>>Expert internet searching

FIFTH EDITION

Every purchase of a Facet book helps to fund CILIP's advocacy, awareness and accreditation programmes for information professionals.

>>Expert internet searching

FIFTH EDITION

Phil Bradley

facet publishing

© Phil Bradley 1999, 2002, 2004, 2013, 2017

Published by Facet Publishing
7 Ridgmount Street, London WC1E 7AE
www.facetpublishing.co.uk

Facet Publishing is wholly owned by CILIP: the Library and Information Association.

The author has asserted his right under the Copyright, Designs and Patents Act 1988 to be identified as author of this work.

Except as otherwise permitted under the Copyright, Designs and Patents Act 1988 this publication may only be reproduced, stored or transmitted in any form or by any means, with the prior permission of the publisher, or, in the case of reprographic reproduction, in accordance with the terms of a licence issued by The Copyright Licensing Agency. Enquiries concerning reproduction outside those terms should be sent to Facet Publishing, 7 Ridgmount Street, London WC1E 7AE.

Every effort has been made to contact the holders of copyright material reproduced in this text, and thanks are due to them for permission to reproduce the material indicated. If there are any queries please contact the publisher.

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library.

ISBN 978-1-78330-247-5 (paperback)
ISBN 978-1-78330-248-2 (hardback)
ISBN 978-1-78330-249-9 (e-book)

First published (as *The Advanced Internet Searcher's Handbook*) 1999
Second edition 2002
Third edition 2004
Fourth edition 2013
This fifth edition 2017

Text printed on FSC accredited material.

Typeset from author's files in 10.5/14 pt Bergamo and Myriad Pro by Flagholme Publishing Services.
Printed and made in Great Britain by CPI Group (UK) Ltd, Croydon, CR0 4YY.

Contents

List of figures **ix**

Preface **xi**

Acknowledgements **xiii**

1 An introduction to the internet **1**
Introduction 1
An overview of the internet 2
What the internet is and is not 3
Search engines 15
Failing resources 17
The information mix 18
Summary 19

2 An introduction to search engines **21**
Introduction 21
Search engine criteria 22
Free-text search engines 23
Data collection 24
Relevance ranking 26
Index- or directory-based search engines 28
Multi- and meta-search engines 29
Visual search engines 30
Hidden, invisible, deep and dark web search engines 31
Multimedia search engines 33
Social media search engines 33
Other search engines 34
Summary 35

3 The world according to Google **37**
Introduction 37

	The Google experience	38
	Google search: the basics	40
	Boolean operators	41
	Phrase searching	43
	Narrowing your search results	44
	Suggestions and searches related to	45
	Advanced search functionality	46
	Other search functions	49
	'One-trick ponies'	52
	The knowledge graph and carousel	54
	URL modifiers	55
	Putting it all together	55
	Other Google search tools	57
	The disadvantages of Google	64
	Summary	68
4	**Other free-text search engines**	**69**
	Introduction	69
	Bing	69
	DuckDuckGo	74
	Yandex	80
	Other free-text search engines	84
	What to look for when exploring a new free-text search engine	87
	Summary	88
5	**Directory, clustering and similarity search engines**	**89**
	Introduction	89
	Directory-based search engines	90
	Clustering search engines	93
	Similarity search engines	94
	Summary	99
6	**Multi- and meta-search engines**	**101**
	Introduction	101
	Multi-search engines	102
	Meta-search engines	107
	Summary	109
7	**Social media search engines**	**111**
	Introduction	111
	General social media engines	113
	Searching Twitter	118
	Hashtags	120
	Searching Facebook	122
	Summary	125

8	**Visual and image search engines**	**127**
	Introduction	127
	Some visual search engines	128
	Major Image search engines	130
	Reverse image searching	136
	Creative Commons image searching	138
	Other image search engines	141
	Summary	143
9	**People-based resources**	**145**
	Introduction	145
	Standard search engines	146
	E-mail search engines	149
	People finders	149
	Social media tools	151
	Finding people that you don't know	153
	A search engine vanishing act	155
	Summary	156
10	**News-based search engines**	**157**
	Introduction	157
	General search engines	160
	Specialised news search engines	161
	Newspapers and traditional media resources	163
	News curation tools	166
	Fake news	170
	Evaluating the news	172
	Summary	175
11	**Multimedia search engines**	**177**
	Introduction	177
	Traditional search engines	178
	Video-specific search engines	179
	Educational video search engines	183
	Audio search engines	188
	Podcast search engines	189
	Summary	191
12	**Specialised search engines**	**193**
	Introduction	193
	Academic search engines	193
	Search engines for children	198
	Other specialised search engines	200
	Summary	207

13 Hints, tips and the future **209**
Domain names, URLs and web pages 209
Get the most out of your browser 210
Searching 213
Finding confidential information 216
Other resources to test your searching skills 218
Mobile search 219
The future of search 221

Index **225**

List of figures

1.1	Trending topics in Facebook, Friday 13 January 2017	9
2.1	The down arrow symbol on Google giving access to the cache	26
2.2	Data provided by Google on cached versions of a page	26
3.1	Search options in Google	44
3.2	The Google on-screen calculator	53
3.3	The author's news preferences on Google News	59
5.1	The circles category of 'Librarian' as defined by Carrot^2	94
6.1	The Soovle results page	106
7.1	Search results from the Socialmention* search engine	114
7.2	Search is Back user interface	123
8.1	The SERPs from Dothop	128
8.2	A set of visual results from Cluuz for the search term CILIP	130
8.3	The Google image search interface	131
9.1	A Google+ reference to an article arising from a search for search engine social networks	148
9.2	Influencers in the subject area of 'CILIP' on Buzzsumo	154
11.1	The transcript of a podcast listing references to the search term library in Audiosear.ch	190
12.1	A search on Knowem	202
12.2	The Wayback Machine's archive of the author's website	204
13.1	The Bookolio browser page	213

Preface

Welcome to this, the fifth edition of *Expert Internet Searching*, which began its life back at the beginning of the century, when it was called *The Advanced Internet Searcher's Handbook*. That very first edition, published towards the end of 1999, didn't even mention Google. Most of the search engines that I wrote about no longer exist and over the years functionality has waxed and waned.

