1 Introduction

Why not just search Google?

Internet searching seems simple. Enter a few words about a topic into a search engine such as Google and you're sure to find something relevant fairly quickly. This is exactly how many people use the Internet as a research resource. A search engine such as Google, Yahoo, MSNSearch or Ask.com will certainly turn up a few sources on almost any topic you can think of. It's a different matter to find the central sources, the most reliable sources and the most relevant sources.

Increased reliance on the big search engines provides a large part of the explanation for a rapid rise in student use of poor quality and unreliable sources in recent years. A 2002 survey of over 2,000 US university students found that almost three-quarters of them used the Internet to search for information more than they used the library. Only 9 per cent of students used the library more than the Internet. This heavy reliance on online sources is a cause of concern to many university lecturers and professors, who fear that students are missing out on high-quality academic sources (Jones and Madden, 2002). In a 2004 survey of over 2,300 US university professors, over 42 per cent agreed that the quality of student work had worsened with use of the Internet. Only 22 per cent thought it had improved (Jones and Johnson-Yale, 2005).

One response to these problems is to ban the use of online sources, as some academics have done. The alternative response, outlined in this book, is to develop the skills to use online sources more effectively and more critically, and to place them at the centre of the research process.

Research is not simply about finding the answer to a question. It involves learning about the main issues in a particular area and identifying the central arguments made by those on all sides of the ongoing debates. It

involves getting to grips with the key works on the topic, the books and articles that set the terms of debate and help you to understand the way in which those debates have developed. It also requires you to become familiar with a range of key resources specific to your subject area, including subject-specific databases and archives, organizations, professional associations, email lists and blogs. In the process of exploring the key debates and the literature, you develop a much more focused understanding of the topic you're researching, preparing the ground for searches on more obscure aspects of your topic, and the identification of very specialized resources and collections of data and documents. The research process is not only about generating a piece of work. It is also about becoming an expert on the sources and on the literature in a specialized area whose boundaries you help to define in the course of your research. This expertise is one of the most valuable outcomes of the research process.

Who this book is aimed at

This book is intended for use as a textbook for courses on Internet research skills both at undergraduate and postgraduate level, courses that introduce students to efficient and critical use of the Internet as a research resource. It will also be of use to researchers in a wide range of areas, from media to government, to the voluntary sector, who are concerned with issues of current debate and controversy. The book assumes that the reader makes regular, if sometimes unproductive and frustrating, use of the Internet.

It will be useful to advanced undergraduates writing an extended essay or pursuing an independent research assignment for the first time, particularly if they hope to continue to postgraduate level. It is aimed particularly at postgraduate students embarking on a thesis or dissertation. It is an ideal companion to the first year of a PhD thesis when a student is writing a literature review, exploring the field and identifying major resources. The book provides focused advice on a range of tasks involved in writing a thesis or dissertation, and associated exercises.

The book will also be valuable to experienced researchers who make regular use of the Internet but who have never systematically explored the resources available in their area and have never developed systematic strategies for carrying out online research.

The exercises at the end of each chapter bring readers through the material covered in the text while allowing them the freedom to concentrate on their own area of research or special interest. By simply doing the exercises readers will work their way steadily through many of the central tasks involved in their own research project. Throughout the book the importance

of taking a critical approach to sources and the need to develop a better understanding of online sources are emphasized.

The book will help all readers to identify major resources in their areas, to flexibly combine a variety of online resources in the pursuit of specific tasks, to use the big search engines more effectively and efficiently, to critically evaluate online resources, and to take much fuller advantage of the huge potential of online resources in their own research.

Understanding the Internet

The Internet is not an organized system. No one is in charge. It is not primarily a network or even a network of networks. Above all, it's a simple fact – the fact that millions of computers across the world can communicate with each other. When you click a link on your computer screen, your computer sends a message to another computer asking it to send you the file the link refers to. The other computer will understand the request and will send the file. Essentially, the Internet consists of files on other people's computers that they allow the outside world to look at. To put their information 'on the Internet' people just move a file to that part of their computer, or to another computer, that is open to the outside world.

