

# KI-gestützte Autovervollständigung und Datenvalidierung

*Fehlervermeidung bei manueller Dateneingabe*



Bild: PSI FLS Fuzzy Logik & Neuro Systeme GmbH

**Die Qualität der Daten ist entscheidend für die Effizienz operativer Prozesse und für die Zuverlässigkeit von Berichten. Manuelle Dateneingabe birgt jedoch hohes Fehlerpotenzial und Vielfach bewährte, regelbasierte Eingabehilfen und Datenvalidierungen stoßen im Hinblick auf eine zunehmende Dynamisierung von Geschäftsprozessdaten oftmals an ihre Grenzen. Die Integration KI-gestützter, datenbasierter Lösungsansätze schafft Abhilfe.**

Branchenübergreifend ist die Erfassung von Daten über Eingabemasken ein zentraler Bestandteil vieler Softwarelösungen, z.B. in ERP- und anderen Verwaltungssystemen. Typisch für industrielle Anwendungen ist dabei unter anderem die große Anzahl verfügbarer Eingabefelder sowie Freiheiten bei der manuellen Dateneingabe. Die händische Bearbeitung kostet Zeit. Das kann dazu führen, dass Anwender vermehrt optionale Felder nicht oder nur teilweise ausfüllen. Hinzu kommt, dass durch die Freiheiten bei der Eingabe – z.B. die Verwendung von Abkürzungen – verschiedene Nutzer für die gleiche Semantik unterschiedliche Schreibweisen verwenden. Die daraus resultierenden Prozessverzögerungen, Dateninkonsistenzen und -fehler führen nicht nur zu einer signifikanten Verschlechterung des Berichtswesens. Da mit der Qualität der eingegebenen Daten sowohl die Effizienz des Prozesses selbst als auch seiner Nachfolger steht und fällt, können sie Folgefehler nach sich ziehen, die wiederum Nacharbeiten erfordern.

## Eingabehilfen als Mittel der Wahl?

Eingabehilfen sollen dieses Problem lösen. Die bisherigen Ansätze können dies aber nur teilweise und können vor allem bei semantisch variierenden, multiplen Abhängigkeiten in den Daten an ihre Grenzen stoßen. Die Mehrheit der bekannten Eingabehilfen basiert auf fest definierten Regeln, die aus der kontextabhängigen Syntax und Semantik abgeleitet sind. Der Vorteil: Sie bilden die jeweiligen Geschäftsprozesse präzise ab. Als solche sind sie eine gute Lösung für gut strukturierte und über die Zeit beständige Vorgänge, etwa bei klar definierten Datenabhängigkeiten oder eingeschränkter Auswahl per Drop-Down. Ihr Nachteil: Für die Eingabeüberprüfung bedienen sie sich aufwändiger Logiken. Je detaillierter die Abbildung von Prozessen also ist, desto komplexer ist das Regelsystem. Hinzu kommt, dass Änderungen immer Anpassungen der Codebasis erfordern. Ungeeignet sind regelbasierte Lö-

sungsansätze in diesem Kontext für die Eingabeunterstützung von Daten, die in multiplen Abhängigkeiten stehen und teilweise semantisch variieren. Das trifft z.B. auf die Datenerfassung in Bestell- und Lieferprozessen zu und damit auf einen wachsenden Anteil relevanter Einsatzszenarien.

## Autovervollständigung per KI

Abhilfe kann eine KI-gestützte, datenbasierte Auto-Vervollständigung bieten. Dieser Lösungsansatz macht sich die Tatsache zunutze, dass für nahezu alle Geschäftsabläufe mit einem Datenerfassungsprozess historisierte Daten zur Verfügung stehen. Unter Zuhilfenahme von Qualitativem Labeln und maschinellem Lernen trainieren und aktualisieren KI-Tools wie Deep Qualicision auf dieser Datenbasis kontinuierlich das relevante Nutzerverhalten. Automatisch werden folglich auch neue Daten einbezogen und die Wissensbasis kontinuierlich angepasst. Das KI-System kann so konfiguriert werden, dass es fortwährend typische Eingabemuster aus Vergangenheitsdaten erlernt – und zwar entweder allgemeingültig oder benutzerabhängig. Als Entscheidungsunterstützung für ein solches System dient ein einfaches Präferieren individuell definierbarer Bewertungskennzahlen (Key Performance Indicators, KPI). Dadurch lassen sich Abweichungen von Vorhersagen bewerten und Eingabedatensätze automatisch validieren. So kann z.B. eine anormale Abweichung einer Bestellmenge bereits bei der Dateneingabe erkannt werden, die gegebenenfalls bei einem Schwellenwertverfahren durch das Erkennungsrastrer fallen würde. Das Ergebnis: Das System vervollständigt die Eingabe von Daten automatisch und individuell angepasst an jeden Geschäftsprozess. Damit ist dieser Ansatz auch für die Abbildung der beschriebenen, semantisch variierenden und multipel abhängigen Daten geeignet.

