

METHODENDOKUMENTATION

Messung der Wörtlichkeit übernommener Feldwerte

Verfahren, Rechenregeln und Ergebnisse aus dem laufenden Betrieb

DOKUMENT-NR.

BS-MD-001

VERSION

1.0

STAND

28.08.2026

HERAUSGEBER

Pixel7 GmbH**Adressiert an**

Gerichtlich zertifizierte Sachverständige, Pflegegutachterinnen und Pflegegutachter, ärztliche Gutachterinnen und Gutachter sowie deren Auftraggeber.

Worum es in diesem Dokument geht

Befundscanner überträgt Inhalte aus medizinischen Dokumenten wortgleich in Word-Vorlagen. Das System interpretiert nicht und formuliert nicht um. Damit diese Zusage nicht nur eine Behauptung bleibt, misst Befundscanner bei jeder Übernahme automatisch, wie wörtlich der übernommene Feldwert im Quelldokument steht. Dieses Dokument legt offen, nach welchen Verfahren gemessen wird, wie die Kennzahlen rechnerisch zustande kommen und welche Werte der laufende Betrieb zum Stichtag zeigt.

Die wichtigste Einschränkung vorweg

Gemessen wird **Wörtlichkeit, nicht Korrektheit**. Ein Wert kann vollständig wörtlich aus dem Quelldokument stammen und trotzdem der falschen Feldposition zugeordnet sein. Kein hier berichteter Wert ersetzt die inhaltliche Prüfung. Die fachliche Freigabe liegt ausschließlich bei der Sachverständigen bzw. beim Sachverständigen.

Inhalt

- | | |
|--|---|
| 1 Zweck und Geltungsbereich | 7 Kategorien und Bänder |
| 2 Was gemessen wird — und was nicht | 8 Erhebung und Datenschutz |
| 3 Ablauf der Messung | 9 Ergebnisse aus dem laufenden Betrieb |
| 4 Vorverarbeitung | 10 Grenzen der Aussagekraft |
| 5 Die vier Maße | 11 Quellen und Kontakt |
| 6 Rechenbeispiele | |

1 Zweck und Geltungsbereich

Befundscanner ist eine Software, die Inhalte medizinischer Dokumente (PDF) in vorbereitete Word-Vorlagen für Gutachten überträgt. Das Produktversprechen ist eng gefasst: **Befundscanner kopiert, es komponiert nicht.** Übernommen wird, was im Quelldokument steht — nicht eine Zusammenfassung, keine Deutung, keine Umformulierung.

Ein solches Versprechen ist für Sachverständige nur dann brauchbar, wenn es überprüfbar ist. Deshalb misst das System jede einzelne Übernahme gegen den erkannten Text des Quelldokuments und schreibt das Ergebnis als Kennzahl fort. Dieses Dokument beschreibt dieses Messverfahren vollständig, damit Dritte es nachvollziehen und im Zweifel auch nachrechnen können.

Geltungsbereich

Die beschriebenen Verfahren gelten für alle Feldwerte, die Befundscanner aus einem hochgeladenen Dokument übernimmt. Sie gelten nicht für Texte, die die Sachverständige oder der Sachverständige selbst im Prüfschritt verfasst oder ändert — solche Eingaben sind Eigenleistung und werden nicht gemessen.

2 Was gemessen wird — und was nicht

Die Trennschärfe dieses Abschnitts ist entscheidend für die Bewertung aller folgenden Zahlen. Die Messung beantwortet genau eine Frage: *Steht der übernommene Wert so, wie er übernommen wurde, auch im Quelldokument?*

Die Messung beantwortet	Die Messung beantwortet nicht
Stammt der Wortlaut aus dem Dokument oder wurde er neu gebildet?	Ist der Wert medizinisch oder rechtlich richtig?
Steht der Wert als zusammenhängender Textblock im Dokument oder wäre er aus verstreuten Einzelwörtern zusammengesetzt?	Wurde der Wert der richtigen Feldposition der Vorlage zugeordnet? Eine Verwechslung zweier Felder bleibt unerkant.
Kommen die Zahlen des Werts zeichengleich im Dokument vor (Schutz gegen Zahlendreher)?	Ist das Quelldokument selbst inhaltlich richtig?
Wie verhält sich das über alle Felder, Praxen und Wochen hinweg?	Hat die Texterkennung den Befund richtig gelesen?

