

Importing Data in a Spreadsheet Format to C-I-SAID.

Applies to C-I-SAID build 3.5.9 - 13 December 2002 and later versions.

Introduction

This document complements the sections in the user guide on adding data from questionnaires.

C-I-SAID allows text and code data to be directly imported using a tab delimited file. However the assumption has been that this file was developed within C-I-SAID rather than externally.

This document demonstrates

- How to modify data from a spreadsheet to fit with C-I-SAID.
- How to add the key data that C-I-SAID requires.
- How to import the files into a C-I-SAID database.

This applies to any form of data but special attention is given to long text responses to open ended questionnaires.

The data has to be in tab delimited format which can be output by most applications.

Setting Up the Data Input File

Preparation

- *Set the C-I-SAID Working Folder*
- *Copy the tab delimited file into that folder (e.g. Tab.txt)*
- *Make a new database in the Document Manager (e.g. Quest)*

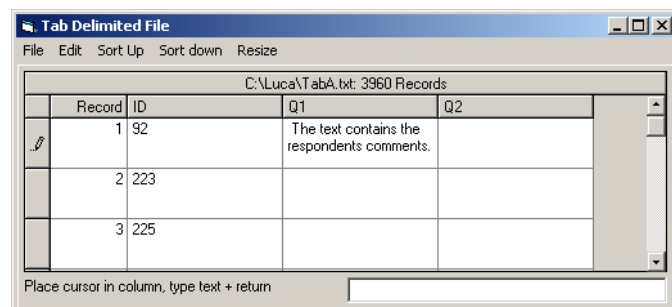
In the Report Manager

- *open the database called Quest.*
- *Select File \ Data Editor and Grid Viewer*

In the Data Editor

- *Select Open and select the file e.g. tab.txt.*
- *Select File \ Read Only – to allow the file to be edited.*

A typical example of such file might look like this in the Data Editor. In this case the data was output from SPSS.



The screenshot shows a window titled 'Tab Delimited File' with a menu bar (File, Edit, Sort Up, Sort down, Resize) and a toolbar. The main area is a grid with the following data:

Record	ID	Q1	Q2
1	92	The text contains the respondents comments.	
2	223		
3	225		

At the bottom of the window, there is a status bar that reads: 'Place cursor in column, type text + return'.

The record field is added by C-I-SAID. ID is the SPSS identifier and Q1 and Q2 are the answers to the questions in text format.

Record	ID	Q1	Q2
1	92	Text	Text
2	223	Text	Text
3	225	Text	Text

A number of changes must be made before the data can be loaded into C-I-SAID.

C-I-SAID requires three fields to be added. These are Document, Speaker and Segment

Add these fields using

- *Edit \ New Field and add the name e.g. Document*

These fields must be the first in the file.

- *Move them to their new position by clicking the field name and dragging them to the appropriate position.*

Record	Document	Speaker	Segment	ID	Q1	Q2
1				92	Text	Text
2				223	Text	Text
3				225	Text	Text

Data has now to be added to these fields.

- *Select the document field by clicking on the caption*
- *Ignore the search for box.*
- *Select Edit \ Find and Replace*
- *In the Replace With box add the name of the document "DocA"*
- *Select Find.*
- *Click Yes to add the replacement value to all the records.*

For the Speaker field add "% "

Segments have to be numbered consecutively.

- *Select the segment field*
- *Select Edit \ Add Numeric Sequence*

The segments will be numbered.

Finally the ID field contains numeric data. This has to be indicated to C-I-SAID by converting it to a numerical scale.

- *Edit \ Change Field Name*
- *Add !! to the name ID to convert it to numerical.*

Your spreadsheet should look like this

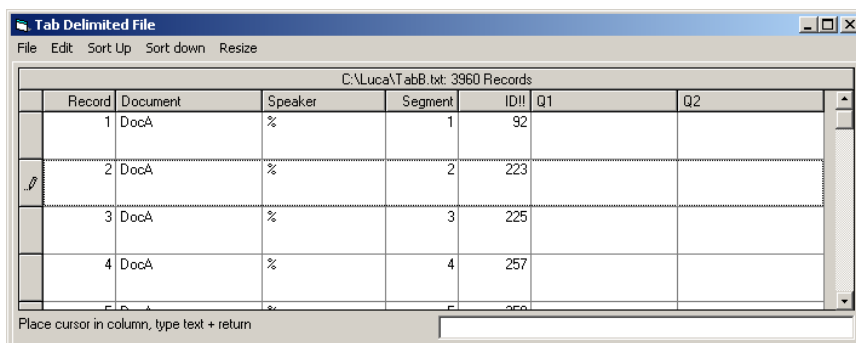
Record	Document	Speaker	Segment	ID!!	Q1	Q2
1	DocA	%	1	92	Text	Text
2	DocA	%	2	223	Text	Text
3	DocA	%	3	225	Text	Text

To use these changes we must now save the file and reopen it.

- *File \ Save Grid As \ Tab Delimited.*
- *Enter the File Name "TabB"*

Now reopen the file.