We're now in a world of user-generated content, access to data not just at the touch of a button but by simply asking the question into our living rooms where our mobile assistants can research and answer the question as soon as we have finished speaking it. 'News' now means things that happened in the last few seconds, instead of the last few hours or the last day or week, and we've seen the rise of 'fake news'. While we have such easy access to information, and everyone can contribute to it, the other side of that coin is that everyone does it, only not for the greater good, but to promote their own political or religious agenda, and in doing so accuracy and truth become immediate casualties.

We all still see the rather silly claim that 'it's all on Google, so why do we need libraries and librarians?' but I firmly believe that information professionals are needed now more than ever. It's in part our role to help stem the tide of fake news, to open people's eyes to the rich abundance of information available in so many different formats, and to assist them in working out what they need to know and the best way of getting it. When I was a child and I told my careers officer that I wanted to be a librarian she said 'is it because you like books?' and I said 'No, it's because I want the power'. It's even more true today than it was rather too many decades ago. The more information that we have at our fingertips, the more useful and powerful we become. However, we don't get that by simply visiting Google at every opportunity. We need to know how Google works inside and out of course, but we also need to understand what

other search engines are better than Google, which engines provide access to material that Google doesn't, and how best to utilise those resources. That's really what this book is all about – providing you with the information that you need to play with other search engines, explore their capabilities, strengths and weaknesses and appreciate when to use one over another. The more we can get the information we need quickly, the more we can help our clients and make their lives easier. While it may be argued that's perhaps something of a 'Pollyanna' view to take, I can balance that out with the rather more cynical view that the more we can find material that other people can't, the more important information professionals become, and the more we will all be valued.

Talking of valuing people, over my career I have been helped by some marvellous, wonderful people, and I'd like to take a moment to recognise them:

Tara Calishain's enthusiasm with her site 'ResearchBuzz' astonishes me with what she is able to find and her generosity in sharing it.

Sarah Houghton (the Librarian in Black) leaves me awestruck by her compassion, her strength, her values and her resilience to never stop fighting for what is right.

Marydee Ojala who knows more about search than I ever will, and is just so enthusiastic about it!

Greg Notess was one of the very first search engine experts I met, and he helped show me what a rich and fascinating area it was.

David Lankes has been my personal 'guru' for many years, and I have been encouraged and excited by his view of librarians, and librarianship.

Gary Price who has always amazed me with his breadth of knowledge in the field of information.

Michael Stephens has a generous and positive view of our profession, and has so often shown me by words and deeds how caring for people benefits us all.

Then there are all of my British colleagues, who over the course of my career helped, aided, assisted and turned me into a better person, and hopefully better information professional! Jo Alcock, Ian Anstice, Barbara Band, Karen Blakeman, CILIP colleagues, Hazel Edmunds, Biddy Fisher, Hazel Hall, Beverley Humphrey, Lisa Jeskins, Alison McNab, Professor Charles Oppenheim, Dave Pattern, Nick Poole, Caroline Roche, Val Skelton, all of the 'Voices of the Library', Arthur Weiss, Nicky Whitsed, and so, so many others. It's little by the way of 'thanks', but this book is dedicated to each and every one of you.

Phil Bradley

Acknowledgements

Working with people in the information community is a real joy, since they are some of the most interesting, helpful and committed professionals on the planet. I have been fortunate enough to work in many countries around the world and have only to walk into the nearest library to feel at home, with new friends waiting to be met. Conferences are a joy to attend since they provide me with inspiration, ideas and unique insights into many different communities. Social media platforms such as Facebook and Twitter allow me to interact with information professionals all day and every day and I can reach out, ask a question and receive a helpful answer within minutes. So thanks to all of the many hundreds of people who have helped increase my knowledge of the world around me.

I'd also like to thank the people – most of whom I don't know and have never (and will never) have chance to meet – who spend their time crafting search engines, tools, apps and helpful websites for the rest of us to make use of. Of course, the staff at Facet Publishing and CILIP are a great group of people to work with – enthusiastic, helpful and informative; it's invidious to name too many names, but particular thanks to Natalie and Damian.

Finally, friends and family, information professionals already mentioned, and you the reader, many thanks for your continual support – even when you don't realise you're providing it!

Every effort has been made to contact the holders of copyright material reproduced in the text, and thanks are due to them all for their kind permission to reproduce the material indicated. If there are any queries, please contact the publisher.

Phil Bradley

CHAPTER 10

News-based search engines

Introduction

There are several questions that arise when looking at the concept of news search. First, what exactly does news mean now? Global and national news is a fairly clear concept, and we can take that down to the next level of local news. However, is it 'news' that your next door neighbour has painted their fence? To you it may well be very interesting and valuable to know what product they used and how effective it is, because you can then go and buy some paint yourself from the same manufacturer if you have been shamed into painting your fence, too. However, your friend at the other end of the country may well not regard this as news at all, though if you paint your fence, it may be – so we could perhaps define 'news' as being fresh information that is of particular interest to us and our friends and colleagues.

