

HW *sufi* library deutsch



**Softwarekomponente zum Suchen und Finden
von Adressen, Dubletten und Daten**

HW-Soft GmbH

Die Spezialisten für Dialog- und
Database-Marketing-Systeme

HWsufi - warum?

In vielen Datenbank-Systemen für Auftragsabwicklung, Customer Relationship Management oder Dialog- und Database Marketing existieren Anschriften zum Beispiel von Kunden und Interessenten mehrfach in unterschiedlichen Schreibweisen. Insbesondere, wenn keine Software zur fehlertoleranten Suche und Erkennung von Daten, Namen und Adressen eingesetzt wird, ist dies der Fall.

Oft ist der gleiche Kunde nur mit unterschiedlichen Ansprechpartnern oder Abteilungsbezeichnungen im Datenbanksystem enthalten.

Bei der Suche von Adressen im Dialog werden diese aufgrund unterschiedlicher Schreibweisen häufig nicht gefunden und daher neu erfaßt, obwohl die Adresse bereits in der Datenbank vorhanden ist.

Insbesondere beim Einsatz von Database Marketing Systemen ist es notwendig, Kunden und Interessenten aus den unterschiedlichsten Quellen mit ihren Zusatzinformationen regelmäßig mit Batchläufen in die Marketing Datenbank zu übernehmen.

Bei all diesen Aufgabenstellungen ist es notwendig und sinnvoll, eine Software einzusetzen, die sowohl bei der Suche im Dialog als auch bei der Suche und Zuordnung im Batch Adressen, Dubletten und Daten effizient und fehlertolerant findet. Dabei ist es wichtig, dass diese Software flexibel einstellbar ist, um die unterschiedlichsten Anforderungen und Aufgabenstellungen erfüllen zu können.

Eine Software, die diese Anforderungen vollständig erfüllt ist die **HWsufi library**.

HWsufi - die Produktfamilie

Die Produktfamilie HWsufi dient zum fehlertoleranten Suchen und Finden von Adressen, von Dubletten und von Daten in Datenbanken und Dateien.

Zurzeit ist HWsufi für den deutschen Sprachraum verfügbar (weitere Sprachräume auf Anfrage).

HWsufi steht als **HWsufi library** zum Implementieren in Dialog- und Batchprogramme und als **HWsufi manager** zur Verfügung.

Der **HWsufi manager** ist ein selbständig ablauffähiges Softwaresystem zur Dublettenerkennung und -bereinigung von Dateien im Batch.

HWsufi library deutsch

Mit der **HWsufi library deutsch** ist es möglich in Datenbanksystemen im Dialog und Batch fehlertolerantes Suchen und Finden von Adressen, Dubletten und Daten durchzuführen.

Für die Ersteinrichtung werden mit einem komfortablen Dialogprogramm die in der Datenbank zu suchenden Informationen definiert.

Mit einem weiteren Programm werden dann einmalig Suchcodes generiert und in einer eigens dafür generierten Datenbank-Tabelle gespeichert.

Diese Suchcodes benötigt die HWsufi library zum Suchen und Finden von Adressen und Daten.

Die **HWsufi library** ist kein selbständig ablauffähiges Softwaresystem. Sie kann in Programme eingebunden werden, die in Programmiersprachen entwickelt wurden, die den Aufruf von C/C++ Funktionen unterstützen. Auch die Implementierung in Internet- und Intranet-Anwendungen ist problemlos möglich.

Über eine Schnittstelle werden der HWsufi library die zu suchenden Adressen und Daten übergeben. Diese werden dann unter Verwendung von phonetischen Näherungs- und anderen Ähnlichkeitsverfahren mit Hilfe der Suchcodes geprüft. Wenn gleiche oder ähnliche Adressen und Daten gefunden wurden, werden diese an der Schnittstelle mit allen ausgewählten Informationen zurückgeliefert.

HWsufi - die Referenzdaten

HWsufi arbeitet mit Referenzdaten, die zur Identifikation von Wörtern dienen, damit diese bei der Erkennung unterschiedlich gewichtet werden können. Dadurch wird insbesondere die Qualität der Dublettenerkennung erheblich verbessert. Das in den Referenzdaten enthaltene umfangreiche Wissen dient zur Erkennung von Titeln, Adelsprädikaten, Vornamen, Nachnamen, Gesellschaftsformen, Branchenbezeichnungen, geografischen Begriffen usw.

Die Referenzdaten sind Bestandteil der **HWsufi library deutsch**. Sie werden durch die HW-Soft GmbH laufend aktualisiert und in unregelmäßigem Rhythmus ausgeliefert.

Die Referenzdaten werden für die Nutzung durch HWsufi von der HW-Soft GmbH speziell aufbereitet und in Form einer Datei ausgeliefert. Diese Datei wird in den Hauptspeicher des DV-Systems komprimiert geladen, auf dem HWsufi eingesetzt wird. Dadurch kann HWsufi die Referenzdaten ohne Plattenzugriffe nutzen.

HWsufi - die Highlights

Die in HWsufi implementierten Ähnlichkeits- und Näherungsverfahren gewährleisten die hohe Qualität und Performance sowie die sichere Erkennungsrate.

Wenn von HWsufi bei der Suche mehrere Datensätze gefunden wurden, werden diese mit allen parametrisierten Informationen an der Schnittstelle zurückgeliefert.

Die Suche mit der HWsufi library ist über maximal 3 Hierarchiestufen möglich (z. B. Mandant, Firma, Mitarbeiter). Dabei können die zu suchenden Informationen auch z. B. nur aus einem Personennamen bestehen, wie das häufig bei der Speicherung von Mit-

arbeitern oder Ansprechpartnern in Firmen der Fall ist.

