

# **HW** *sufi* manager deutsch

Softwaresystem zur Dublettenerkennung  
und -bereinigung von Dateien



**HW-Soft GmbH**  
Die Spezialisten für Dialog- und  
Database-Marketing-Systeme



## HWsufi - warum?

In Adressbeständen sind häufig Adressen mehrfach in unterschiedlichsten Schreibweisen vorhanden, insbesondere, wenn die Erfassung und Pflege dieser Adressen in Systemen durchgeführt wurde, in denen keine Dubletten-erkennung bei der Dialogerfassung und -pflege integriert ist.

Meist wird das erst bemerkt, wenn Briefe und Mailings mehrfach an die gleichen Adressen versandt werden. Kunden und Interessenten reagieren darauf negativ.

Wenn Adressbestände unterschiedlicher Herkunft für Mailings eingesetzt werden oder aus anderen Gründen zusammengeführt werden sollen und dabei eventuell noch bestimmte Adressen (Robinsons, Nixies usw.) vom Versand oder von der Zusammenführung ausgeschlossen werden sollen, ist der Einsatz einer fehlertoleranten Dublettenabgleichsoftware mit effizienten Ähnlichkeits- und Näherungsverfahren zwingend erforderlich.

Dabei ist es wichtig, dass diese Software flexibel einstellbar ist, um die unterschiedlichsten Anforderungen und Aufgabenstellungen erfüllen zu können.

Um diese Anforderungen effizient und mit hoher Qualität durchführen zu können, wurde der **HWsufi manager** entwickelt.

## 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 manager** und als **HWsufi library** zur Verfügung.

Mit der **HWsufi library** ist es möglich in Datenbanksystemen im Dialog und Batch fehlertolerantes Suchen und Finden von Adressen, Dubletten und Daten durchzuführen. Die **HWsufi library** kann in Dialog- und Batchprogramme eingebunden werden, die die Funktionalitäten von HWsufi nutzen sollen. Auch die Implementierung in Internet- und Intranet-Anwendungen ist problemlos möglich.

## HWsufi manager deutsch

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

Die Satzaufbauten der Dateien werden mit der „Dialogsoftware zur Verwaltung der Satzbeschreibungen“ erfasst und verwaltet.

Mit der „Dialogsoftware zur Verwaltung der Feldzuordnungen“ werden die HWsufi-Felder, die zur Dublettenerkennung und -bereinigung herangezogen werden sollen, den entsprechenden Feldern einer Satzbeschreibung zugeordnet.

Mit der „Dialogsoftware zur Verwaltung der Dublettenbereinigungsläufe“ werden die Jobs zur Dublettenbereinigung erfasst, verwaltet, überwacht und freigegeben. Mit diesem Programm können auch alle Protokolle und die Dublettenliste in einer Vorschau angezeigt und, wenn gewünscht, ausgedruckt werden.

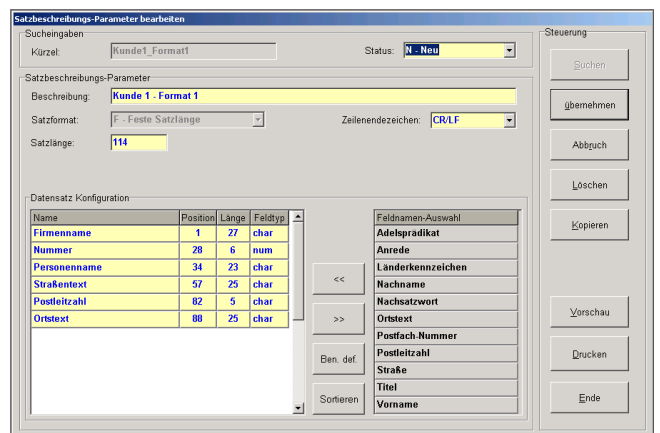
## Das HWsufi manager Menüsystem

Alle Dialogprogramme des HWsufi managers werden über das HWsufi manager Menüsystem aufgerufen.



## Die Dialogsoftware zur Verwaltung der Satzbeschreibungen

Die Parametrierung der Satzaufbauten für die Ein- und Ausgabedateien erfolgt mit diesem Dialogprogramm. Es können Satzaufbauten für Dateien im festen oder variablen Format sowie für dBASE-Dateien erfasst und verwaltet werden.