Given the ease of use of social media-based tools where we can share updates so easily it is perhaps understandable that people will view the concept of 'news' in a very different way now than they did prior to this ease of access. Indeed, with cameras – both still and video – in most people's pockets in the form of their smartphones, everyone can suddenly become a journalist, and it's quite usual these days to see video footage appear on news channels which started off in a smartphone, or images in newspapers that have been taken (sometimes with, sometimes without, permission) from Flickr and other photograph-sharing websites. Tools such as Periscope, at www.periscope.tv, allow you, in its own words to 'explore what the world is seeing'. I visited the website and I was able to see live streaming of a person showing me the traffic passing an intersection in New York City, the sunrise in California and a musician singing in a bar in Dublin. I could also follow along with the speeches at the Scottish Nationalist Party Spring Conference, learn how to bake cream

cheese tarts and watch a protest in respect of immigration rights. All of this is news to some, of mild interest to others and boring to yet another group!

Did you know?
2017 is the first year in which half the world's population was connected to the internet (https://thenextweb.com/insights/2017/01/24/digital-trends-2017-report-internet).

On a professional level, speeches and workshop discussions can also easily be shared, so while it's of no international or even national interest, what a speaker thinks about a subject is still going to be important news to the people who are dedicated to that particular subject. So 'news' now has to be seen rather more in the context of what an individual defines as news.

Secondly, what sort of time period are we looking at here? Books can have a production cycle of several months, magazines and journals perhaps a month, newspapers will define as news anything that happened since they last published, so let's say that's roughly a day, and search engines can update their indexes as often as needed to reflect the changing content of news sites. We then have the news sites themselves, which try their best to keep up with the breaking stories of the day, and closely allied to that we have blogs; once again, we're talking about hours or minutes. Finally we have the greatest news pipe of all – Twitter, which can present a news story that is seconds old.

Did you know?
The word 'Internetted' was first used in 1849, meaning interconnected or interwoven (https://en.wikipedia.org/wiki/Internet).

Thirdly, we have to consider the idea of authority as well. In 'the old days' it was relatively easy to work this out – we could assume that an academic article had been peer-reviewed prior to publication, a book had been edited, and that newspapers had a particular bias that everyone was aware of. However, if I get a news item off a blog, or as a tweet, it doesn't necessarily have the same type of authority that it would have had otherwise. In the early days of Twitter, the only way that one could be certain that a celebrity was actually the famous person you thought was tweeting was if Stephen Fry rang them up, asked them and confirmed it in a tweet. While it's slightly amusing that he was acting as a *de facto* validation source, the rather more important point is that we now need to consider the context of the news item – not just who said it, when and where they said it, but who their friends are, and the extent to which their

extended network can validate them. Unfortunately, it's not really that easy; the need to get a story out as quickly as possible can sometimes take precedence, as a daily tabloid found out some time ago. It ran a story based on a tweet from Steve Jobs, which stated that he was going to have to withdraw the iPhone 4 due to technical problems. The journalist working for the newspaper had built up a story based on a single tweet and hadn't checked the actual biographical details of the account, which made it perfectly clear that it was a fake/parody account. At the other end of the spectrum, someone can tweet something which they don't even realise is news. Sohaib Athar started tweeting descriptions of events that happened close to his home in Pakistan, and it was only later that the world realised that he was live-tweeting the attack on Bin Laden by US forces. So – can we, or should we, trust the established newspaper and professional journalist, or should we trust someone who is reporting exactly what they are seeing in front of them? Of course, it's not really that simple, but the concept of authority, validity and trustworthiness becomes more important every day. End-users probably will not realise this, and indeed they may unwittingly perpetrate a myth – in January 2012 there were reports that Cuban leader Fidel Castro had died. This story was tweeted and retweeted hundreds of thousands of times, but it wasn't true. So simply because a large, indeed huge, number of people say something, it's really meaningless.

Finally, we have to consider how we can deal with news that we're not supposed to have. Wikileaks, for example, is a not-for-profit media organisation which provides a way for sources that wish to remain anonymous to leak information to the Wikileaks journalists. This may sometimes be highly sensitive information that should not be in the public domain. There are times when in the UK superinjunctions have been used by celebrities to keep details of legal cases that they're involved with out of the press, but details of which have appeared in social media resources such as Facebook or Twitter. What are our ethical and moral duties here? If we are faced with a request for information, and we're able to get the information which has appeared in the public domain, should we use it, provide it with a caveat, regard it as entirely unreliable and ignore it, or something else?

The concept of 'news' is therefore fraught with many and varied difficulties, and there are very few answers, so perhaps all that we can do is to keep an open mind, educate our members and just do the best we can. Having raised several spectres I will now take a step back from them and turn my attention to the actual tools that you can use in order to find newsworthy information.

General search engines

Many of the major search engines such as Google, Bing, Yahoo! and DuckDuckGo will provide access to news stories. Consequently, it's very easy to use your preferred search engine as a one-stop shop when it comes to news information, but it does help to have an indication as to the sources that the engine uses, how the stories are rated, if there are any inherent biases towards the subject matter, and the extent to which you can get a rounded and comprehensive view of a particular story.

