



## **CRM De-dupe Wizard**

Interactive and batch mode de-duplication of CRM data

The CRM De-Dupe Wizard is fully integrated with Sage CRM, which allows interactive as well as batch mode de-duplication of CRM data.

CRM De-dupe Wizard now has multi browser compatibility, when used in Sage CRM 7.1 service pack 2 or greater. Not only is Internet Explorer supported, but, also Firefox, Chrome, and Safari.

### **Key Features of CRM De-dupe Wizard**

- Company de-duplication
- Company phone de-duplication
- Person de-duplication
- Person phone de-duplication
- Custom de-duplication
- De-duplication Statistics
- Customisable and configurable

When a company has duplicated records the system identifies the first record as the master record, and each duplicated record is first merged into the master record and deleted right after.

The merge is done field by field and this means that if the master record has an empty field e.g. company source while one of the duplicated records has a value for this field, the system will merge its value into the master record.

Main Menu allows for De-duplication at a Company and Person level, Data fixes within de-normalised Phone Numbers, and removal of redundant data within Phone Numbers.

It also includes a general statistics option, which keeps check of your data duplication levels.

To create an efficient method of de-duplicating data, an algorithm has been designed around ranking the data in such way that a "master" and "target" model can be achieved around duplicates.

This allows for the most suitable candidate within the duplicates to become the "master" record and all other duplicates (targets) to be merged into this data.

*Example of a de-dupe scan in progress.*

**Status**

**Progress Status**

Please wait, still scanning..

Remaining Time to Complete : 0h 1m 27s

0 matches found and merged so far.

150 1207

Main menu screenshot.

Company Deduplication Company Phone Dedupe Person Deduplication Person Phone Dedupe Custom Deduplication Info Config ...

**Dedupe Filter Criteria**

To restrict the range of data that the Dedupe Wizard will include in the Deduplication process, you may set the filter criteria here before pressing Go.

**Type:** Customer  
**Status:** Active  
**Mail Restriction:** Yes

**Source:** Tradeshow  
**Country:** --All--  
**Industry:** --All--

**Enable Live Run:**

Go

Statistics screenshot.

### Deduplication Status

**Stat as of Monday, February 18, 2013 3:16:16 PM**

Companies Merged in Current Loop (i.e. not deleted yet)	0
Percentage Companies Merged in Current Loop	0
All Companies Merged in Database (i.e. now deleted)	0
All Companies in Database (i.e. not deleted)	1207
Percentage Companies Merged in Database	0
Number of Companies with empty denormalised Phone Numbers	7
Number of Companies with empty denormalised Fax Numbers	808
<hr/>	
Persons Merged in Current Loop (i.e. not deleted yet)	0
Percentage Persons Merged in Current Loop	0
All Persons Merged in Database (i.e. now deleted)	0
All Persons in Database (i.e. not deleted)	1763
Percentage Persons Merged in Database	0
Number of Persons with empty denormalised Phone Numbers	7
Number of Persons with empty denormalised Fax Numbers	1113
<hr/>	
All Phone numbers merged in Database (i.e. now deleted)	0

### Log File

Today's Log File

Configuration

screenshot.

### Dedupe Configuration Settings

Log File: /log/dedupewizard.txt \*

Person Weight Factor: 500

Company Source Query: select \* from vDedupeCompany

Person Source Query: select \* from vDedupePerson

Delete Flag Value: 201 \*

Maximum Process Loops: 25 \*

Comm Weight Factor: 1 \*

Company Duplication Query: select \* from vDedupeCompany where (comp\_dedupe\_timestamp is null or comp\_dedupe\_timestamp <> #DEDUPE\_TIMESTAMP#) and comp\_companyid <> #comp\_companyid# and ( ( difference (comp\_name, #comp\_name#) = 4) and addr\_country = #addr\_country# and addr\_postcode = #addr\_postcode# )

Person Duplication Query: select \* from vDedupePerson where pers\_companyid = #comp\_companyid# and (pers\_dedupe\_timestamp is null or pers\_dedupe\_timestamp <> #DEDUPE\_TIMESTAMP#) and pers\_personid <> #pers\_personid# and ( difference (pers\_firstname, #pers\_firstname#) = 4 and difference (pers\_lastname, #pers\_lastname#) = 4)

New Data Score: 1005 \*

Interactive

Loop Delay Time (msecs): 750 \*

Oppo Weight Factor: 1000 \*

Case Weight Factor: 1 \*

Company Phone Query: select comp\_name, comp\_companyid, comp\_phonenumber, comp\_faxnumber from vCompany where comp\_phonenumber is null or rtrim(comp\_phonenumber) = '' or comp\_faxnumber is null or rtrim(comp\_faxnumber) = ''

Person Phone Query: select \* from vDedupePerson where pers\_phonenumber is null or rtrim(pers\_phonenumber) = '' or pers\_faxnumber is null or rtrim(pers\_faxnumber) = ''

There are a variety of configuration settings that are fully customizable, so that the balance of how the de-duplication runs can be adjusted according to the client's data requirements.

For more information on Greytrix Products and Services, write to us at [crmproducts@greytrix.com](mailto:crmproducts@greytrix.com)  
**Website: [www.greytrix.com](http://www.greytrix.com) | Blog: <http://www.greytrix.com/blogs/sagecrm/>**