How to Tidy and Merge addresses in your SIMS database

Introduction

Why do we need to Tidy & Merge Addresses?

Prior to the Autumn Census 2017, schools had been using SIMS for several years and were able to enter any address format/data in wrong fields/Incorrect postcodes etc. As the DfE has since then collected the UPRN codes from each student reported on in the Census; it has been necessary to run the Tidy and merge addresses routines in SIMS. The Tidy routine highlights any addresses with such issues, therefore schools should already be running the Tidy and Merge Routines outlined in this document each time they are preparing their data for using the Address validation service.

What is address Validation

Address Validation can be used when adding any new addresses to your system, this validates the address across the internet. Only correctly formatted addresses with the correct postcodes can be validated, this prevents incorrect or duplicate addresses being added to your system.

Address Validation also introduced the use of Unique Property Reference Numbers (UPRN's). The DfE collect the address UPRN numbers from each Census; therefore all schools need to make sure that their UPRN records are up to date.

The Address Tidy & Merge routine

These Housekeeping routines can be used to review addresses stored in SIMS. The Tidy routine identifies addresses that contain erroneous or incorrectly formatted information, and wherever possible, make applicable changes. A report identifying the proposed changes is then produced, which can be checked before proceeding with the Address Tidy. This process can be run at any time and should be part of your good system housekeeping. The Merge routine will tidy any duplicated addresses.

ICT service Desk: 233 8600

Table of Contents

Introduction

The Address Tidy routine is used to review addresses stored in SIMS, identify those that contain incorrect information or are incorrectly formatted and where possible make applicable changes to ensure addresses are accurate and uniformly formatted.

Address tidying can involve moving data between fields, (e.g. a post code that has been incorrectly entered in the **Town** field is moved to the **Post Code** field), and correcting the formatting in various ways.

Tidying addresses increases the likelihood of success when searching for an address and reduces the risk of the same address being added twice. The tidy process also increases the chances of identifying duplicate addresses (e.g. moving a post code incorrectly stored in the **Town** field, to the **Post Code** field, may then match another address based on that post code).

Such addresses can then be merged using the Address Merge routine. When tidying all addresses held in the address database that have not yet been validated are considered for tidying, including school, staff, contacts and agency addresses. (This may also include addresses for leavers as **all** addresses held need to be tidied as may be required in the future).

The Address Tidy routine should be run twice, because some of the rules used in the routine depend on a comparison with other addresses. Therefore, any addresses tidied during the first run may enable further improvements on the second run of the routine to be made.

It is therefore recommended that the Address Tidy routine is run twice in succession. You should make any necessary further manual corrections to addresses **only after the second run of the routine.** This ensures that any addresses edited manually, are not affected by subsequent tidy routines. Any new addresses added to the database since the tidy routine has been run are available for tidying twice.

Running the Tidy Routine produces a report which identifies the proposed changes. This can be reviewed before proceeding to tidy or addresses manually amended.

Overview of Merging Duplicate Addresses

Duplicate addresses may exist in SIMS, because they have been double entered, or imported from a **CTF/ATF file**. These duplicate addresses can be merged into a single address, and each resident at that address is then assigned to the same single address.

This may also sort the issue experienced in some schools where the house does not show against pupil contacts to indicate the students are living at the same address as parents, due to duplicate addresses being created, in the same way it may help to ensure siblings are linked to each other.

The same as the Tidy Routine, The Merge Routine produces a report which identifies the proposed changes. This can be reviewed before proceeding with the actual merge.

There are 4 steps to Tidy and Merging your addresses as outlined below.

Step 1 - Delete Unlinked Contacts

This routine enables you to delete contacts who are no longer linked to any pupil/student or applicant in SIMS. This will save on the address included in the address tidy routines to follow.

Unlinked contacts might exist because:

- They are duplicate contacts created in error. True contact will still exist linked to the correct pupil record

- The pupil/student has been deleted from SIMS.

- The link to any pupil has been broken; this usually occurs when a contact is removed; as the contact is no longer required, e.g. they no longer have links to a pupil

Contacts who are linked to pupil/students who have left the school are not deleted using this routine, even if they have no links to current pupil/students.

