

X



(https://dataladder.com/de/)



BLOG

# Grundlagen der Datenbereinigung – Wie man mit schlechten Daten einfach umgeht

·donnees-comment-traiter-

»-datos-como-tratar-los-



Written by **lbarrera**(<https://dataladder.com/de/author/lbarrera/>) Mai 29, 2020

Ein Datenbereinigungstool ist vielleicht die leistungsfähigste und doch am meisten unterschätzte Lösung. Unternehmen geben Millionen von Dollar für die Anschaffung von Cloud-Lösungen und teuren Datenbanken aus, schrecken aber zurück, wenn es um den Kauf einer Datenbereinigungslösung geht.

Wir haben die Erfahrung gemacht, dass über 80 % der Unternehmen, mit denen wir zusammengearbeitet haben, nicht wussten, welche Funktionen ein Datenbereinigungstool hat und wie es ihnen bei der Bewältigung ihrer Datenqualitätsprobleme helfen kann.

Daher werden wir in diesem kurzen Beitrag heute eine kurze Erklärung geben und häufig gestellte Fragen beantworten:

- Was sind Datenbereinigungstools?
- Wie führen Sie die Datenbereinigung durch?
- Was verstehen Sie unter Datenbereinigung?
- Warum ist das wichtig?

Fangen wir an.

## Was ist eine Datenbereinigungssoftware oder ein Tool?

Bevor wir über das Tool sprechen, sollten wir über das Problem der schlechten Daten sprechen.

Wenn Sie jemals einen Blick auf die Kundendaten Ihres Unternehmens geworfen haben, werden Sie feststellen, dass diese buchstäblich unordentlich sind.

Hier ist ein Beispiel:

X

✓

✓

✓

✓

Zu dem Zeitpunkt, an dem Sie diese Daten für Marketing-, Werbe- oder strategische Entscheidungszwecke benötigen, verschwenden Sie wichtige Zeit mit der Behebung dieser Probleme. Schlimmer noch: Vielleicht können Sie die Daten nicht einmal nutzen.

Es ist nicht möglich, diese Fehler manuell zu beheben. Natürlich können Sie es der IT-Abteilung überlassen, aber diese wird das Problem nur noch schlimmer machen. Die IT-Abteilung kennt die Sensibilität und die Natur der Daten nicht so gut wie die Geschäftsanwender. Letztendlich verbringen Sie mehr Zeit als nötig mit der Koordinierung mit der IT-Abteilung, der Überprüfung von Änderungen und verzögern Ihre Ziele.

Hier kommt der Bedarf an einer Datenbereinigungssoftware (<https://dataladder.com/de/datenbereinigungssoftware-schnelles-und-kostenguenstiges-crm-datenbereinigungstool/>) ins Spiel.

Ein Datenbereinigungstool (<https://dataladder.com/de/der-vollstaendige-leitfaden-fuer-datenbereinigungstools-loesungen-und-bewaehrte-praktiken-fuer-die-unternehmensebene/>) ist eine einfach zu bedienende Lösung für Geschäftsanwender. Es handelt sich um eine wichtige Software, mit der Sie alle oben genannten Probleme der Datenqualität beheben können. Eine erstklassige Datenbereinigungssoftware wie DataMatch Enterprise (<https://dataladder.com/de/produkte/datamatch-enterprise-bewertetes-datenqualitaetsmanagementprodukt-nr-1/>) kann jedoch viel mehr als nur bereinigen – sie ermöglicht es Ihnen, Duplikate aus mehreren Datenquellen (<https://dataladder.com/de/zusammenfuehrung-von-daten-aus-mehreren-quellen-herausforderungen-und-loesungen/>) zu entfernen, Daten miteinander abzugleichen (<https://dataladder.com/de/datenabgleichssoftware-als-best-in-class-mit-96-match-genauigkeit-bewertet/>), Daten zusammenzuführen (<https://dataladder.com/de/merge-purge-software-integrierte-und-benutzerdefinierte-ueberlebensregeln-verwenden/>), zu standardisieren (<https://dataladder.com/de/datenstandardisierungssoftware-schnelles-und-kostenguenstiges-datenstandardisierungstool/>) und Ihre aktuellen Daten zu optimieren.

All dies können Sie **\*ohne\* die** Hilfe von IT-Ressourcen, direkt auf Ihrem Desktop erledigen. Auch das Bereinigen und Sortieren von einer Million Datenzeilen wird weniger als eine Stunde in Anspruch nehmen. Das spart Ihnen Zeit, Mühe und Geld.

