

1 INTRODUCTION

In our roles as consultants, we often receive calls or email to discuss the development of an online research project. On one occasion, an individual, let's call her Jill, told me (TJG) that she was going to apply for a particular type of grant, and that she was interested in developing a web-based screening and intervention tool for the core component of the grant application. She expressed an interest in attending any courses I could recommend and was looking for some general advice regarding how to proceed with the development of her study.

First, how did she know to contact one of us? Why contact one of us? Jill contacted me through the usual word-of-mouth method. She had a friend, who knew a friend, who knew me. The friend of a friend knew that I had experience with online research, and so suggested that she get in touch with me for advice. In other words, Jill found me through professional networking. The networking, in itself, should say something to you. Information about developing online research isn't readily available, so researchers tend to track down colleagues who have some experience developing online studies, to seek their advice.

Jill contacted me for advice, and I suspect my responses were less than satisfying. In fact, they were probably a little discouraging. I noted that any courses I was aware of would typically cover technical tools such as PHP, which means they are designed and offered for computer programmers. I indicated that it might be useful to attend such courses, but that they might be difficult, and much of the material covered might not be directly applicable to her study. In addition, I indicated that there might be a course in the area of electronic research methods, but those courses were few and far between. Furthermore, while methodologically they will be interesting and useful, they are likely to lack the practical application of various technological tools to research methods. These latter courses are also unlikely to include hands-on assistance in the actual construction of an online study.

For some, my response would be a show-stopper. The lack of assistance available in developing online research often leads people back to more traditional forms of research. In the case of this particular inquiry, Jill was persistent, suggesting that she might seek out a consultant to help her with her project.

Jill's next question to me was whether or not I had any advice for her about hiring a consultant. She was aware that I had experience managing technology consulting teams. In fact, both of us have worked with a number of individuals

2 A GUIDE TO CONDUCTING ONLINE RESEARCH

who have hired technical consultants to help them with their research projects, most often being asked to get involved after a consultant left the project or something had gone seriously wrong.

My suggestion to Jill was to find someone with technical skill who also had some research experience or suitable educational background, so that he or she would understand what she was trying to accomplish with her research. The ideal candidate for her, I surmised would be, someone with technical skill, research experience and academic credentials. This typically would mean a graduate student or intern who has programming or other technical experience. I cautioned, however, that this type of person is usually difficult to find, may already be employed, and/or is likely to be transient. This last point of being transient, meaning only available for a limited period of time, I noted, should not be considered lightly. A project can run into considerable snags when the main technical architect is no longer available. Replacement staff may not understand what was initially developed, source files may not be well organized or, worse, files may be missing. Ultimately, everything ends up on hold until the project can be put back on track.

It is very common for researchers to find themselves in a quandary about how to proceed with an online research project. They may be clear that their research population is best reached online. They may have colleagues encouraging them and making it all look so easy. They may have pressure to use technical tools to keep down project costs, for example by limiting the need for data entry. So they now face the question of what to do next. Who can help? What resources are available to them? Like this friend of a friend, they want to develop an online study, but need assistance thinking through the various challenges related to their particular needs, before they start developing the study.

About this book

It is out of the preceding type of experience that this book was developed. Our focus and purpose is to provide you, the reader, with a research consultant. Using our consulting experience and knowledge of best practices, we have developed a guide to help answer your questions. This book isn't meant to be read cover to cover like a novel or a textbook. It is more like a manual, though not your typical manual. Hopefully, you'll find a section that meets your particular needs. In some cases, you will discover practical step-by-step advice that is useful when embarking on a research study. When coverage is limited, either due to space or the depth of the issue, we'll also offer advice and direct you to resources that we hope will address your questions.