## Die Dialogsoftware zur Verwaltung der Feldzuordnungen

Um dem **HWsufi manager** bekannt zu geben, welche Felder einer Eingabedatei für den Dublettenabgleich herangezogen werden sollen, und welche Daten in diesen Feldern enthalten sind, wird mit der „Dialogsoftware zur Verwaltung der Feldzuordnungen“ eine Zuordnung der HWsufi-Felder zu den entsprechenden Feldern der Satzbeschreibung durchgeführt.

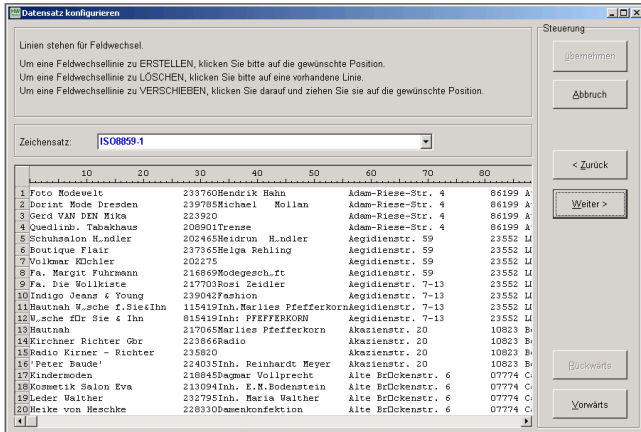
## Der Import-Assistent

Wenn eine Eingabedatei, die dublettenbereinigt werden soll, bereits auf dem Serversystem vorliegt, kann der Anwender deren Satzbeschreibung zusammen mit den HWsufi-Feldzuordnungen mit Unterstützung eines komfortablen Import-Assistenten, ähnlich wie bei bekannten Office-Produkten, erfassen.

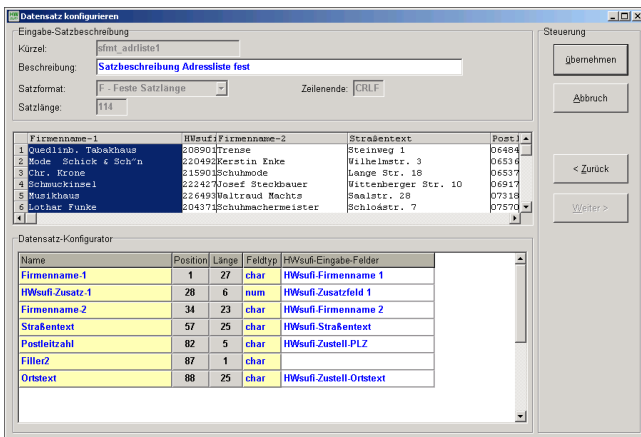
In diesem Fall kann die Erfassung von Satzbeschreibungen und Feldzuordnungen mit den Programmen „Dialog-



software zur Verwaltung der Satzbeschreibungen“ und „Dialogsoftware zur Verwaltung der Feldzuordnungen“ entfallen.



Der Import-Assistent analysiert die vorliegende Eingabedatei und unterbreitet dem Anwender Vorschläge über den Aufbau der Datensätze. Anschließend kann über eine Vorschau geprüft werden, ob die Satzbeschreibung für die Eingabedatei korrekt ist.



Dann werden über den Import-Assistenten die Felder der Eingabedatei festgelegt, die zur Dublettenbereinigung herangezogen werden sollen.

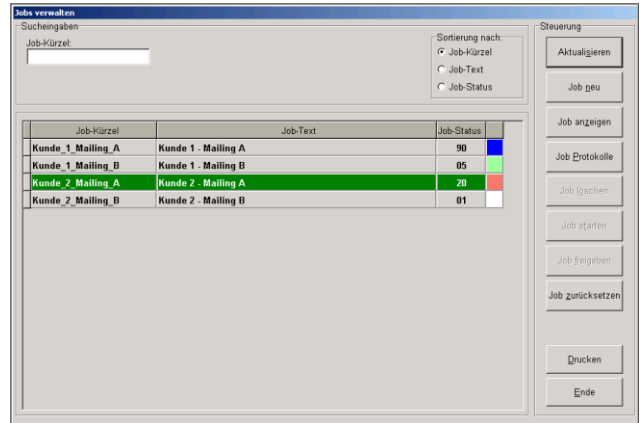
## Die Dialogsoftware zur Verwaltung der Dublettenbereinigungsläufe

Mit der „Dialogsoftware zur Verwaltung der Dublettenbereinigungsläufe“ werden die Jobs zur Dublettenbereinigung erfasst, verwaltet, überwacht und freigegeben. Nach Aufruf des Programms werden die vorhandenen Jobs mit ihren Status-Informationen angezeigt.