## Wie führen Sie die Datenbereinigung durch?

Traditionell wurde die Datenbereinigung manuell durchgeführt. In einigen Unternehmen gibt es auch heute noch Mitarbeiter, die sich ausschließlich damit beschäftigen, Daten zu extrahieren, sie in mehrere Segmente aufzuteilen und Excel-Funktionen zum Filtern und Aussortieren von Duplikaten oder Ungenauigkeiten auszuführen.

Unternehmen, die SQL verwenden, haben spezielle Programmierer und Ressourcen, die die Sprache kennen, um Einträge zu bereinigen. Dadurch bleiben die Geschäftsanwender außen vor und sind von der Zeitplanung des SQL-Programmierers abhängig.

Während eine Datenbereinigungssoftware Ihnen hilft, zwei Datensätze zu verbinden, müssen Sie in SQL Abfragen wie die folgenden ausführen, um zwei Tabellen zusammenzuführen.

X

✓

✓

✓

✓

**select id, name, abt\_name from****student\_metadata s join department\_details d****on s.abt\_id = d.abt\_id;***Quelle: DataCamp (<https://www.datacamp.com/community/tutorials/cleaning-data-sql>)*

Das ist nicht nur extrem zeitaufwändig, sondern auch unpraktisch, wenn es darum geht, komplexe Datenprobleme zu lösen.

### **Wie man mit schlechten**

Im Zeitalter von Big Data sind die Dinge kompliziert.

Ihre CRM-Daten bestehen nicht mehr nur aus ein paar hundert Zeilen mit grundlegenden Kontaktinformationen. Heute können sie aus Verhaltensdaten, sozialen Daten, digitalen Daten, Transaktionsdaten und vielem mehr bestehen. Je nach den Anforderungen Ihres Unternehmens verfügen Sie möglicherweise über noch komplexere Datensätze, die mit den herkömmlichen Methoden nicht bereinigt werden können und fortschrittliche CRM-Datenbereinigungstechniken (<https://dataladder.com/de/datenbereinigungssoftware-schnelles-und-kostenguenstiges-crm-datenbereinigungstool/>) erfordern. Lesen Sie mehr darüber, wie Sie Ihre Kundendatenbank aufräumen

Große multinationale Marken verwenden teure Lösungen wie Informatica, Oracle oder IBM, um ihre Datenqualitätsprobleme zu lösen – aber nicht jedes Unternehmen hat das Budget, um eine Lösung von diesen Anbietern zu kaufen. Allein die Lösung von Oracle kostet über 200.000 Dollar, und Sie müssen mehrere Tools verwenden, um die gewünschten Ergebnisse zu erzielen.

Für mittelständische Unternehmen gibt es also kaum andere Möglichkeiten, als zu versuchen, die Daten manuell zu bereinigen oder Dritte damit zu beauftragen, die Excel-Tabellen und deren Funktionen zur Bereinigung grundlegender Datenprobleme verwenden. Aus diesem Grund ist eine Lösung wie DataMatch Enterprise die perfekte Lösung für Unternehmen, die ein Tool auf ihrem Desktop- oder Cloud-Server haben möchten, um die Arbeit zu erledigen, ohne all die Komplikationen und den unnötigen Schnickschnack, der mit anderen Lösungen einhergeht.

## **Was bedeutet Datenbereinigung genau?**

Bereinigung Ihrer Daten von Rechtschreibfehlern, Lösung von Formatproblemen, Deduplizierung von doppelten Daten (<https://dataladder.com/de/datendeduplizierungssoftware-verwenden-sie-integrierte-und-benutzerdefinierte-regeln-fuer-die-crm-deduplizierung/>) und Sicherstellung fehlerfreier Daten.

Inzwischen wissen Sie, dass Sie unordentliche Daten haben und dass diese bereinigt werden müssen. Sie wissen auch, wie traditionelle Methoden zur Datenbereinigung eingesetzt werden. Aber wie wird die Datenbereinigung bei einer Software durchgeführt?

Lassen Sie mich dies anhand des Rahmens von DataMatch Enterprise erklären.

