder des Arbeitnehmers fördern, darf aber weder unwahr noch unvollständig sein; es soll standardisiert und effizient erstellt werden, ohne in formelhafte, kaum überprüfbare Textbausteинlogik abzugleiten. Auf Basis von Desk Research und Study-Aggregation entwickelt dieses Paper den Ansatz Kompass by ZeugnisPilot als evidenzbasierten Standard für moderne Schweizer Arbeitszeugnisse. Der Ansatz integriert die gesetzlichen Mindestanforderungen nach Art. 330a OR, die verwaltungs- und gerichtliche Praxis zu Wahrheit, Vollständigkeit und Wohlwollen, datenschutzrechtliche Anforderungen an arbeitsbezogene Personendaten sowie arbeitspsychologische Modelle individueller Arbeitsleistung. Die Synthese zeigt, dass ein zukunftsfähiges Arbeitszeugnis weder als rein narrativer Gefälligkeitstext noch als numerisches Bewertungsdashboard verstanden werden sollte. Vorgeschlagen wird vielmehr ein strukturiertes Vertrauensdokument mit vier obligatorischen Leistungsfeldern und einem optionalen Führungsfeld, ergänzt durch Evidenzsätze, beobachtbares Verhalten und eine dokumentierte Beurteilungsbasis. Kompass soll dadurch Aussagekraft, Fairness, rechtliche Plausibilität und Automatisierbarkeit zusammenführen.

Schlüsselwörter

Arbeitszeugnis; Schweiz; Art. 330a OR; Leistungsbeurteilung; individuelle Arbeitsleistung; Evidenzsätze; Kompass Standard; HR-Tech; Legal-Tech; ZeugnisPilot

1. Einleitung

Das Arbeitszeugnis nimmt im Schweizer Arbeitsmarkt eine besondere Stellung ein. Es ist gleichzeitig Rechtsanspruch, Reputationsträger, Selektionsartefakt und biografisches Dokument. Diese Mehrfachfunktion erklärt, weshalb Zeugnisse trotz wachsender Kritik an ihrer Aussagekraft weiterhin mit erheblichem Aufwand erstellt, geprüft und interpretiert werden. Die gesetzliche Grundlage ist bewusst knapp formuliert: Nach Art. 330a OR kann die arbeitnehmende Person jederzeit ein Zeugnis verlangen, das sich über Art und Dauer des Arbeitsverhältnisses sowie über Leistungen und Verhalten ausspricht; auf besonderes Verlangen kann das Dokument auf Art und Dauer beschränkt werden [1].

Die praktische Herausforderung besteht darin, dass gerade diese knappe gesetzliche Norm eine hochkomplexe Dokumentpraxis erzeugt. Ein qualifiziertes Arbeitszeugnis muss informativ, verständlich und vollständig sein, gleichzeitig aber wohlwollend formuliert bleiben. Das SECO beschreibt das Vollzeugnis als Dokument, das neben Identität, Beginn und Ende des Arbeitsverhältnisses sowie prägenden Tätigkeiten auch eine aussagekräftige Bewertung von Leistung und Verhalten enthalten soll [2]. Die Rechtsprechung präzisiert diese Spannung: Das Zeugnis soll das berufliche Fortkommen fördern, muss aber zugleich ein möglichst getreues Bild von Tätigkeit, Leistung und Verhalten vermitteln; erhebliche negative Tatsachen dürfen nicht verschwiegen werden, wenn sie für die Gesamtbeurteilung wesentlich sind [3].

Kompass by ZeugnisPilot setzt an dieser Schnittstelle an. Der Ansatz fragt nicht primär, wie Arbeitszeugnisse schneller formuliert werden können, sondern welche Struktur ein modernes Arbeitszeugnis benötigt, um rechtlich plausibel, arbeitspsychologisch anschlussfähig, lesbar und automatisierbar zu sein. Damit verschiebt sich der Fokus von der Textbausteinproduktion zur Gestaltung eines methodisch kontrollierten Vertrauensdokuments.