Tabelle 1: Abgrenzung des Messgegenstands.

Warum diese Trennung so streng gezogen wird

Ein Maß, das mehr verspricht, als es leisten kann, ist im gutachterlichen Kontext gefährlicher als gar kein Maß. Wörtlichkeit lässt sich rein rechnerisch und ohne jede Bewertung feststellen. Korrektheit lässt sich das nicht — sie erfordert fachliches Urteil. Befundscanner misst deshalb ausschließlich das Erste und trifft über das Zweite keine Aussage.

3 Ablauf der Messung

Schritt	Vorgang	Ergebnis
1	Texterkennung	Das hochgeladene PDF wird per OCR in Text mit Zeilen- und Positionsangaben überführt. Dieser Text ist ab hier die alleinige Bezugsgröße („Quelltext“).
2	Übernahme der Feldwerte	Für jede Felddefinition der Vorlage wird der zugehörige Wert aus dem Quelltext übernommen.
3	Einordnung in eine Kategorie	Jeder Wert wird als <i>messbar</i> , <i>Datum (ausgenommen)</i> oder <i>leer</i> eingeordnet (Abschnitt 7).
4	Messung	Messbare Werte werden gegen den vollständigen Quelltext des Dokuments gerechnet — nicht gegen den Abschnitt, aus dem sie stammen sollen. Vier Maße werden berechnet (Abschnitt 5).
5	Fortschreibung	Die Ergebnisse werden als reine Zähler und Summen auf die Tagesscheibe der jeweiligen Praxis und Felddefinition addiert. Inhalte werden dabei nicht gespeichert (Abschnitt 8).

Tabelle 2: Ablauf je übernommenem Feldwert.

Die Messung läuft bei jedem Dokument mit. Sie ist keine Stichprobe und keine nachgelagerte Prüfung ausgewählter Fälle: Jeder übernommene Feldwert jedes verarbeiteten Dokuments geht in die Kennzahlen ein. Zugleich ist die Messung vom Übernahmelauf technisch entkoppelt — ein Fehler in der Messung kann die Übernahme selbst nicht beeinflussen oder abbrechen.

4 Vorverarbeitung

Vor jedem Vergleich werden der übernommene Wert und der Quelltext in dieselbe Schreibweise gebracht. Andernfalls würde eine bloße Formatierungsdifferenz — etwa ein Zeilenumbruch mitten im Diagnosetext — fälschlich als Abweichung gezählt.

Kleinschreibung	„Chronische“ → „chronische“
Umlaute und ß	ä→ae ö→oe ü→ue ß→ss
Diakritika	entfernt (Unicode-Zerlegung NFKD)
Leerraum	Umbrüche und Mehrfach-Leerzeichen → ein Leerzeichen
Rand	führender und folgender Leerraum entfällt

Ziffern, Komma und Punkt bleiben unverändert erhalten — sie werden für die Zahlen-Abdeckung (Abschnitt 5.4) benötigt. Entscheidend ist, dass Wert und Quelltext **durch dieselbe Programmfunktion** laufen, nicht durch zwei getrennte, ähnlich gemeinte Implementierungen. Eine auseinanderlaufende Vorverarbeitung wäre die naheliegendste Fehlerquelle für geschönte Werte; sie ist hier konstruktiv ausgeschlossen.

5 Die vier Maße

Der aufbereitete Wert wird in Wörter zerlegt. **S** bezeichnet die Wortfolge des übernommenen Werts, **|S|** ihre Länge in Wörtern, **A** den aufbereiteten Quelltext des gesamten Dokuments.