In my opinion, of the general search engines Google provides the best news coverage, and you can find it at https://news.google.com. This was my view in the last edition of this title, and I haven't changed my mind since, even though other news resources have come along. It provides some really good functionality and users of the service have a great deal that they can customise. As well as the 'look and feel' of the page, users can choose what resources they have displayed on the page by personalising it with various sliders, so that for example if World News is really important the slider can be shifted to the right, resulting in more stories and the Entertainment slider moved to the left to decrease information on celebrities. However, it's also possible to add in any news topic that is of interest, and Google also offers the option of adding in ready-made sections such as social networking, space, physics and so on. Therefore, users can create custom sections to locate news items that are local to them, they can add country editions and can even decide how the news is presented – sorted by most users (of a section), highest rated or newest for example.

News stories are displayed clearly, with different news sections to the left of the main stories and recent items, local items and editors' picks to the right, forming an interesting and effective triptych. Each news item has its origins clearly marked and related stories can quickly and easily be pulled up, to add in-depth coverage to a subject. If available and appropriate Google will also display video news clips from major news sources such as BBC News. Google News has an advanced search function, and search terms can be limited to particular elements of the news coverage, such as the title, the source or the country the article is published in. However, it's only available as an option when a search has been run, which is quite irritating if you know exactly what you want, but have to run a search in order to get you to a position where you can do the exact search that interests you.

When searches have been run, the searcher can create an e-mail alert for the subject that they have searched for. They can choose the terms of the query, the type of information that's returned to them, such as news, blogs, video or discussions, how often they get an e-mail (as it happens, once a day or once a

week), how many results, and the e-mail address the alert should be sent to. There is no limit to the number of alerts that can be set up, and they can be as simple as an individual word, or a much more complex search query. Searchers can also save the search as an RSS feed, which can be incorporated into their newsreader.

Bing's options, at www.bing.com/news, are a little less 'full on' than Google's, and the news page is far less busy, with a search box, some leading headlines and then below that leading headlines in a variety of different subject areas such as World, UK, Business, Politics and so on. Those subject areas can be expanded if preferred so that the engine just displays all the leading stories from a subject section. Finally, users can of course search for whatever interests them, with results displayed either as the best match or the most recent. However, it's not then possible to limit further without more work, since stories on the subject of community libraries, for example, will return all stories on the subject, regardless of country, which is an annoying limitation of the service.

Yahoo! News, at https://uk.news.yahoo.com/, is another busy screen of news. Under the search bar the engine presents a variety of news categories such as UK, World, Technology and so on, and then presents news stories, with an emphasis on image rather than text. Added to that is a series of advertisements which are not always clearly indicated as such, leading to some confusion. The rest of the news screen is a mixture of images and textual summaries for different categories, 'most popular' news items, latest videos, a poll on a current news story, editor's picks, slideshows, weather and so on. The Yahoo! news resource can obviously be searched, with results being displayed and filtered by either relevance or time. Searchers can then further filter by source and/or time. The search can also be saved as an RSS feed as well.

DuckDuckGo has a small dedicated news section, with half a dozen stories displayed in tile format culled from various global sources. However, it also uses the *!n bang* option, which pulls results directly from Google's news feed. Due to the nature of the search functionality that DuckDuckGo offers, there are a wide variety of other news-based search functions. Users can search via key broadcasters, such as *!bbc*, and they can search international resources, by magazine, newspapers, speciality sources and weather channels.

Specialised news search engines

There are a small number of search engines that simply report on the news – gathering data from news websites such as television stations, newspaper sites and so on. They have a single focus, which is to get news to you as quickly as

possible. You may therefore find them to be a rather more appropriate choice if you just need global or national news which is regularly updated.

Infonary

Infonary, at http://infonary.com, is described as a 'real time news digest'. It has a very basic home page, with a small number of headlines, and it also takes headlines from leading UK resources such as television news and newspaper sites. There are a variety of video news items from different video sites such as YouTube. As well as the normal search box several categories of news are available, such as Politics, Sports and Technology. It is also possible to view news via country or alphabetically by subject. Once users move away from the category approach and into personalised searching for key terms Infonary finds results from news sites, search engine news, YouTube, blogs and Twitter. Infonary is updated regularly and provides a good service, although it's rather let down by the very sparse interface.

Newsmap

Newsmap, at www.newsmap.jp, is at first glance a very confusing page, because it has a large amount of headline content contained in different sized squares across the page, all in different colours. However, given a few minutes perusal it all becomes clear. First, users need to locate the tab across the very top of the screen which gives them a choice of country – about 16 options, heavily biased towards the West. The next thing to check is the bottom of the screen, which has a colour-coded index, so headlines with red backgrounds are world news, purple backgrounds are health news and so on. Finally, the lighter the colour, the more recent the news, and the larger the square, the more important the news. This sounds terribly convoluted and messy, but it's a very quick concept to pick up, and the system does provide a considerable amount of granularity, so that with a couple of mouse clicks a user can visually focus on technology stories from the USA, or health and sports stories from Argentina. The colour and size focus really does come into play when running a keyword search, however, as the main stories can leap out of the page at the searcher, helping them to quickly isolate lead stories.

News Now

News Now (www.newsnow.co.uk/h) was launched in 1998, and is currently

visited by over 14 million users a month. It's an aggregation service, so it matches breaking news stories in real time against keyword-based topics. In common with other news services it provides small 'boxes' of information on particular topics, which can be clicked on as appropriate, with the original story opening in a new browser tab. News Now also offers a menu of subject areas such as hot topics, science, sports and arts and culture which break down into narrower subjects. It's therefore really easy to go from Sport to Football to the Premier League, to a club and then to transfer news and details of particular players. At every point in this focusing process the searcher is shown how many new stories are available, which is very effective. Searchers can simply type in their topic(s) of interest and the engine will provide information arranged by time; the last hour, two hours, back to the last seven days – if preferred, other date ranges can be chosen. News stories are taken from over 40,000 sources and auto-refresh every five minutes.