1.1 Relevanz

Die Relevanz des Themas ergibt sich aus drei Entwicklungen. Erstens verändert sich Arbeit: Rollen werden dynamischer, Projekte und Schnittstellenarbeit gewinnen an Bedeutung, und Leistungsbeiträge entstehen zunehmend in interdisziplinären Kontexten. Zweitens verändern sich HR-Prozesse: Digitale Workflows, KI-gestützte Textgenerierung und standardisierte Freigabeprozesse ermöglichen neue Formen der Zeugniserstellung, erhöhen aber zugleich die Verantwortung für Datenqualität und Governance. Drittens verändert sich die Erwartung an Nachvollziehbarkeit: Mitarbeitende und künftige Arbeitgeberinnen wollen nicht nur wohlklingende Formulierungen, sondern verständliche Hinweise auf Aufgaben, Beiträge und beobachtbares Verhalten.

Die Forschung zur individuellen Arbeitsleistung stützt diese Entwicklung. Koopmans et al. zeigen in einer systematischen Übersicht, dass Arbeitsleistung in unterschiedlichen Disziplinen verschieden konzeptualisiert wird, sich jedoch häufig um Aufgabenleistung, kontextuelle Leistung, kontraproduktives Verhalten und adaptive Leistung gruppiert [7]. In einer späteren Arbeit zur Messung individueller Arbeitsleistung werden diese Dimensionen weiter operationalisiert und über Literatur, Fragebogeninstrumente sowie Expertinnen- und Experteninterviews verdichtet [8]. Daraus folgt, dass ein modernes Arbeitszeugnis seine Aussagekraft erhöht, wenn es Leistung nicht eindimensional, sondern anhand klarer, arbeitsbezogener Felder beschreibt.

1.2 Problemstellung

Die traditionelle Zeugnispraxis leidet unter einem strukturellen Kommunikationsproblem. Einerseits sind Zeugnisse in vielen Organisationen hoch standardisiert; andererseits bleibt unklar, ob die verwendeten Formulierungen tatsächliche Beobachtungen, rechtliche Vorsicht, interne Textbausteinlogik oder implizite Codes abbilden. Dieses Problem wird verschärft, wenn vermeintlich moderne Formate lediglich grafische Elemente wie Balken, Sterne oder Prozentwerte einführen. Solche

Darstellungen können den Eindruck empirischer Präzision erzeugen, ohne dass die zugrunde liegenden Beobachtungen, Skalen oder Bewertungsanker offengelegt werden.

Aus wissenschaftlicher Perspektive ist dies problematisch, weil Leistungsbeurteilung nicht nur ein Mess-, sondern auch ein sozialer und organisationaler Prozess ist. Murphy und Cleveland betonen, dass Performance Appraisal nicht isoliert als technische Bewertung verstanden werden kann, sondern in soziale, organisatorische und zielbezogene Kontexte eingebettet ist [12]. DeNisi und Murphy zeigen in ihrer Übersicht über ein Jahrhundert Forschung, dass Leistungsbeurteilung insbesondere von Skalenformaten, Beurteilungszwecken, Beurteilungsquellen, Reaktionen auf Beurteilungen, kognitiven Prozessen und organisationaler Einbettung geprägt ist [13]. Ein Arbeitszeugnis, das diese Komplexität ignoriert, läuft Gefahr, formell korrekt, aber inhaltlich wenig belastbar zu sein.

1.3 Forschungsfragen und Beitrag

Dieses Paper verfolgt drei Forschungsfragen. Erstens: Welche rechtlichen, datenschutzrechtlichen und arbeitspsychologischen Anforderungen muss ein moderner Schweizer Arbeitszeugnisstandard erfüllen? Zweitens: Welche Struktur erlaubt es, Leistung und Verhalten differenziert, wohlwollend und wahrheitsorientiert darzustellen, ohne eine unzulässige Scheingenauigkeit zu erzeugen? Drittens: Wie kann ein solcher Standard so gestaltet werden, dass er in digitalen HR-Prozessen automatisierbar bleibt und dennoch eine menschlich verantwortete Beurteilung sicherstellt?