5.1 N-Gramm-Abdeckung (N = 2, 3, 4)

Ein N-Gramm ist eine Folge von N unmittelbar aufeinanderfolgenden Wörtern. Die Abdeckung ist der Anteil der N-Gramme des Werts, die im Quelltext vorkommen. Das Maß entspricht der in der Textforschung gebräuchlichen *N-Gram Novelty* — hier umgekehrt berichtet, also als Anteil des

Belegten statt des Neuen.

$$\text{Abdeckung}(N) = \frac{\text{Anzahl der N-Gramme von } S, \text{ die in } A \text{ vorkommen}}{\text{Anzahl der N-Gramme von } S}$$

Sonderfall: Ist der Wert kürzer als N, wird die gesamte Wortfolge als ein einzelnes N-Gramm geprüft. Ein Einwort-Wert verschwindet damit nicht aus der Messung, sondern wird weiterhin geprüft.

Drei Längen werden getrennt berichtet, weil sie unterschiedlich streng sind: N = 2 erkennt bereits, ob Wortpaare zusammen vorkommen; N = 4 verlangt, dass jeweils vier Wörter in genau dieser Reihenfolge im Dokument stehen. Ein Wert, der bei N = 2 hoch und bei N = 4 deutlich niedriger liegt, ist ein Hinweis auf eine umgestellte oder gekürzte Wiedergabe.

5.2 Fragment-Abdeckung (Extractive Fragment Coverage)

Nach Grusky et al. (2018) wird die Menge der *extraktiven Fragmente* gebildet: die längsten zusammenhängenden Wortfolgen, die sowohl im Wert als auch im Quelltext vorkommen. Sie werden von links nach rechts über den Wert und ohne Überlappung gebildet. Die Abdeckung ist der Anteil der Wörter des Werts, die durch solche Fragmente belegt sind.

$$\text{Fragment-Abdeckung} = \text{Summe der Fragmentlängen} / |S|$$

5.3 Fragment-Dichte (Extractive Fragment Density)

Dasselbe Fundament, aber mit quadrierten Längen. Dadurch schlagen lange, zusammenhängende Übernahmen stark durch, während viele kurze Treffer kaum Gewicht bekommen. Die Dichte ist **nicht** auf 0 bis 1 normiert; ihr Höchstwert ist die Wortanzahl des Werts.

$$\text{Fragment-Dichte} = \text{Summe der quadrierten Fragmentlängen} / |S|$$

Wertebereich: 0 ... |S| (|S| = ein einziger zusammenhängender Block)

Die Dichte ist das eigentlich aussagekräftige Maß gegen die Sorge, ein Sprachmodell könnte einen Satz aus im Dokument vorhandenen Einzelwörtern *zusammenbauen*. Abschnitt 6 zeigt genau diesen Fall.

5.4 Zahlen-Abdeckung

Aus dem Wert werden alle Zahlen gelesen (Ziffernfolgen, optional mit einem inneren Komma oder Punkt). Die Zahlen-Abdeckung ist der Anteil dieser Zahlen, die **zeichengleich** im Quelltext vorkommen.

$$\text{Zahlen-Abdeckung} = \frac{\text{Anzahl der Zahlen aus } S, \text{ die zeichengleich in } A \text{ stehen}}{\text{Anzahl der Zahlen in } S}$$

Enthält ein Wert keine Zahl, gibt es für ihn keine Zahlen-Abdeckung. Er geht in diese Kennzahl nicht ein — daher ist die Messanzahl hier kleiner als bei den übrigen Maßen.

Dieses Maß bildet einen Risikofall ab, der in Gutachten besonders schwer wiegt: einen Zahlendreher. Wird aus 12,5 im Dokument ein Wert 21,5 in der Übernahme, bleibt der Satz sprachlich unauffällig, die Zahlen-Abdeckung fällt jedoch unmittelbar auf 0.

6 Rechenbeispiele

Beide Beispiele sind Teil der automatisierten Testfälle des Systems und lassen sich mit Papier und Bleistift nachrechnen.

Beispiel A — wörtlich übernommener Block

Quelltext A : ... der patient hat eine chronische bronchitis links basal ...
 Wert S : chronische bronchitis links $|S| = 3$
 2-Gramme von S : „chronische bronchitis“, „bronchitis links“
 beide in A enthalten $\rightarrow$ Abdeckung(2) = $2/2 = 1,00$
 Fragmente : ein Fragment der Länge 3
 Abdeckung = $3/3 = 1,00$
 Dichte = $3^2 / 3 = 3,00$ (= $|S|$, also der Höchstwert)

Ergebnis: vollständig wörtlich, als ein zusammenhängender Block. Die Dichte erreicht ihren Höchstwert $|S|$.

