

Customer Identity Resolution



„Customer Identity Resolution“ hat zur Aufgabe, Unternehmen bei den Herausforderungen im Umgang mit Kunden-Identitätsdaten zu helfen: Das sind die Daten aus unterschiedlichen Quellen, die spezifisch und korrekt einen Kunden, einen Lieferanten, einen Interessenten, einen Meinungsmacher, einen Patienten, einen Steuerzahler, einen Kriminellen etc. identifizieren.

Die alte Weisheit „der Kunde ist König“ ist nun definitiv Wahrheit: Wir sind im Zeitalter des Kunden angekommen. Das Internet und ganz besonders die fortschreitende Nutzung sozialer Medien haben die Machtverhältnisse im Markt zu Gunsten der Kunden und Verbraucher gedreht. Sie beziehen Information überall und jederzeit mittels des mobilen Internets über viele verschiedene Kanäle sowie Meinungsmacher und Trendsetter hinweg. Kunden wollen Produkte sofort, in bester Qualität, günstig und mit einem hohen Maß an Service. Das Wissen um und über den Kunden ist wettbewerbsentscheidend geworden, um im Rahmen von „Customer Experience Management (CEM)“ positive Kundenerfahrungen zum Aufbau einer emotionalen Bindung zwischen Kunden und Produkt, Marke, Service und/oder Anbieter zu schaffen. Ziel von CEM ist es, aus zufriedenen Kunden treue Kunden und aus treuen Kunden „begeisterte Kunden“ zu machen.

Unter „Customer Experience Management“ versteht man die Strategie, Methodologie und/oder die Prozesse, um in einer umfassenden Weise das Multi-Kanal-Erscheinungsbild, die Interaktionen und Transaktionen von Kunden mit einem Unternehmen, Produkt, Marke oder Service zu managen.¹

Das klingt sehr einsichtig, ist aber gar nicht so einfach anzuwenden. Die Probleme beginnen mit dem Multi-Kanal-Erscheinungsbild: Eine Kundin, die in der Kundendatenbank im Unternehmen mit dem Namen Ruth-Hanna Friese eingetragen ist, könnte beispielsweise in einem sozialen Netz Ruth Anne Friese heißen oder als Ruth Friese reklamieren. Sind das jetzt drei Personen oder ist das eine Person mit drei Identitätsbezeichnungen? Solche Probleme mit der Kundenidentität sind nicht selten, sondern fast die Regel. Ursachen sind eine natürliche Variabilität wie im Beispiel von Frau Friese, aber auch unerwartete Fehler durch Schreib- oder Transkriptionsfehler wie Spitznamen, Abkürzungen und Schreibweisen in unterschiedlichen Schriftsätzen (wie arabisch, chinesisch, griechisch, kyrillisch, lateinisch etc.) oder sogar professionell gemachte Lügen, die eine falsche Identität vortäuschen sollen. Noch schwieriger wird es, Identitäten zu finden, wenn der Kunde anonym auftritt. Im Endeffekt

haben wir es bei einer solchen „Customer Identity Resolution“ mit einer der zähsten Herausforderungen in der Datenvorbereitung für CEM oder dem traditionellen CRM zu tun.

Das Kern-Thema von Customer Identity Resolution ist nicht neu. Als Methode im Datenqualitäts-Management kommt es beispielsweise seit langem bei der Einrichtung von Data Warehouse zum tragen – oder im analytischen CRM beim Aufbau einer einzigen Sicht auf den Kunden. Auch im Direktmarketing ist Customer Identity Resolution nichts neues – sorgt doch die Dublettenbereinigung hier für reibungslose Prozesse. Ob bei der Konsolidierung von Adressbeständen, Bestandsbereinigung, Fremdbereinigung, Listen-Mischung, Cluster-Abgleich, Negativ-Abgleich (insbesondere bei der Robinsonliste), Positiv-Abgleich zur Datenanreicherung und im internationalen Abgleich bei unterschiedlichen Schriftsatz-Räumen – der richtige Umgang mit Kunden-Identitätsdaten ist hier erfolgsentscheidend. Customer Identity Resolution greift auf all diesen Gebieten und ist im Datenqualitätsmanagement, wie es beispielsweise bei Uniserv umgesetzt wird, daher nicht mehr wegzudenken.

Heute im Zeitalter des Kunden ist Customer Identity Resolution noch wichtiger geworden. Die Nutzung von sozialen Medien – wie Facebook, LinkedIn, Xing, Foursquare, Twitter, Pinterest etc. – durch den Kunden hat ja zur bereits genannten Machtverschiebung im Markt zugunsten des Kunden geführt. Jetzt ist es für Unternehmen entscheidend und wettbewerbskritisch zu wissen, was die Kunden in den sozialen Netzen sagen und meinen. Für das Marketing bieten die Kundendaten in den Netzwerken eine bisher nicht gekannte und gekonnte Möglichkeit, das Kundenwissen anzureichern und entsprechend zu nutzen.² Beim Abgleich und der Anreicherung der Unternehmens-Kundendaten mit sozialen Daten ist Customer Identity Resolution erfolgskritisch. Indem man jede Information in den Gesamtkontext eines Kunden stellen und dann kumulieren kann, erhält man ein besseres Verständnis und vor allem Kenntnis über seine Kunden. Aus zusammengeführten Informations-Puzzleteilen entsteht so ein Kundengesamtbild: Auf Basis von Customer Identity Resolution erhält man ein präziseres Multikanal-Erscheinungsbild des Kunden, kann so bessere Kundenmodelle im Rahmen von prädiktiver Analyse aufbauen und im Endeffekt die Geschäftsergebnisse verbessern.

