

Data Quality Score: Messgröße für die Zuverlässigkeit von Stammdaten im Einkauf

Der Data Quality Score ist eine Kennzahl, die die Qualität von Datenbeständen – insbesondere von Artikelstammdaten, Lieferantendaten und Transaktionsdaten – auf einer Skala bewertet. Für Einkäufer und Supply-Chain-Verantwortliche ist er ein zentrales Steuerungsinstrument, denn fehlerhafte Stammdaten führen zu Fehlbestellungen, falschen Lieferdaten und unnötigen Nachfragen. Ein systematisch erhobener Score macht Datenprobleme sichtbar, quantifizierbar und damit priorisierbar. Im Lexikonbeitrag erklären wir Definition, Berechnungsmethoden, KPIs und Praxismaßnahmen rund um den Data Quality Score.

Das Wichtigste in Kürze

- Der Data Quality Score quantifiziert die Datenqualität typischerweise in Prozerten oder auf einer Skala von 0 bis 100.
- Häufig verwendete Dimensionen sind Vollständigkeit, Korrektheit, Konsistenz, Aktualität, Eindeutigkeit und Plausibilität.
- In der Praxis gelten Werte über 95 Prozent als guter Richtwert für kritische Stammdatenfelder wie Einkaufspreise oder Materialnummern.
- Fehlerhafte Stammdaten verursachen typischerweise Nachfragen, Fehlbestellungen, Mahnlauffee und Retouren entlang der gesamten Source-to-Pay-Kette.
- Ein wirksamer Score benötigt klare Feldprioritäten, definierte Validierungsregeln und einen Verantwortlichen je Datenobjekt.
- Data Quality Scores lassen sich automatisiert ermitteln und in Dashboards laufend überwachen.

Definition: Data Quality Score

Der Data Quality Score ist eine aggregierte Kennzahl, die die Qualität einzelner Datensätze, Datenfelder oder gesamter Datenbestände bewertet. Er fasst mehrere Qualitätsdimensionen zu einem Wert zusammen und ermöglicht so einen schnellen Vergleich über Objekte, Teams oder Systeme hinweg.

Abgrenzung zu Datenqualität im Allgemeinen

Datenqualität beschreibt den fachlichen Zustand von Daten anhand von Kriterien wie Vollständigkeit oder Korrektheit. Der Data Quality Score macht daraus eine messbare, vergleichbare Zahl. Während Datenqualität ein Zielzustand ist, ist der Score ein Steuerungsinstrument: Er zeigt, wo der Ist-Zustand liegt, wie er sich über die Zeit entwickelt und wo Maßnahmen ansetzen müssen. Ohne Score bleibt Datenqualität häufig eine diffus wahrgenommene Unzulänglichkeit; mit Score wird sie zu einem steuerbaren Kennzahlenbereich neben Liefertreue oder Beschaffungskosten.

Relevante Datenobjekte im Einkauf

Im Fokus stehen für Einkäufer vor allem Artikelstammdaten mit Materialnummern, Maßeinheiten,

technischen Merkmalen und Einkaufspreisen sowie Lieferantenstammdaten mit Anschriften, Bankverbindungen und Zertifizierungen. Dazu kommen Transaktionsdaten aus Bestellungen, Lieferavisen und Rechnungen. Fehler in diesen Objekten wirken sich unmittelbar aus: Eine falsche Maßeinheit führt zur Überbestellung, eine veraltete Bankverbindung gefährdet die Zahlung und ein unvollständiger Lieferantenstatus erschwert die Risikobewertung. Gerade bei der Lieferantenbankdatenprüfung und im Kreditoren-Onboarding ist eine hohe Datenqualität Voraussetzung für reibungslose Prozesse.

Methoden und Vorgehen zur Score-Ermittlung

Die Berechnung eines Data Quality Scores folgt in der Regel einem festen Ablauf: Definition der Qualitätsdimensionen, Festlegung der Bewertungsregeln je Feld, Gewichtung nach Geschäftsrelevanz und Aggregation zu einem Gesamtwert. Je nach Reifegrad erfolgt die Bewertung manuell, regelbasiert oder automatisiert mit Data-Profiling- und Matching-Tools.

Qualitätsdimensionen festlegen

Bewährt ist die Orientierung an sechs Dimensionen: Vollständigkeit (alle Pflichtfelder gefüllt), Korrektheit (Werte stimmen mit der Realität überein), Konsistenz (keine Widersprüche zwischen Systemen), Aktualität (Zeitstempel und Revisionen stimmen), Eindeutigkeit (keine Dubletten) und Plausibilität (Werte in sinnvollen Bereichen). Für das Dubletten-Management ist die Eindeutigkeit zentral, während die Aktualität etwa bei Lieferantenstandort-Risiko-Bewertungen entscheidend ist, da Standortdaten regelmäßig geprüft werden müssen. Jede Dimension wird je Objekttyp einzeln bewertet.