Beispiel B — aus Einzelwörtern zusammengesetzt

Quelltext A : alpha xxx beta yyy gamma zzz delta
 Wert S : alpha beta gamma delta $|S| = 4$
 2-Gramme von S : „alpha beta“, „beta gamma“, „gamma delta“
 keines in A enthalten $\rightarrow$ Abdeckung(2) = $0/3 = 0,00$
 Fragmente : vier Fragmente der Länge 1
 Abdeckung = $(1+1+1+1)/4 = 1,00$
 Dichte = $(1^2+1^2+1^2+1^2)/4 = 1,00$ (statt 4,00)

Warum beide Maße berichtet werden

In Beispiel B kommt jedes einzelne Wort des Werts im Dokument vor. Die Fragment-Abdeckung allein zeigt daher 100 % — und wäre hier irreführend. Erst die Fragment-Dichte macht den Unterschied sichtbar: 1,00 statt des Höchstwerts 4,00. Zusammen mit der N-Gramm-Abdeckung von 0,00 ist ein zusammengesetzter Wert eindeutig als solcher erkennbar. Genau deshalb weist Befundscanner nie eine einzelne Verdichtungszahl aus, sondern immer Abdeckung *und* Dichte nebeneinander.

7 Kategorien und Bänder

7.1 Kategorien

Kategorie	Bedeutung	Begründung
Gemessen	Werte mit Textinhalt. Alle vier Maße werden berechnet.	Bezugsgröße aller Prozentangaben zu den Maßen.
Datum (ausgenommen)	Felder vom Typ Datum. Es wird nur gezählt, nicht gemessen.	Datumsangaben werden auf die Schreibweise TT.MM.JJJJ vereinheitlicht. Ein Dokument, das „04/93“ oder „4. April 1993“ schreibt, würde sonst als unbelegt gelten, obwohl nichts hinzuerfunden wurde.
Leer	Es wurde kein Wert übernommen.	Ohne Wert gibt es nichts zu vergleichen. Solche Fälle als „nicht belegt“ zu zählen, würde die Kennzahl verzerren.

Tabelle 3: Kategorien der Messung.

Wichtig für die Lesart aller Prozentwerte: Die Kennzahlen der vier Maße beziehen sich auf die *gemessenen* Werte, nicht auf alle Werte. Datums- und Leerwerte sind ausgenommen. Die zugehörige Messanzahl wird deshalb zu jeder Kennzahl mit ausgewiesen.

7.2 Bänder

Für die Übersichtsdarstellung werden Einzelwerte in Bänder einsortiert. Die Bänder sind rein beschreibend: Sie sind keine Qualitätsschwellen, keine Ampel und lösen keine Aktion aus.

Abdeckungs-Bänder (N-Gramm und Fragment)	Dichte-Bänder	Wortanzahl des Werts
Sehr hoch — ab 0,95	ab 10	16 Wörter und mehr
Hoch — 0,80 bis unter 0,95	5 bis unter 10	5 bis 15 Wörter
Mittel — 0,50 bis unter 0,80	2 bis unter 5	2 bis 4 Wörter
Niedrig — unter 0,50	unter 2	1 Wort

Tabelle 4: Bandgrenzen. Die Wortanzahl wird mitgeführt, weil die Dichte durch die Wortanzahl des Werts nach oben begrenzt ist und ohne diese Angabe nicht einzuordnen wäre.

8 Erhebung und Datenschutz

Die Messung verarbeitet Gesundheitsdaten im Sinne von Art. 9 DSGVO nur flüchtig im Arbeitsspeicher. Was dauerhaft gespeichert wird, sind ausschließlich Aggregate.