Es ist möglich Adressen, die sowohl eine Zustellanschrift (Straße mit Hausnummer) als auch eine Postfachanschrift haben, zu verarbeiten. Wie die Dublettenerkennung dabei erfolgen soll, ist einstellbar.

Es können zusätzlich weitere Informationen zur Suche und Erkennung verwendet und speziell parametrisiert werden, wie E-Mail-Adresse, Geburtsdatum, Abteilung usw.

Durch die in HWSufi implementierte Worterkennungslogik und das in den Referenzdaten enthaltene umfangreiche Wissen ist es HWSufi möglich festzustellen, ob es sich bei einer Adresse um eine Person oder eine Firma handelt. Dadurch kann die Software bei der Gewichtung darauf unterschiedlich reagieren, z. B. durch die Trennung von Firmen- und Personenadressen.

HWSufi verfügt auch über weitere Möglichkeiten zur Einstellung der Dublettenerkennung wie Familientrennung, Gesellschaftsformtrennung, Branchenbegriffstrennung, Straßen- und Hausnummerngewichtung usw.

HWsufi - Beispiele

In den folgenden Beispielen zur Dublettenerkennung sind die Suchergebnisse für unterschiedliche Parametereinstellungen dargestellt.

Firmen-Dubletten

CSI Computer + Software Ingenieur-GmbH 70771 Leinfelden	Karlsruher Str. 11/1
Csi Ing.Gesellsch. Computer&Software 70771 Leinfelden	Karlsruher Str. 11
Verlag Karl Hofmann GmbH & Co. 73614 Schorndorf	Steinwasenstr. 6-8
Karl Hoffmann Verlag GmbH & Co 73614 Schorndorf	Steinwasenstr. 6-8
Alfons W. Gentner Verlag GmbH & Co. KG 70193 Stuttgart	Forststr. 13
Gentner, A.W. Verlag GmbH & Co. KG 70193 Stuttgart	Forststr. 131
Bosch GmbH & Co. 74572 Blaufelden	Engelstr. 49/55
Bosch Tiernahrung GmbH & Co 74572 Blaufelden	Engelstr. 55
SpVgg 07 Ludwigsburg e.V. 71638 Ludwigsburg	Fuchshofstr. 60
Sportvereinigung 07 Ludwigsburg e.V. 71638 Ludwigsburg	Fuchshofstr. 60
Drogerie Hamdorf 07980 Berga	Bahnhofstr. 1
Hamdorfer Apotheke u. Drogerie 07980 Berga	Bahnhofstr. 1b

Michetschläger Sport 94157 Perlesreut	Marktplatz 29
Sport und Schläger 94157 Perlesreut	Marktplatz 29

ZF Lenksysteme GmbH & Co. KG 73650 Winterbach	Peter-Hahn-Str. 13
ZF Lenksysteme GmbH 73650 Winterbach	Peter-Hahn-Str. 13

Diese Erkennung erfolgt bei der Parametrierung Gesellschaftsform-Trennung = nein. Wenn die Parametrierung auf Gesellschaftsform-Trennung = ja geändert wird, werden diese beiden Adressen nicht als Dublette erkannt.

Bausparkasse Schwäbisch Hall AG 74523 Schw. Hall	Crailsheim Str. 52
Schwäbisch Hall Service AG 74523 Schw. Hall	Crailsheim Str. 52

Diese Erkennung erfolgt bei der Parametrierung Branchen-Trennung = nein. Wenn die Parametrierung auf Branchen-Trennung = ja geändert wird, werden diese beiden Adressen nicht als Dublette erkannt.

Personen-Dubletten

Wolfgang Winkler 79194 Gundelfingen	Bundesstr. 100
Wolfgang Winter 79194 Guntelfingen	Bundesstr. 100

Lars Schur 73230 Kirchheim	Brunenstr. 39
Bettina Schur 73230 Kirchheim	Brunenstr. 39
Gustav Schur 73230 Kirchheim	Brunenstr. 39

Diese Erkennung erfolgt bei der Parametrierung Familientrennung = nein

B. Schur 73230 Kirchenheim	Brunenstr. 39
Bettina Schur 73230 Kirchheim	Brunenstr. 39
Betina Schur 73230 Kirchheim	Brunenstr. 39

Diese Erkennung erfolgt bei der Parametrierung Familientrennung = ja

Hans Link 78076 Villingen	Lichtensteinstr. 89
Dr. Hans Link 78056 Villingen-Schw.	Lichtensteinstr. 83

Diese Erkennung erfolgt bei der Parametrierung Hausnummern-Trennung = nein. Wenn die Parametrierung auf Hausnummern-Trennung = ja geändert wird, werden diese beiden Adressen nicht als Dublette erkannt

Rudolf Wagner 07980 Berga	Rob.-Guezou-Str. 8
Wagner, Rudi 07980 Berga	Guezou-Str. 8

Firmen- und Personen-Dubletten

H. Hermann GmbH 70327 Stuttgart	Ulmer Str. 300
Heinr. Hermann 70327 Stuttgart	Ulmer Str. 300

Diese Erkennung erfolgt bei der Parametrierung Personen- und Firmen-Trennung = nein. Wenn die Parametrierung auf Personen- und Firmen-Trennung = ja geändert wird, werden diese beiden Adressen nicht als Dublette erkannt.

HW-Soft GmbH

Güterbahnhofstraße 3 • D-37186 Moringen

Telefon: +49 (0) 5554 9952-0 • Telefax: +49 (0) 5554 9952-55

eMail: info@hw-soft.de • Internet: www.hw-soft.de