In short, the focus of this book is to support the use of technologies for research purposes. Our goal is to offer a collection of resource material to get the researcher started. We hope you'll think of it as an inexpensive consultant. As

a resource, this book is meant to be the place to which a researcher turns when beginning to think through the practical aspects of developing an online research project, and where to seek information regarding the use of a particular technology for an online study. We believe that we accomplish our mission by providing you with technical explanations, instructions, self-help tips, useful links to other resources, references and case examples.

This book is not a methods text. There are a number of substantive online research methods texts available (e.g. Best & Krueger, 2004; Blank, Fielding & Lee, 2008; Jones, 1999). In addition, there is also a diverse body of research literature on a number of technology-related topics. Some books, for example, cover using computers for qualitative analysis (Miles & Huberman, 1994; Richards & Richards, 1991; Walker, 1993; Weitzman & Miles, 1995). Several authors have focused on the possibilities and subtleties of participant observation and online field work (Clodius, 1994; Hine, 2000; Ito, 1996; Markham, 1998; Nardi, 1996). In recent years this body of literature has been complemented by several books on online methods and the research of online social phenomena (e.g. Howard & Jones, 2004; Mann & Stewart, 2000). Adding to the growing body of literature are works that assist researchers in understanding the nuances and subtle differences in researching online (Hewson et al., 2003; Hine, 2000; Seale, 2004).

In addition, this book is not a software manual. People often think in terms of the “how-to for Dummies” books such as *Windows XP for Dummies* (Harvey, 2004), or a technical manual such as *Adobe Photoshop CS Classroom in a Book*, a training workbook produced by a software developer or a licensee (2003). While there are several examples of how to do particular tasks and illustrations to highlight particular actions, none of the illustrations is meant to provide comprehensive training on a given subject or application. In all cases, they serve to make specific points about a technology and to provide some basics for a new user, so that a user can make educated decisions as he or she embarks on a new online research study.

We are aware that some readers may suggest that many software applications are intuitive today and relatively easy to use. They may also note that online services such as SurveyMonkey, a web-based survey service provider, make developing an online survey a breeze. And on those points they would be correct. While we may not be able to convince everyone to stick it out, we’ll offer one thought. Just because there are statistical software packages that make it easy to run the numbers doesn’t mean we’re now all statisticians. While software gets easier to use, it’s still necessary to have certain skills to be able to make informed choices. We believe this book can help the reader make many of those decisions.

At the risk of offending, it has been our experience that many users don’t know what they don’t know ... that is, until they find themselves in a bind with their project. They say things like, “Why didn’t anyone warn me?” or “If I’d only known when I started this project.” If you’re unsure of the utility of this book, we encourage you to flip through a few chapters before putting it down. The

4 A GUIDE TO CONDUCTING ONLINE RESEARCH

most it will cost you is a few minutes of your time. On the other hand, if you learn something new, you may discover that we can save you considerable time.

Introduction to the content

Computer technology has greatly enhanced the ability to communicate, or interact, with others globally. Computer technology has also exposed people to new public and private spaces which constitute cyberspace, where humans and computers coexist (Fahey, 1994). When people discuss these spaces, they might be talking about any number of computer-mediate spaces. Electronic mail, or email, is one of the most basic of these “spaces.” Email “is a store and forward method of composing, sending, storing, and receiving messages over electronic communication systems” (<http://en.wikipedia.org/wiki/Email>). A listserv discussion list enables people to send messages to a server that then distributes the message to all of the members. Then there are newsgroups (Usenet) where people participate in a kind of online conference by reading messages posted in a particular location and then posting their own contributions to a discussion. More recently, this type of interaction is seen in blogging.