## Ideales Zusammenspiel

Als ideal ist letztlich das Zusammenspiel zwischen regel- und datenbasierter Eingabehilfe zu betrachten. Das zeigt sich unter anderem bei der Dynamisierung von Formulare: Soll ein System kontextabhängig irrele-

vante Eingabefelder automatisch ausblenden, sind strikte Regeln sinnvoll. Für ihren Einsatz sprechen nicht zuletzt auch feste Abhängigkeiten zwischen Attributen. Bei einem Handelsvertrag hängen die auswählbaren Positionen z.B. immer davon ab, ob es sich um Kauf oder Verkauf handelt.

### Nicht vertauschen

Auto-Vervollständigung allein sichert die Richtigkeit eines Datensatzes aber noch nicht vollumfänglich ab. Eine weitergehende Prüfung ist notwendig, um auszuschließen, dass z.B. Einheiten verwechselt oder Zellen vertauscht und fehlerhaft in eine Datenbank übertragen wurden. **Ein Praxisbeispiel:** Wird bei der Erfassung eines Verkaufsvertrags bei der Eingabe der vereinbarten Menge anstatt Kilogramm Tonnen ausgewählt und das Komma bei der Preisangabe um eine Stelle nach vorn verrutscht, wird ohne Datenvalidierung eine viel zu große Menge (Faktor 1.000) zu einem viel zu geringen Preis (Faktor 0,1) verkauft. Einige Unternehmen setzen aus diesem Grund auf Kontrollstellen, die vor der Übertragung in eine Datenbank Formulare auf ihre Richtigkeit prüfen. Dieser Vorgang nimmt viel Zeit in Anspruch und ist auch

selbst fehleranfällig. Für die meisten Unternehmen ist es wirtschaftlich ohnehin nicht realisierbar. Daher kommt ein vorgeschalteter Filter zum Einsatz. **Das Ziel:** Die Prüfer sollen sich auf Fehler mit großen Folgewirkungen konzentrieren. Die wiederum vorrangig regelbasierten, pflegeaufwändigen Systeme suchen auf Basis von Schwellenwertprüfungen nach fest definierten Anomalien in den Eingabedatensätzen. Überschreitet z.B. ein eingetragenes Gewicht den Schwellenwert von einer Tonne, überprüft die Kontrollstelle die vorliegenden Eingabedaten. Problematisch wird es jedoch, wenn Prüfregeln fehlen oder bezogen auf ihre Abhängigkeiten unvollständig sind. Fehlt im geschilderten Beispiel eine Prüfregele für das Mengen-Preis-Verhältnis, werden mehrere hundert Datensätze knapp unter einer Tonne mit zu geringen Preisen an der Kontrollstelle vorbei in die Datenbank übertragen. Die Lösung für eine effiziente, ergänzende Datenvalidierung liegt auch hier in dem beschriebenen Mechanismus. Auf der breiten Basis historisierter Daten kann das KI-System selbstständig Anomalien in den prozessspezifischen Strukturen der gesamten Datensätze erkennen und passt sich kontinuierlich an die aktuellen Gegebenheiten an. **Das Ergebnis:** Das System

leitet für die manuelle Prüfung ausschließlich Dateneingaben mit entsprechenden Anomalien weiter, ohne auf starre Regeln oder Eigenschaften festgelegt zu sein.

### Zeit sparen

Die funktionale Komplexität von Business-IT-Lösungen wird zukünftig weiter zunehmen, auch wenn die Entwicklung von modularen Architekturen zu offeneren und agileren Strukturen führt. Umso wichtiger sind schnelle und korrekte Dateneingaben – unterstützt und geprüft durch intelligente Logiken. Dafür eignen sich integrierte KI-Systeme. Sie lernen die notwendigen Muster und Strukturen selbstständig auf Basis historisierter Daten und passen diese automatisiert und kontinuierlich an. Dadurch lässt sich sowohl die Anwenderfreundlichkeit als auch die Datenkonsistenz messbar erhöhen. Das spart Zeit, Aufwände und Kosten. ■

**Autor:** Dr. Jonas Ostmeyer,  
PSI FLS Fuzzy Logik & Neuro Systeme GmbH  
[www.fuzzy.de](http://www.fuzzy.de)