Der Beitrag des Papers besteht in einer theoretisch fundierten Verdichtung zum Kompass-Ansatz. Kompass wird als Standardmodell vorgeschlagen, das die gesetzlichen Elemente des Volzeugnisses mit einer arbeitspsychologisch informierten Leistungsstruktur verbindet. Zentral ist die These, dass moderne Arbeitszeugnisse nicht primär durch längere Texte, sondern durch bessere epistemische Ordnung gewinnen: klare Felder, überprüfbare Bezüge, evidenznahe Sätze und transparente Beurteilungsgrundlagen.

1.4 Methode

Methodisch folgt die Arbeit einem qualitativen Desk-Research- und Study-Aggregation-Design. Desk Research bezeichnet hier die strukturierte Auswertung öffentlich zugänglicher Rechtsquellen, behördlicher Orientierungshilfen, gerichtlicher Leitentscheide und wissenschaftlicher Literatur. Study Aggregation meint keine statistische Meta-Analyse, sondern eine konzeptionelle Verdichtung mehrerer Forschungsstränge: schweizerisches Arbeitszeugnisrecht, Datenschutz im Arbeitsverhältnis, arbeitspsychologische Leistungsdimensionen, Forschung zur Leistungsbeurteilung sowie Grundlagen zur strukturierten und narrativen Beurteilung.

Die Methode ist für die Forschungsfrage angemessen, weil der Kompass-Ansatz keine einzelne empirische Hypothese testet, sondern ein normativ-praktisches Artefakt begründet. Ähnlich wie konzeptionelle Management- und Design-Science-Arbeiten zielt die Analyse darauf, aus etablierten Quellen Gestaltungsprinzipien abzuleiten, die in einem produktiven HR-Tech-Kontext operationalisiert werden können. Die Grenzen dieses Vorgehens werden in Abschnitt 7 explizit diskutiert.

2. Verwandte Arbeiten

2.1 Rechtlicher Rahmen des Schweizer Arbeitszeugnisses

Der rechtliche Ausgangspunkt ist Art. 330a OR. Die Norm unterscheidet implizit zwischen dem qualifizierten Zeugnis und der Arbeitsbestätigung: Während das qualifizierte Zeugnis Art und Dauer des Arbeitsverhältnisses sowie Leistung und Verhalten umfasst, beschränkt sich die Arbeitsbestätigung auf Art und Dauer [1]. Das SECO übernimmt diese Differenzierung und konkretisiert, dass ein Vollzeugnis unter anderem die prägenden Tätigkeiten sowie eine aussagekräftige Bewertung von Arbeitsqualität, Arbeitsquantität und Verhalten enthalten sollte [2].

Die schweizerische Praxis verlangt darüber hinaus eine Balance zwischen Wahrheit und Wohlwollen. BGE 136 III 510 hält fest, dass das Zeugnis das berufliche Fortkommen begünstigen soll, aber nicht unwahr oder unvollständig sein darf. Wesentliche negative Tatsachen sind aufzunehmen, wenn ihr Verschweigen ein falsches Gesamtbild erzeugen würde [3]. Für Kompass folgt daraus, dass der Standard nicht auf reine Positivprosa reduziert werden darf. Vielmehr muss er eine Sprache ermöglichen, die kritisch relevante Sachverhalte sachlich, verhältnismässig und arbeitsbezogen ausdrückt.

Der Kanton Zürich beschreibt ergänzend formale und inhaltliche Anforderungen, etwa Kennzeichnung, Arbeitgeberidentifikation, vollständige Personalien, Tätigkeitsbeschreibung, Leistungs- und Verhaltensbeurteilung, Datierung sowie optische und sprachliche Sorgfalt [4]. Diese Hinweise zeigen, dass Dokumentqualität nicht nur ein ästhetisches Thema ist, sondern Teil der Glaubwürdigkeit des Zeugnisses.

2.2 Datenschutz und arbeitsbezogene Daten

Arbeitszeugnisse enthalten Personendaten und potenziell sensible arbeitsbezogene Bewertungen. Deshalb ist der Grundsatz der Zweckbindung und Verhältnismässigkeit zentral. Art. 328b OR beschränkt die Bearbeitung von Daten durch den Arbeitgeber auf Informationen, die die Eignung für das Arbeitsverhältnis betreffen oder zur Durchführung des Arbeitsvertrags erforderlich sind [6]. Der EDÖB betont ebenfalls, dass Arbeitgeber Personendaten im Arbeitsverhältnis nur im erforderlichen Umfang bearbeiten sollen [5].

