



social research UPDATE

This *Update* describes ingenious uses of software and websites including:

- The use of autocorrect to create a shorthand that makes it quicker and easier to transcribe qualitative data
- The use of desktop search engines and their application as a basic tool for analysing uncoded transcripts
- Intelligent search strategies that improve the effectiveness of Internet and desktop searches
- The use of ATnotes as virtual post-it notes

Exploiting freely available software for social research

Ruth Rettie

This *Update* covers a range of readily available resources which are useful in social research¹. Although some of these are particularly useful for qualitative researchers, they are also relevant to writing and researching academic papers. The first section explains the use of the Word autocorrect function, which can be used to create a predictive text that is particularly useful for relieving the chore of transcription. The second section describes and compares three desktop search engines, and suggests how they can be used for the analysis of uncoded qualitative data. The final section briefly mentions several other useful programs and websites.

Autocorrect

In Microsoft Word the autocorrect option can be used to make transcription quicker and easier. It enables one to create a personal shorthand for any words or phrases, potentially reducing keystrokes by two-thirds or more. Autocorrect codes can be used for any text including standardised questions, phrases, and respondent names. For example, the keystrokes 'yk' can be used for 'you know' and 'I don't know' can be reduced to 'idk'. The autocorrect option is on the standard Microsoft Word² Tools menu (the shortcut keys are AltT, shiftA). First make sure that the 'replace text as you type' box is ticked, then type the desired shorthand in the column headed 'replace' and the desired text in the column headed 'with', then click 'add'. To use autocorrect,

one simply types the shorthand followed by the space bar or enter key. You do not have to use the shorthand: the word can still be typed in the normal way, because the autocorrection only takes place when the space or enter keys are pressed.

Autocorrect can be developed to cover all the most common words and phrases, and tailored to cover specific research topics and respondents' language. The shorthand keystrokes are chosen by the user, but it obviously makes sense to keep single letters for the most common words, for instance, 'b' for 'but', 'n' for 'not', 'c' for 'can' etc. In addition, I adopt a basic grammar, so that if 'ph' is phone, 'phs' is 'phones', 'phd' is phoned, 'phg' is 'phoning' and so on. It makes sense to add these derived codes when defining the original code. When choosing the shorthand codes it is important not to use whole words, for example using 'bus' for 'business' creates a problem when one wants to refer to a bus! Similarly one should avoid using codes such as 't' or 've' or 'll' which often occur after an apostrophe, because the program treats these as new words and autocorrects. It is easy to change autocorrections which are not wanted, simply backspace and retype normally. Codes can be removed or changed from the autocorrect menu.

Phrases can include punctuation, which is useful for common transcription phrases. I use 'clt' for '((clears throat))' and, for inaudible sound, I type 'inau' which is autocorrected to '()'. Autocorrect can

Ruth Rettie is a senior lecturer at Kingston University teaching Internet marketing and qualitative market research. She is also doing a part-time PhD in sociology at the University of Surrey, researching social interaction on mobile phones. Her research interests include qualitative research methodology and social studies of technology.

social research UPDATE

also personalise the keyboard, replacing frequently used awkward keys with easier ones. For instance, when transcribing, I change the semicolon to the '/' sign used to indicate overlapping speech.

It takes a while to enter the phrases and learn the shorthand codes but this can be done gradually, entering new autocorrect entries as the work progresses. The reduction in keystrokes relieves and may protect against Repetitive Strain Injury. Once one knows the shorthand, it can also be used to take field and observation notes, which can then be typed up directly; autocorrect automatically translates the codes into the relevant words and phrases.

Autocorrect works for any of the Microsoft Office programs: any shorthand entered will be automatically available in Word, Excel, Powerpoint and, for emails, in Outlook. It can be switched off for any program by changing the options on the autocorrect tab of the Tools menu. Autocorrect can even be used for whole paragraphs, tables and figures. This is very useful when editing documents, because one can rearrange several paragraphs at the same time. To do this, type the passage normally, select the text or figure, and go to autocorrect. The selected text will automatically appear and one merely has to enter an appropriate code. For editing, autotext (also on the autocorrect tab of the Tools menu) is sometimes preferable. For autotext one has to enter the replacement by selecting text in the document; autotext usually uses the first four keystrokes of the relevant text as the code, but this can be changed. The replacement phrase is flagged above the text and is inserted only when the 'enter' key is pressed. Usually autocorrect is more convenient because the 'enter' key is less accessible than the space bar when typing, but the advantage of autotext when editing is that it displays the suggested replacement.

A potential problem is that the autocorrect feature is language specific. If it is created when the document language is set to UK English, it will not be available in US English. This can create a problem when copying and pasting direct from other documents as the inserted text may include a change of language. In this case the language can be

reset by selecting the whole document and using the 'set language' option on the Tools menu. Autocorrect is also responsible for the automatic correction of spelling mistakes as one types. To maximise this feature make sure that the 'automatically correct spelling mistakes' box in the autocorrect menu is ticked, and always try to correct spellings by using spell checker suggestions (rather than retyping or correcting the word) so that the system learns to correct your typos automatically.