1. Select **Tools | Housekeeping | Delete Unlinked Contacts** to display the Find Unlinked Contacts browser.

   Click **Search** button to display a list of all unlinked contacts. - You are able to view this list to see what is to be removed. If there are any names of contacts in the list you do not wish to remove they can be deselected.
2. To indicate the contacts you wish to delete, select **All** from the **Select** drop-down button at the top of the screen. The selected contacts are then displayed in the *Unlinked Contacts to be deleted* panel at the bottom half of the page.

   You may notice a lot of the contacts only have partial details entered –indicating they have never really been in use.

3. Contacts can be removed from this list by highlighting their name in the Contacts to be deleted panel then clicking the Remove button to the right of this panel. These contacts will then not be deleted.

4. A list of contacts you intend to delete can be produced for reference purposes, if required, by clicking the **Print** button.

5. Click the **Delete** button to delete the selected unlinked contact

6. Click the Yes button when prompted to confirm the deletion.

Step 2 - Unmatchable Addresses Report

It is advisable to run this report before running the Tidy and Merge Addresses routine as it identifies addresses that are not complete enough to be included in the tidy Routine. Such addresses should be manually edited before proceeding with the Tidy & Merge routine.

The Unmatchable Address report identifies any addresses that are too incomplete to represent a unique residence or have data in incorrect fields (e.g. post code data in the Town field or contain superfluous data such as phone numbers or other contact details).

NOTE: *This is particularly important because any phone numbers or email addresses found in address fields are removed during the tidy process, and assigned to **all** the residents living at that address. You may therefore wish to manually move this information before proceeding with the tidy routine.*

Running the Unmatchable Address Report

Select **Tools | Housekeeping | Bulk Address Tools | Unmatchable Address Report** to display the report in your web browser (the report may take a short while to display). The report lists any unmatched addresses together with the name of the people who are residents at the address

Unmatchable Address Report

Unmatchable Address	Residents
House 17 Hampshire	[ApRel] May, Mrs Sandra for May, Shaun
House 29 Holborn	[ApRel] Stewartby, Mr Mark for Cummins, Patrick
East Town Bus Depot Eastshire	[NoK] Smyth, Mrs Linda for Smyth, Mr Gregory
C C Instruments East Town Eastshire	[NoK] Johnson, Mr T for Johnson, Ms Sylvia
East Town Council	[NoK] Chaplin, Mr D. for Chaplin, Mrs Anne
Address Not Known Market Blandings Barset	[Cnt] EDWARDS, Mrs Debbie for Edwards, Helen Marie
23 Peak Drive	[Cnt] Yoo, Mr Charlie for Yoo, Aiden

Above is a sample of what the report will look like. The Residents field gives you an indication of where to go to correct the address. E.g.

[St] – Indicates the address is linked to a Student

[xSt] – indicates the address is linked to an ex-student (or leaver)

[Cnt] – Indicates the address is linked to a contact for a student

[Em] – indicates the address is linked to a member of staff

[NoK] – indicates the address is linked to the next of Kin for member of staff

[App] - Indicates the address is linked to an applicant

[ApRel] – indicates the address is linked to an applicant contact

It is recommended that these addresses are manually edited where possible before proceeding further. Don't worry if there are addresses you cannot sort at this point – you can still proceed to the Tidy Routine (these addresses will not be included in the tidy).

The report is saved to the following folder on your PC -in both .html and .txt format: - so can be accessed again at a later date
C:\Documents and Settings\<Username>\My Documents\SIMS Address Reports

Step 3 - Tidying Address Data

IMPORTANT NOTE: Please carry out a full backup of the SIMS database before running either of the tidy or merge routines.

1. Select **Tools | Housekeeping | Bulk Address Tools | Tidy and Merge Addresses** to display the **Address Tidy and Merge** wizard.

2. You must run the Tidy report before the actual tidy process can be run.

   Click the **Tidy Report** button to produce the list of Address details to be tidied. (This may take several minutes to complete).

The report lists any incorrectly formatted addresses – Before (1st row) together with the suggested reformat –After (2nd row) which shows how the address will be displayed after the tidy process has been run. It also shows the residents recorded as living at the address.