There are of course plenty of other alternatives available if you're unhappy with any of the ones that I've discussed above. Try:

- Rocket News, www.rocketnews.com
- World News, http://wn.com
- Daily Earth, http://dailyearth.com
- Headline Spot, www.headlinespot.com
- News Lookup, www.newslookup.com.

Newspapers and traditional media resources

An obvious place to look for news stories is where we've always looked for news in the past – newspaper and television websites. There are quite simply thousands of news sites which go from the tiny local newspaper site that's updated once a week to the huge news resources from organisations such as the BBC which are updated minute by minute. Obviously each of these resources has their own bias, but that is to be expected. Nowadays we can expect each resource to update continually throughout the course of the day with new information and breaking stories, and to include video reports as and when necessary, and even live reporting as needed.

Did you know?
Print-only is still the most common way of reading newspapers in the USA (www.journalism.org/2016/06/15/newspapers-fact-sheet).

Some newspaper sites

Having apparently dismissed these tools in a paragraph, there are some sites that I would like to particularly draw your attention to. The Newseum site, at www.newseum.org/todaysfrontpages, shows over 2000 front pages from 91 countries. It's enlightening and frustrating by turn, since the front page IS all that you get, but to be fair, there's a link to the newspaper's website itself. Consequently it has limited value, but it's a good way to see how subjects are treated around the world.

Google has a newspaper archive of hundreds of broadsheets, some going back to the early 1800s, and it is international in scope. It is arranged alphabetically, so the searcher needs to have a fairly good idea of exactly what they are looking for, but there is a search function that allows users to jump straight to a particular article in a journal. Google also provides a very handy 'related articles' link to similar stories. Perhaps the most interesting aspect of this collection is that the newspapers are provided visually, so it's possible to move around a newspaper, looking at the advertisements, for example. Unfortunately Google has decided not to continue developing this resource, but they are leaving it online for researchers, so if you are looking for archival information, this could be a good starting point.

If your needs are more current, The Paperboy, at www.thepaperboy.com, has a searchable collection of over 6000 online newspapers and e-papers. Content is displayed in magazine format style, but users can click on a world map, choose by country or simply enter a search term.

If you are serious in a desire to read news, the British Library provides access to the British Newspaper Archive at www.britishnewspaperarchive.co.uk and thousands of new pages are scanned every day – the current figure is at 18,729,440 pages. However, this is a commercial product, and does not have a free option.

The RefDesk site, at www.refdesk.com/paper.html, also has a useful listing of American and national newspapers from around the world, neatly divided into US state and country. It provides excellent coverage and links to almost 50 UK newspapers, for example. They are, however, slightly outdone by the Internet Public Library site, at www.ipl.org/div/news, which lists almost 120 UK resources, as well as magazines. Unfortunately the ipl.org site is no longer being maintained and updated, but is still available with links, so it's worth taking a look at, although as time moves forward 'link rot' will firmly take hold as sites change addresses, so look at it sooner rather than later!

Finally, the Library of Congress has a newspaper and current periodical reading room at www.loc.gov/rr/news/oltitles.html, which links to newspaper

archives, indexes and morgue resources. As you might expect, there is a strong US bias to the content, but there are also plenty of international references and links.

Good news sites

On a lighter note, you may be interested to know that there are a large number of news sites that are devoted to positive news stories. When talking and teaching about news search I often have discussions about how grim all of the news appears to be, so it's quite nice to be able to highlight a few sites that focus on good news. I certainly wouldn't consider replacing any of these with the tools that I have previously mentioned, but if for some reason you're feeling a little down, take a brief look.

Good News Network, at www.goodnewsnetwork.org, has various news sections, just like regular news sites, with top stories, RSS, a subscription service, inspirational quotes and so on. Good News Broadcast, at http://goodnewsplanet. com, is a very busy site with an emphasis on YouTube videos, and very feel-good, though in my opinion not quite so many 'news' items as the previous one. Gimundo, at http://gimundo.com, provides interesting and uplifting stories on a regular basis. The Great News Network, located at www.greatnewsnetwork.org/index.php, has news by category, region, submit news, and an RSS feed. In its 'About' section it really does clearly define exactly what the role of this type of news resource is all about: 'The Great News Network is meant to supplement your daily news sources – not replace it. Its role is to show that there is hope, people are making a difference, and that a lot of things are getting better.' Positive News is a UK-based site at http://positivenews.org.uk. It's the web-based version of a paper-based quarterly. Not so much current affairs, more feel-good, sustainable-future type of information. Yes! Magazine, at www.yesmagazine. org, again looks like a website of a paper-based newsletter/magazine; more on political issues than current affairs from the look that I took. As a brief aside, when researching this section I looked at a lot of different sites, as you can imagine, and I followed up links to find resources that were quite often several years old. While several more 'traditional' news sources had closed, all of these 'positive' news sites were still all in existence, so there's a message in there for us all somewhere!

News curation tools

Of course, the definition of what is 'news' is going to be different for each and every one of us. Up until now I have looked at national and international news resources, but there are two other types that we should consider. The first is news relating to individual subject areas, which while not of national or international interest may still be very important for each of us in our jobs. The second type of news is going to be rather more personal; what our friends and colleagues are looking it, highlighting and sharing.