Erstellen Sie **ein Profil Ihrer Daten, um Probleme zu erkennen**: Sie können Daten nicht bereinigen, wenn Sie nicht wissen, was mit ihnen los ist. Sie können Excel-Funktionen verwenden, um unvollständige Zeilen und Spalten zu identifizieren, aber Sie können sie nicht verwenden, um bestimmte Spalten zu markieren, die nicht druckbare Zeichen enthalten oder die Zahlen in Namensfeldern und Buchstaben in Zahlenfeldern enthalten. Dies sind Probleme, die oft nicht ins Auge fallen, aber zu großen Hindernissen werden, wenn Sie die Daten verwenden müssen.

Der erste Teil der Datenbereinigung besteht also darin, die Probleme zu identifizieren, die Ihre Daten betreffen. Sobald Sie in der Lage sind, Probleme zu erkennen, können Sie mit der Reinigung oder dem

Schrubben (<https://dataladder.com/de/datenbereinigungssoftware-bewertetes-nr-1-tool-zum-bereinigen-von-daten/>) beginnen.

**Umgang mit Duplikaten:** Unordentliche Daten sind ein leicht zu bewältigendes Problem. Das eigentliche Problem sind tiefe Duplikate. Mit zunehmender Komplexität der Daten nehmen auch die Duplikate zu. Jedes Mal, wenn ein Kunde eine andere E-Mail, Nummer oder Adresse eingibt, wird ein Duplikat erstellt. Wenn Sie außerdem Kundendaten aus verschiedenen Abteilungen kombinieren, erhalten Sie Duplikate.

Einige Duplikate lassen sich leicht aussortieren, die meisten jedoch nicht. Sie benötigen Lösungen, die eine Kombination von Fuzzy-Matching-Algorithmen (<https://dataladder.com/de/fuzzy-matching-101-bereinigung-und-verknuepfung-ungeordneter-daten/>) verwenden, um wahrscheinliche Duplikate zu identifizieren (Cath und Catherine können dieselben Personen sein, aber sie haben nicht die exakten Namen und werden als wahrscheinliche Duplikate betrachtet). DataMatch Enterprise verwendet eine Kombination aus Fuzzy-Algorithmen und bewährten proprietären Algorithmen, um Ihre Datenquellen zu durchsuchen und Duplikate mit einer Genauigkeit von 95 % zu identifizieren. Keine andere Lösung kann Ihnen dieses Maß an präziser Duplikaterkennung bieten – nicht einmal IBM oder SAS.

**Bereinigung und Standardisierung von Daten:** Hier findet der Hauptbereinigungsprozess statt. Während Sie früher Regeln zur Datenbereinigung manuell erstellen mussten, können Sie jetzt einfach auf vordefinierte Regeln klicken.

**donnees-comment-traiter-**

Werfen Sie einen Blick auf das folgende Bild.

**-datos-como-tratar-los-**

Sehen Sie, wie einfach es ist, alle Großbuchstaben in Kleinbuchstaben umzuwandeln, negative Abstände zu entfernen, Zeichen zu ersetzen usw.? Diese Art von Aufgabe würde Monate dauern, wenn sie manuell durchgeführt würde – ganz zu schweigen davon, dass Sie für jede Aufgabe Abfragen und Codes ausführen müssen. Außerdem gibt es die WordSmith-Funktion, mit der Sie Namen und Wörter in Ihrem Verzeichnis automatisch ersetzen können.

Sobald Sie all diese unordentlichen Daten bereinigt haben, können Sie Standardisierungsregeln erstellen. So können Sie z. B. in Ihren Lead-Formularen festlegen, dass die Eingabe der Postleitzahl obligatorisch ist oder dass nur geschäftliche E-Mails als Kontaktadressen zulässig sind. Wenn Sie eine Vorschau auf die Probleme erhalten, die Ihre Daten betreffen, können Sie dort, wo es möglich ist, Kontrollen einrichten, um zu verhindern, dass die Daten wiederholt von denselben Problemen betroffen sind.

Und nun endlich zur letzten Frage.

## Warum ist sie wichtig?

Das ist zwar offensichtlich, aber schlechte Daten beeinträchtigen Ihre Berichte, Erkenntnisse, Analysen und betriebliche Effizienz. Tatsächlich wirken sich schlechte Daten auf jeden Aspekt Ihres Unternehmens aus. Sie

sind sich dessen nur nicht bewusst.

