

QualAnon

The ICPSR Qualitative Data Anonymization Program

Purpose

QualAnon is web-based program to anonymize qualitative data – such as interview transcriptions, typed field notes, or other records. QualAnon is a result of a collaborative partnership among ICPSR, DSDR, and The University of Pennsylvania Population Studies Center.

The program is designed as a tool to streamline the process of removing and/or changing identifiers to maintain respondent confidentiality. Instead of a data producer reviewing each document for identifiers, QualAnon uses a name key created by the user with a list of names and corresponding pseudonyms to change identified names to specified pseudonyms. Such a tool decreases the amount of time connected to reviewing each document. QualAnon can be used to anonymize identifiers other than proper names, such as names of schools, universities, businesses, city parks or other identifiers.

The description provided here outlines the steps in using the QualAnon program.

Input:

- Documents to be anonymized need to be in Rich Text Format or ASCII text.
- User generated Name Key needs to be in tab-delimited ASCII text or Excel spreadsheet.

Output:

- Anonymized document: [document name]_anon.rtf or [document name]_anon.txt (output will be in the same format as the input document).
- Anonymizer summary report: anonymizer_rep.txt.

Running QualAnon

Step 1: Create a name key.

- The name key should include all the names that the researcher would like to anonymize.
- The name key must be a tab-delimited ASCII text file or an Excel spreadsheet file. The Excel file should have the name key on only one sheet.
- The first column includes original names. The second column specifies the replacement names.
- First and last names and nicknames need separate entries. In the example below, Farai Zelda is the first and last name, each are entered separately and assigned pseudonyms
- Users can include multiple spelling variations of the same name to be changed to anticipate typographical errors. For example, Phillip is assigned the pseudonym John, however, the user can anticipate that Phillip might be misspelled as Philip, and can enter Philip as another variation to be changed to John.

Example:

Agatha	Abigelo
Ajusa	ALFESO
Amadi	ABATA
Fei	FAUSTA
Sevi	SAINET
Deborah	DAILES
DeClerks	D.
Edie	ELENI
Emma	Rose
Ernest	ELIAM
Fanizo	Fred
Farai	FOUNTAIN
Zelda	ZUNDA
Hanock	HELLINGS
Jambo	JAFIFI
Lovemore	LONJEZANI
Mwenyemanzi	MAJONGOTI
Malikita	MABVUTO
Ngaurye	NAMBEWE
Oga	OLINI
Patricia	Eveness
Phillip	RAPHAEL
Philip	RAPHAEL
Runes	REDESI
Muhone	NDEGE
Seleman	SAIDI
Sharai	SAWSYEJE
Simba	SEREVASI
Wilford	Geoffrey
Yuda	YOUNGSON

Step 2: Upload files to be anonymized.

QualAnon can either anonymize files individually or in batch mode. To anonymize multiple files in batch mode, files must be uploaded in a [Zip file](#).

To upload an individual or Zip file, select the “Browse” button next to the “**Transcript** file to upload”

Transcript file to upload:

Step 3: Upload NameKey file.

Upload the NameKey file to be used to anonymize the files from Step 2 by selecting the “Browse” button next to the “**Name Key** file to upload” selection.

Name Key file to upload

Step 3: Choose options for anonymization.

There are two output options to QualAnon.

Annotate anonymized output with original names

This option produces anonymized documents with notations where changes were made. For example:

Shawn [Patrick] added, 'Indeed, and from rich families'and begin naming them like Becky [Annie] , Ngawechinyalo [Hawa] , Luwiza [Beatrice] , both Beatrices, that's Luwiza [Beatrice] Maulida [Daffen] and Luwiza [Beatrice] Jafeti [Chande] , etc. (I can comment that both girls mentioned indeed come from a rich family)'He said he was selecting the young and healthy schoolgirls, more [carefully] than anyone else when he was at primary.

Allow duplicates across columns in name key -- use with caution!