The After row includes the rule that has been applied to tidy the addresses. A list of these rules can be found at Appendix A or on your SIMS System **Tools | Setup | Address Tidy and Merge Set up**

Address Tidy Report

Address Type	Apartment	House Name	House Number	Street Description	District	Town	County	Post Code	Country	Residents
Before	-	-	20	East Street,	-	East Town	-	-	United Kingdom	<Agency>
After a3 n9	-	-	20	East Street	-	East Town	Eastshire	-	United Kingdom	as above
Before	-	Health Centre	1	Kelham Close	-	Swindon	Wiltshire	SN3 2JF	United Kingdom	<none, ever>
After b0	<<DELETE>>	-	-	-	-	-	-	-	-	as above
Before	-	Arundel Consultants	21	Queens Drive	-	Swindon	Wiltshire	SN3 1AL	United Kingdom	<none, ever>
After b0	<<DELETE>>	-	-	-	-	-	-	-	-	as above
Before	-	-	2	Wimpole Close	-	Swindon	-	SN3 2HD	215	<none, ever>
After b0	<<DELETE>>	-	-	-	-	-	-	-	-	as above

*WARNING: Please check the report carefully before clicking the **Tidy** button. The tidy process will permanently change the address data stored in SIMS. Any addresses that are incorrectly tidied will need to be manually amended.*
NOTE: Any phone numbers or email addresses found in address fields are removed during the tidy process, and assigned to ALL the residents living at that address. You may therefore wish to manually move this information before proceeding with the tidy routine.

3. Once the Address Tidy report has been produced, review the content and make any required manual changes. Click the Tidy button to initiate the tidy process (this may take several minutes to complete).

If you do not wish to complete the Tidy process at this point you can choose **Cancel** or **Skip to Merge**

All the reports are saved to the following folder in both .html and .txt format.
C:\Documents and Settings\<Username>\My Documents\SIMS Address

Reports The text document also includes tables of abbreviations that are used to identify for example, that 'Drive' might be abbreviated to 'Dr' or 'Drv'. Also shows all the validation rules that are applied by this tidy routine

Once the tidy process has been completed, it is recommended that you run the process a second time and then run and check the *Unmatchable Addresses* report again. Any addresses that still feature in the list probably require some manual amendment. They may not be complete enough to identify a unique residence, or may contain non-address information that needs to be moved.

Step 4 - Merging Addresses

It is recommended that you run the **Unmatchable Address** and the **Address Tidy routine** (as described in the previous section) before merging address data because the tidy process increases the chances of duplicate addresses being identified.

IMPORTANT NOTE: Please carry out a full backup of the SIMS database before running tidy or merge routines.

1. Select **Tools | Housekeeping | Bulk Address Tools | Tidy and Merge Addresses** to display the **Address Tidy and Merge** wizard.

2. Click the **Skip to Merge** button and then the **Merge Report** button to produce the Address Merge report (this may take several minutes to compile).

The report lists all duplicate addresses that will be merged into a single address, together with the residents who are recorded as living at the addresses about to be merged. (if a few duplicates exist for one address there may be several addresses listed as original).

Address Merge Report

Status	Apartment	House Name	House Number	Street Description	District	Town	County	Post Code	Country	Residents
Original	-	-	7	Bevan Place	-	Cardiff	-	CF14 3UX	-	[Cand] Harlin, Louise Emma
Original	-	-	7	Bevan Place	-	Cardiff	-	CF14 3UX	-	[St] Western, Ryan
Merged	-	-	7	Bevan Place	-	Cardiff	-	CF14 3UX	-	as above
Original	Flat 6	-	21	Millbrook Road	-	Bedford	-	MK42 9HJ	-	[Cand] Evans, Hu Lun
Original	Flat 6	-	21	Millbrook Road	-	Bedford	-	MK42 9HJ	-	[Cand] Evergreen, Thomas
Merged	Flat 6	-	21	Millbrook Road	-	Bedford	-	MK42 9HJ	-	as above

NOTE: *Please check the report before clicking the Merge button. The merge process will permanently change the address data stored in SIMS. Any addresses that are incorrectly merged will have to be manually recreated.*
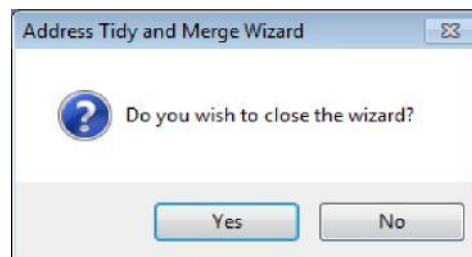
3. Once the Address Merge report has been produced, reviewed and printed, click the **Merge** button to start the merge process. This may take several minutes to complete.