Having developed a personalised autocorrect dictionary, it is useful to be able to use it on different computers. The easiest way to do this is to download 'autocorrect.dot' from the MPV³ site, unzip the file, open the resulting Word file (enabling macros is necessary) and then click 'backup' to create an autocorrect backup file. Transfer this file and the autocorrect.dot program to the second computer, open the autocorrect.dot file and select 'restore', navigating to the saved autocorrect backup file. The autocorrect entries on the second computer will now be the same as on the first computer.

The next section describes the use of desktop search engines for the preliminary analysis of uncoded qualitative data.

Desktop Search Engines

Desktop search engines work like Internet search engines and are very fast because they have pre-indexed all the words in the computer. Although it is not their main application, they can also be used for a preliminary analysis of qualitative data. There are three main alternatives: Google³, Copernic or Yahoo. They all search most file formats including email attachments, Acrobat⁵ files, music files, etc. Google integrates with the Google Internet search engine, so that when searching the Internet it is possible to see if one already has relevant documents on the computer, and it is very quick. However, unlike Copernic and Yahoo, the Google desktop search engine does not currently include a file viewer or allow one to restrict a specific search to a particular folder, which means that it is not suitable for the analysis of qualitative research. Yahoo uses less computer

social research UPDATE

resources, but I find the Copernic interface and viewer generally better for qualitative analysis.

Desktop search means that one can quickly identify the responses to specific questions in uncoded qualitative data, and instantly look through these without opening the files. First, it is necessary to put all the transcribed files in a single folder, so that one can limit the scope of the search to those files. In Copernic this is done by specifying the folder in the 'refine search' column⁶. One can then search the folder, entering appropriate search terms, and immediately locate all the transcripts that contain those terms. Having found these, one can then instantly examine all the occurrences of these terms within each transcript. Select the name of the file and then click on the highlight button for each search term or phrase. This can be used to collate an overview of sample characteristics, which is useful when deciding whether to extend a sample. One can quickly search on each question and read the answer, without having to open and search each transcript separately. Desktop search can also pick out those interviews in which specific words and combinations of words occur, using Boolean search. For example, I recently used Copernic to find the relative salience of words such as 'personal' and 'sensitive' in 32 transcripts. I quickly found transcripts which contained both terms (search term 'personal + sensitive'); those with only one of the words (search term: 'personal OR sensitive') and those with 'personal' but not 'sensitive' (search term: 'personal -sensitive'). To find exact quotations use double quotes e.g. "phoning is more personal". Although some CAQDAS programmes such as Qualrus include quite sophisticated search engines, they are not pre-indexed and so are much slower. Desktop search is also useful as a precursor to CAQDAS coding; looking at the frequency of particular terms in the data helps in the choice of appropriate codes.

Whether searching the desktop or the Internet the selection of search terms is crucial and this is discussed in the next section.

Intelligent Searching

The choice of search terms is undoubtedly more important than the search engine used, especially since the differences between search engines have declined over time. These suggestions are based on Google.com which is the most popular, partly because the Google toolbar⁷ is so convenient. I use two strategies that increase search effectiveness. Firstly it is important to remember that search engines (unlike search directories) are not categorised by topic, so that rather than searching with a phrase that describes what the sought after web page is *about*, one should search for words that are likely to occur *on* the target pages. It helps to think of a search program mindlessly collecting text. I try to imagine phrases that are likely to be in the pages I'm looking for. For example, when searching for papers on ethnomethodology an effective search term is 'Garfinkel Studies in Ethnomethodology 1967.' This is because academic papers on this topic would inevitably cite this as a reference. A similar tactic is to use a quotation relevant to the area. Secondly, to avoid having to look beyond the first few pages, the search terms should be as specific as possible. The perfect search term is something precise and unique, such as a phone number. To make the search more specific I would add the term 'pdf' (as most academic papers would be in this format) the term '-outline' (to exclude course outlines) and 'site:ac.uk'⁸ which restricts the search to UK universities. Sometimes I restrict the search by going to Google advanced search (from the normal Google search page) and select only sites updated in the past year⁹. This particular search resulted in just two pages of results (about 100 sites) compared to over 57,000 when the search term was just 'ethnomethodology'. From the Google toolbar one can quickly scan search results (and web pages) by typing appropriate terms directly into the toolbar and clicking 'highlight' (and not the 'enter' key). Searching is more effective when clearly directed in this way, but it is unsystematic and can introduce bias: it is all too easy to find a paper making any particular claim (by searching on that claim) while neglecting

social research UPDATE is distributed without charge on request to social researchers in the United Kingdom by the Department of Sociology at the University of Surrey as part of its commitment to supporting social research training and development.