- Vollständigkeit: Anteil gefüllter Pflichtfelder
- Eindeutigkeit: Ergebnis der Dublettenprüfung
- Plausibilität: Wertebereichs- und Referenzprüfung

Regelwerk und Gewichtung

Jedes Datenfeld erhält eine Validierungsregel (z. B. Pflichtfeld, Format, referenzielle Integrität, Schwellenwert) und eine Priorität. Kritische Felder wie Einkaufspreis, Maßeinheit oder Lieferantenummer erhalten ein höheres Gewicht als optionale Beschreibungsfelder, weil Fehler dort direkte operative Folgen haben. Üblich ist eine Gewichtung etwa im Verhältnis 70:20:10 zwischen Kern-, Standard- und Zusatzfeldern. Die Gewichtung sollte von den Fachbereichen Einkauf, Controlling und IT gemeinsam festgelegt und periodisch überprüft werden, da sich Geschäftsrelevanz verschieben kann.

Automatisiertes Profiling und Audits

In der Praxis kombinieren Unternehmen automatisierte Regelprüfungen mit stichprobenartigen fachlichen Audits: Software erkennt Lücken, Formatfehler und Dubletten tageslich, während sachliche Korrektheit (etwa ob ein technisches Merkmal zum Bauteil passt) nur durch Fachprüfung verifizierbar ist. Analog zur Messsystemanalyse Gage R&R aus der Qualitätssicherung empfiehlt sich gelegentlich zu prüfen, wie verlässlich die Prüfmethode selbst ist, also ob verschiedene Prüfer zum selben Ergebnis kommen. Nur so ist der Score selbst belastbar.

Wichtige KPIs rund um den Data Quality Score

Der Data Quality Score ist selbst eine Kennzahl, entfaltet aber erst in Kombination mit weiteren Metriken seine Steuerungswirkung. Sinnvoll ist ein kleines Kennzahlenset, das Score-Werte, Fehlerquoten und Geschäftsauswirkungen verbindet.

Score-Werte je Objekttyp und Dimension

Berichtet wird typischerweise getrennt nach Artikelstammdaten, Lieferantenstammdaten und Transaktionsdaten sowie nach Dimensionen. Ein Gesamt-Score für das Unternehmen ist kommunikativ nützlich, steuernd sind aber die Detailwerte: So kann eine hohe Vollständigkeit von 98 Prozent neben einer schwachen Aktualität von 82 Prozent stehen – die Maßnahmen unterscheiden sich deutlich. Als Richtwerte gelten in der Praxis für kritische Felder Werte über 95 Prozent, für Nebendaten über 85 bis 90 Prozent als akzeptabel.

Verknüpfung mit Prozesskennzahlen

Wirkung zeigt der Score erst im Zusammenhang mit Prozesskennzahlen: Anteil von Bestellungen mit Preis- oder Mengenabweichung, Rate abgelehnter Rechnungen wegen Stammdatenfehlern, Anteil doppelter Lieferanten oder Kreditoren sowie Nachfragen infolge unvollständiger Anfragedaten. Auch die Lieferterminverfolgung leidet unter unpräzisen Termin- und Mengenangaben aus Bestelldaten. Wer diese Kennzahlen gemeinsam auswertet, kann den wirtschaftlichen Nutzen von Datenbereinigung belegen und Investitionen in Datenpflege begründen.

- Fehlerrate in Bestellungen je 1.000 Positionen
- Anteil abgelehnter Rechnungen wegen Stammdatenfehlern
- Dublettenquote in Lieferanten- und Kreditorenbeständen

Trend- und Benchmark-Betrachtung

Einzelwerte sagen wenig; entscheidend sind Entwicklungen. Ein monatlich oder quartalsweise berichteter Trend zeigt, ob Bereinigungsprojekte wirken oder ob neue Fehlerquellen – etwa durch Fusionen, Systemwechsel oder schnell wachsende Kataloge – den Score drücken. Ein interner Vergleich zwischen Standorten, Warengruppen oder Teams schafft gesunden Wettbewerb um Datenqualität. Ein externer Lieferantenbenchmark der Datenlieferqualität, etwa bei Lieferavis-Verarbeitung (ASN), ergänzt das Bild um die Sicht auf Geschäftspartner.