Die Verwaltung der Jobs erfolgt über Status-Kezis, die zu jedem Zeitpunkt den aktuellen Fortschritt des Jobs anzeigen.

Alle notwendigen Aktivitäten, wie starten des Batchprogramms, zurücksetzen von Jobs und anzeigen des Zustands eines laufenden Jobs mit ständiger Aktualisierung der verarbeiteten Stückzahlen und des voraussichtlichen Programmes, werden mit diesem Dialogprogramm durchgeführt.

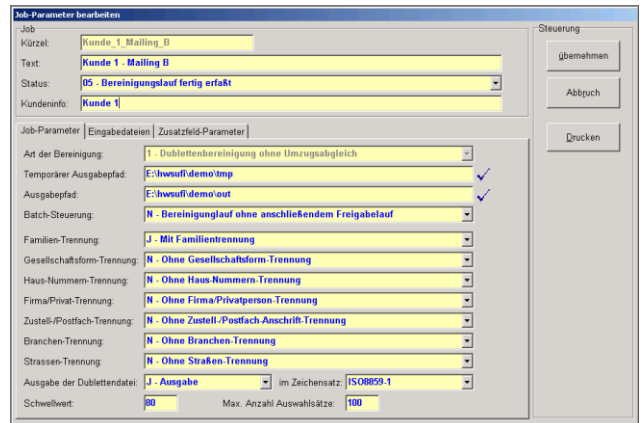
Mit dem HWsufi manager können mehrere Dateien in unterschiedlichen Satzaufbauten, Satzformaten und mit unterschiedlichen Zeichensätzen in einem Programmablauf in sich und gegeneinander dublettenfrei gemacht werden.



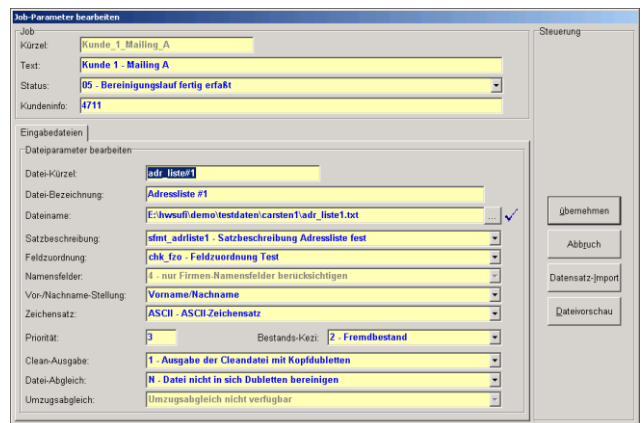
Bei der Erfassung eines Jobs lassen sich unter anderem folgende Parametrierungen für den gesamten Dublettenabgleichjob einstellen:

- Festlegen der Dateipfade für die Eingabe- und Ausgabedateien sowie für die temporären Dateien
- Festlegen der Eingabedateien mit Dateinamen, Satzbeschreibung, Feldzuordnung, Zeichensatz usw.
- Festlegen der Parametrierungen für Familientrennung, Gesellschaftsformtrennung, Trennung von Firmen- und Privatpersonenadressen usw.
- Einstellen der Parametrierungen für Umzugsabgleiche, wenn die HWsufi manager Erweiterung **HWumzug Deutschland Post Adress** lizenziert wurde

Das Erfassen und Ändern der Informationen für einen Job wird mit dem folgenden Dialogbildschirm über den Karteireiter „Job-Parameter“ durchgeführt.



Das Erfassen und Ändern der Informationen für die Eingabedateien wird mit dem folgenden Dialogbildschirm über den Karteireiter „Eingabedateien“ durchgeführt.