Für Kompass hat dies zwei Konsequenzen. Erstens dürfen im Zeugnis keine unnötigen Persönlichkeitsprofile, privaten Informationen oder intransparenten Datenquellen verarbeitet werden. Zweitens sollte der Standard beobachtbares Arbeitsverhalten statt psychologisierende Persönlichkeitszuschreibungen erfassen. Begriffe wie ruhig, verbindlich oder lösungsorientiert sind nur dann angemessen, wenn sie arbeitsbezogen konkretisiert und durch beobachtbare Situationen gestützt werden.

2.3 Arbeitspsychologische Grundlagen individueller Leistung

Die arbeitspsychologische Literatur liefert ein differenziertes Verständnis von Leistung. Campbell, McHenry und Wise untersuchten im Rahmen eines umfangreichen Projekts 19 militärische Eingangsfunktionen, entwickelten über 200 Leistungsindikatoren und analysierten Daten von 9'430 Stelleninhabenden; daraus wurde die Mehrdimensionalität

von Arbeitsleistung deutlich [9]. Dieses Ergebnis ist für Arbeitszeugnisse bedeutsam, weil es die Unzulänglichkeit einer einzigen pauschalen Leistungsformel unterstreicht.

Borman und Motowidlo unterscheiden zwischen Aufgabenleistung und kontextueller Leistung [10]. Aufgabenleistung beschreibt Beiträge, die unmittelbar zur technischen oder fachlichen Kernaufgabe gehören. Kontextuelle Leistung umfasst Verhalten, das organisatorische, soziale und psychologische Umfeld unterstützt, etwa Kooperation, Hilfeleistung, Initiative oder Ausdauer. Motowidlo, Borman und Schmit definieren Arbeitsleistung als aggregierten Wert diskreter Verhaltensepisoden über einen Zeitraum und betonen, dass Leistung verhaltensbezogen, episodisch, evaluativ und multidimensional ist [11].

Adaptive Leistung bildet eine weitere wichtige Dimension. Pulakos et al. untersuchten eine Taxonomie adaptiver Leistung und fanden Unterstützung für ein mehrdimensionales Modell, das unter anderem kreatives Problemlösen, Umgang mit Unsicherheit, Lernen neuer Aufgaben und interpersonelle Anpassung umfasst [14]. In modernen Organisationen ist diese Dimension für Arbeitszeugnisse besonders relevant, weil Rollen, Technologien und Prozesse häufiger wechseln.

2.4 Leistungsbeurteilung und Beurteilungsqualität

Die Forschung zur Leistungsbeurteilung zeigt, dass Bewertungsqualität nicht allein von der Skala abhängt. Vielmehr sind Bewertungszweck, Beobachtungsgrundlage, Beurteilerschulung und organisationale Einbettung entscheidend [12,13]. Woehr und Huffcutt zeigen in ihrer meta-analytischen Integration, dass unterschiedliche Formen von Rater Training - etwa Frame-of-Reference-Training, Beobachtungstraining und dimensionsbezogenes Training - jeweils mindestens moderat wirksam sein können, wenn sie auf spezifische Beurteilungsfehler ausgerichtet sind [15].

Für Arbeitszeugnisse bedeutet dies: Ein Standard muss nicht nur Ausgabertexte definieren, sondern auch den Beurteilungsprozess strukturieren. Wenn Führungskräfte lediglich Freitext liefern, entstehen Inkonsistenzen. Wenn sie hingegen entlang klarer Felder, Beobachtungsfragen und Evidenzkategorien geführt werden, verbessert sich die Qualität der Grundlage, auf der das Zeugnis erstellt wird.

3. Theoretische Synthese: Vom Textbaustein zum Vertrauensdokument

Die Aggregation der Literatur legt nahe, dass ein modernes Arbeitszeugnis als strukturiertes Vertrauensdokument verstanden werden sollte. Dieser Begriff bezeichnet ein Dokument, das rechtliche Mindestanforderungen erfüllt, den beruflichen Beitrag einer Person nachvollziehbar darstellt und für unterschiedliche Lesergruppen anschlussfähig bleibt. Der Begriff Vertrauensdokument ist bewusst gewählt: Das Zeugnis soll Vertrauen in die Fairness des Verfahrens, die Plausibilität der Aussagen und die Integrität der ausstellenden Organisation schaffen.

