Quality & Quantity (2005) 39:581–614
DOI 10.1007/s11135-005-3149-z 

# More than Typewriters, More than Adding Machines: Integrating Information Technology into Political Research

KEN COUSINS and WAYNE MCINTOSH
*Government and Politics, University of Maryland, College Park, MD 20742, U.S.A.*
*E-mail: kcousins@gvpt.umd.edu*

**Abstract.** It has long been recognized that information technology (IT) can facilitate (or even permit) forms of empirical analysis unimagined just a generation ago. However, political scientists have generally experimented with only a narrow range of the possibilities the new technologies offer – easier writing, editing and communications processes, and more sophisticated statistical analyses. While these have undoubtedly increased productivity and rigor within the discipline, other IT applications have largely gone unnoticed. The growing pool of Computer-Aided Qualitative Data Analysis Software (CAQDAS) technologies presents researchers with new ways to conceptualize, perceive, and communicate their research, with the potential of revolutionizing social science research. This paper offers an overview of those IT and CAQDAS solutions with the most potential to facilitate political inquiry, and offers a series of practical steps by which technological novices might begin to apply the new technologies in their research. In closing, we discuss the benefits and perils of using such technology, and suggest the ways by which IT might strengthen current research techniques. The appendix catalogues several dozen applications for data acquisition, organization, processing, analysis and presentation, with contact and pricing information.

**Key words:** analytical software; qualitative analysis; quantitative analysis; social science research methods.

> *"By 'augmenting human intellect' we mean increasing the capability ... to approach a complex problem situation, to gain comprehension ... and to derive solutions ... more-rapid comprehension, better comprehension, the possibility of gaining a useful degree of comprehension ... We do not speak of isolated clever tricks that help in particular situations. We refer to a way of life in an integrated domain where hunches, cut-and-try, intangibles, and the human 'feel for a situation' usefully co-exist with powerful concepts, streamlined terminology and notation, sophisticated methods, and high-powered electronic aids."*
>
> (Engelbart, 1962)

## 1. Introduction

A half-century into the information revolution, we are faced with a new Malthusian dilemma:[1] while the volume of information explodes exponentially, our ability to produce genuine knowledge from that information appears limited to only arithmetic increases. In past decades, advances in computing technologies have been applied throughout the physical sciences, enabling researchers to address questions of greater scale and complexity with better accuracy, consistency, and transparency. Similarly, linguists have found ways to utilize the processing and analytical power of computer technology to assess large quantities of text in ways not feasible manually. Other social scientists, however, have been relatively late in utilizing new technologies to help them sift through and analyze text-based and other non-quantitative datasets.

Early attempts to harness computing power for political research were deductive efforts to model political and military decisionmaking (Druckman, 1994, cited in Starkey and Blake, 2001). When these dreams proved unattainable, interest in computer-facilitated political research receded, at least until the personal computer became commonplace, and then computing power was often applied only to assess statistical correlation. Until the past decade, data to test or inform such models were rare; researchers usually had to create the datasets themselves, often coding from (hardcopy) print sources. Because such data principally served quantitative analyses, coding tended to focus on "thin" characteristics (e.g., legal or economic variables) which could be generated at the scale needed for statistical inference to be possible. Even after information sources became widely available in electronic form, quantitatively minded researchers have still largely treated them the same, even to the point of printing hardcopies of electronic sources for later coding. However, in the past decade and a half, a number of European social scientists have devised innovative means to analyze and compare text-based data (Gibbs et al., 2002), and several legal services firms have created proprietary systems for organizing, searching, analyzing, and synthesizing vast quantities of documents to support litigation projects of epic proportions (Emery, 2003). Social scientists are gradually beginning to realize Engelbart's dream – to use information technologies to strengthen, expand, and transform intellectual effort.

### 1.1. FROM DATA TO KNOWLEDGE

In December of 2000, the US Library of Congress initiated the National Digital Information Infrastructure and Preservation Program, a $99.8 million effort to develop a comprehensive strategy to "save America's cultural and intellectual heritage in digital formats" (LOC, 2002). The BBC recently announced that it would be make its entire television and radio program archive (the world's largest) available via the internet (BBC, 2003). Since 1996, the Cyberspace Policy Research Group (a joint project of the University of Arizona and George Mason University) has identified over 190 governments with national-level public agency websites (La Porte et al., 2001). Although Howard Dean became a temporary frontrunner in the race for the US Presidency through the ability to mobilize over 600,000 email supporters, the Bush campaign is said to maintain a list *ten times* as large (PoliticsOnline, 2004). Non-state political actors had discovered the Internet's potential long before this election cycle (Finel and Lord, 2002; McLaughlin, 2003). On all fronts, the volume of politically relevant digital content is exploding, providing a body of evidence for political scholars that has already become overwhelming.

While "information glut" is not a new phenomenon (Emery, 2003), the exponential growth of digital content shifts the challenge of identifying (and understanding) critical data into an entirely new realm.[2] As this process deepens, social scientists must develop more efficient and rigorous means of dealing with the scale of available data. Tracking new and existing data sources, and establishing quality controls for their use and storage, have become critical issues. This is true not only of textual sources, but increasingly of audio, video, and multiple-media forms, as well. Just as "digital convergence" is changing the way we communicate, it poses new opportunities and challenges for researchers, with the potential to transform analytical practices, individually and as a community of scholars (Brown, 2002; Gibbs et al., 2002).

### 1.2. CHANGING RESEARCH OPPORTUNITIES

Applied social research (e.g., law, marketing, public health, intelligence) has generally led in the use of computer-supported analysis (Fielding, 2002), a trend which began long before electronic content became widely available (Heesterman, 1993). While much of the data was generated on a by-project basis (e.g., surveys), some of the more ambitious efforts have worked with monumental volumes of paper-based evidence which need to be analyzed by teams of researchers, often working for several years, from multiple locations. Such "knowledge management" systems, supporting both large-scale legal projects (e.g., class-action lawsuits) and military intelligence analyses, have integrated text files of digitized source documents (via optical character recognition software) with "fuzzy" search algorithms and electronic notation systems (Emery, 2003). Practitioners are able to search the full corpus of evidence (and comments by other researchers) to develop analyses tied directly to original records, often located in disparate and unstructured databases. Projects such as these raise fascinating questions about the potential scale, scope and speed of research. Though the methodologies and rules-of-evidence may differ, the information management tasks of such projects are not fundamentally different from that of social science.

Unfortunately, most political science research training (especially in the United States) has yet to catch up with the possibilities of a world where information is increasingly digitized (Fielding and Lee, 2002). While most researchers are now familiar with email and Internet searches, relatively few American political scientists have learned to integrate other technological capacities into their research (apart from the growing use of statistical packages). Few examples of Computer-Assisted Qualitative Data Analysis Software (CAQDAS) fully exploit the capacity of the newer technologies to support research (Kerlin, 2000). Most are idiosyncratic, driven by scholars' particular skills and methodological focus (Burton, 2002). Most often, these solutions appear to be motivated by the potential to improve research efficiency, through semi- or fully-automated data processing (Brown, 2002). Few projects have tried to integrate the efforts of multiple researchers, methods or disciplines, applied to multiple source media (Bourdon, 2002). To the degree that such projects have been successful, they have emphasized the importance of establishing clear, formal guidelines for the interpretation of data, as well as maintaining *audit trails* for each participant (by which each instance of the interpretation of evidence can be made explicit) (Bong, 2002). Many would argue that formalism and transparency have long been held to be central pillars of the scientific process regardless of whether research is qualitatively or quantitatively oriented (King et al., 1994; Morse et al., 2002).

### 1.3. CUMULATIVE SOCIAL KNOWLEDGE

Anyone who has worked on one of the larger quantitatively focused database projects is poignantly aware of the potential for "conceptual slippage" during the coding process. Not only are there challenges of inter-coder reliability, but individual coders face problems of consistency with themselves, even over relatively short time intervals (Marshall, 2002). One response is to develop clear, formal coding guidelines (Nagler, 1995), but this can sometimes limit researchers' ability to fine-tune theory, as they shift between inductive to deductive modes in the course of interrogating the evidence. The demands of rigor and transparency are constant challenges for researchers, whether they work alone or in teams.