Risiken, Ursachen und Kontrollen

Schlechte Datenqualität entsteht selten böswillig, sondern durch unklare Verantwortlichkeiten, fehlende Validierung, Schnittstellenverluste und Organisationsveränderungen. Der Data Quality Score macht diese Ursachen messbar –Controllen müssen sie systematisch begegnen.

Typische Ursachen im Einkaufsumfeld

Häufige Ursachen sind manuelle Pflege ohne Pflichtfeldlogik, mehrere quellenführende Systeme ohne klare Führung, fusions- und migrationsbedingte Dubletten sowie unzureichendes Anforderungsmanagement bei neuen Materialien und Lieferanten. Gerade im Kreditoren-Onboarding entstehen Fehler, wenn Eingangsdaten nicht strukturiert validiert werden. Auch Unternehmenswachstum wirkt: Neue Standorte, zusätzliche Lieferantenstandorte und expandierende Lohnbearbeitung vergrößern den Datenbestand schneller als die Pflegekapazität.

- Fehlende eindeutige Datenverantwortung (Data Ownership)
- Schnittstellenverluste zwischen ERP, SRM und Katalogsystemen
- Migrations- und Fusionsdubletten ohne laufende Dublettenprüfung

Kontrollen und Prozessmaßnahmen

Wirksame Kontrollen setzen an der Quelle an: Validierungsregeln bereits bei der Anlage von Material

und Lieferant, Pflichtfeldkonzepte im Anforderungswesen, Vier-Augen-Prinzip bei kritischen Feldern wie Bankverbindungen und periodische Bereinigungsläufe. In den Source-to-Pay (S2P)-Prozess integrierte Qualitätschecks – etwa beim Bestelleingang und vor Freigabe von Rechnungen – verhindern, dass Fehler weiter propagieren. Data Quality Scores gehören auf die Agenda regelmäßiger Einkaufsreviews, ähnlich wie Lieferantenstatus und Risikobewertungen.

- Validierungsregeln direkt in den Anlageprozessen
- Vier-Augen-Prüfung kritischer Zahlungs- und Preisdaten
- Quartalsweise Bereinigungs- und Dublettenprüfungsläufe

Organisation und Verantwortung

Dauerhaft gute Datenqualität erfordert klare Rollen: Ein Data Owner je Datenobjekt entscheidet über Definitionen und Prioritäten, Data Stewards pflegen operativ, die IT betreibt Werkzeuge und Schnittstellen. Ein Notfallteam Einkauf kann bei akuten Datenproblemen – etwa nach Systemausfällen oder Fehlmigrationen – kurzfristig priorisieren und korrigieren. Die Schulung der beteiligten Mitarbeiter, verankert etwa in einer Kompetenzmatrix Einkauf, stellt sicher, dass Regeln verstanden und angewendet werden.

Data Quality Score in der Praxis des industriellen Einkaufs

Der Score entfaltet seinen Nutzen entlang konkreter Einkaufsprozesse – von der Anfrage über das Lieferantenmanagement bis zur Zentralisierung von Stammdaten im Lieferantennetzwerk. Ein hoher Score ist dabei Voraussetzung für weitere Digitalisierungsprojekte.

Anfragen, Vergabe und Lieferantenmanagement

Vollständige und korrekte Stammdaten beschleunigen Anfrageerstellung und Vergabe: Zeichnungsreferenzen, Stücklisten und Lieferbedingungen stimmen, Rückfragen sinken. Im Bieterregister und bei der Lieferantenauswahl helfen gepflegte Zertifikats- und Kapazitätsdaten, geeignete Anbieter schneller zu identifizieren. Bei Make-or-Buy-Entscheidungen und in der Formenbau-Beschaffung stützen sich Rechenmodelle auf korrekte Stücklisten und Prozessdaten – Datenfehler führen hier zu verzerrten Kostenvergleichen und falschen Vor- oder Nachteilen.

Operativer Betrieb und Logistik

Im operativen Geschäft wirken sich Datenfehler direkt auf Lieferfähigkeit und Terminen aus: Falsche Wiederbeschaffungszeiten verfälschen die Safety Stock Calculation, unpräzise Abrufdaten führen zu Fehlmengen beim Abrufauftrag und inkonsistente Verpackungs- und Avisdaten erschweren die Lieferavis-Verarbeitung (ASN). Auch der Lieferplan und die Lieferterminverfolgung setzen voraus, dass Mengen, Einheiten und Termine in den Systemen stimmen. Hohe Datenqualität reduziert hier Mahnlaufee, Feuerwehractionen und teure Expresslieferungen.