A blog is often a mixture of what is happening in a person’s life and what is happening on the Web, a kind of hybrid diary/guide site, although there are as many unique types of blogs as there are people. People maintained blogs long before the term was coined, but the trend gained momentum with the introduction of automated published systems, most notably Blogger at blogger.com. Thousands of people use services such as Blogger to simplify and accelerate the publishing process. *Blogs* are alternatively called *web logs* or *weblogs*. However, “blog” seems less likely to cause confusion, as “web log” can also mean a server’s log files. (<http://www.marketingterms.com/dictionary/blog/>)

Another space often inhabited in cyberspace is a Multi-user dungeon (MUD). MUDs have attracted the interest of academic scholars from many fields, including communications, sociology and law. They also have synthetic economies in different environments or what they would call different worlds. As stated in Wikipedia, “MUDs often have a fantasy setting, while many others are set in a science-fiction-based universe or themed on popular books, movies, animations, history, etc. Still others, especially those which are often referred to as MOOs, are used in distance education or to allow for virtual conferences” (<http://en.wikipedia.org/wiki/MUD>). The environments have become increasingly sophisticated and now can include virtual reality using 3-D animations in places such as Second Life (<http://secondlife.com>).

Many today are familiar with Internet Relay Chat (IRC) or instant messaging (IM). Both allow for synchronous interaction between individuals. IRC functions

as an online conference call whereby many people can participate in a discussion simultaneously. While IM functions in a similar way, typically IM is thought of as a one-on-one type of interacting, while IRC is most often thought of as a multi-user environment. Instant communication is nowadays also available in the comment area of blogs, and at social sites such as Myspace, Facebook and Twitter.

As these various computer protocols enable individuals to interact in new ways, they open new spaces and forms of interaction that warrant research. Likewise, they make it possible to conduct research in new ways. An interview or focus group can be conducted asynchronously through email or synchronously in a chat room using an instant messaging or internet chat relay application. Any place where text is available on the internet represents the possibility for any number of qualitative studies such as narrative analysis and content analysis. Any place where people interact online represents a potential place where interactants can be observed and discussions can be analyzed. The locations for posting a survey or accessing potential study participants are virtually unlimited. In addition, accessible data abounds online from organizations such as US Center for Disease Control (<http://www.cdc.gov/datastatistics/>), the International Monetary Fund (<http://www.imf.org/external/data.htm>) and the World Bank (<http://www.worldbank.org/>).

Blurring the boundaries of traditional research

Before getting into the meat of our topic, it is worth making a few additional introductory remarks. Online researching has opened new environments to researchers that move beyond traditional research and challenge some of our notions of what it means to research, how people engage online, and so forth. A body of literature relative to assessing the value and experience in online learning environments has expanded to include games and simulations (Gibson, Aldrich & Prensky, 2007; Prensky, 2006). These environments not only offer new ways for learning, but also new ways in which to conduct research, creating simulations and testing conclusions (e.g. Gibson, Aldrich and Prensky, 2007; Seo and Barrett, 2007).

One area often overlooked when reviewing the literature, but available nonetheless, is a growing body of work on ethical and legal considerations related to online research. For example, there are an increasing number of books on confidentiality (applied to research participants in this context) (e.g. Smedinghoff, 1996b). In the United States, any kind of health or medical related research requires adherence to federal regulations such as the Health Insurance Portability and Accountability Act (HIPAA) (<http://www.hhs.gov/ocr/hipaa/>), as well as adherence to copyright laws (e.g. Imparl, 2006; <http://memory.loc.gov/learn/start/cite/index.html>). It is safe to

6 A GUIDE TO CONDUCTING ONLINE RESEARCH

assume that many countries will have some kind of privacy requirement in the coming years, if they do not already have one. For example, in the UK, certain privacy matters are regulated by the governing body of the General Medical Council and the Health and Social Care Act. So, before beginning a study that may bring into question the management of personal data, it is a good practice to become familiar with current privacy regulations and legislation. A suggested starting point is discipline-specific literature.