Ultimately, the subjective nature of inquiry and interpretation is inescapable, even where efforts to achieve accuracy and consistency are relatively successful. This truth is not limited to qualitative research (nor social science generally), and need not limit our ability to develop shared or cumulative knowledge. The keys to achieving inter-subjective understanding are transparency and replicability, two goals facilitated by digital media and computing technologies (Brown, 2002). Thus, a shared and cumulative social knowledge need not be limited to quantitative interpretations, but may emerge from an explicit stating of hypotheses, assumptions, and interpretative rules, combined with the ability to store original data and record each step of an analysis (Thompson, 2002). While there are no silver bullets, information technologies (IT) can relieve the logistical burden of maintaining audit trails and collaborative research (Bong, 2002).

## 2. Modes of Inquiry

Because we are concerned here about the ways by which IT can facilitate social science research, our scope extends beyond the world of analytical instruments. Indeed, one of the main reasons the rise of digital technology is considered "revolutionary" has been the explosive growth in the volume of available data, linked to a parallel growth in our capacity to communicate and collaborate, regardless of physical distance (NRC, 2003). Even those who have yet to connect to the "virtual world" are still likely to have experienced a shift in their personal, social, cultural, or political worldviews, as both content and communications technologies have become increasingly fluid throughout the world (Kinnevy and Enosh, 2002). While it has been common to focus merely on the technical aspects of these changes, the digital world presents enormous opportunities and challenges for social science, as well (Burton, 2002).

In this paper, we present an overview of those IT solutions with the most potential to facilitate the various tasks of political research, including: collecting evidence (*acquisition/organization*); developing knowledge from data and building theories (*analysis*); and facilitating inter-subjective research (*collaboration*) (Brown, 2002). Rather than focusing on the types of research these applications support (i.e., qualitative vs. quantitative), we will instead emphasize the instrumental[3] functions they can perform for users (e.g., archiving, data management, text searching, text coding, code retrieval). Appendix A summarizes the currently available applications and categorizes them according to functionality. Appendix B presents a catalog of applications, with their respective analytical capability, location, and cost.

### 2.1. ACQUISITION AND ORGANIZATION

Most scientists are by now familiar with a variety of electronic means of acquiring and organizing secondary research materials, through library catalogues and online databases and journals (Mann and Stewart, 2000).

Many have also understood that the tools they use to access such content can serve analytical goals (e.g., "hit counts" on keyword searches). Both traditional (e.g., newspaper and television) and "new media" sources (e.g., websites, weblogs, listservs, USENET) can be automatically searched, monitored and archived, and these files can be manipulated as data (List, 2001). Yet, as the volume and variety of digital content explodes, researchers are faced with new questions about the means of collecting, processing and publishing data, in order to produce quality research (Zhang 1999; Gibbs, Friese et al. 2002). The factors which have facilitated the growth in digital content have also often made that content malleable and ephemeral, subject to partial or wholesale change in an instant.

Despite appearances, our means of identifying content (and changes to content) remain quite limited – search engines are still only able to index less than one-fifth of the content available on the World Wide Web[4] (Thelwall, 2001). It should come as no surprise that each of these tools systematically excludes some content, while emphasizing others. Although such biases are both intentional and accidental (varying from tool to tool), researchers who draw on Internet-based content need to be aware of these limitations (Introna and Nissenbaum, 2000).

It can also be tempting for regular Internet users to project beliefs about "real world" communities onto the virtual domain. While in some cases this may be an accurate expectation (Hill and Hughes, 1997), understanding the context (e.g., membership and norms) of digital interaction is not always a straightforward exercise (Jones, 1999). This may be a critical issue for some researchers – many of the applications which facilitate online communication (e.g., email, instant messaging, and chat rooms) can also support "remote interviewing." Although these can sometimes appear functionally similar to more traditional means (e.g., telephone interviews), participants may perceive both the applicable rules and even their own identities differently in a virtual context (e.g., online focus groups) than in the "real world." Indeed, the medium of communication can strongly condition the content of the exchange. Researchers need to consider the potential effects of these differences on the interaction between themselves and their interview subjects, between the subjects themselves, and between subjects' actions and attitudes as expressed online and in the real world (Bampton and Cowton, 2002).

#### 2.1.1. *Non-text media and transcription*

The rising popularity of digital audio formats has led to an enormous variety of affordable portable recording devices. The compression technology which facilitated this popularity now means that field researchers are able to record hours (even days) of interview material, which can then be transferred to a personal computer for archiving and analysis (Stockdale, 2002).[5]

Depending on the purpose which interviews serve in an analysis, transcription (conversion of voice to text) may or may not be worth the effort. There are several applications now available which enable researchers to attach memos to segments of audio and video files, reducing the need for transcription (Ford et al., 2000). For projects where interviews are more useful in text form, several applications are available to improve transcription efficiency. Unfortunately, it is still not possible to fully automate the conversion of recordings to text – current voice recognition software must be trained to each new voice. However, some users have achieved very high conversion speeds (up to 130 w.p.m.) by listening to the audio feed and repeating (even translating) that dialogue into voice recognition software themselves.

### 2.1.2. *Surveying*

Just as the Internet and the Web present opportunities to produce richly detailed qualitative data, they can also be used to gather survey data from geographically dispersed respondents, with functionalities not easily achieved through traditional paper-based approaches (Fricker Jr and Schonlau, 2002). Beyond the potential to reduce costs, they also offer researchers the ability to track variables such as latency (i.e., timed responses), and to produce focused information more efficiently, posing different questions to respondents based on their previous answers (i.e., interactive surveys) (Schonlau et al., 2001). As with all research, users need to be aware of potential sampling biases in the distribution of Internet and Web users, as well as their responses rates, as compared to traditional mail surveys (Solomon, 2001).

### 2.1.3. *Databases*

Anyone who has produced a table or spreadsheet has created a database. The Internet itself can be thought of as a vast, distributed database, with heterogeneous pockets of organized data amidst a sea of unstructured, unstable content. The diversity of ways to structure information presents a challenge to every user, even more so for researchers who want to compare or combine data from multiple sources. However, despite sometimes daunting initial costs, databases can be enormously useful tools, providing steadily increasing returns to scale. The range of potential research applications include: time-series and event data (Thomas, 2000; Hochheiser, 2003); the actors and institutions relevant to a given issue area (Corti, 2000; Corti et al., 2000); and geographically referenced data (Eagles et al., 2000; Sui and Hugill, 2002).

Many scholars have also begun to realize the usefulness of citation management software, most of which can fully integrate with word processors (enabling authors to automatically change citation formats), and some which can directly access library databases, eliminating the need for data entry.

Given the high cost of generating database content, there have been attempts to automate acquisition processes. These typically draw on sources with a relatively consistent means of presenting information (e.g., wire reports), parsing that descriptive text to identify a select range of variables (Bond and Jenkins, 1997; Bond et al., 2003). Although the range of source data which can be parsed by these applications is often restricted, some claim to have achieved reliability and accuracy levels equivalent to human coders (King and Lowe, 2003).

### 2.2. ANALYSIS

As we have already suggested, political scientists have generally experimented with only a narrow range of the possibilities these new technologies offer – easier writing, editing and communications processes, as well as more sophisticated statistical analyses. While these applications have undoubtedly increased productivity or rigor within the discipline, other IT solutions have largely gone unnoticed.

Debate over the relative merits of qualitatively or quantitatively oriented research has raged off and on for years (Creswell, 1994; King et al., 1994; Leydesdorff and Wouters, 1996; Brace-Govan, 2002). It is not the purpose of this essay to throw another log on that fire, but rather to discuss the potential of information and computing technologies to support social science research generally. Given that the technologies supporting statistical analysis are familiar to most social scientists, we focus here instead on applications which are less commonly known. These so-called CAQDAS[6] technologies offer scholars additional ways to conceptualize, manage and communicate their research (Lee and Esterhuizen, 2000; Gibbs et al., 2002). Whether a researcher approaches her topic from the viewpoint of grounded theory, phenomenology, narrative, ethnography, or case study, computer applications have been developed to *facilitate* such analyses[7] (Fielding, 2002; Merriam, 2002). Just as some methods are based on similar beliefs about what constitutes "relevant" or "appropriate" evidence, many applications are capable of serving multiple methodological modes. Here, we focus on these functions in isolation for the sake of clarity.

#### 2.2.1. *Text retrieval, concordance*

The analysis of text (broadly conceived) is the foundation of traditional social science research (Roberts, 1997). As a principal means of communication, text provides a record from which theories can be developed, and against which they can be tested. Methodologies can focus on vocabulary (lexical analysis), sentence construction (syntactic analysis), and the interpretation of meaning (semantic and hermeneutic analysis). Probably the earliest use of computers to support text analysis was the simple retrieval of text strings,[8] thereby facilitating researchers' ability to quickly identify specific elements within one or many documents (Heesterman, 1993). This simple functionality remains useful, as many of those who regularly read online content know.