I don't expect to find a great deal on the BBC news about developments with search engines, new and interesting Web 2-based utilities or things that are happening within the library world. I could of course quite easily set up my own RSS feeds and news alerts to keep me informed, but once again, because much of this information will not appear on the news, or via news sites, these tools can only provide me with a snapshot. It's at this point that my colleagues and their activities come into their own. I have already looked at the value of social media search engines in Chapter 7 and the role that colleagues can play with them, and this is a subject that I'm returning to again.

We are all familiar with the Facebook 'Like' button, the Google +1 button, blogging about resources, adding them to bookmarking services and so on. Even if we don't always do this ourselves, friends and colleagues do. It can be very useful to see what others in your social circles are looking at, commenting on and sharing, because if they like something or find it interesting, there's a high chance that you will experience something similar yourself. However, it's simply not possible to spend all day checking out these resources, looking through friends' profiles and so on – there simply isn't enough time. This is where news curation services become useful. Simply put, you can provide a resource with access to your Twitter account, or perhaps your Facebook account, and let it gather data on the likes and references to resources that contacts find and make. These can then be presented to you in an easy-to-use format, allowing you to very quickly skim through that highly personalised news. While there are a few of these available for the browser, they have particularly come into their own in various apps for smartphones and tablets, though in some cases they are available for both. It's worth pointing out that although a lot of these products ask for your Facebook or Twitter details and then access data produced by your contacts, they are really only acting as a mechanism from drawing together information that these contacts are already sharing with you.

Addict-o-matic

Addict-o-matic, at http://addictomatic.com, has as its strapline 'Inhale the web; instantly create a custom page with the latest buzz on any topic'. It pulls in results from a variety of news resources such as Bing News, Silicon Alley Insider and Cnet's Tech Blog. This is a slightly different product to the other engines in this category in that it doesn't provide a home page of global or national news; it has a category approach but the user is expected to know what they're after, and to create a search for themselves. The user is in control of the resources used, since a pulldown menu of sources is available and it works on a toggle basis, so that different sources can be turned on and off as necessary. News items are displayed in small news boxes which can be moved around or even deleted as necessary, thus creating a very personalised version of the news, which can then be shared with other people if necessary.

Flipboard

Flipboard, at http://flipboard.com, is the one that you'll probably be aware of, if you are at all familiar with this type of resource. It's made up of small subject cubes, so that when you open the app you see a display of squares, which you can tap on to take you into more detail, with stories displayed in a similar cube or magazine format. When you swipe across the screen, you do get a quite pleasing flip effect as the page 'turns'. Clicking on a story opens up the original tweet or status, displays the story on the screen, and links to the full story on the website. An upward swipe loads it directly for you to read. New sections can be added easily, and they're just a tap or two away. It's a very well regarded resource, and it makes excellent use of the visual and touch screen elements of tablet devices. You can link your Flipboard account to your Facebook and Twitter accounts, for example, as well as news feeds and search results for tweets covering specific subject areas.

Paper.li

The Paper.li service, at http://paper.li, is a content curation service. Users can publish and share their own newspapers based on the topics that interest them, and to which others can share. It can automatically find and curate stories of interest based on feedback from users, and from the stories shared by the people that an individual user follows. It provides social network content aggregation from the major sites such as Twitter, Facebook and YouTube, and also has a professional paid-for version.

My newspaper is 'The Phil Bradley Daily' and you can find it at https://paper.li/Philbradley#. It includes headline news, photographs, videos and information on science, technology and business, as well as other categories. In a matter of seconds I can catch up on what the people I'm following on social media platforms are looking at, sharing and commenting. As the paper is freely available for other people to see, you too can read the same content. It isn't the easiest resource in the world to search but if you use your favourite search engine to add in some keywords and run a *site:paper.li* filter you will very quickly pull up newspapers that may interest you. I ran a search for *librarian site:paper.li* and got over 6500 results, and the first 50 were all exactly on topic.

Scoop It!

A web-based tool, Scoop It!, at www.scoop.it, is designed to act as a personal aggregation tool that is designed to let users bookmark stories of interest to them using the bookmarklet, and to share these stories with friends and colleagues. It's simple to create a topic, and when a user finds a story that would fit well into their magazine they can 'scoop it' and add in any commentary that they wish, or simply leave it as it stands. I have two separate pages on the site; 'Internet search' at www.scoop.it/t/internet-search and 'Privacy and search' at www.scoop.it/t/privacy-and-search (two pages are the limit if you have a free account; a commercial account currently allows people to create five pages). I can use Scoop It! in a variety of ways. First of all, when I'm going around the web and I see a page or article that's of interest to me, I can 'scoop it' with the use of a bookmarklet on my browser bar in a matter of seconds. That's then added to the appropriate page, and I can go back and view it again at any time I wish; it's rather like a bookmarking service. Secondly, as you have seen, I can make the page(s) publicly available for other people to view and make use of. Finally, Scoop It! provides me with suggestions for good content; it learns what I am interested in (I can like or discard its suggestions, so that in time it becomes more accurate) and so keeps me up to date in my subject areas. I can filter suggestions by type (articles, documents, pictures, videos or social networks) and I can manually add sources to further focus my interest.

Scoop It! also has a general search box, so people can obviously just search for what is of interest to them, and once they have found a good resource, visit it again in the future.

Storify

If you wish to curate and share stories by bringing together content that is scattered across various social media sites, then Storify, at http://storify.com, is a useful tool to consider. Users can build a consistent creative narrative, pulling in content from sites as diverse as Twitter, Facebook, YouTube and Flickr. Status updates can be dragged and dropped, reordered and edited as necessary. The story can then be embedded anywhere on the web.