Considering the literature

What is lacking in the growing body of literature on online research is a general self-help book. There is limited written guidance on how to bring together the technologies and the research methods. Researchers may be able to gain an appreciation for the subtleties of the online environment and the ways in which research questions need to be challenged, rethought and reshaped when researching online phenomena. They can also gain an appreciation of the kinds of challenges their methods will receive both in the research process and in academic research discourse. But what is not readily available to the researcher is how to actually conduct the research and apply the technologies. What technology should be used for conducting an interview? In what ways will digesting features support or challenge the research effort? Is there anything that can be done to guarantee the anonymity and security of a research participant in a chat room being used to conduct a focus group? How should firewall issues be handled? How to organize, manage, analyze and present data in online research? These are all questions that remain unaddressed in the current body of literature.

In our highly rationalized society, what material is available is specialized. You can find books and manuals that go into detail on specific topics such as making graphs (*Charts & Graphs in Microsoft Office Excel*, 2007), or how to use Internet Relay Chat (IRC) (*The Ultimate Guide to Internet Relay Chat*, Charalabidis, 2000). The books, like our professions, are presented as the experts on their respective topics. But what if you don't even know what technology you should use? Is it necessary to pick up every book on every different type of technology to determine which technology is right for your research purposes? We don't think so. We believe that there are times when the generalist is helpful.

We suggest you think of this book as your generalist. It focuses on addressing technical rather than methodological questions. It will provide technological advisement for conducting online research. This book will supply a link between various technologies and research methods, enabling the reader to embark on a study and address questions and concerns of his/her own particular research approach. Some of these topics, will be covered several times, with a differing emphasis on administrative, technological and potential ethical challenges.

This book is probably best used as a reference book. As noted previously, it was not created to be read from cover to cover like a manual. It focuses first on what questions to ask and how to prepare to conduct online research in the study design process. The subsequent chapters are organized around particular types of technologies such as email, chat and databases. As we cover these topics, there will be some discussion of specific products. We try to be current, but software and the World Wide Web are very fast changing fields, and new versions and entirely new products appear all the time, seemingly within the time it takes to describe a particular product. The final chapter touches upon some additional thoughts unique to research in the online environment, with the suggestion that there are topics beyond the scope of this book that are worth further consideration.

Cyberspace, the internet, the world wide web and other definitions

We often hear a number of words bandied about regarding the online environment. The three most prevalent are internet, cyberspace and World Wide Web. While not knowing their specific meanings might have limited impact on your research, it's possible you'll be more convincing as an online researcher if you can articulate their differences.

Cyberspace was coined by the novelist William Gibson and used to reflect the storage, modification and exchange of data. As noted in Wikipedia, it is often used synonymously with the internet, but it is not. Cyberspace, more accurately, reflects "objects and identities that exist largely within the communication network itself" (<http://en.wikipedia.org/wiki/Cyberspace>, accessed on 21 August 2007). The **internet**, on the other hand, refers to a network of networks through which the data of cyberspace are transmitted.

The **World Wide Web**, often referred to as 'the web', is a system of inter-linked hypertext documents, the web sites and hot link references that reside on the internet. Communication of many sorts (text, still and motion pictures, sound) can be put on a web site, which is on a web server. One of the main innovations of the web (and the reason for the name) was the use of embedded hyperlinks from one document to another. The network of links between documents becomes the web. Web documents are viewed with a web browser. As a form of communication, the web is more like a broadcast in that there is no guarantee that the web site will be viewed by an intended recipient, or not viewed by an unintended recipient.

It was originally true that communication by web site was one way – from the author to the reader – but this is becoming less accurate as new web technologies allow for two-way exchanges. Web technologies have evolved from

8 A GUIDE TO CONDUCTING ONLINE RESEARCH

essentially one-way (server to browser) connections, to increasingly more interactive exchanges. First came forms, where end-users could enter a small amount of text into a field on the web page, or set a button or checkbox, and have that information returned to the server for processing, or storage. In parallel, technologies were developed for the visual (WYSIWYG) editing of web pages; and then for the synchronization of local copies of web documents with the server copies. At this point these have been combined, so that the user can be on a certain web site, and edit the appearance of the web page or create new web pages. In this way the web server/client relationship has become very interactive, and has allowed for the development of technologies such as weblogs and wikis; and of social networking sites like Facebook, MySpace, Flickr, YouTube and others. The impact on the nature of research and the evolving roles of participant and researcher remains to be seen.