An early lexical analysis solution was *concordance*, by which all unique words within a document (or corpus) were tallied, along with their incidence rates. Based on inferred relationships between word occurrence and document content or authorship, researchers can map semantic differences and author characteristics, even to the point of being able to identify gender (Corney et al., 2002; Koppel et al., 2002). While in theory these methods might be calculable by hand, computing power offered great improvements in speed and reduced error (Raber, 1997). Key-Word In Context (KWIC), another form of concordance, displays search strings amid the sentences in which they occur. Many CAQDAS applications support KWIC as a means of displaying search results. Concordance methods as a whole are used by literary scholars, linguists, translators, historians, political scholars, and anyone who uses the Google search engine.

### 2.2.2. *Quantitative linguistics*

Quantitative linguistics tools also track text strings, but support more complex statistical inferences by comparing such lexical *themes* (word or phrase occurrences), *semantics* (themes within grammatical contexts), and *networks* (themes within conceptual contexts) within one or more documents (Bolden and Moscarola, 2000; Roberts, 2000). While such techniques have been successfully applied to the study of political dialogue (Laver and Garry, 2000; Laver and Benoit, 2002; Benoit and Laver, 2003; Laver et al., 2003), other social science disciplines such as Communications Studies (Stephen, 1999) and Education (Burnley et al., 2002) have experimented more aggressively with such techniques. Typically, users must identify *reference documents*, to which others are compared; although this enables researchers to identify changes in the tone or content of a corpus over time, it also restricts such analyses to relatively homogenous blocks. Cross-validation with other coding methods suggests that this approach is capable of matching the reliability and validity of manual coding, with the added benefit of providing margin of error estimates (Laver and Garry, 2000). Moreover, because quantitative linguistics methods compare text strings apart from their semantic value, they can be applied to any script recognized by the software – apart from identifying the reference documents (itself a critical task), researchers need not even understand the languages they analyze (Laver et al., 2003).

**Example: Automated Content Analysis**

Michael Laver, Kenneth Benoit and others (Laver and Garry, 2000; Laver et al., 2003) have worked extensively with a computer-supported approach to content analysis that draws quantitative linguistics. The process enables researchers to infer semantic content, based on statistical comparison (e.g., word counts) with so-called "reference documents," for which interpretation of their content (e.g., subject, bias) is established *a priori*. This automated approach to content analysis is based on the assertion that (more-or-less) consistent patterns will exist between the use of particular words and the semantic "dimensions" of text (Benoit and Laver, 2003). Those "key" words are identified and "scored" according to their occurrence in documents that represent a range of positions along a given semantic dimension (Laver and Benoit, 2002).

For instance, somewhat distinct vocabularies may be associated with alternative perspectives on say, civil rights. Where such differences exist, probabilities can be assigned to those words, which then enable automated content analysis of additional documents. Although the ultimate reference points for this form of analysis are ultimately grounded in subjective interpretation, the purely statistical basis of subsequent analyses enables researchers to not only automate content analytical processes, but also to provide confidence intervals for those assessments.

To the degree that word use does reflect semantic content, such approaches may improve efficiency and coding consistency (Roberts, 2000). Yet these gains are limited by the degree to which natural variations in language patterns diverge from the reference documents over time. The technique is perhaps best applied to contexts where language use is relatively normalized (e.g., legal arguments), or within narrower timeframes, where changes are expected to be minimal. Finally, though coding the reference documents still requires content and language expertise, researchers can analyze the additional documents without knowledge of either (Laver et al., 2003).

### 2.2.3. *Coding*[9] *and notation*

Traditionally, social scientists might write notes on document margins, and use tape flags or paper clips to identify passages or concepts central to their analyses. While this enables researchers to extract content from (or more properly, *add* content to) source materials, the mechanical demands of the technique imposes limits on both the complexity and scale of analysis (Fielding and Lee, 1991). Code-and-retrieve software effectively removes such limitations, enabling researchers to: develop inductive categorical structures; automate repetitive tasks (e.g., auto-coding lexicons of relevant terms); test deductive theories against a corpus (by exporting code and keyword occurrence to statistical packages); and even integrate analyses across multiple media types (Alexa and Zuell, 2000). Codes can be structured hierarchically or horizontally; many allow Boolean search tools to be applied to both strings and codes, to test for possible relationships within the corpus (Alexa and Zuell, 1999). To facilitate theory building and consistent coding practice, most packages also enable researchers to attach notations to text segments (Carmichael, 2002).

### 2.2.4. *Hyperlinking*

Another approach to the inductive "discovery" of meaning in a text is to link segments within one or more documents, such that researchers may embed their understandings directly within narratives or other textual sources. While some caution that such free-form structures can be disorienting, and difficult to summarize or otherwise communicate (Fielding and Lee, 1998), others argue that this freedom offers the best means of supporting the principles of grounded theory (Coffey and Atkinson, 1996; Coffey et al., 1996). For subjects that appear to be outside of existing theory, hyperlinks may be a useful means of developing preliminary hypotheses, or to highlight key features and relationships within source narratives.

### 2.2.5. *Network analysis*

Social network theory focuses on variations in the structure and function of group interactions. It is a well-established technique in Sociology (Wasserman and Faust, 1994; Scott, 2000), but has not been widely adopted by political scientists (see Ansell, 2000 for an application to political networks). Once researchers have identified the matrix of relationships (in terms of directionality and magnitude) between individuals and organizations (known as *nodes*), they can apply network theory tools to identify the centrality of "key" actors, equivalent structures within or across organizations, and "structural opportunities" (known as *holes*) (Fisher, 2003). Although many of the underlying theories of social network analysis are quantitatively based, social networks can also be represented graphically (known as *sociograms*). While quite complex visual models have been developed manually (Kimmelman, 2003; Min, 2003), software enables users to develop networks with thousands of nodes, or to map such relationships automatically from electronic sources (e.g., email, address books) (Fisher, 2003). Because histories can be thought of as directed networks (Everton, 2004), time series data can also be graphically represented, enabling researchers to visually and interactively compare trends and events (Hochheiser, 2003).

### 2.3. COLLABORATION

Collaboration can be understood in two ways – spatial and temporal. Just as promoting intercoder reliability between individuals poses significant challenges to researchers, so does ensuring that one's own interpretations are consistent through time. The means of guaranteeing consistency, explicitness and transparency, are also foundations of the scientific process (Crawford et al., 2000). By formally stating assumptions, hypotheses and conclusions, and making evidence available for testing and replication, we facilitate inter-subjective understanding (Brown, 2002).

Many CAQDAS applications support team-based research, generating audit trails for each researcher, and enabling their codes or notations to be viewed individually, or in combination with others (Alexa and Zuell, 2000; Ford et al., 2000). Moreover, because most current applications maintain these codes and notations separate from the source data, it is also possible to integrate multiple theoretical perspectives and methodologies within a single project. Researchers may thus be able to explore analytical approaches which they may never have developed, working alone (Bourdon, 2002). The ability to "drill down" from concepts to source data not only promotes intercoder reliability, but also allows later researchers to assess the overall theoretical perspective of the research (Bong, 2002). It has been suggested that the "proper" way to present qualitative research is to include such focused "snapshots" of individual research protocols, alongside more general conclusions drawn from that analysis (Thompson, 2002). Perhaps it is the difficulty of implementing such practices (without computer support) that has made it hard for qualitatively focused researchers to establish a strong record of transparent and replicable research.

## 3. First Steps

Fortunately for novices, there are several scholarly communities actively concerned with the development and use of computing solutions in social science research. Most are aware of the resources (e.g., user groups, journals) which have coalesced around the more popular statistical packages (e.g., STATA, SPSS, SAS). However, as we have already emphasized, our goal here is to draw attention to some of the less well-known information technologies. Many of these can support both quantitatively and qualitatively oriented research.