- Korrekte Wiederbeschaffungszeiten als Basis für Safety Stock Calculation
- Sauberer Avisdatenfluss zur stabilen ASN-Verarbeitung
- Stimmige Abrufdaten für zuverlässige Lieferpläne

Finanzprozesse und Zahlungsverkehr

Im Rechnungs- und Zahlungswesen ist Datenqualität unmittelbar messbar: Dublettenrechnungen, falsche Bankverbindungen oder abweichende Kreditorennamen führen zu Prozessstopps und Compliance-Risiken. Beim Gutschriftverfahren (ERS), bei dem ohne klassische Rechnung aus

Wareneingangsdaten fakturiert wird, sind korrekte Preis- und Mengendaten zwingende Voraussetzung – hier schlägt jeder Datenfehler direkt in Zahlungsdifferenzen durch. Auch bei Themen wie Forfaitierung und Saldomanagement setzen Finanzprozesse auf saubere Kreditoren- und Transaktionsdaten auf.

Aktuelle Entwicklungen: Automatisierung, KI und Datenökosysteme

Data Quality Scores werden zunehmend automatisiert erhoben und laufend überwacht. Maschinelles Lernen erkennt Anomalien in Stammdaten, Matching-Verfahren identifizieren Dubletten über Systemgrenzen hinweg, und Dashboards machen Scores für Fachbereiche ohne Datenbankkenntnisse nutzbar.

KI-gestützte Prüfung und Anreicherung

Sprachmodelle und Klassifikationsverfahren helfen, freitextliche Materialbeschreibungen zu vereinheitlichen, Artikel zu kategorisieren und fehlende Attribute aus Dokumenten zu extrahieren. Damit sinkt der manuelle Pflegeaufwand deutlich, und Dubletten, die sich nur über Bedeutungsähnlichkeit erkennen lassen, werden auffindbar. Wichtig bleibt die Kontrolle: KI-Vorschläge sollten als Review-Vorlage dienen und von Fachbereichen freigegeben werden, damit sich keine systematischen Fehlannahmen in die Stammdaten einschleichen.

Datenqualität über Unternehmensgrenzen hinweg

Im Lieferantennetzwerk und in integrierten Source-to-Pay (S2P)-Plattformen entsteht ein gemeinsames Datenverständnis zwischen Kunde und Lieferant: Stammdaten werden einmal gepflegt und mehrfach genutzt. Das verschiebt die Verantwortung für Datenqualität teilweise in die Wertschöpfungskette. Einkäufer sollten daher Datenqualitätsanforderungen in Lieferantenvereinbarungen verankern – etwa beim Dienstvertrag mit einem Entwicklungs- oder Logistikpartner oder bei langfristigen Rahmenverträgen – und die Einhaltung über Score-ähnliche Kennzahlen regelmäßig prüfen.

Regulatorische Anforderungen

Zunehmende Berichtspflichten, etwa zu Lieferketten und Nachhaltigkeit, erhöhen den Anspruch an Lieferantendaten: Zertifikate, Standorte, Herkunftsnachweise und Kennzahlen müssen aktuell, vollständig und prüfbar sein. Ein niedriger Data Quality Score bei diesen Feldern wird damit zum Compliance-Risiko. Rechtliche Anforderungen im Detail sind stark kontextabhängig und sollten im Einzelfall – etwa mit der Rechts- oder Compliance-Abteilung – geprüft werden. Der Score liefert hier ein Frühwarninstrument, das Lücken rechtzeitig sichtbar macht.

Typische Kennzahlen

Data Quality Score Kernfelder: > 95 %

Richtwert für kritische Stammdatenfelder wie Preise, Einheiten, Bankdaten

Data Quality Score Nebendaten: > 85–90 %

Akzeptabler Bereich für optionale Beschreibungs- und Zusatzfelder

Dublettenquote: < 1 %

Zielwert für Lieferanten- und Kreditorenbestände nach Bereinigung

Rechnungsablehnungen wegen Stammdatenfehlern: < 2 %

Indikator für den Prozessnutzen hoher Datenqualität

Score-Trend: monatlich steigend

Trend wichtiger als Einzelwerte; Rückschritte zeigen neue Fehlerquellen

Praxisbeispiel

Ein mittelständischer Maschinenbauer mit rund 4.000 aktiven Lieferanten bemerkte bei der Einführung eines automatisierten Rechnungsabgleichs zunehmend Prozessstopps: Preisabweichungen, Dublettenrechnungen und falsche Bankverbindungen. Eine erste Score-Messung zeigte für Lieferantenstammdaten nur 78 Prozent Qualität, vor allem wegen veralteter Bankdaten und Dubletten aus einer früheren ERP-Migration. Das Unternehmen etablierte ein strukturiertes Bereinigungsprogramm mit Data Ownership im Einkauf und quartalsweisen Regelprüfungen.