Use this option with caution. The user must verify that the Name Key does not include pseudonyms that are the same as original identifiers. For example:

Russ --> Brian
Brian --> Joe

yields Russ --> Joe

One way to get around this is to pay close attention to the order of the name key and select the “allow duplicates.” For example:

Brian --> Joe
Russ --> Brian

yields Russ --> Brian

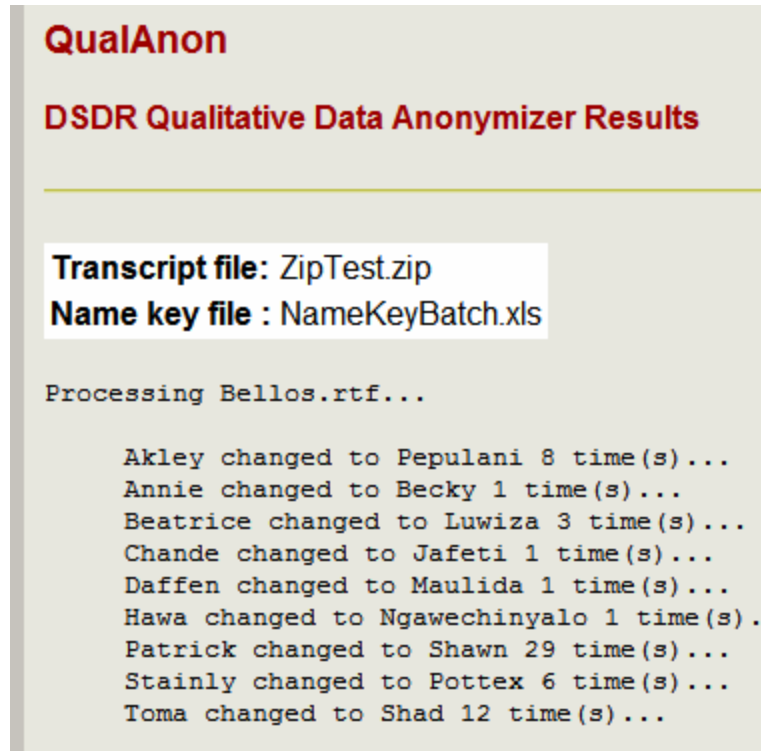
Step 4: Terms of Use

Click on “**I agree** to the QualAnon terms of use below (*required*)” to indicate that you have read and agreed to the QualAnon terms of use. QualAnon will not run unless this has been clicked/selected.

Step 5: Start QualAnon

Click the “Upload Files” button to start the program.

As QualAnon runs, you will see a running report of the changes being made to your uploaded documents. QualAnon will provide a complete change report at completion.



At the completion of the QualAnon run, select “Download anonymized file” at the bottom of the running report to download your anonymized files and the change report.



Important Note: QualAnon will only hold the anonymized files for download for FIVE (5) hours. If you do not select “Download anonymized files” within five hours of QualAnon completing the anonymization process, the files will be deleted from the server and you will have to re-run QualAnon.

QualAnon will create a Zip file for download that includes a report file (e.g., anonymizer_rep.txt) as well as fully anonymized or annotated anonymized files.

Name	Type	Modified	Size	Ratio	Packed	Path
anonymizer rep.txt	Text Document	9/28/2007 9:33 AM	4,166	80%	816	
Bellos_anon.rtf	Rich Text Format	9/28/2007 9:33 AM	74,298	73%	20,402	
George2_anon.rtf	Rich Text Format	9/28/2007 9:33 AM	81,910	74%	21,510	

Example Output

Annotated Document

Here is an excerpt from an annotated transcript:

YOUNGSON [Yuda] was saying this while shaking his head disappointedly and that was showing that he had already started recalling some of the things he was doing with girls at that time. Then MABVUTO [Malikita] said, “I am sure that none of the guys from Muriel Mine who were at our age at that time will be spared with that AIDS because that was the time when we shared that disease at that place.”

The original names are enclosed in [] preceded by the pseudonym as assigned in the name key.

Fully Anonymized Document

Here is an excerpt from a document without name annotations:

YOUNGSON was saying this while shaking his head disappointedly and that was showing that he had already started recalling some of the things he was doing with girls at that time. Then MABVUTO said, “I am sure that none of the guys from Muriel Mine who were at our age at that time will be spared with that AIDS because that was the time when we shared that disease at that place.”

Summary Report

The summary report identifies the number of times a name has been changed in the transcript:

=====

ANONYMIZATION RESULTS

Date : Fri Sep 28 09:33:40 EDT 2007
Original transcript : Bellos.rtf
Name Key : namekey_23421
Anonymized transcript: Bellos_anon.rtf
Annotated : Yes

<u>NUMBER</u>	<u>ORIGINAL NAME</u>	<u>PSEUDONYM</u>
8	Akley	Pepulani
1	Annie	Becky
3	Beatrice	Luwiza
1	Chande	Jafeti
1	Daffen	Maulida
1	Hawa	Ngawechinyalo
29	Patrick	Shawn
6	Stainly	Pottex
12	Toma	Shad

Questions

If you have any questions about QualAnon, please contact Russel Hathaway at rhataway@umich.edu or 734.615.9525.