The CAQDAS Networking Project (hosted by the Sociology Department of the University of Surrey) acts as a clearinghouse for information about such software and to "encourage debate about methodological and epistemological issues raised by the use of such software."[10] The same department also publishes the quarterly online journal Social Research Update, which strives to keep social scientists informed about methodological innovations, including the innovative use of technology in research (Fielding, 1993). Forum: Qualitative Social Research (FQS) is a peer-reviewed online journal for qualitative research, established in 1999 to "promote discussion and cooperation between qualitative researchers from different nations and social science disciplines." The journal has been developed as an open community, with all articles available free of charge; readers are encouraged to contact authors directly, and to participate in periodic online forums. Several issues have highlighted the use of IT in social science research, revealing how advanced this process is in Europe (Crawford et al., 2000; Gibbs, 2002).

Although scholars in the United States have been slower to adopt CAQDAS technologies (Kerlin, 2000), this pattern does appear to be shifting. The National Center for Supercomputing Applications (NCSA) at the University of Illinois at Champaign-Urbana has worked with the (apparently defunct) Social Science Computing Association (SSCA) to produce the multi-media Wayfarer CD, which was packaged with the print book *Computing in the Social Sciences and Humanities* (Burton, 2002). The Consortium for Qualitative Research Methods (CQRM), which holds workshops every winter at the University of Arizona, has addressed the use of software (NUD*IST) to support research needs (2002, presented by Susan Clarke, University of Colorado). The University of Georgia's College of Education is home to the Qualitative Interest Group (QUIG), which has hosted annual conferences on interdisciplinary qualitative research since 2000.[11] The International Association for Social Science Information Service and Technology[12] (IASSIST) has had an active US Secretariat for nearly a decade (longer in Canada), although this group generally focuses on issues concerning quantitatively oriented research.

### 3.1. THINKING OF EVIDENCE AS DATA

Perhaps the first big hurdle to thinking about ways to use computing technology in social science (at least for qualitative oriented research) is learning to think of text[13] as data – which can be processed, transformed and amended, according to the user's research goals (Laver et al., 2003). While it may also be that many social scientists feel uncomfortable with the technology itself, the ubiquity of word processors, spreadsheets, email and the Internet suggests that actually using these tools is less a barrier than knowing how to use them more effectively. Much of the functionality already described in this essay can be achieved using these common applications.

The first step to thinking about text as data is to learn to identify the "meta-patterns" within sources – whether the order of title, authorship, abstracts, etc. is consistently structured. Virtually every publisher uses style guides (e.g., MLA, Chicago Style) to organize content, databases generally produce regularized output,[14] and many Internet sources often present text in stable meta-formats. Often, key text elements are "tagged" according to their "meta-content" (e.g., TI: *title,* AU: *author*). While qualitatively minded researchers may tend to gravitate to the unique features within texts, locating regularities in the way data are presented can greatly assist researchers in making comparisons between and among texts.

### 3.2. LEARNING TO MANAGE DATA

While it may be tempting for researchers to simply download existing CAQDAS applications and begin using them straight away, the potential for misunderstanding and misuse of these tools is quite high (Brown, 2002). Just as good quantitatively minded scholarship is based on a solid understanding of statistical operations, social scientists who learn how to create simple versions of CAQDAS functionalities are likely to have a much greater appreciation for the possibilities and limitations of computer-supported research.

The simplest way to begin "recreating the wheel" is also likely the easiest and most familiar: text retrieval. As every Internet user is well-aware, search tools can greatly reduce the time needed to locate documents containing words or phrases (*strings*); many also regularly use search tools to jump directly to where such strings occur within a document (e.g., Adobe Acrobat's "binocular" tool). These tools can dramatically reduce the amount of time needed to assess the relevance or usefulness of sources.

A similar approach can be used to modify the formatting of those documents which have been selected as relevant and stored (or generated) locally. *Faux-coding* uses the "find-and-replace" function common to most word processors to modify alpha-numeric strings (e.g., keywords and phrases) within one or more documents. A couple of simple yet effective modifications include replacing a string with the same **string** in bold, or the same STRING in all caps. By experimenting with word roots or lemmas,[15] users can highlight keywords and phrases throughout a document. This technique can greatly improve reading efficiency, as well as reduce errors common to repetitive tasks. To apply any given faux-code (or other repetitive data processing sequences) consistently across many documents, users may wish to record custom *macros*, which save user-defined command sequences (within a single program) as a single menu option or keystroke.

Most researchers also have considerable experience with the *Boolean search*[16] features common to database and Internet search engines. Within those databases which also support *nested* searches (e.g., the "search within results" feature of *Lexis-Nexis*), users can identify the rates at which additional terms occur within the larger pool, and the specific records containing them. Such functionality can help researchers to identify general patterns quickly, at very low-cost.

If it is possible to identify (relatively) consistent meta-formatting in a source, such regularities can be exploited to improve the usefulness of those sources for researchers. The find-and-replace tools of most word processing software also support formatting tags (e.g., *hard returns*, *tabs*), giving users the ability to restructure documents semi-automatically.[17] For instance, randomly compiled event reports can be re-ordered by date, location or source, often by simply replacing *hard returns* with *tabs* and copying those files into spreadsheet software for sorting. In essence, such processing converts document files into tab-delimited databases.

As more content becomes available in digital form, researchers may also find it useful to be trained in the basics of Information Science.[18] Knowing the ways data are organized and managed in various contexts (e.g., legal databases), or how information management tools actually work can greatly improve our efficiency and effectiveness as researchers. A deeper understanding of the limitations of such technologies – as informed by the practice of social science research – may ultimately help us to develop applications that more closely fit our needs.

**Example: Litigation Support Systems**

The practice of law (a.k.a. "the paper chase") is fundamentally based on the analysis of text. The first automated litigation support systems (LSS) of the mid-1970s primarily served as a more efficient way to index large volumes of paper records (DuBowe, 1982). Since then, applications have been developed which can not only organize text so vast as to be unmanageable by traditional, paper-based means (Keeva, 1990), but also to convert documents from paper to electronic formats, via digitization and optical character recognition (OCR) (Emery, 2003).

The scale of many of these projects is staggering – the data management firm CACI International has supported class-action cases involving over a *billion* pages of potential evidence and related documentation (CACI International, 2001). To facilitate the identification of meaningful patterns within these oceans of text, IT developers have also created sophisticated tools which combine Boolean search methods with "fuzzy" algorithms (e.g., allowing for variations in spelling) (Emery, 2003) or concept-based algorithms (e.g., based on research contexts) (Krause, 2003). Because many of the mega-projects also require large, often geographically dispersed staffs, documentation techniques (e.g., annotation, code-and-retrieve) and collaborative technologies (e.g., instant messaging) are also becoming increasingly common (Juhnke, 2003).

Many of these technologies have also been used by other text-centric professions, such as law enforcement and intelligence. While these communities may differ in their means (e.g., rules of evidence) and ultimate aims, they share many of the challenges of empirical social science.

### 3.3. KNOW WHERE YOU'RE GOING BEFORE YOU PICK THE VEHICLE

The range of CAQDAS applications currently available[19] is highly diverse; while some programs perform very narrow tasks, others support much broader functionalities. Fortunately, most commercial software vendors offer downloadable demonstration versions through their websites. Moreover, a striking variety of tools are freely available. As researchers learn how to use the tools they already have more effectively (as described in the preceding section), they can also experiment with more specialized applications. Decisions about which application(s) may be most appropriate for a given project should be informed by users' technical skills, the format (or media) of the source data, the software's flexibility and user-friendliness, and how well those tools support the research methodology (Parker, 1996; Lee and Esterhuizen, 2000). Many experienced CAQDAS users integrate multiple tools in their research, applying different approaches and techniques as necessary (Alexa and Zuell, 2000; Carvajal, 2002). In the end, informed users are more likely to find the most satisfying fit between these tools and their own research needs (Lee and Esterhuizen, 2000).

## 4. Conclusions

### 4.1. POTENTIAL GAINS FROM USING IT

Barring an unforeseen leap in biological evolution, Engelbart's dream may be the only hope social scientists have of remaining responsive – even relevant – in an era of information glut. Those who remember using file cards and typewriters are likely very aware of the capacity of technology to change our habits as researchers. Yet as far as we have come over the past 50 years, we are still only beginning to realize the ways in which computing power can strengthen, expand, and transform intellectual work. Of course, these shifts are not purely technological – as scientific disciplines mature, our expectations about what constitutes quality research changes, as well. Achieving cumulative social knowledge certainly requires us to be more explicit in our assumptions and hypotheses, as well as our rules for interpreting data. What Engelbart understood (and many are just beginning to grasp) is that communications technologies can not only help us to filter oceans of data to produce useful information, but may also relieve a great deal of the burden of research, making it easier to maintain the conditions of transparency and replicability (Bong, 2002; Brown, 2002).