- Score-Messung je Objekttyp mit gewichteten Feldprioritäten, Start mit Lieferanten- und Artikelstammdaten
- Laufende Dublettenprüfung und Lieferantenbankdatenprüfung mit Vier-Augen-Prinzip bei kritischen Feldern
- Validierungsregeln und Pflichtfelder direkt im Kreditoren-Onboarding verankert
- Monatliches Score-Reporting an Einkauf und Controlling, Zielwert über 95 Prozent für Kernfelder

Fazit

Der Data Quality Score übersetzt einen abstrakten Anspruch – gute Daten – in eine steuerbare Kennzahl. Er macht Fehlerquellen sichtbar, ermöglicht Priorisierung nach geschäftlicher Relevanz und liefert den Nachweis, dass Bereinigungsmaßnahmen wirken. Im Einkauf zahlt sich das unmittelbar aus: weniger Nachfragen, stabilere Liefertermine, reibungslosere Rechnungsprozesse und belastbare Entscheidungen bei Make-or-Buy oder Lieferantenauswahl. Entscheidend sind klare Verantwortlichkeiten, Validierung an der Quelle und regelmäßiges Reporting. Wer Datenqualität messbar macht, schafft die Grundlage für Digitalisierung, Automatisierung und belastbare Beschaffungskennzahlen.

Häufige Fragen

Was ist ein Data Quality Score?

Ein Data Quality Score ist eine Kennzahl, die die Qualität von Daten – etwa Artikel- oder Lieferantenstammdaten – als aggregierten Wert ausdrückt, typischerweise in Prozent oder auf einer Skala von 0 bis 100. Er fasst Dimensionen wie Vollständigkeit, Korrektheit, Konsistenz, Aktualität und Eindeutigkeit zusammen. Damit lassen sich Datenbestände vergleichen, Entwicklungen verfolgen und Bereinigungsmaßnahmen priorisieren.

Wie wird ein Data Quality Score berechnet?

Zunächst werden Datenfelder mit Validierungsregeln definiert (Pflichtfeld, Format, Wertebereich, Referenz) und nach Geschäftskritikalität gewichtet. Anschließend wird je Feld geprüft, ob es die Regel erfüllt, und die Ergebnisse werden zum Feld-, Datensatz- und Bestandsscore aggregiert. Die Gewichtung legen Fachbereich und IT gemeinsam fest; automatisiertes Data Profiling übernimmt die laufende Prüfung.

Welche Datenqualitätsdimensionen gibt es?

Üblich sind sechs Dimensionen: Vollständigkeit, Korrektheit, Konsistenz, Aktualität, Eindeutigkeit und

Plausibilität. Vollständigkeit und Eindeutigkeit lassen sich gut automatisiert prüfen, etwa durch Dublettenprüfung. Korrektheit im fachlichen Sinn erfordert dagegen meist menschliche Prüfung durch die zuständigen Fachbereiche.

Ab wann gilt ein Data Quality Score als gut?

In der Praxis gelten für kritische Stammdatenfelder wie Einkaufspreise, Materialnummern oder Bankverbindungen Werte über 95 Prozent als guter Richtwert. Für Nebendaten sind häufig 85 bis 90 Prozent akzeptabel. Wichtiger als Einzelwerte ist der Trend: Sinkende Scores zeigen neue Fehlerquellen, steigende Werte belegen wirksame Maßnahmen.

Welche Folgen haben schlechte Stammdaten im Einkauf?

Fehlerhafte Stammdaten führen zu Fehlbestellungen, falschen Maßeinheiten, abgelehnten Rechnungen und unnötigen Nachfragen. Sie verfälschen auch die Sicherheitsbestandsberechnung und die Lieferterminverfolgung, was Expresslieferungen und Mahnlaufee nach sich ziehen kann. Darüber hinaus entstehen Compliance-Risiken, etwa bei veralteten Bankdaten oder fehlenden Zertifikaten im Lieferantenmanagement.

Wie kann man die Datenqualität im Einkauf verbessern?

Bewährt sind Validierungsregeln direkt bei der Anlage von Material und Lieferant, klare Data Ownership, Pflichtfeldkonzepte im Anforderungswesen und periodische Bereinigungsläufe einschließlich Dublettenprüfung. Ergänzend helfen automatisiertes Profiling, Vier-Augen-Prüfung kritischer Felder und ein monatliches Score-Reporting. Schulungen der beteiligten Mitarbeiter sichern, dass Regeln im Alltag angewendet werden.