Currently, the protocols for web sites are fairly standardized: HTTP (Hypertext Transport Protocol) for the transmission of data; HTML (Hypertext Markup Language) and in some cases CSS (Cascading Style Sheets) for the appearance of documents in a browser. HTML is most often written to a document with a text editor or a visual web site creation tool, but in either case it is stored as a static document on the web site. In some cases, for web sites that are created in response to some event or feedback from a user, a program is used to generate dynamically the HTML that is presented to the web browser. The communication from a web form to the web server is based on the CGI (Common Gateway Interface) standard. Information entered into text fields, or checkboxes is returned via the CGI to the program running on the web server, which can then use the information to create a new web page. These programs can be written in any programming language, but are very often written in scripting languages such as perl, python, PHP and ruby. These languages have strong built-in tools for parsing text (in the data processing sense of separating continuous text into individual words or symbols of interest) and recognizing patterns, which is one of the main tasks of a CGI program, and additionally have tools for interacting with database software, which will be seen to be useful.

Other computer-based communications methods, such as Usenet and Chat, are increasingly being presented on top of a web-based interface as well. The Usenet is one of the oldest internet communication technologies. The server component is called a "news" server; it stores posted messages, and downloads them to any subscribed user who checks for new messages. It is organized by news groups, which have internet domain-like names, such as rec.food.cooking, or soc.culture.thai, etc.... The user interfaces to a Usenet server with a news reader, which is often built into mail programs, as is the case for Outlook, Thunderbird, and others. Users subscribe to individual Usenet groups, and the news reader checks for any new messages in the group and downloads the header and contents to the client. Users may post to the group by composing a message and sending it. The technology and experience is very similar to an

email listserver. The difference is that Usenet groups have no restrictions on who may subscribe and download. The Usenet first became searchable in the 1990s with DejaNews. Google has since purchased DejaNews and reformulated it as Google Groups. Now, other internet portals like Yahoo! provide group interaction as well.

Locating the internet in technological change

The transformation of communicative technologies has had a significant impact on society throughout the ages. From the early uses of art as a method of communication to the first written language of the Sumerians in 3200 BC, humans have changed, and been changed, through their ability to communicate. The introduction of the alphabet by the Greeks around 700 BC moved communication of ideas from the spoken to the written word (Castells, 2000: 355). Likewise, the development of the Gutenberg printing press in 1440 facilitated the eventual implementation of universal education and the mass production of literature (Buchanan, 1992: 173). In a similar fashion, the personal computer, a present-day technology, and the internet, the publicly accessible network for computer interaction, have impacted our understanding of work and leisure activities. People can work from home instead of going to an office space, conduct personal banking from home, and gather information on a wealth of topics with relative ease. In addition, they have given rise to new forms of social interaction and changed the ways in which people interact.

Interacting on the internet raises questions about subtle nuances in the way people communicate, shifting mannerisms and both challenging and enabling individuals to choose new ways in which to present themselves to others. These types of subtle shifts are fertile ground for research. Online interacting raises complications for researchers, in that they are unable to physically see a research participant's reactions to questions, for example, or, have a context in which to appreciate and understand sarcasm. It is much easier for individuals to lurk in an online environment, limiting a researcher's ability to know who is actually "present" in the research process. While these topics raise important questions about the online environment, and challenges for conducting research online, they will not be addressed in depth in this book. Instead, we will direct you to some resources that will enable you to follow up on those issues of greatest importance to your particular work.

It is our hope that the following pages will support you, the researcher, with the technical information you need to utilize the various technologies we discuss in your research endeavors. We wish you the best in your research, and hope you find the resources you seek as you flip through the following pages.