### 4.2. CAVEATS: IT IS NOT A PANACEA

Clearly, automating trivial tasks and maintaining electronic audit trails does not threaten to displace researchers from our central role as the producers of knowledge – "technology is the servant and not the expert" (Parker, 1996). At least for the foreseeable future, humans will remain obliged to assess whether the concepts and evidence for knowledge claims is fundamentally sound (Thompson, 2002). Moreover, given that access to information infrastructures is not evenly distributed, and that our current tools for indexing and retrieving even the available information are woefully inadequate to the task, we must always question whether biases have crept into our analysis (O'Muircheartaigh and Campanelli, 1998; Clarke, 2000; Introna and Nissenbaum, 2000). Of course, this has always been a problem for scholars – and a key reason why the scientific process itself developed.

Because computers cannot think about or experience the world for us, they cannot resolve essential dilemmas of inquiry, nor eliminate the important role of creativity in scientific investigation (Smith and Hesse-Biber, 1996). The task of assigning meaning to evidence remains a fundamental human art (Marshall, 2002). Although information technology may reduce some of the more tedious aspects of research management – thus enabling scholars to devote time and resources to more interesting questions – it will not ultimately make the work less challenging (Smith and Hesse-Biber, 1996).

### 4.3. LEARNING THE CRAFT AND THE TOOLS

It is not uncommon in the literature to find CAQDAS tools discussed as if they represent a "new, distinctive kind of analytic procedure" (Fielding and Lee, 2002). Yet although some applications facilitate methods which are impracticable by manual means, we have yet to see computing solutions which can truly be described as new methodologies in their own right (Fielding, 2002). While social scientists may find it useful to develop a level of technological expertise, they will still need to be well-trained in the traditional tools and methods of their fields (Carvajal, 2002). Just as quantitative analytical tools can be poorly understood and misapplied, uninformed use of CAQDAS solutions can confuse our analysis (or worse). In a collaborative context (to which all scientific efforts should ultimately aspire), the need for transparency may help to limit such abuses, but only as far as the larger community of scholars is methodologically well-trained, and informed about the limitations of these new tools.

As Doug Engelbart was able to foresee more than four decades ago, sophisticated networking and analytical tools now enable researchers and practitioners to tackle problems of increasing scope and scale, with greater efficiency and transparency, than was conceivable merely a decade ago. Whether – and how – we take advantage of such opportunities may not only impact our responsiveness as social scholars, but our relevance, as well.

*Appendix A*: IT functionalities by mode of inquiry

| MODE/ SOURCE | DESCRIPTION | IT FUNCTIONALITY |
|---|---|---|
| Surveying/Polling | Formal/semi-formal inquiry (in-person or remote), ranging from narrow to open responses | Time responses (latency), remote delivery (internet/email), dynamic inquiry (interactive surveys), etc. |
| In-person interviewing | Recorded (audio/video) interviews (or focus-groups) | Record/archive audio, video or text, manipulate as data |
| Remote interviewing | Interviewing via electronic media (telephone, videophone, e-mail, instant messaging) | Record/archive audio, video or text, reduce response bias (blinding/masking), manipulate files as data |
| Remote focus-groups | 3+ participant discussion/interviews (generally more open-ended) via instant messaging programs | Record/archive, time responses (latency), remote participation (internet), manipulate as data |
| Internet media | "New media" (e.g., usenets, weblogs, websites and other internet-based applications) | Automate searching/monitoring or retrieval, archive audio, video or text, manipulate as data etc. |
| Traditional media, secondary literature | Traditional media (e.g., newspaper, television) and analytical literature (e.g., articles, reports) distributed electronically | Automate searching/monitoring or retrieval, archive audio, video or text, manipulate as data etc. |
| Network analysis | Graphical and statistical mapping of connections between actors, events, concepts, etc. | Study cases/attributes as relationships, with or without controls |
| Process tracing | Synchronous (sequential) and asynchronous (parallel) event tracking, based on interviews or supporting literature | Analysis, theory building and testing of greater complexity or subtlety, automating repetitive tasks; integrate sources from multiple media |

*Appendix A*: (Continued)

| MODE/ SOURCE | DESCRIPTION | IT FUNCTIONALITY |
|---|---|---|
| Cognitive mapping | Representing and comparing perceptions, understandings or beliefs, based on interviews, text or experiment responses | Analysis, theory building and testing of greater complexity or subtlety, by automating repetitive tasks; integrate sources from multiple media |
| Time-diary/ blog studies | Tracking individuals' activites, interests, etc. | Establish temporal, spatial context; more open-ended, personal responses; "back-casting" error reduction |
| Experi-ments/simu-lations | Role-playing scenarios, puzzles, etc | Dynamic multimedia simulations, monitor observations/responses (latency), remote delivery/participation (internet), dynamic inquiry (interactive), etc. |
| Socio-economic mapping | Spatial (geographic) or pseudo-spatial (e.g., boundary-level) variable comparison | Analysis, theory building and testing of greater complexity (e.g., layering), speed or subtlety, data visualization, automate repetitive tasks |
| Content/tex-tual analysis | Lexical (vocabulary), syntactic (sentence construction), and semantic (hermeneutic) | Analysis and testing of great complexity or speed, reduced error, integrate sources from multiple media, data visualization, automate repetitive tasks |
| Code and retrieve | Label content by theoretical/experimental categories, Boolean comparison of coded corpus | Analysis and testing of greater complexity or speed, integrate sources from multiple media, data visualization, automate repetitive tasks |
| Boolean analysis | Logic testing via discrete and fuzzy-set theory (e.g., Venn diagrams) | Automating logical reduction processes increases speed, reduces error; increased conceptual complexity (e.g., fuzzy logic analysis) |
| Theory building | Mapping of causal/symbolic relationships via hyperlinks, Forester diagramming | Test/tie theory directly to underlying data, data visualization, increased control of variables/linkage representation |

**Appendix B: Descriptions of individual applications**

What follows is a lengthy (but by no means exhaustive) catalogue of currently available IT solutions with the potential to support social science research. The general price range (free, $$, $$$) of each is identified, except where such applications are commonly available in a modern office environment.

**ACQ1** Acquisition of primary research materials (e.g., interviews, surveys).
**ACQ2** Acquisition of secondary and tertiary materials (e.g., news reports, analytical papers).
**ORG** Organization and management of recorded data and analytical materials (e.g., audit trails).
**ANA** Support of analysis and theory-building.
**COL** Support of local and distributed collaboration.
**TCH** Facilitates pedagogical aims.
**DIST** Distribution of research and analytical materials.

| **AnnoTape** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| $$$ | * | | * | | | | |

Computer audio recording and annotation (www.annotape.com)

| **AnSWR** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| free | | | * | * | * | | |

The Center for Disease Control has developed its own freeware code-and-retrieve software (www.cdc.gov/hiv/software/answr.htm).

| **AQUAD** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| $$$ | | | * | * | * | | |

Code and retrieve, theory building (www.aquad.de)

| **ArcView** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| $$$ | * | | * | * | * | * | |

Entry-level Geographical Information System (GIS) software facilitates analysis of spatially referenced data. This data can be represented graphically, or output to statistical packages for further analysis. A great number and variety of datasets are publicly available.

| **ATLAS/ti** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| $$$ | | | * | * | * | | |

Code and retrieve, theory building (www.atlasti.de)

| **CASES** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| $$$ | * | | * | * | * | | |

"Computer-Assisted Survey Execution System" collect and process-data through structured questionnaires (http://cases.berkeley.edu:7504)

| **CATPAC** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| $$ | | | | * | | | |

"Summarizes text without pre-coding" (www.galileoco.com)

| **CDC EZ-Text** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| Free | | | * | * | * | | |

Text management software to analyze semi-structured qualitative data (www.cdc.gov/hiv/software/ez-text.htm)

| **C-I-SAID** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| $$$ | | | * | * | | | |

Code and retrieve, theory building, code AV files without transcription (www.scolari.co.uk/cisaid/cisaid.htm)

| **Code-A-Text Multi Media** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| $$$ | | * | * | * | | | |

Code and retrieve, theory building applied to text or AV files (www.code-a-text.co.uk)

| **CodeRead** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| free | | | * | * | | | |

Automated text-coding (www.unc.edu/~aperrin/CodeRead)

| **Concordance** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| $$ | | | | * | | | |