Die verwendeten Methoden basieren auf landesspezifischen Regel- und Begriffstabellen, auf sprachraumspezifischer Phonetik und adressenspezifischer Fuzzy-Logik (Mehr Details zu den Methoden finden Sie

am Ende des Beitrages). Informationstechnisch wird Customer Identity Solution am besten als Service angeboten. Dann lassen sich Customer Identity Resolution Services in Batch-Läufen zur Massenverarbeitung einsetzen, beispielsweise bei der Bereinigung großer Datenbestände, als auch in Geschäftsprozesse einbetten, wo sie in Echtzeit angewendet werden können.

Beispiel:

Transaktionen im Handel. Manche Händler betreiben unterschiedliche Webshops. Wenn ein Neukunde in einem Shop eine Bestellung aufgibt, dann sollte man wissen, ob er vielleicht bereits ein guter Kunde in einem anderen Shop ist, um im Sinne von CEM die Kundenbindung weiter zu steigern. Vielleicht ist es aber auch ein „fauler“ Kunde, der bereits auf der schwarzen Liste des Unternehmens oder bei Kreditbewertungsorganisationen steht, und sich nun unter einer falschen Identität in einem anderen Shop bedienen will. Mit Customer Identity Resolution in Echtzeit kann noch vor dem Abschluss der Transaktion die wahre Identität festgestellt und entsprechend gehandelt werden.

Das Beispiel zeigt auch sehr deutlich, dass Customer Identity Resolution nicht nur wie früher im Direktmarketing der Kostensenkung dient, sondern auch im Sinne von CEM die Kundenbindung und Kundenprofitabilität steigern kann oder im Sinne von Risiko-Management betrügerische Transaktionen vermeiden kann.

Customer Identity Resolution Management Services haben weiterhin den Vorteil, dass sie nicht nur als On-Premise-Services, sondern auch als SaaS im Rahmen von Cloud Computing genutzt werden können. Mit anderen Worten: Customer Identity Resolution Services lassen sich schnell installieren, testen und in einem Pilotprojekt auf Kosten und Nutzen prüfen.

Customer Identity Resolution ist also ein wichtiger und erfolgskritischer Baustein in einem Kunden-Datenmanagement, das die unterschiedlichen Aufgaben und Funktionsbereiche der Datenvorbereitung für CEM und dem traditionellen CRM darstellt:

1. Customer Data Management beginnt mit der Datenintegration, um Daten aus unterschiedlichen Quellen zusammenzuführen. Das kann traditionell mit ETL-Prozessen³ gemacht werden oder mit Hilfe von Datenvirtualisierung, einem neueren Verfahren, bei dem nur ein logisches Datenmodell des Gesamtdatenbestandes erstellt wird, die Daten aber physikalisch nicht bewegt werden.

2. Im zweiten Schritt erfolgt eine Datenbereinigung mit den typischen Services eines Datenqualitäts-Management.
3. Der dritte Schritt ist dann Customer Identity Resolution.
4. So erfolgt dann im vierten Schritt die 360° Grad Sicht auf den Kunden.
5. Schließlich folgen noch die Aufgaben der Information Governance, die die Sicherheitskonzepte umsetzt und die Compliance sicherstellt.

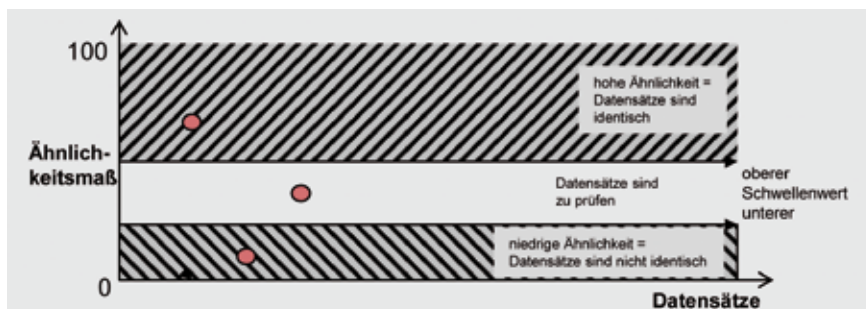
So erhält man eine solide Basis für CEM und CRM.

