

Fallbeispiel: Normieren von Kontaktdaten als Managed Data Service

Die Herausforderung

Der Kunde, ein großes bundesweit tätiges Getränkegroßhandelsunternehmen erfasst und verarbeitet Daten von Restaurants und Märkten. Das Problem: Eine schlechte Datenqualität, doppelte und zum Teil widersprüchliche Einträge machen eine adäquate Auswertung schier unmöglich.

Die fast 70.000 Datensätze bestehen aus zig Feldern, welche nicht einheitlich benutzt und gepflegt wurden. Eine Suchmöglichkeit besteht aufgrund mangelnder Normierung nicht.

Die Lösung

- Benutzer** In gemeinsamen Gesprächen wurden für die Normierung einheitliche Regelungen/Formatierungen erarbeitet. Die Mitarbeiter wurden im Umgang mit den Daten geschult.
- Prozess** Strukturierte Aufarbeitung (orthografische Korrektur, Normierung) der Objektnamen durch versierte Datenredakteure. Automatisierte Verarbeitung und Anpassung von Adressen und Telefonnummern.
- Daten** Die Daten wurden in Token (Datensegmente) zerlegt und analysiert. Nach der Aufbereitung mittels spezieller Werkzeuge wurden die finalisierten Datensätze als csv. zurückgespielt.
- System** Die Aufarbeitung der Daten erfolgte im Editor TokenEdit. Dieser verfügt über interne Analyse und Abgleichtools, welche für eine erfolgreiche Datenaufbereitung unerlässlich sind.

Das Ergebnis

Von den knapp 70.000 Daten waren in 65.125 Datensätzen Änderungen erforderlich. Die Adressen und Telefonnummern wurden auf Richtigkeit geprüft und einheitlich formatiert. Falls möglich wurden die Daten mit zusätzlichen Informationen angereichert.

Durch einheitliche Schreibweisen sind nun eine Suche und Filtermöglichkeit gegeben.