Nuzzel

The Nuzzel site, http://nuzzel.com, is another that provides access to 'top news stories and feeds'. Users have to sign in using their Twitter credentials and Nuzzel can then show you news from your friends – that is to say, links to what people you are following are tweeting about and linking to. It's updated every hour, so the news is always fresh, and the headlines are arranged by the number of friends who respond to particular stories. When I checked my account there were articles on how physical books are outperforming digital titles, books for girls, reading lists and information on fake news. It's also possible to respond directly to the tweets from inside Nuzzel, making it quick and easy to get involved with a story and the person you're following.

Nuzzel can also create news collections based on the lists that you have created for yourself on Twitter; collections of people that you follow that you have tagged to specific collections. You may therefore be interested in taking a look at the top news stories from 'Search engines' at http://nuzzel.com/Philbradley/searchengines. Nuzzel also has a search function and links to stories that it has found on social media, based on the keywords that you have searched on. Finally, Nuzzel can send you a daily digest of news stories, so it's really easy to keep up to date with what's going on in your world.

Other tools

As I am sure you expect by now, these are far from the only news curation tools that are available on the web. While they work in a variety of ways, they all exhibit characteristics similar to those previously mentioned. So if none of those attract you, it's worth exploring resources such as Bagtheweb (http://bagtheweb.com), Feedly (www.feedly.com) or Summify (http://summify.com).

Fake news

There has always been incorrect, inaccurate and just plain wrong information on the internet. There's a famous cartoon of a man furiously typing away on his computer and his wife's voice comes from upstairs 'Are you coming to bed?' and as he continues to type he says 'No! Someone said something wrong on the internet.' For many years we've had sites such as Snopes, at www.snopes.com, as the definitive reference source for urban legends, folklore, myths, rumours and misinformation, and before that 'The Straight Dope', a question-and-answer column published in the *Chicago Reader* since 1973, now with a website at www.straightdope.com. However, these sites have always tried to take a story that may perhaps have once had a grain of truth in them and sort the wheat from the chaff. Many of these stories have been handed down, passed on as 'a friend of a friend said' and shared on social media. While these stories may well be fake they are the result of misunderstandings and entertainment, rather than anything more sinister.

However, in the last couple of years we've seen an increasing reference to 'fake news', particularly around events such as Brexit (Britain's referendum on leaving the European Union) and the 2016 US Presidential Election, although there have been plenty of fake stories that cover many other subject areas. The current definition of fake news, which I'm taking from an article in *The Guardian* newspaper in December 2016, says:

> Fake news is a type of hoax or deliberate spread of misinformation, be it via the traditional news media or via social media with the intent to mislead in order to gain financially or politically.
>
> (Hunt, E., 'What is Fake News? How to spot it and what you can do to stop it', *The Guardian*, 17 December 2016; retrieved 15 January 2017)

There are two key points here: financial or political gain, neither of which are found by spreading stories surrounding mystery hitchhikers or ghostly apparitions, which sites such as the aforementioned Snopes take delight in debunking.

Financial gain

'For the love of money is a root of all kinds of evil' is the often misquoted sentence that can easily be turned into 'money is the root of all fake news'. As we've seen already, Google makes billions of dollars every year from advertising clicks; a company contracts with Google to pay the search engine a certain sum

of money every time a paid for advert is clicked on. If the link is actually on another website, but supplied by Google, the web author will get a proportion of the money, rather like a finder's fee. If lots of people visit a website, perhaps 1–2% will click on an advert, which admittedly isn't much, but 8000 visits may lead to 80 clicks, which soon adds up to a reasonable income. Consequently, if the link to a story is contentious or salacious enough, people will click to visit it, and the fake news provider is one step closer to making some money. These stories can be about, quite literally, anything at all. I've seen headlines saying that the British monarch Queen Elizabeth is dead, that a certain celebrity is moving to a town close to where I live, that certain health foods are actually cancerous (when they're obviously not) and so on. There is absolutely no truth in any of these stories, but people can't but help click on the link, look at the story, and then perhaps click on an advert before leaving the site.

Unscrupulous people will create websites that claim to have been in existence for years, or that they are the online version of a newspaper that has apparently been publishing for years (when in actual fact the newspaper doesn't even exist), and they'll create Facebook pages to bolster the site. They can then visit a celebrity Facebook page, post a story, link to their own Facebook page, then link back to their bogus website in the hope of tricking people into believing the story, visiting and clicking advertising links.

Malware attacks

Another reason for getting people to click onto websites by providing them with fake news headlines is to try and get them to install malware. Malware is short for 'malicious software' and it's designed to disrupt your computer, get sensitive information such as bank details, get access to private files and so on – rather different to a virus, which is more likely designed to replicate itself and destroy your software and erase your data. People can be tricked by a site asking them to view a video, but in order to do that it's necessary to install another piece of software (the malware) in order to view the salacious material.

Biased information

There are plenty of fake news stories that appeal to people who have particular political or religious viewpoints. The *New York Times* reported in November 2016 that Facebook, Twitter and other social media outlets have exposed millions of Americans to false stories asserting that Hillary Clinton's campaign

pollster, Joel Benenson, wrote a secret memo detailing plans to 'salvage' Hillary Clinton's candidacy by launching a radiological attack to halt voting, the Clinton campaign senior strategist John Podesta practised occult rituals, and that Mrs Clinton was paying public pollsters to skew results. There are many other examples, from all sides of the political spectrum, and just as many that attack different religious groups for various reasons.