Contributions to social research UPDATE that review current issues in social research and methodology in about 2,500 words are welcome. All UPDATE articles are peer-reviewed.

social research UPDATE

papers with alternative findings.

Google has two other search engines which are sometimes useful. Google Scholar (<http://scholar.google.com>) has a more academic focus. Its ranking takes into account how often a paper has been cited in 'scholarly' literature, so it indicates the relative popularity of different papers or authors. Google Book Search (<http://books.google.com>) searches the full text of books rather than websites and, subject to the compliance of publishers, is potentially an excellent resource for academic research. Best for recent books, it can be used to check the page number of a quotation. Google is also useful for checking spelling (just search on the word and Google will suggest an alternative spelling if it is incorrect); definitions (type 'define:' followed by the word); synonyms (thesaurus from the toolbar); calculations (just type in the calculation) and topic specific news (set up alerts at <http://www.google.com/alerts>). If you register for Personalized Search (<http://www.google.com/psearch>) Google will store and learn from your previous searches. For detailed advice on using Google see Dornfest and Calishain (2005).

Other Resources and Suggestions

Obviously it is not possible to give an exhaustive description of the vast array of software and useful research websites¹⁰ available, but this section suggests a few other resources. **Furl.net** (<http://www.furl.net>) is useful for keeping bookmarks on the Internet and therefore accessible from any computer. Cite-u-like (<http://www.citeulike.org>) stores academic references. These are publicly available and searchable, creating a useful source of references in a particular area. Wikipedia (<http://en.wikipedia.org>) is a quick source of information. Although it can be edited by anyone, it is increasingly cited.

When searching the Internet one sometimes comes across undated pages, where the status of the information depends on when it was updated. An easy way to date a web page is to type 'javascript:alert (document.lastModified)' directly into the web browser

address bar. In both Internet Explorer and Firefox this brings up a panel showing the date when the web page was last updated.

In addition to the resources above, ATnotes¹¹, a personal favourite, allows one to create virtual 'post-it' notes which can be displayed, hidden, or set as audiovisual alarms. With ATnotes one can keep notes (e.g. points to include or check, code descriptions or lists, etc.) on screen at the same time as the active document. Unlike paper notes they cannot be lost and can be searched.

Caveat

The software solutions suggested here are subject to two limitations. Firstly, although the information here is currently correct, the affordances of the various programs and websites will undoubtedly change. Secondly, response to software is individual and often polarised. Although some of those who have tried these suggestions share my positive views, others find them frustrating and irritating.

Notes

- 1 I thank Nick Allum for suggesting I write this article and for his helpful comments or an earlier draft.
- 2 Autocorrect is in available in all versions of Microsoft Word for PC and Mac since 2000, and in Microsoft Works. Star Office also has an autocorrect function.
- 3 Available at <http://word.mvps.org/FAQs/Customization/ExportAutocorrect.htm>
- 4 Google (<http://desktop.google.com>) and Yahoo (<http://desktop.yahoo.com>) only work with **Windows XP or 2000**. Copernic (<http://www.copernic.com>) works on all versions of Windows since 98. The latest Mac systems include the desktop search Spotlight.
- 5 Desktop search engines only index pdf files if they are in renderable text; these can be converted using the 'paper capture' feature on the Document menu of Acrobat Professional.
- 6 In Yahoo click the 'refine' button and then type a distinctive part of the file name under 'path'. Select any file and then scroll through. Search terms are highlighted but difficult to read on a small screen.

- 7 The Google Toolbar is downloadable free from <http://toolbar.google.com> for Windows with Internet Explorer, or from <http://toolbar.google.com/firefox/> for Window, Mac and Linux with Firefox.
- 8 'Site:edu' restricts a search to US university web sites.
- 9 See Dornfest and Calishain (2005) for how to restrict a search to a specific date range.
- 10 Other sites worth exploring are <http://www.sosig.ac.uk> ; <http://www.ssm.com>; <http://cogprints.org>; <http://citeseer.ist.psu.edu> and <http://plato.stanford.edu>
- 11 This is a very small program for Windows PC only, downloadable free from <http://atnotes.free.fr/introduction.html> or from my website <http://www.kingston.ac.uk/~ku03468>. The scratch pad on the latest Google Desksearch side panel is somewhat similar.

Reference

- Dornfest, R., and Calishain, T., (2005) *Google Hacks: Tips and Tools for Smarter Searching* 2nd Edition. O'Reilly, Sebastopol, CA.

social research UPDATE

(ISSN: 1360-7898)

is published by

Department of Sociology
University of Surrey
Guildford GU2 7XH
United Kingdom.
Tel: 01483 300800
Fax: 01483 689551

Edited by Nigel Gilbert
(e-mail: n.gilbert@surrey.ac.uk)

This issue and all back issues are also available on the web at <http://www.soc.surrey.ac.uk/sru/>