Dabei werden dateibezogene Informationen wie Datei-Kürzel, Dateiname, Satzbeschreibungs-Kürzel, Zeichensatz, Feldzuordnung, Priorität, Cleandatei-Steuerung usw. festgelegt.

Nachdem der Programmlauf zur Dublettenbereinigung durchgeführt wurde, können die Protokolle und die Dublettenliste in einer Vorschau angezeigt und, wenn gewünscht, ausgedruckt werden.

Anschließend wird der Job zur Endverarbeitung freigegeben und beendet. Dabei werden alle Ausgabedateien erstellt und alle temporären Dateien werden gelöscht.

## Die Dialogsoftware zur Verwaltung der Zeichensätze

Wenn Eingabedateien nicht im ISO 8859-1-Code vorliegen können die dafür notwendigen Code-Umsetzungen mit der „Dialogsoftware zur Verwaltung der Zeichensätze“ erfasst und gepflegt werden.

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0																
1																
2	40	5A	7F	7B	38	6C	50	7D	4D	5D	5C	4E	6B	60	4B	61
3	F0	F1	F2	F3	F4	F5	F6	F7	F8	F9	7A	5E	4C	7E	6E	6F
4	7C	C1	C2	C3	C4	C5	C6	C7	C8	C9	D1	D2	D3	D4	D5	D6
5	D7	D8	D9									3D	E0	3C		6D
6	B5	81	82	83	84	85	86	87	88	89	91	92	93	94	95	96
7	97	98	99	A2	A3	A4	A5	A6	A7	A8	A9				B9	
8	9F															
9																
A			4A					5B	8B	75			5F		76	78
B						B6							39	3A	3B	
C						31	B0		59							
D								30	BA					32		36
E	B2	B1			34								B3			
F							33	63	BB	B4				35		

## HWsufi manager - die Highlights

Die in HWsufi implementierten Ähnlichkeits- und Näherungsverfahren gewährleisten eine hohe und sichere Erkennungsrate.

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, Straßen- und Hausnummerngewichtung, Branchenbegriffstrennung, usw.

Durch das Laden der aufbereiteten und komprimierten Referenzdaten in den Hauptspeicher des Server-Systems kann der HWsufi manager die Referenzdaten ohne Plattenzugriffe lesen.

Durch die Client-Server-Architektur des HWsufi managers können die Batchprogramme und Server-Komponenten auf einem leistungsfähigen Server-System ablaufen.

Der Anwender des HWsufi managers braucht keine Kenntnisse über das Betriebssystem des Servers zu haben, auf dem die Batchprogramme und Server-Komponenten ablaufen.

Durch die auf PC-Systemen ablaufenden Dialogprogramme wird eine einfache Bedienung und Steuerung des HWsufi managers erreicht.

## HWumzug Deutschland Post Adress - die Erweiterung für den HWsufi manager

Mit der HWsufi manager Erweiterung HWumzug Deutschland Post Adress ist es möglich, zusätzlich zur Dublettenbereinigung mit dem HWsufi manager auch Umzugsabgleiche mit der Umzugsdatenbank der Deutschen Post Adress GmbH & Co. KG durchzuführen. Diese Erweiterung kann zusätzlich lizenziert werden.

Siehe auch Produktblatt: HWumzug Deutschland Post Adress, die Erweiterung für den HWsufi manager deutsch.

## HWsufi - die Referenzdaten

Der HWsufi manager deutsch arbeitet mit Referenzdaten, die für die Nutzung durch den HWsufi manager von der HW-Soft GmbH speziell aufbereitet und in Form einer Datei ausgeliefert werden.

Die Referenzdaten sind Bestandteil des HWsufi managers. Sie werden durch die HW-Soft GmbH laufend aktualisiert und in unregelmäßigem Rhythmus ausgeliefert.

## HWsufi manager - technische Details

Als Server-System für den HWsufi manager werden zurzeit die Betriebssysteme Windows 2000/XP/2003, AIX, UNIX oder Linux unterstützt (weitere Betriebssysteme auf Anfrage).

Die Dialogsoftware des HWsufi managers ist auf PCs mit den Betriebssystemen Windows 98/ME/2000/XP ablauffähig.

Die Kommunikation zwischen den Dialogprogrammen (Clients) und der auf dem Server ablaufenden Software erfolgt über TCP/IP auf der Basis von Socket/Port-Techniken.