It doesn't help that Google (for example) has linked to fake news sites. It was heavily criticised in the wake of the US Presidential election for linking to a site called 70news (https://70news.wordpress.com). This site headlined the news that Donald Trump had won both the popular and electoral college vote, when in actual fact he lost the popular vote by about 3 million votes. It's still not clear why Google linked to this site, but they did. Google responded by saying that they were going to fact-check and label sites with a 'Fact check' link to a reputable resource, but I have only ever seen this once or twice. Facebook has also stated that it is going to add warning labels to stories to say that they are 'disputed', but again, there's been little evidence of this actually happening.

In summary therefore, I don't believe that users can trust 'the news' that they find, and that they never should have done. A survey in 2015 by Ofcom (the UK Office of Communications) (https://nakedsecurity.sophos.com/2015/11/23/1-in-5-kids-believe-search-engine-results-are-always-true) found that only 60% of adults think that some websites will be accurate or unbiased and some won't be, 23% think that information returned by search engines is true and unbiased and 14% simply don't think about it.

Evaluating the news

When working with other people – and indeed when doing our own research – we really do need to stay on the ball. This is particularly the case when we find information that confirms our own bias towards something; we may well be far more inclined to believe it without question, since we live in our own 'filter bubbles', which are explained in more detail in a TED talk by Eli Pariser (www.ted.com/talks/eli_pariser_beware_online_filter_bubbles).

There is a simple, and easily remembered acronym that we can use when it comes to evaluating what we find, and it's CRAAP. Currency, Relevance, Authority, Accuracy and Purpose. I'll briefly go through each of these in turn, but an online search for the acronym will return many useful sites that are able to go into more detail than I have space for here.

Currency

It's not always clear when something has been published. The page may well have today's date on it, but that's simply due to a computer program designed to do exactly that. The actual article may well be weeks, months or even years old. This is particularly the case when you see news items on Facebook; they can be regurgitated stories that surface to corroborate some other event that's happening in the news. An easy way to check is to go to a search engine that allows you to search on, or rank results based on, time periods. Filter by material that's been added to the engine's database in the last 24 hours, or perhaps the last week. That should hopefully start to weed out more dated material. Run a search on the exact phrase that's used in the headline or the main body of the text to see if it's been used elsewhere and when. If there are any particular names or places mentioned in the article run separate searches for those as well, to see when and where they are mentioned.

Relevance

Is the information that you're finding relevant to the query at hand? Does the information make sense to you, or is it far too technical, using terminology that you don't understand? Does the information that you're looking at match information that you've found on other sites? Explore the website in some detail to see if it has a particular bias towards any one view, and look at the contributors to the site. Check their own background to see if they have published material elsewhere or, to be really cynical, do they actually exist, or are they completely fictitious?

Authority

This leads us neatly onto a proper investigation of the author of the piece or article that you're looking at. What are their credentials, and where do they work? Are they cited anywhere else? Take a look at their social media presence and see how many people they follow on Twitter, and perhaps more importantly, how many people follow them back. If you're really concerned, see if any of their followers are fake; sites such as Status People, at https://fakers.statuspeople.com, will let you run some free searches to see how many followers are fake. It's not a perfect tool by any means, but if any more than 10% of followers turn out to be fake it's worth proceeding with caution. Once again, check the website that is hosting the information, and take a look at the URL; is it one of the top-level domains, such as .ac, .gov or .edu? If it

isn't, dig more deeply into the site; the Wayback Machine, which I talk about in Chapter 12, will be helpful here.

Accuracy

Check some of the statements on the page to see if they are accurate by cross checking with other sites. Cut and paste some of the text to see if it has been copied wholesale from somewhere else. Are there any spelling errors or problems with punctuation or grammar? As a rule of thumb, the more errors you find, the less likely the information is to be accurate, although it's worth suspending some level of disbelief if the author's original language isn't English.

Purpose

Why has someone produced the information that you're looking at? Is the information there to entertain or amuse, in which case you may well be looking at one of the many satirical sites that exist and their purpose isn't to spread fake news. However, many people can get taken in by the information found in sites such as The Onion (www.theonion.com) and in good faith spread the information that they find far and wide – usually to their friends on Facebook, who send it onto their friends and so on. Does the article reference or present different viewpoints, and is there any obvious bias or prejudice, perhaps around the choice of words used to represent individuals, racial or religious groups? Can you find other information that hasn't been included which might weaken the case or claim of the author?

Check, check and check again!

One advantage of social media and user-generated content is that anyone can write anything that they want to quickly and easily. However, this is a major disadvantage as well. There are plenty of reasons why fake news is on the increase, and however much search engines or social media platforms try and do to stop this, they are doomed to failure; fake news will always get through. Trust no one, not even your friends (even if they come from an information background!), and keep checking your sources until you're absolutely convinced that the information you have got is sound.

Summary

Searching for news is already becoming one of the most interesting and fastest-moving areas of internet search – mainly because no one really knows what it means anymore! Consequently we're seeing the development of news search engines, the collation of content, the delivery of content and even the definition of that content changing continually. People are becoming journalists, so as well as reading the news we're more capable than ever of recording and publishing it, which brings the associated difficulties of authority and validity. With the ever-increasing use of mobile phones we will see news being brought to us all of the time, and it's going to be tailored down to where we physically might be located, with a search engine working in tandem with the GPS option on the smartphone. News really is all around us now.