Gespeichert wird	Nicht gespeichert wird
Zähler je Band und Kategorie	der übernommene Feldwert selbst
Summe und Anzahl je Maß (für den exakten Mittelwert)	jeder Textausschnitt aus dem Dokument
Kennung der Praxis	Dokument- oder Auftragskennung, Zeilenverweise
Name der Felddefinition (z. B. „diagnosis“)	jeder Bezug zu einer Patientin oder einem Patienten
Kalendertag (UTC)	Uhrzeit oder Reihenfolge der Einzelmessungen

Tabelle 5: Umfang der gespeicherten Messdaten.

Aus den gespeicherten Zählern lässt sich eine einzelne Übernahme nicht rekonstruieren. Die Fortschreibung erfolgt additiv auf die Tagesscheibe der jeweiligen Praxis und Felddefinition; Wochen-, Monats- und Gesamtwerte entstehen ausschließlich durch Aufsummieren dieser Tageszeilen. Der Zugriff auf die Aggregate ist auf die Betreiberrolle beschränkt und über Mandantentrennung auf Datenbankebene abgesichert.

Die Rechenregeln der Abschnitte 4 bis 6 sind durch automatisierte Tests abgesichert, die bei jeder Änderung am System mitlaufen. Die Testfälle arbeiten mit erfundenen Beispieltextrn und enthalten keine Patientendaten.

9 Ergebnisse aus dem laufenden Betrieb

Die folgenden Werte stammen unverändert aus der Betriebsübersicht des Systems, Stand **28.08.2026**. Berichtszeitraum sind die Kalenderwochen 33 bis 35 des Jahres 2026. Berücksichtigt sind die sieben Felddefinitionen, für die im Zeitraum Messungen vorliegen.

9.1 Messbare Werte

Kategorie	Anteil	Messungen	Bezug
Gemessen	79 %	837	Bezugsgröße für alle Maße in 9.2 und 9.3
Datum (ausgenommen)	21 %	227	nur gezählt, nicht gemessen
Leer	0 %	0	kein Wert übernommen
Gesamt	100 %	1.064	alle übernommenen Feldwerte

Tabelle 6: Verteilung der Kategorien, KW 33–35/2026.

9.2 Mittelwerte je Maß

Exakter Mittelwert aus Summe und Anzahl über alle gemessenen Werte.

Maß	Mittelwert	Messungen	Lesart
N-Gramm-Abdeckung, N = 2	97 %	837	Wortpaare des Werts stehen im Dokument
N-Gramm-Abdeckung, N = 3	94 %	837	Dreierfolgen des Werts stehen im Dokument
N-Gramm-Abdeckung, N = 4	93 %	837	Viererfolgen des Werts stehen im Dokument
Fragment-Abdeckung	97 %	837	Anteil der durch Fragmente belegten Wörter
Fragment-Dichte	16,85	837	Verhältniszahl, nicht Prozent. Werte über 10 liegen im obersten Band und weisen auf lange zusammenhängende Übernahmen hin.
Zahlen-Abdeckung	98 %	449	nur Werte, die überhaupt eine Zahl enthalten

Tabelle 7: Mittelwerte, KW 33–35/2026.

9.3 Werte je Felddefinition

Die letzten drei Spalten geben den Anteil der gemessenen Werte an, die im obersten Abdeckungsband („Sehr hoch“, ab 0,95) liegen.

Felddefinition	Messungen	gemessen	N=2 sehr hoch	N=3 sehr hoch	N=4 sehr hoch
doctor_or_institution	228	100 %	81 %	81 %	81 %
document_date	227	0 %	—	—	—
diagnosis	217	100 %	74 %	74 %	74 %
therapien	158	100 %	96 %	89 %	72 %
anamnese	114	100 %	100 %	100 %	100 %
aktuell	82	100 %	60 %	60 %	60 %
zusammenfassung	38	100 %	71 %	71 %	71 %

Tabelle 8: Kennzahlen je Felddefinition, KW 33–35/2026. „document_date“ ist ein Datumsfeld und daher von der Messung ausgenommen (Abschnitt 7.1).

9.4 Verlauf

Anteil der messbaren Werte je Kalenderwoche.

Kalenderwoche	Messungen	Anteil messbar	Anmerkung
2026-W33	894	79 %	—
2026-W34	54	78 %	geringe Fallzahl, entsprechend unsicher
2026-W35	116	78 %	Woche zum Stichtag noch nicht abgeschlossen

Tabelle 9: Verlauf über den Berichtszeitraum.