Daraus ergeben sich vier Gestaltungsprinzipien. Erstens muss das Zeugnis gesetzlich vollständig sein: Art und Dauer des Arbeitsverhältnisses, Aufgaben, Leistung und Verhalten müssen erkennbar abgebildet werden [1,2]. Zweitens muss es wahrheits- und wohlwollensorientiert sein: Es darf weder schädigende Details ohne Relevanz

noch beschönigende Unrichtigkeiten enthalten [3]. Drittens muss es verhaltens- und evidenznah sein: Beurteilungen sollen aus beobachtbaren Beiträgen abgeleitet werden [7,8,11]. Viertens muss es prozessual kontrollierbar sein: Automatisierung darf die menschliche Verantwortung nicht ersetzen, sondern muss sie strukturieren.

4. Der Kompass-Ansatz

4.1 Strukturmodell: vier Pflichtfelder und ein optionales Führungsfeld

Kompass operationalisiert die Synthese in einem 4+1-Modell. Vier Leistungsfelder sind grundsätzlich für nahezu alle qualifizierten Arbeitszeugnisse relevant: fachlicher Beitrag, Umsetzung und Verlässlichkeit, Zusammenarbeit und Wirkung sowie Lernen und Anpassung. Das fünfte Feld Führung und Entwicklung anderer wird nur verwendet, wenn Führungs-, Fachführungs-, Projektleitungs- oder Mentoringverantwortung tatsächlich vorlag.

Der fachliche Beitrag entspricht der Aufgabenleistung im Sinne der arbeitspsychologischen Forschung [10,11]. Umsetzung und Verlässlichkeit bündelt Aspekte wie Priorisierung, Termintreue, Sorgfalt und Stabilität unter Belastung. Zusammenarbeit und Wirkung übersetzt kontextuelle Leistung in die Zeugnissprache [10]. Lernen und Anpassung nimmt adaptive Leistung auf [14]. Führung und Entwicklung anderer ist ein kontingentes Feld, das nur bei entsprechender Rolle aktiviert wird, um weder Führung zu erfinden noch relevante Führungsbeiträge zu unterschlagen.

Dieses Modell ist bewusst schlank. Es vermeidet eine Überfrachtung des Zeugnisses mit psychologischen Dimensionen und hält die Anschlussfähigkeit an die schweizerische Zeugnislogik aufrecht. Gleichzeitig ist es differenzierter als pauschale Gesamtformeln, weil es die Beurteilung entlang beobachtbarer Leistungsdomänen ordnet.

4.2 Evidenzsätze statt Codes und Scheingenaugigkeit

Ein zentrales Element des Kompass-Ansatzes ist der Evidenzsatz. Ein Evidenzsatz verbindet eine Beurteilung mit einem konkreten Bezug: Welche Aufgabe, welches Projekt, welches Verhalten oder welches Ergebnis stützt die Aussage? Der Satz ersetzt keine interne Dokumentation, erhöht aber die Nachvollziehbarkeit im Zeugnis.

Kompass verzichtet im finalen Dokument bewusst auf Sterne, Balken, Prozentwerte oder numerische Scores. Der Grund ist nicht Technikfeindlichkeit, sondern wissenschaftliche und rechtliche Vorsicht. Leistungsbeurteilung ist zwar strukturierbar, aber selten mit der Präzision quantifizierbar, die visuelle Skalen suggerieren. Eine Prozentangabe im Zeugnis würde nur dann Sinn ergeben, wenn Skala, Beobachtungszeitraum, Bewertungsanker, Beurteilende und Messfehler transparent wären. Andernfalls entsteht Scheingenaugigkeit.

Evidenzsätze sind demgegenüber epistemisch bescheidener und zugleich nützlicher. Sie zeigen nicht vor, dass die Organisation eine exakte Zahl besitzt, sondern legen offen, worauf sich die qualitative Beurteilung stützt. Dies ist mit der Forschung vereinbar, wonach Arbeitsleistung aus konkreten Verhaltensepisoden über Zeit besteht [11].