Fazit

Customer Identity Resolution ist die Voraussetzung für erfolgreiches CEM und CRM. Heute verstehen wir darunter nicht mehr allein Dublettenbereinigung im Rahmen von Datenqualitäts-Management: Die Möglichkeit, alle Information über Kunden in den Gesamtkontext zu stellen und die Informationsbausteine aus den unterschiedlichen Kanälen zu einem präziseren Multikanal-Erscheinungsbild zu kumulieren, bringt einen großen Mehrwert an Information. Dublettenbereinigung spart Kosten, durch den Gewinn an Information über den jeweiligen Kunden steigert man zusätzlich den Umsatz und senkt Risiken.

Methoden des Customer Resolution Management.

Traditionelle Verfahren in der Customer Identity Resolution setzen Zeichenkettenvergleiche und Match-Codes ein. Heute werden mehr und mehr mathematische Verfahren insbesondere aus der Fuzzy-Logik eingesetzt, die durch landesspezifische Wissensbasen (Abbildung) ergänzt werden. Diese Wissensbasen sind offen und können daher im Lauf der Zeit mittels Lernverfahren verbessert werden. So können zunächst allgemeine Wissensbasen problemspezifisch angepasst werden. Eine umfassende Zusammenstellung der gängigen und auch fortgeschrittenen Verfahren findet man beispielsweise bei Uniserv. www.uniserv.com/de/products/data-quality-technology/methods-algorithms.php



Qualitätslevel von Identity Resolution

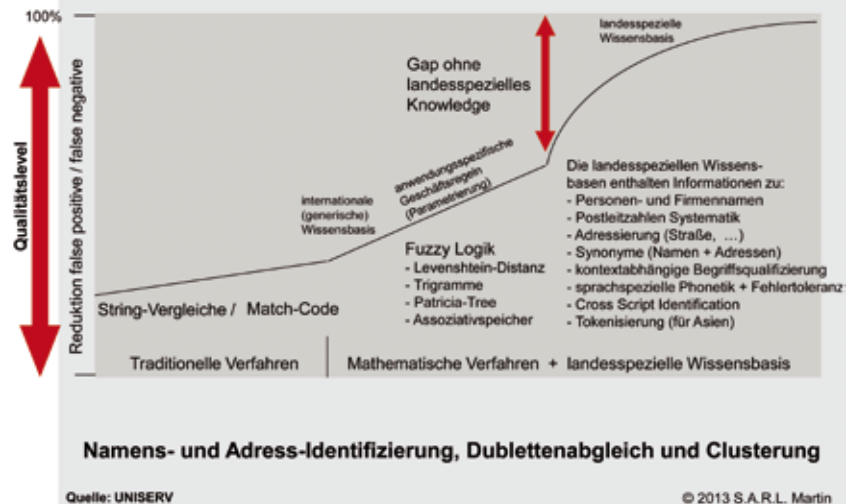


Abbildung. Die Qualität von traditionellen Verfahren wie Zeichenkettenvergleichen und Match-Codes ist in der Regel nicht ausreichend, kann aber sehr gut durch mathematische Verfahren der Fuzzy-Logik verbessert werden. Um schließlich auf ein Qualitätslevel von nahezu 100% zu kommen, sollten landesspezifische Wissensbasen eingesetzt werden.

Beim Customer Resolution Management können natürlich auch Fehler auftreten. Man spricht von einem Fehler der ersten Art (oder auch „falsch positiv“), wenn zwei Datensätze, die zu verschiedenen Kunden gehören, einem Kunden zugeordnet werden. Vom Fehler zweiter Art (oder auch „falsch negativ“) spricht man, wenn zwei Datensätze, die den gleichen Kunden bezeichnen, nicht zugeordnet werden. Das Schaubild unten erläutert die Situation:

Im Schaubild unten bedeutet der „obere Schwellenwert“ die Mindestübereinstimmung, damit verschiedene Datensätze einer Identität zugeordnet werden. Entsprechend bedeutet der „untere Schwellenwert“, dass alle Datensätze mit einem kleineren Ähnlichkeitsmaß verschiedenen Identitäten zugeordnet werden. Datensätze mit Ähnlichkeitsmaßen, die zwischen dem oberen und unteren Schwellenwert liegen, sind manuell

zu prüfen. Setzt man den oberen Schwellenwert zu niedrig an, dann erhöht man die Anzahl der falsch positiven Entscheidungen. Setzt man den unteren Schwellenwert zu hoch an, dann erhöht man die Anzahl der falsch negativen Entscheidungen. Im Laufe der Zeit gewinnt man Erfahrung und kann den oberen und unteren Schwellenwert empirisch optimieren.

Dr. Wolfgang Martin

Wolfgang Martin Team

Mitglied im CRM-Expertenrat

E-Mail: info@wolfgang-martin-team.net

¹ Bernd H. Schmitt.: Customer Experience Management: A Revolutionary Approach to Connecting with Your Customers. Wiley; 1 edition 2003.

² Wir setzen hier voraus, dass die Beschaffung und Nutzung von Kundendaten aus sozialen Netzen konform mit der Datenschutz-Gesetzgebung erfolgt.

³ ETL = Extraktion, Transformation, Laden. ELT ist eine Variante dazu.