The networks that carry your request include phone lines, satellite dishes, and cables. Such lines of communication have existed for a long time. Many of them were not built specifically to carry Internet traffic and they are not the essence of the Internet.

The Internet has existed since the 1960s. That is, computers have been able to communicate over long distances since then. When the World Wide Web came into existence in 1993 it became far easier for the wider public to take advantage of this fact. The Web is based on http, the Hyper Text Transfer Protocol, and the major innovation it introduced was hypertext, the links in web pages that link you to another web page. Embedded in every hyperlink is an Internet address, a URL. It might point you to another document on the computer you are connected to or it might point you to a document on a computer on the other side of the world. It allowed people to create very elaborate documents online for the first time, including links to other documents, to graphics, to sound and to video.

The Web is not to be confused with the information people make available on their computers. Everybody puts up his or her own information. The Web 'protocol' just means that all the machines understand each other and display documents from other machines in a standard way. Web browsers such as Safari, Netscape, Firefox or Internet Explorer are not responsible for that information. They simply provide an easy way to view it.

Summary of chapters

The early chapters in this book focus on academic books and articles online, directing users towards the best organized and the most authoritative resources available, before moving on to deal in later chapters with the open Web. Search strategies illustrated by examples in each chapter combine a variety of online resources in pursuit of a series of research tasks. These strategies bring readers from the initial exploration of their topic, through the literature search, to the identification of major sources, the development of specialized search queries and the search for primary sources of data and documentation. The wide range of online resources drawn on in these strategies emphasize the fact that the big search engines, on which many people rely entirely, provide access to only a fraction of the research resources available online.

Chapter 2 focuses on academic books online. A range of online services, including bookshops such as Amazon, electronic book services such as Net Library and search services such as Google Book Search provide new ways to access the full text of academic books. This chapter outlines strategies for combining different online resources to conduct the kind of book searches that were impossible a few years ago. It outlines the way in which 'needle in the haystack' searches can now direct you to books that you would never have been alerted to through their title or list of contents, how to use these services to identify key works in your area and how to incorporate novel resources, such as reader reviews, into your search strategies.

Chapter 3 deals primarily with academic articles. It explains how to combine the databases of abstracts, full text databases, library catalogues, Google Scholar and similar tools to make full use of the available resources. It explains the extent to which these services overlap and the ways in which they differ, explaining their comparative strengths and weaknesses. It deals too with the increased access to articles, theses, dissertations, working papers, conference papers and other kinds of academic material that have been made possible through the open access movement and the growth of institutional repositories.

Chapter 4 on subject guides is the first chapter to deal with the 'open Web', a realm of chaos in comparison with the databases of academic books and articles. It seeks to build an understanding of the way in which websites are organized and to set out strategies for identifying the key online resources in your area. It explains enough about the technology to allow readers to develop a better understanding of the materials they are looking at and the context in which these materials are located.

The big keyword search engines such as Google, Yahoo, Ask.com and MSNSearch are dealt with in Chapter 5. The chapter discusses advanced search options, provides advice on analysing results, explains how and

what the search engines search and provides advice on building queries. It explains how to make effective and efficient use of these search engines.

Web 2.0 is the term used to describe a new generation of online services and resources characterized in part by their interactive nature. Many of these services blur the distinction between writer and reader, producer and consumer, incorporating the contributions of viewers and readers into online services to create powerful and novel means of classifying and locating information. Chapter 6 on 'Interaction, news and multimedia' deals with these services, with news sources and with a range of visual and audio resources available online.

Chapter 7 considers government sources, archives and statistics. Perhaps the most dramatic effect of the new technologies has been to open up access to these primary sources, allowing us to search millions of pages of legislative debates, to access not only statistical reports but in many cases the raw data on which they are based as well, and to make archives much more transparent. This chapter identifies key resources for searching these primary sources and suggests strategies for approaching them.

Chapter 8 sets out detailed guidelines for evaluating and understanding online sources, providing a set of questions that can usefully be asked of any materials found online. It also provides detailed advice on citing a variety of online materials.

This book focuses primarily on English-language resources, and on multilingual resources that include the English language. In every language there will be a range of additional key resources uniquely important to those searching for resources in that language but they are beyond the scope of this book.