4.3 Beurteilungsbasis und Governance

Kompass empfiehlt, die Beurteilungsbasis im Zeugnis oder zumindest im internen Erstellungsprozess sichtbar zu machen. Dazu können strukturierte Leistungsdialoge, Projektresultate, Zielvereinbarungen, dokumentierte Rückmeldungen, direkte Führungsbeobachtungen oder relevante Stakeholder-Interaktionen gehören. Entscheidend ist, dass die Beurteilungsbasis arbeitsbezogen, verhältnismässig und nachvollziehbar bleibt [5,6].

Governance ist besonders wichtig, wenn KI-gestützte Formulierung eingesetzt wird. Die KI darf keine arbeitsrechtliche Entscheidung treffen und keine Leistung eigenständig bewerten. Sie kann jedoch strukturierte Eingaben sprachlich konsistent verdichten, Risiken markieren, unklare Aussagen hervorheben und alternative Formulierungen vorschlagen. Die Verantwortung verbleibt bei Führung und HR. Damit wird Automatisierung als Qualitätssicherung verstanden, nicht als Delegation der Beurteilungsverantwortung.

5. Ergebnisse der Studienaggregation

Die Desk-Research- und Study-Aggregation führt zu fünf verdichteten Ergebnissen. Erstens ist die rechtliche Mindeststruktur des Schweizer Arbeitszeugnisses relativ stabil, aber offen genug, um eine moderne Dokumentarchitektur zuzulassen [1,2]. Weder Art. 330a OR noch die behördliche Praxis verlangen klassische Floskelprosa; verlangt wird vielmehr eine aussagekräftige Darstellung von Art, Dauer, Leistung und Verhalten.

Zweitens bestätigt die Rechtsprechung die Notwendigkeit eines Gleichgewichts zwischen Wohlwollen und Wahrheit. Ein Zeugnis, das negative wesentliche Tatsachen verschweigt, kann ebenso problematisch sein wie ein unnötig schädigendes Zeugnis [3]. Kompass adressiert dieses Problem durch sachliche, arbeitsbezogene und evidenznahe Formulierungen.

Drittens stützt die arbeitspsychologische Literatur die Mehrdimensionalität von Leistung. Die Unterscheidung zwischen Aufgabenleistung, kontextueller Leistung und adaptiver Leistung ist für Zeugnisse besonders anschlussfähig [7-11,14]. Das 4+1-Modell übersetzt diese Forschung in eine praxistaugliche und rechtlich unaufgeregte Struktur.

Viertens zeigt die Leistungsbeurteilungsforschung, dass Prozessqualität entscheidend ist. Skalen, Texte und Dokumentlayouts sind nur so gut wie die Beobachtungs- und Bewertungsgrundlage, aus der sie gespeist werden [12,13,15]. Kompass muss daher als Prozessstandard und nicht bloss als Layout verstanden werden.

Fünftens begrenzt das Datenschutzrecht die Versuchung, Arbeitszeugnisse zu psychologisieren. Persönlichkeitsprofile, irrelevante private Informationen oder intransparente Datenquellen sind mit den Grundsätzen arbeitsbezogener Datenbearbeitung nicht vereinbar [5,6]. Kompass bleibt deshalb bei beobachtbarem Arbeitsverhalten.

6. Diskussion

Der Kompass-Ansatz unterscheidet sich von zwei verbreiteten Modernisierungsrichtungen. Die erste Richtung setzt auf Textbausteinoptimierung. Sie erhöht Effizienz und Konsistenz, löst aber das Problem der Aussagekraft nur

begrenzt, wenn die zugrunde liegende Beurteilung unstrukturiert bleibt. Die zweite Richtung setzt auf Visualisierung durch Skalen oder Dashboards. Sie wirkt modern, riskiert aber Scheingenauigkeit und kann den rechtlich sensiblen Charakter des Arbeitszeugnisses unterschätzen.