The report is saved to the following folder in .html format:

C:\Documents **and Settings\<Username>\My Documents\SIMS Address Reports**

4. Click the **Finish** button to close the wizard.

5. Click the **Yes** button to complete the process.

Should you experience any difficulties, please contact the
ICT service Desk: on **233 8600** and your call will be logged for assistance.

Appendix A - Address Tidy Rules

Code	Descripti
a0	Address is empty or <unknown>: delete it.
a1	No Postcode, address line ends BFPO nnn: move this to postcode.
a2	Address is <as above>: if address is of a close contact of a student or employee next of kin then copy the address from that ...
a3	Address line has leading or trailing spaces or punctuation: remove them.
a4	Field other than Apartment starts <flat>: move to Apartment if empty, or if it is House name and Apartment ends with a house...
a5	Extract telephone numbers, and move them to the telephone table (except when other data is present that we are unsure ho...
a6	Extract email addresses.
a7	No House number, Street starts with a number: transfer this number to House number.
a8	Convert embedded multiple spaces to commas.
a9	No District, and Street has a comma and does not end with <street>: split at comma into Street and District.
b0	Address is not that of a school or agency and has no past, present or future residents: delete it.
b1	No apt. or House name, street ends <house>, town starts with number: move number to apt, street to name, start of town to ...
b2	County is empty and a <county> appears at end of another address line: move it to County.
b3	No Town, District has a comma: split at comma into District and Town.
b4	No Apartment, House number or House name, Street is one word or ends with <house>: move Street to House name.
b5	Street starts 'reg ' or 'rgt' (regiment) and House number exists: add House number to start of Street.
b6	No Apartment or House name, District ends with <street>, Street does not: move Street to House name and District to Street.
b7	Street has three words, second is <street>, third is not <house> or <street>: if no District move third word to District else if no...
b8	House number but no Street, House name exists but does not end <House>: move it to Street.
b9	Town has comma, Street or District is empty: split into District and Town if District empty, otherwise move District to Street be...
c0	No district, Country is <town>,<county>: Move Town to District and first part of County to Town.
c1	No District, London postcode, Town not London: move Town to District, make town London, delete London if elsewhere in ad...
c2	No District, Town and County are not empty, County appears more often as a town and Town more often as a district: move ...
c3	No Town, House name non empty, and appears more often as a town: move House name to Town.
c4	No Town, County not recognised and more usually appears in as a town: move County to Town.
c5	Expand and/or move items in the District/Town abbreviation table.
c6	UK address: tidy Postcode, and correct misuse of letters O/I/Z and digits 0/1/2.
c7	No District, Town is more than one word, last part is not <don't orphan> and appears elsewhere as a town: move first part to ...
c8	No House name, Apartment starts with a digit and ends with <house>: split into Apartment and House name.
c9	No House Number, Last word of house name is all digits: move it to House Number.
d0	No district, Apartment ends <street>, Street does not: move Street to District, Apartment to Street.
d1	No Apartment or House name, House number starts with a digit and ends with <house>: move the fist part to Apartment and ...
d2	No House name, House number with no digits ends with <house>: move House number to House name.
d3	No Postcode, UK Postcode appears at end of Street, District, Town or County: move it to Postcode.
d4	No Town: copy it if possible from a validated record with same District and first four characters of Postcode.
d5	No House number or Street, House name contains a number inside it and ends with <street>: split into House name, House ...
d6	No House name, Apartment ends <house>: move it to House name.
d7	House name ends <flat>, apartment ends with <house>: switch them.
d8	District has a comma, word before comma is <street>: add first part to Street.
d9	No District, House name is a UK town and Town is not: move Town to District, House name to Town.
e1	House Name is a number, no House number: move it to House number.
e2	No space or comma after House Number: extract the rest of the word and add it to start of Street, or move to Apartment / Ho...