Word counts, usage comparison, phrases, proximity search (www.rjcw.freeserve.co.uk)

| **CyberTracker** | | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|---|
| | free | * | | * | | | | |

Flexible forms-based PDA-based data entry, potential GIS referencing (www.cybertracker.co.za)

| **Diction** | | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|---|
| | $$$ | | | | * | | | |

Analyzes writing "tones" (e.g., certainty, optimism, realism) (www.scolari.com)

| **Digital audio** | | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|---|
| | $$$ | * | | | | | * | * |

.mp3, .wav, .wma format recordings for archiving or analysis

| **Email** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| | * | | | | * | | * |

| **EndNote** | | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|---|
| | $$$ | | * | * | ~ | ~ | | * |

Records (bibliographic) management (www.endnote.com)

| **Ethno2** | | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|---|
| | free | | | * | * | * | | |

Supports Event Structure Analysis (www.indiana.edu/~socpsy/ESA/ethno2.html)

| **Ethnograph** | | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|---|
| | $$$ | | | * | * | * | | |

Code and retrieve, theory building (www.scolari.com)

| **Form2Data** | | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|---|
| | $$$ | * | | * | * | | | |

Web survey software (www.verbi.de/form2data)

| **FS/QCA .93** | | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|---|
| | free | | | ~ | * | | | |

Fuzzy set / Boolean logic software (www.u.arizona.edu/%7Ecragin)

| **General Inquirer** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| free | | | | * | | | |

Internet-based automated text analysis (e.g., concordancing, natural language processing) (www.webuse.umd.edu:9090)

| **Google hacks** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| free | * | * | | * | | | |

Search syntax for specialized web searching / analysis

| **HyperRESEARCH** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| $$$ | | | * | * | | | |

Code and retrieve, theory building (www.researchware.com)

| **IHMC Concept Map** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| free | | | | ~ | ~ | * | |

Theory mapping (http://cmap.coginst.uwf.edu)

| **Internet / html** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| free | * | * | * | | * | * | * |

| **Intext** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| free | | | | * | | | |

DOS version of TextQuest (www.intext.de/eindex.html)

| **Keyplayer** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| free | | | | * | | | |

Identifies critical nodes in networks (www.analytictech.com/keyplayer.htm)

| **Kwalitan** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| $$$ | | | | * | | | |

Code and retrieve, theory building (www.kun.nl/methoden/kwalitan)

| **MaxQDA** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| $$$ | | | * | | | | |

Code and retrieve, theory building (www.maxqda.com)

| **Microconcord** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| free | | | | * | | | |

Word counts and usage comparison (the predecessor of WordSmith) (www.lexically.net/downloads/_freebies/mconcord1.zip)

| **Minnesota Contextual Content Analysis** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| free | | | | * | | | |

Word counts and usage comparison (www.clres.com/cata)

| **MonoConc Pro** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| $$ | | | | * | | | |

Concordancing, text analysis (www.athel.com)

| **Nvivo / N6 (NUD*IST)** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| $$$ | | | * | * | ~ | | |

Code and retrieve, theory building (www.qsr.com.au)

| **Office software suites** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| | ~ | ~ | * | ~ | ~ | * | * |

Macros automate pseudo-coding, VBL enables broader functionality, Perl could be used to integrate applications

| **Pajek** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| free | | | | * | | | |

Social network analysis and visualization (http://vlado.fmf.uni-lj.si/pub/networks/pajek/)

| **PC Matchmaker** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| free | | * | * | | * | * | * |

Manages one-to-one links between records from independently generated sources (www.vanderbilt.edu/Econ/faculty/Atack/atackj.htm)

| **PCAD** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| $$$ | | | | * | | | |

Gottschalk-Gleser (emotion) text scoring (www.gb-software.com)

| **Perl** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| free | * | * | * | | * | | * |

Language to integrate and run almost any program

| **Personal Librarian** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| free | | | ~ | * | ~ | | |

A freeware full-text search engine that can support natural language and Boolean options (www.pls.com/pl.htm)

| **PolyAnalyst** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| $$$ | * | * | | * | | | |

Natural language analysis software (www.megaputer.com)

| **Profiler+** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| $$$ | | | | * | * | | |

Code and retrieve, theory building (www.socialscience.net)

| **Pro Tools** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| free | ~ | ~ | | | | | |

Edit or process audio files (www.digidesign.com)

| **QDA Miner** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| $$$ | | | * | * | * | | |

Concordance, code and retrieve (www.qdaminer.com)

| **Qualifiers** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| $$ | * | | * | * | ~ | | |

Multimedia data organizer, audio recorder, code and retrieve - MAC only, Windows in development (http://hisii.hawaii.edu/qualifiers/qualifiers.html)

| **Qualitative Media Analysis** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| $$$ | | | * | * | * | | |

Apply codes directly to AV files without transcription (www.cvs.dk/qma.htm)

| **Qualrus** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| $$$ | | | * | * | ~ | | |

Code and retrieve, theory building (www.ideaworks.com)

| **Readware** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| $$$ | * | * | | * | | | |

Automated content analysis (www.readware.com)

| **RSS** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| free | * | * | | | | | * |

Format for syndicated news and web content updates

| **SALT** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| $$ | ~ | | * | * | | | |

Transcription and language analysis (www.waisman.wisc.edu/salt/index.htm)

| **SDA** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| $$$ | * | | * | * | * | | |

Survey Documentation and Analysis (http://sda.berkeley.edu:7502/index.htm)

| **Shoebox** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| NA | * | | * | * | * | * | * |

Code, retrieve and display visual images (currently in development at: www.uk.research.att.com/dart/shoebox)

| **Sonar** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| $$$ | | | * | * | * | | |

Rapid text search across multiple formats (www.virginiasystems.com)

| **SphinxSurvey Lexica** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| $$$ | * | | * | * | * | | |

Survey design, statistical processing, lexical analysis (www.scolari.com)

| **SQL** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| $$$ | | | * | * | | | |

Structured Query Language – an ANSI standard computer language to access and manipulate databases

| **Stella** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| $$$ | | | | * | * | * | |

Dynamic modeling / Forester diagrams (www.hps-inc.com/Science/science.htm)

| **Storyspace** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| $$$ | | | * | * | * | * | * |

Hypertext authoring tool (www.eastgate.com)

| TABARI | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| free | | | | * | | | |

The Windows version of the "Kansas Event Data System" for automated coding of event data from wire service reports or chronologies (www.ukans.edu/~keds/tabari.html)

| TACT | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| $$ | | | * | * | | | |

Text coding, retrieval and analysis (www.chass.utoronto.ca/cch/tact.html)

| TEXTPACK | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| $$$ | | | | * | | | |

Word counts and usage comparison (www.gesis.org/zuma)

| TextQuest | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| $$$ | | | | * | | | |

Word counts and usage comparison (www.textquest.de/tqe.htm)

| T-LAB | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| $$$ | | | | * | | | |

Automated content analysis - multilingual (www.tlab.it)

| TextAnalyst | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| $$$ | | | | * | | | |

Automated analysis and summary of natural language text (www.megaputer.com)

| TimeSearcher | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| free | | | * | * | * | | |

Visual exploration of time-series data (www.cs.umd.edu/hcil/timesearcher)

| Tosmana | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| free | | | * | * | | | |

Boolean logic software (www.staff.uni-marburg.de/~cronqvis/tosmana/index2.html)

| **Transana** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| free | ∼ | | * | | | | |

Transcription software (www.transana.org)

| **Transcriber** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| Free | ∼ | | * | | | | |

Transcription software (www.ldc.upenn.edu/mirror/Transcriber/)

| **UCINET** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| xx | | | * | * | | | |

Social network analysis and visualization (www.analytictech.com/ucinet.htm)

| **VBPro** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| free | | | * | * | | | |

Automates word counts and supports text coding and analysis (http://mmmiller.com/vbpro/vbpro.html)

| **Vensim PLE** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| free | | | | * | * | * | |

Dynamic modeling / Forester diagrams (www.vensim.com/venple.html)

| **WordNet** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| free | | | | * | | | |

Lexical analysis (www.cogsci.princeton.edu/∼wn)

| **WordStat** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| $$$ | | | | * | | | |

Supports a range of quantitative linguistics techniques (e.g., tracking phrase-incidence, lexical comparison). Automated content analysis (www.simstat.com/wordstat.htm)

| **XML** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| free | ∼ | ∼ | ∼ | ∼ | ∼ | | ∼ |