Kompass schlägt eine dritte Richtung vor: strukturierte Narrativität. Das Zeugnis bleibt ein Textdokument, aber es wird nicht mehr als freie Prosa verstanden. Es folgt einer klaren Architektur aus Rolle, Kontext, Gesamtwürdigung, Leistungsfeldern, Evidenzsätzen, beobachtbarem Arbeitsstil, Beurteilungsbasis und Schlussformel. Diese Struktur macht das Dokument lesbarer, konsistenter und für digitale Prozesse anschlussfähig.

Gleichzeitig ist Kompass kein Versuch, die Ambivalenz des Arbeitszeugnisses vollständig aufzulösen. Jede Beurteilung enthält Ermessensanteile. Der Standard kann jedoch die Bedingungen verbessern, unter denen dieses Ermessen ausgeübt wird: durch klare Felder, transparente Fragen, kontrollierte Sprache, Quellenbewusstsein und HR-verantwortete Freigabe.

7. Grenzen der Arbeit

Diese Arbeit ist konzeptionell und aggregierend. Sie basiert nicht auf einer eigenen empirischen Feldstudie mit Zeugnislesenden, HR-Fachpersonen oder Gerichten. Die vorgeschlagene Struktur ist deshalb als theoretisch begründeter Standardentwurf zu verstehen, nicht als bereits validiertes Messinstrument. Zudem kann das Paper die Vielfalt schweizerischer Branchen, Unternehmensgrössen und kantonaler Verwaltungspraxen nur begrenzt abbilden.

Eine weitere Grenze liegt in der Übertragung arbeitspsychologischer Forschung auf Arbeitszeugnisse. Modelle individueller Arbeitsleistung wurden überwiegend für Auswahl, Performance Management oder Forschung entwickelt, nicht spezifisch für schweizerische Arbeitszeugnisse. Die Übertragung ist jedoch plausibel, weil das Zeugnis ebenfalls Leistung und Verhalten über einen Zeitraum darstellen muss.

Schliesslich bleibt die rechtliche Bewertung einzelfallabhängig. Kompass kann rechtliche Plausibilität und Struktur fördern, ersetzt aber keine juristische Prüfung in strittigen oder risikobehafteten Fällen.

8. Weiterer Forschungsbedarf

Zukünftige Forschung sollte den Kompass-Ansatz empirisch prüfen. Erstens wäre zu untersuchen, ob Kompass-Zeugnisse von HR-Fachpersonen, Mitarbeitenden und Recruiterinnen als verständlicher, fairer und nützlicher wahrgenommen werden als klassische Zeugnisse. Zweitens sollte experimentell getestet werden, ob Evidenzsätze die Interpretationssicherheit erhöhen und implizite Code-Lesarten reduzieren. Drittens wäre zu analysieren, ob Führungskräfte mit strukturierter Eingabe und gezieltem Beurteilungstraining konsistentere Zeugnisse erstellen.

Viertens sollte untersucht werden, welche Rolle KI-gestützte Formulierung in der Zeugniserstellung sinnvoll einnehmen kann. Besonders relevant sind Fragen der Bias-Kontrolle, Nachvollziehbarkeit, Datenminimierung und menschlichen Letztverantwortung. Fünftens könnte eine juristische Validierungsstudie prüfen, in welchen Fallkonstellationen Kompass-Formulierungen als besonders robust oder risikobehaftet gelten.

9. Schlussfolgerung

Das Schweizer Arbeitszeugnis bleibt relevant, aber seine Form muss weiterentwickelt werden. Die Kombination aus rechtlicher Bedeutung, geringer praktischer Aussagekraft vieler klassischer Zeugnisse und steigendem Automatisierungsdruck verlangt nach einem Standard, der mehr leistet als Textbausteine oder visuelle Modernisierung.

Kompass by ZeugnisPilot wird in diesem Paper als evidenzbasierter Standard für moderne Arbeitszeugnisse vorgeschlagen. Der Ansatz verbindet die gesetzlichen Anforderungen nach Art. 330a OR, die Wahrheit-Wohlwollen-Spannung der Rechtsprechung, datenschutzrechtliche Verhältnismässigkeit und arbeitspsychologische Leistungsmodelle. Das Ergebnis ist ein strukturiertes Vertrauensdokument: vier Pflicht-Leistungsfelder, ein optionales Führungsfeld, Evidenzsätze, beobachtbares Verhalten, Beurteilungsbasis und HR-verantwortete Automatisierung.