e3	No House number or House name, District or Town ends with <house>: move it to House name.
e4	House name starts <flat> or is all digits: if no House number move House name to House number and Apartment to House n...
e5	No Apartment or House name, Street ends with <house>, District ends with <street>: move House number to Apartment, Str...
e6	Re-unite <orphaned word> with next non-empty address line.
e7	No district, County is a recognised UK Town, and not a county: move Town to District and County to Town.
e8	No House number, House name starts with a digit and ends with <street>: move it to House Number and Street.
e9	No House name, comma in Street, first word is 'the' or word before comma is <house>: split into House name and Street.
f1	House number includes final word which is <street>, Street or District is empty: if Street non empty and does not end with a s...
f2	No Apartment, House name has fewer than three characters: move House name to Apartment.
f3	No House number or Street, House name begins with a number and contains 'the': split into House number and Street.
f4	Apartment is just <flat>, House name starts with a number: move this number to Apartment.
f5	Town is XXX, YYY or XXX YYY where YYY is a UK town): add XXX to District (if existing District ends <street> add it to Stre...
f6	Move P O BOX nnn / PO BOX nnn / P.O.BOX nnn / P.O. Box nnn to House name.
f7	No house name, Apartment starts with a number, does not include 'flat': move the rest to House name.
f8	No House name, Apartment does not contain <flat> or any digits: move it to House name.
f9	No House name, Street has comma, word before comma is not <street>: move first part to House name.
g1	No Street, Apartment ends <street>, house name does not end <street>: move Apartment to Street.
g2	BT postcode, no country: make it UK (else it will become NI).
g3	No House name, Street has comma and does not start with digit, word before comma not <street>: move first part to House ...
g4	Town appears more often as a county and vice versa: switch them round.
g5	No District, County is a UK town: move Town to District, County to Town.
g6	No Street, House name has comma: move part beyond comma to Street.
g7	No Apartment, Street and House number start with digits: move House number to Apartment, first part of Street to House nu...
g8	Street contains comma, word left of comma is not <street>, part after comma matches validated address street: move left par...
g9	House name and Street both appear in street of validated address: move House name to start of Street.
h1	No District: copy if possible from a validated address with same town and same first four characters of postcode.
h2	House name or Apartment is 'St': move it to front of Street.
h3	No House number or Street, House name begins with digit and ends with street word: move it to House number and Street.
h4	House number ends ' St': move 'St' to start of Street.
h5	No House name, Apartment like 'Flat nn XXX <house>': move last two words to House name.
h6	No Street, Apartment like 'Flat nn XXX <street>': move last two words to Street.
h7	No House number, House number like 'XXX <house>': move to House name.
h8	Delete repeated postcode from end of street, district or town
h9	No House name, House number is one word and has no digits: move it to House name.
i1	Apartment ends with <house>, House name contains <flat>: switch them.
i2	No District, Street like 'XXX <street> YYY ZZZ' where YYY ZZZ is a District: move YYY ZZZ to District.
i3	No District, Street like 'XXX YYY <street> ZZZ' where ZZZ is a District: move ZZZ to District.
i4	No District, House name contains comma, part after comma is a district: move it to District.
i5	No District, Town like 'XXX YYY ZZZ' where XXX YYY appears as a district and ZZZ as a town, and ZZZ is not <street>: split...
i6	Street like 'AAA <street> BBB <street>': split with a comma.
i7	No House name, street like 'AAA <house> BBB <street>', or 'AAA BBB <house> CCC <street>': split into House name and St...
i8	No House no or Street, House name like 'AAA <house> nn DDD <street>': split into House name, House number and Street.
i9	No Street, District like 'AAA <street> BBB' where BBB is a district: move first part to Street.
j1	No Apartment, House name like 'nn AAA <house>: move nn to Apartment.
j2	No House number or Street, House name like 'nn XXX <street>': move to House number and Street.
j3	District and Town are the same: delete the one which appears least often in its address line.
j4	No District or Town: county like 'AAA, BBB, CCC': move AAA to District and BBB to Town.