X

Der Kreislauf sieht folgendermaßen aus:

Fehlerhafte Daten gelangen ins System >> Geschäftsanwender können Daten nicht nutzen >> Kontaktaufnahme mit der IT-Abteilung zur Lösung von >> Die IT folgt nicht den Zeitplänen der Unternehmen >> Business und IT rangeln miteinander >> Ziele werden verzögert >> Die Mitarbeiter sind mit der Bereinigung von Daten belastet, indem sie diese manuell in Tabellenkalkulationen vornehmen >> Fehler werden übersehen >> Daten werden genutzt, weil Ziele erreicht werden müssen >> Kunden sind die Leidtragenden von schmutzigen Daten >> Beschwerden werden eingereicht >> Geld ist verloren >> Arbeitnehmer werden beschuldigt >> Arbeitsplätze gehen verloren >> Zurück zum Anfang.

Die Moral Ihrer Mitarbeiter, Ihre Ziele, die Zufriedenheit Ihrer Kunden und Ihre betriebliche Effizienz werden ALLE durch schlechte Daten beeinträchtigt. Dabei sind ROI, Jahresberichte und Erkenntnisse sowie die Kosten schlechter Geschäftsentscheidungen aufgrund fehlerhafter Daten noch gar nicht berücksichtigt.

Wie man mit schlechten

Und was ist das Gegenmittel für all diese Probleme? Ein leistungsstarkes Werkzeug zur Datenbereinigung.

**Bereinigen Sie Ihre Daten. Sprechen Sie mit uns.**

[WD\_Button id="7841"]

donnees-comment-traiter-

datos-como-tratar-los-

**Try data matching today**

No credit card required

Geben Sie Ihre Arbeits-E-Mail ein

**HERUNTERLADEN**

WANT TO KNOW MORE?

**Check out DME resources**

X

✓

✓

✓

✓

**ie-man-mit-schlechten-**

(<https://dataladder.com/de/eine-kurzanleitung-zur-adressstandardisierung-und-ueberpruefung/>)

**Eine Kurzanleitung zur Adressstandardisierung und -überprüfung (<https://dataladder.com/de/eine-kurzanleitung-zur-adressstandardisierung-und-ueberpruefung/>)**

**·donnees-comment-traiter-**

Ibarrera - Juli 25, 2024

**»-datos-como-tratar-los-**

(<https://dataladder.com/de/8-bewaehrte-verfahren-zur-gewaehrleistung-der-datenqualitaet-auf-unternehmensebene/>)

**8 bewährte Verfahren zur Gewährleistung der Datenqualität auf Unternehmensebene (<https://dataladder.com/de/8-bewaehrte-verfahren-zur-gewaehrleistung-der-datenqualitaet-auf-unternehmensebene/>)**

Ibarrera - September 19, 2022

X

✓

✓

✓

✓

**Wie man mit schlechten Daten umgeht** (<https://dataladder.com/de/leitfaden-zum-musterabgleich-was-bedeutet-er-und-wie-geht-er/>)

**Leitfaden zum Musterabgleich: Was bedeutet er und wie geht er?** (<https://dataladder.com/de/leitfaden-zum-musterabgleich-was-bedeutet-er-und-wie-geht-er/>)

**Comment-traiter** Hubert - September 12, 2022

**Como tratar los datos** (<https://dataladder.com/de/datenqualitaets-hub-datenleiter-ressourcen/>)

Resource Center

## Try datamatch unternehmendatenbereinigungdatenbereinigungdatenqualität today.

Get started with our fully-functional free trial.

No credit card needed.

„\*“ zeigt erforderliche Felder an

Geben Sie Ihre Arbeits-E-Mail ein

HERUNTERLADEN

(<https://dataladder.com/de/>)

Data Ladder offers an end-to-end data quality and matching engine to enhance the reliability and accuracy of enterprise data ecosystem without friction.

Quick Links

- X Über uns
- Arbeiten Sie mit uns zusammen
- ✓ Dienstleistungen
- Hilfedokumente
- ✓ Änderungsprotokoll
- ✓ Alle Ressourcen
- ✓ Contact

✉ +1 (516) 468 6879 (tel: +1 (516) 468 6879)

☎ sales@dataladder.com (mailto:sales@dataladder.com)

📍 68 Bridge St Suite 307 Suffield CT 06078

Wie man mit schlechten

Privacy Policy (<https://dataladder.com/de/datenschutzrichtlinie/>)

Données comment traiter

¿datos como tratar los