Meta-language, organizes by data-type, not format

| **ZySCAN** | ACQ1 | ACQ2 | ORG | ANA | COL | TCH | DIST |
|---|---|---|---|---|---|---|---|
| $$$$ | * | * | * | | | | |

Highspeed automated digitization (www.zylab.com)

## Notes

1. Such a conundrum is named for economist Thomas Malthus, who first posited the argument in 1798. Malthus reasoned that giving more money to the poor would likely not diminish the problem of hunger. Instead, because of the overall scarcity of food, such a policy would probably merely drive up the price of food resources, thus having virtually no effect on poverty or hunger (see e.g., Malthus, 1926).
2. While it is common to use the terms *data*, *information*, and *knowledge* as synonyms, it is more useful to identify the latter as a function of the contextualization of data or information within a conceptual framework (e.g., categorization schemes, causal stories) (READWARE 2002). Data without understanding is mere noise – its conversion to knowledge is based (in part) on its expected meaning.
3. See Alexa and Zuell , 2000 for an earlier example of this approach, applied to a review of text analysis software.
4. Hereafter referred to simply as "the Web."
5. This technology is changing *very* rapidly – researchers will need to inform themselves about available hardware options before purchase. However, one constant is the need for an external microphone – the internal microphones of most recorders produce low-quality recordings.
6. The abbreviation CAQDAS is somewhat poorly defined; it is commonly used to refer to applications which support both qualitatively and quantitatively-oriented research. In this essay, we also use it in this more general sense.
7. While most of these solutions reduce the tedium of managing research practices, few can be understood as methodological innovations themselves.
8. *Strings* are sequences of alphanumeric characters varying from single letters or numbers to entire documents. Most often, text retrieval is used to search for phrases, words, word fragments (e.g., *stems*) or conjugations (i.e., *lemmas*).
9. Following convention within the CAQDAS literature, we use the term *coding* to refer to the practice of associating qualitative categories (or attributes) to specific text elements (e.g., words, phrases, passages). These coded associations can be either quantitative or qualitative in nature, or both.
10. See http://caqdas.soc.surrey.ac.uk.
11. See www.coe.uga.edu/quig.
12. See www.iassistdata.org.
13. While we use the term "text" to refer to printed language (including transcriptions), many the principles described here may also be applied to other media forms (e.g., audio, still or moving images).
14. Users should be cautioned that some commercial databases (e.g., Lexis-Nexis) regularly alter the meta-formatting of their output, as a means of safeguarding proprietary rights.
15. The full variety of conjugations a given word might take (e.g., do, does, did).
16. These enable users to define whether specific terms can (or must) be included (or excluded) in retrieved records, based on set theory operators (e.g., AND, OR, NOT).
17. As already suggested, users may also wish to record *macros* to fully automate those formatting operations they use on a regular basis.
18. Library Science programs are increasingly offering digital information management courses, as well.
19. An extensive (but by no means comprehensive) sample of which is presented in Appendix B of this essay.

## References


Alexa, M. & Zuell, C. (1999). *A review of Software for text Analysis*. Mannheim, ZUMA.

Alexa, M. & Zuell, C. (2000). "Text Analysis Software: Commonalities, Difference and Limitations: The Results of a Review". *Quality and Quantity* 34(3): 299–321.

Ansell, C. (2000). *Community Embeddedness and Collaborative Governance*. Social Movement Analysis: The Network Perspective, Ross Priory, Loch Lemond, Scotland.

Bampton, R. & Cowton, C. J. (2002). The E-Interview. *Forum: Qualitative Social Research*. 3:2 May.

BBC, (2003). Dyke to open up BBC archive, British Broadcasting Corporation.

Benoit, K. & Laver, M. (2003). "Estimating Irish Party Positions Using Computer Wordscoring: The 2002 Elections." *Irish Political Studies* 17(2).

Bolden, R. & Moscarola, J. (2000). "Bridging the Quantitative-Qualitative Divide: The Lexical Approach to Textual Data Analysis." *Social Science Computer Review* 18(4): 450–460.

Bond, D. & Bond et al. J. (2003). "Integrated Data for Events Analysis (IDEA): An Event Typology for Automated Events Data Development." *Journal of Peace Research* 40(6): 733–745.

Bond, D. & Jenkins, J. C. (1997). "Mapping Mass Political Conflict and Civil Society: Issues and Prospects for the Automated Development of Event Data." *Journal of Conflict Resolution* 41(4): 533–579.

Bong, S. A. (2002). Debunking Myths in Qualitative Data Analysis. *Forum: Qualitative Social Research.* 3:2 May.

Bourdon, S. (2002). The Integration of Qualitative Data Analysis Software in Research Strategies: Resistances and Possibilities. *Forum: Qualitative Social Research*. 3:2 May.

Brace-Govan, J. (2002). "Qualitative Data for a Quantitative Audience: An Evaluative Framework for Analysis." *Qualitative Research Journal* 2(1): 47.

Brown, D. (2002). Going Digital and Staying Qualitative: Some Alternative Strategies for Digitizing the Qualitative Research Process. *Forum: Qualitative Social Research.* 3:2 May.

Burnley, P. C. & Evans et al., W. (2002). "A Comparison of Approaches and Instruments for Evaluating a Geological Sciences Research Experiences Program." *Journal of Geoscience Education* 50(1): 15–24.

Burton, O. V., Ed. (2002). *Computing in the social sciences and humanities.* University of Illinois Press, Urbana, IL.

CACI International (2001). CACI Awarded "Mega 2" Contract With Department of Justice. Arlington, VA, CACI International, Inc: 2.

Carmichael, P. (2002). Extensible Markup Language and Qualitative Data Analysis. *Forum: Qualitative Social Research.* 3:2 May.

Carvajal, D. (2002). "The Artisan's Tools. Critical Issues When Teaching and Learning CAQDAS." *Forum: Qualitative Social Research* 3(2).

Clarke, P. (2000). *The Internet as a medium for qualitative research*. 2nd Annual Conference on World-Wide Web Applications, Centre for Research in Web-based Applications, Rand Afrikaans University, South Africa, Department of Information Studies.

Coffey, A. & Atkinson, P. (1996). *Making sense of qualitative data: complementary research strategies*. Sage Publications, Thousand Oaks.

Coffey, A., & Holbrook et al., B. (1996). Qualitative Data Analysis: Technologies and Representations, Sociological Research Online. 1.

Corney, M. & de Vel et al., O. (2002). *Gender-Preferential Text Mining of E-mail Discourse*. 18th Annual Computer Security Applications Conference, Applied Computer Security Associates, Las Vegas, NV.

Corti, L. (2000). Progress and Problems of Preserving and Providing Access to Qualitative Data for Social Research: The International Picture of an Emerging Culture. *Forum: Qualitative Social Research*. 1:3 December.

Corti, L. & Day et al., A. (2000). Confidentiality and Informed Consent: Issues for Consideration in the Preservation of and Provision of Access to Qualitative Data Archives. *Forum: Qualitative Social Research*. 1:3 December.

Crawford, H. K. & Leybourne et al., M. L. (2000). How we Ensured Rigour in a Multi-site, Multi-discipline, Multi-researcher Study. *Forum: Qualitative Social Research*. 1:1 January.

Creswell, J. W. (1994). *Research Design: Qualitative and Quantitative Approaches*. Sage Publishers, London.

DuBowe, J. S. (1982). "Automated Litigation Support: A Litigator's Primer." *ABA Journal* 68(9): 1118–1123.

Eagles, M. & Katz et al., R. S. (2000). "Controversies in political redistricting: GIS, geography, and society: editorial." Political Geography 19(2): 135–139.

Emery, P. (2003). Capturing Terrorists and Criminals: Speeding Up the Intelligence Gathering Process. McLean, VA, ZyLAB Technologies: 16.

Emery, P. (2003). Document and Records Management: Understanding The Differences and Embracing Integration. McLean, VA, ZyLAB Technologies: 4.

Engelbart, D. C. (1962). *Augmenting Human Intellect: A Conceptual Framework.* Stanford Research Institute, Menlo Park, CA.

Everton, S. F. (2004). A Guide for the Visually Perplexed: Visually Representing Social Networks. Stanford, CA.

Fielding, N. (1993). "Analysing qualitative data by computer." *Social Research Update*(1): 7.

Fielding, N. (2002). Automating the Ineffable: Qualitative software and the meaning of qualitative research. *Qualitative Research in Action*. T. May. Sage. London, pp.161–178.