10 Grenzen der Aussagekraft

Die folgenden Einschränkungen gehören untrennbar zu den Zahlen in Abschnitt 9. Wer sie zitiert, sollte sie mitzitieren.

Nr.	Einschränkung
1	Wörtlichkeit ist nicht Korrektheit. Ein Wert kann alle vier Maße mit Höchstwerten durchlaufen und dennoch an der falschen Stelle des Gutachtens stehen. Feldverwechslungen sind mit diesem Verfahren grundsätzlich nicht erkennbar.
2	Bezugsgröße ist der erkannte Text, nicht das Papier. Gemessen wird gegen das Ergebnis der Texterkennung. Liest die Texterkennung eine Passage falsch, gilt dieser Text für die Messung als Quelle. Solche Fehler senken die Kennzahlen tendenziell, sie erzeugen aber keine falsch hohen Werte.
3	Datums- und Leerwerte sind ausgenommen. Rund ein Fünftel aller Werte im Berichtszeitraum sind Datumsfelder und gehen nicht in die Maße ein. Alle Prozentwerte der Abschnitte 9.2 und 9.3 beziehen sich auf die gemessenen Werte.
4	Es handelt sich um einen Betriebsstand, nicht um eine Studie. Der Berichtszeitraum umfasst drei Kalenderwochen und 1.064 Messungen aus dem Beta-Betrieb. Werte für Felddefinitionen mit kleiner Messanzahl — etwa „zusammenfassung“ mit 38 Messungen — sind statistisch entsprechend unsicher.
5	Keine externe Validierung. Die Kennzahlen entstehen im System selbst. Ein Vergleich gegen einen von Menschen erstellten Referenzdatensatz hat nicht stattgefunden. Ein solcher Vergleich wäre die Voraussetzung dafür, Aussagen über Korrektheit zu treffen — er ist nicht Gegenstand dieses Dokuments.
6	Die Bänder sind keine Qualitätsurteile. Die Bezeichnungen „Sehr hoch“ bis „Niedrig“ beschreiben Zahlenbereiche. Sie definieren keine Freigabeschwelle und ersetzen keine Prüfung.

Die Rolle der Sachverständigen bleibt unberührt

Befundscanner nimmt der Sachverständigen und dem Sachverständigen das Abtippen ab, nicht das Prüfen. Jeder übernommene Wert wird im Prüfschritt neben dem Originaldokument angezeigt und kann geändert oder ausgeblendet werden. Die Verantwortung für den Inhalt des Gutachtens liegt unverändert und vollständig bei der bestellten Sachverständigen bzw. dem bestellten Sachverständigen.

11 Quellen und Kontakt

Verwendete Verfahren

- Grusky, M., Naaman, M., Artzi, Y. (2018): *Newsroom — A Dataset of 1.3 Million Summaries with Diverse Extractive Strategies*. Proceedings of NAACL-HLT 2018, S. 708–719. Quelle der Extractive Fragment Coverage und Density; die Fragmentbildung folgt dem dort beschriebenen Algorithmus.
- N-Gram Novelty: in der Textzusammenfassungsforschung etabliertes Maß für den Anteil neu gebildeter Wortfolgen. In diesem Dokument umgekehrt als Abdeckung berichtet.

Kontakt und Rückfragen

Herausgeber	Pixel7 GmbH, Wien, Österreich
Produkt	Befundscanner · app.befundscanner.at
Fachliche Rückfragen	piere@pixel7.at
Dokument	BS-MD-001 · Version 1.0 · Stand 28.08.2026
Gültigkeit	Die Verfahrensbeschreibung (Abschnitte 1-8) gilt bis zu einer ausgewiesenen neuen Version. Die Zahlen in Abschnitt 9 sind ein Stichtagswert und werden mit jeder Version neu erhoben.

Dieses Dokument beschreibt ein Messverfahren und dessen Ergebnisse. Es ist keine medizinische, rechtliche oder gutachterliche Stellungnahme und begründet keine Zusicherung inhaltlicher Richtigkeit übernommener Werte.