Damit kristallisiert sich die zentrale These heraus: Das Arbeitszeugnis der Zukunft ist weder ein längerer Floskeltext noch ein numerisches Dashboard. Es ist ein methodisch geführtes, rechtlich plausibles und arbeitspsychologisch anschlussfähiges Dokument, das Orientierung schafft - für Arbeitgebende, Mitarbeitende und den Arbeitsmarkt.

Literatur

- [1] Schweizerisches Obligationenrecht (OR), Art. 330a: Zeugnis. Bundesgesetz betreffend die Ergänzung des Schweizerischen Zivilgesetzbuches, SR 220.
- [2] Staatssekretariat für Wirtschaft SECO. Arbeitszeugnis: FAQ zum privaten Arbeitsrecht. Bern: SECO.
- [3] Bundesgericht. BGE 136 III 510: Arbeitszeugnis; Wahrheit, Vollständigkeit, Wohlwollen und erhebliche negative Tatsachen.
- [4] Kanton Zürich. Handbuch Personalrecht: Arbeitszeugnis. Zürich: Kantonale Verwaltung.
- [5] Eidgenössischer Datenschutz- und Öffentlichkeitsbeauftragter EDÖB. Datenbearbeitung durch den Arbeitgeber. Bern: EDÖB.
- [6] Schweizerisches Obligationenrecht (OR), Art. 328b: Bearbeitung von Personendaten im Arbeitsverhältnis.
- [7] Koopmans, L., Bernaards, C. M., Hildebrandt, V. H., Schaufeli, W. B., de Vet, H. C. W. und van der Beek, A. J. (2011). Conceptual frameworks of individual work performance: A systematic review. *Journal of Occupational and Environmental Medicine*, 53(8), 856-866. <https://doi.org/10.1097/JOM.0b013e318226a763>
- [8] Koopmans, L., Bernaards, C. M., Hildebrandt, V. H., de Vet, H. C. W. und van der Beek, A. J. (2014). Measuring individual work performance: Identifying and selecting indicators. *Work*, 48(2), 229-238. <https://doi.org/10.3233/WOR-131659>
- [9] Campbell, J. P., McHenry, J. J. und Wise, L. L. (1990). Modeling job performance in a population of jobs. *Personnel Psychology*, 43(2), 313-333. <https://doi.org/10.1111/j.1744-6570.1990.tb01561.x>
- [10] Borman, W. C. und Motowidlo, S. J. (1993). Expanding the criterion domain to include elements of contextual performance. In N. Schmitt und W. C. Borman (Hrsg.), *Personnel Selection in Organizations*, 71-98. San Francisco: Jossey-Bass.
- [11] Motowidlo, S. J., Borman, W. C. und Schmit, M. J. (1997). A theory of individual differences in task and contextual performance. *Human Performance*, 10(2), 71-83. https://doi.org/10.1207/s15327043hup1002_1
- [12] Murphy, K. R. und Cleveland, J. N. (1995). *Understanding Performance Appraisal: Social, Organizational, and Goal-Based Perspectives*. Thousand Oaks: Sage.
- [13] DeNisi, A. S. und Murphy, K. R. (2017). Performance appraisal and performance management: 100 years of progress? *Journal of Applied Psychology*, 102(3), 421-433. <https://doi.org/10.1037/apl0000085>
- [14] Pulakos, E. D., Schmitt, N., Dorsey, D. W., Arad, S., Borman, W. C. und Hedge, J. W. (2002). Predicting adaptive performance: Further tests of a model of adaptability. *Human Performance*, 15(4), 299-323. https://doi.org/10.1207/S15327043HUP1504_01
- [15] Woehr, D. J. und Huffcutt, A. I. (1994). Rater training for performance appraisal: A quantitative review. *Journal of Occupational and Organizational Psychology*, 67(3), 189-205. <https://doi.org/10.1111/j.2044-8325.1994.tb00562.x>
- [16] Motowidlo, S. J. und Van Scotter, J. R. (1994). Evidence that task performance should be distinguished from contextual performance. *Journal of Applied Psychology*, 79(4), 475-480. <https://doi.org/10.1037/0021-9010.79.4.475>