- *File \ Open – Select TabB.*
- *File \ Read Only – if you don't uncheck this you cannot edit the file.*



There is no practical limit on the amount of text in a response that is analysed in C-I-SAID. However, to gain this advantage the responses too each question have to be analysed either in separate documents or if they are to be compared as separate segments within each speech unit. The demonstration in the User Guide shows how these responses can be automatically analysed.

This means that you can only enter one text question at a time into C-I-SAID – so the file needs to be divided into two separate files. To do this we delete the questions in turn.

- *Select Q2*
- *Select Edit \ Delete Field*

Now we have to rename Q1 to "Text"

- *Select Q1*
- *Select Edit \ Change Field Name*
- *Add the name Text*

Record	Document	Speaker	Segment	ID!!	Text
1	DocA	%	1	92	Text
2	DocA	%	2	223	Text
3	DocA	%	3	225	Text

Save the file as

TabQ1

Repeat the process above only this time deleting Q1 and making Q2 text.

You have two choices when entering data – you can either have each question as a separate document or have them in a combined document. If you are going to compare or use the questions together they need be in the same document. Otherwise it is easier to have two separate documents. If you have to have two separate documents then you need to change the document name for this file.

If you wish to create separate documents

- *Select the Document Field*

- *Select \ Edit Find and Replace*
- *Enter DocA in the Search For Box*
- *Enter DocB in the Replace with Box.*
- *Press Find*
- *Press Replace All*

Now you must save the file

- *Save this as TabQ2.*
- *Close the Data Editor.*

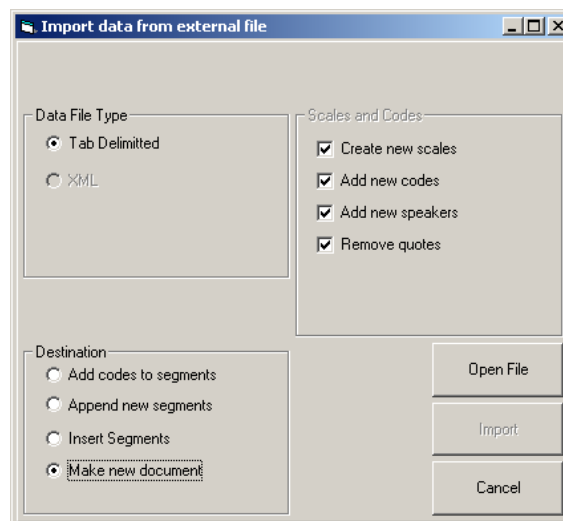
These Files Can be input into C-I-SAID.

Entering the Data Into C-I-SAID

The files are entered into C-I-SAID from the Rating Scale Manager window.

- *Open the Database*
- *File \ Open C-I-SAID Database.*
- *Select Rating \ Scale Management*
- *Click the Radio Button by Import Data From External File.*
- *Press Next*

To create a new document



- *In the destination box press the Make new Document Radio Button*
- *Leave all the scales and codes checked.*
- *Select the TabQ1.txt file from the list*
- *C-I-SAID will check the validity of the file. This includes checking the scale name, codes and speakers. New values will automatically be added to the database. It will also tell you of any problems in the text. A common problem is unbalanced brackets in the text which can be corrected later. Another is failing to set numerical scales to the correct type or trying to enter long text fields into normal text scales. In each case a message will appear saying their are too many codes.*
- *Press continue.*
- *Press Import and the New Document DocA will be created.*

If you wish to make a second document from TabQ2 (you must have changed the name in the document field as described above) then repeat the process.

If you want to insert the data into the existing file then

- *In the destination box press the Insert Segments Radio Button (This is a new button added in this version)*
- *Select the TabQ2.txt file from the list*
- *C-I-SAID will check the validity of the file and describe what will happen next. It will also tell you of any problems in the text. The most common problem is unbalanced brackets which can be corrected later.*
- *Press continue. This will take much longer than creating a new document.*

Checking the File Validity

Before proceeding you must check the file in the Document Manager. The Document Manager makes some check which cannot be made when importing data files.

- *Open the Document Manager.*
- *Select Documents \ Open Document*
- *Select the Database*
- *Select the Document.*

If yours is a large database then this may take some time to open and save.

When the document opens you will see that if you inserted segments then each Speech Unit will have two segments (or more) and the responses to each question will be against the segment number.

There are two potential problems to be aware of with questionnaire data.

First. Interviewers often put comments in brackets. If anything is in brackets it is ignored by C-I-SAID. Brackets must also be balanced within each segment.

Second. A series of responses is often separate by /. Answer1/Answer2. If there is not a space either side of a / the two words will be treated as one.

To correct this.

- *Select Search \ Find and Replace*
- *Enter "/" in the search box – add the quotes.*
- *Enter " / " replace box*
- *Press Find*
- *Press Replace All*

If you want all brackets removed – repeat this for the brackets.

Now check the text

- *Document \ Check Text.*

Correct any identified errors.

Now Save the Text.

- *Document \ Save Text.* With a big file (the one I have worked on has 8000 segments) this may take a little time. **Do not do anything in C-I-SAID until the data has been saved.**

If it finds unusual forms of apostrophe it will notify you that these are being changed before saving. The reason being that apostrophes can cause problems in processing. Processing data within the

Report Manager is not effected by file size in the same way because the process of formatting the file has already been done.