Fielding, N. & Lee, R. (2002). "New Patterns in the Adoption and Use of Qualitative Software." Field Methods 14(2): 197–216.

Fielding, N. & Lee, R. M. (1991). Using Computers in Qualitative Research. Sage, London.

Fielding, N. & Lee, R. M. (1998). Computer Analysis and Qualitative Research. Sage, London.

Finel, B. I. & Lord Eds., K. M. (2002). *Power and Conflict in the Age of Transparency*. Palgrave Macmillan, New York.

Fisher, D. (2003). Social Networks for End Users. Irvine, CA: 20.

Ford, K. & Oberski et al., I. (2000). "Computer-Aided Qualitative Analysis of Interview Data: Some Recommendations for Collaborative Working". *The Qualitative Report.* 4:3–4.

Fricker Jr, R. D. & Schonlau, M. (2002). "Advantages and Disadvantages of Internet Research Surveys: Evidence from the Literature." *Field Methods* 14(4): 347.

Gibbs, G. (2002). *Qualitative data analysis: explorations with NVivo*. Open University, Philadelphia.

Gibbs, G. R. & Friese et al., S. (2002). The Use of New Technology in Qualitative Research. Introduction to Issue 3(2) of FQS. *Forum: Qualitative Social Research*. 3:2 May.

Heesterman, W. (1993). "Text Retrieval Software: AskSam, Idealist, Isys, Zyindex." *Law Technology Journal* 3(1): 7.

Hill, K. & Hughes, J. (1997). "Computer-Mediated Political Communication: The USENET and Political Communities." *Political Communication* 14: 3–27.

Hochheiser, H. (2003). Interactive Graphical Querying of Time Series and Linear Sequence Data Sets. *Department of Computer Science*. University of Maryland, College Park, MD, p. 313.

Introna, L. & Nissenbaum, H. (2000). "Shaping the Web: Why the Politics of Search Engines Matter." *The Information Society* 16(3): 169–186.

Jones, S. (1999). *Doing Internet research: critical issues and methods for examining the Net*. Sage Publications, Thousand Oaks, CA.

Juhnke, D. J. (2003). "Electronic Discovery in 2010." *Information Management Journal* 37(6): 34–40.

Keeva, S. (1990). "Why Litigation Support?: The Sky Is the Limit." *ABA Journal* 76(5): 76.

Kerlin, B. (2000). Qualitative research in the United States. *Forum: Qualitative Social Research*. 1:1 January.

Kimmelman, M. (2003). Webs Connecting the Power Brokers, the Money and, Ultimately, the World. *New York Times*, New York, p. 2.

King, G. & Lowe, W. (2003). "An Automated Information Extraction Tool for International Conflict Data with Performance as Good as Human Coders: A Rare Events Evaluation Design." *International Organization* 57: 617–642.

King, G., Keohane, R. O. & Verba, S. (1994). *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton University Press, Princeton, NJ.

Kinnevy, S. C. & Enosh, G. (2002). "Problems and promises in the study of virtual communities." *Journal of Technology in Human Services* 19(2–3): 119–134.

Koppel, M., & Argamon et al., S. (2002). "Automatically Categorizing Written Texts by Author Gender." *Literary and Linguistic Computing* 17(4): 401–412.

Krause, J. (2003). "What a Concept!" *ABA Journal* 89(9): 60.

La Porte, T. M. & Demchak et al., C. C. (2001). "Webbing Governance: Global Trends across National- Level Public Agencies." *Communications of the ACM* 44(1): 63–67.

Laver, M. & Benoit, K. (2002). "Locating TDs in Policy Spaces: Wordscoring Dáil Speeches." *Irish Political Studies* 17(1).

Laver, M. & Benoit et al., K. (2003). "Extracting Policy Positions from Political Texts Using Words as Data." *American Political Science Review* 97(2): 311–331.

Laver, M. & Garry, J. (2000). "Estimating policy positions from political texts." *American Journal of Political Science* 44(3): 619–634.

Lee, R. M. & Esterhuizen, L. (2000). "Computer Software and Qualitative Analysis: Trends, Issues and Resources." *International Journal of Social Research Methodology* 3(3): 231–243.

Leydesdorff, L. & Wouters, P. (1996). "Quantitative Measuring or Qualitative Understanding: Is it possible to bridge the divide in STS?" *EASST Review* 15(3): 5.

List, D. (2001). *Know Your Audience: A Practical Guide to Media Research*. Original Books, NZ.

LOC (2002). Preserving Our Digital Heritage: Plan for the National Digital Information Infrastructure and Preservation Program. US Library of Congress, Washington, DC, p. 74.

Mann, C. & Stewart, F. (2000). *Internet communication and qualitative research: a handbook for researching online*. Sage Publications, Thousand Oaks, CA.

Marshall, H. (2002). "What Do We Do When We Code Data?" *Qualitative Research Journal* 2(1): 56–70.

McLaughlin, W. S. (2003). "The use of the Internet for Political Action by Non-state Dissident actors in the Middle East." *First Monday* 8(11): 44.

Merriam, S. B., Ed. (2002). *Qualitative Research in Practice: Examples for Discussion and Analysis*. Jossey-Bass, San Francisco.

Min, S. (2003). *Mark Lombardi: Global Networks*. The Drawing Center, New York, 2.

Morse, J. M. & Barrett et al., M. (2002). "Verification Strategies for Establishing Reliability and Validity in Qualitative Research." *International Journal of Qualitative Methods* 1(2).

Nagler, J. (1995). "Coding Style and Good Computing Practices." *PS: Political Science and Politics* 6(28): 488–493.

NRC (2003). *Preparing for the Revolution: Information Technology and the Future of the Research University*. National Acadamies Press, Washington, DC.

O'Muircheartaigh, C. & Campanelli, P. (1998). "The Relative Impact of Interviewer Effects and Sample Design Effects on Survey Precision." *Statistics in Society* 161(1): 63–77.

Parker, D. R. (1996). *Emerging Uses of Computer Technology in Qualitative Research*. Annual Conference of the Association of Louisiana Evaluators, Association of Louisiana Evaluators, New Orleans, LA.

PoliticsOnline (2004). 2004 Special Report on the Year of Internet and Politics, PoliticsOnline, Inc.

Raber, L. (1997). "Toxics Data: How Much is too Much?" *Chemical and Engineering News* 75(22): 26.

Roberts, C. W. (1997). *Text Analysis for the Social Sciences: Methods for Drawing Statistical Inferences from Texts and Transcripts*. Lawrence Erlbaum, Mahwah, NJ.

Roberts, C. W. (2000). "A Conceptual Framework for Quantitative Text Analysis." Quality and Quantity 34(3): 259–274.

Schonlau, M., Fricker, J. & Ronald D et al. (2001). *Conducting Research Surveys via E-mail and the Web*. RAND, Santa Monica, CA.

Scott, J. (2000). *Social network analysis: A handbook*. Sage, London.

Smith, B. A. & Hesse-Biber, S. (1996). "Users' Experiences with Qualitative Data Analysis Software: Neither Frankenstein's Monster Nor Muse." *Social Science Computer Review* 14(4): 423–432.

Solomon, D. J. (2001). "Conducting Web-Based Surveys." *ERIC Digest*.

Starkey, B. A. & Blake, E. L. (2001). "Simulation in International Relations Education." *Simulation and Gaming* 32(4): 537–551.

Stephen, T. (1999). "Computer-Assisted Concept Analysis of HCR's First 25 Years." *Human Communication Research* 25(4): 498–513.

Stockdale, A. (2002). Tools for Digital Audio Recording in Qualitative Research, Social Research Update. Autumn 2002.

Sui, D. Z. & Hugill, P. J. (2002). "A GIS-Based Spatial Analysis on Neighborhood Effects and Voter turn-out:-A Case Study in College Station, Texas." *Political Geography* 21(2): 159–173.

Thelwall, M. (2001). "The Responsiveness of Search Engine Indexes." *Cybermetrics* 5(1): 8.

Thomas, G. D. (2000). *The Machine-Assisted Creation of Historical Event Data Sets: A Practical Guide*. International Studies Association Annual Meeting, Los Angeles, CA.

Thompson, R. (2002). Reporting the Results of Computer-assisted Analysis of Qualitative Research Data. Forum: Qualitative Social Research. 2003.

Wasserman, S. & Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge University Press, New York.

Zhang, Y. (1999). "Scholarly use of Internet-Based Electronic Resources: A Survey Report." *Library Trends* 47(4): 746–770.