

3-2016

How Readers Discover Content in Scholarly Publications

Tracy Gardner
Simon Inger Consulting Ltd

Simon Inger
Simon Inger Consulting Ltd

Follow this and additional works at: <http://digitalcommons.unl.edu/scholcom>



Part of the [Intellectual Property Law Commons](#), [Scholarly Communication Commons](#), and the [Scholarly Publishing Commons](#)

Gardner, Tracy and Inger, Simon, "How Readers Discover Content in Scholarly Publications" (2016). *Copyright, Fair Use, Scholarly Communication, etc.*. 13.

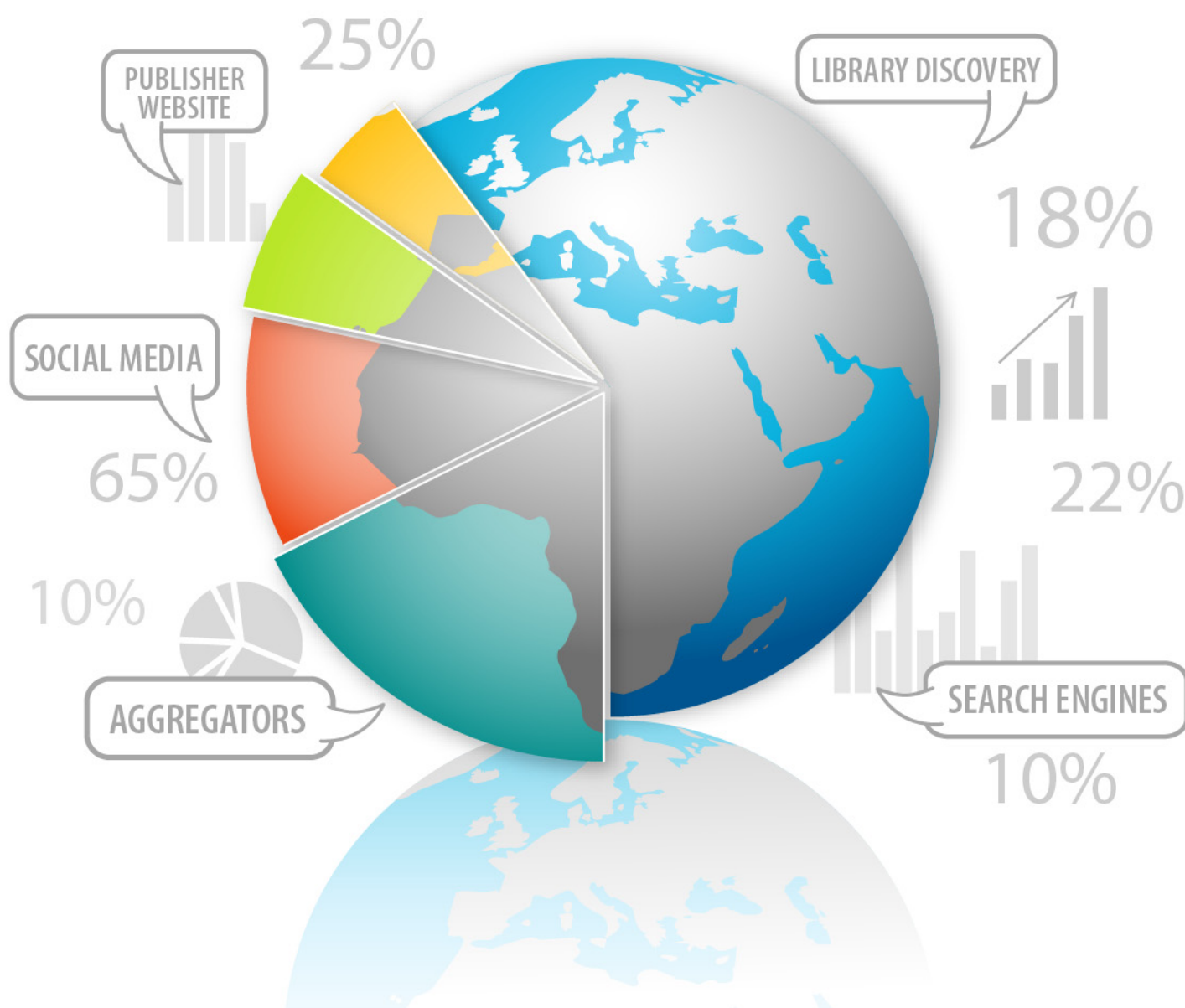
<http://digitalcommons.unl.edu/scholcom/13>

This Article is brought to you for free and open access by the Libraries at University of Nebraska-Lincoln at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Copyright, Fair Use, Scholarly Communication, etc. by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

[sic]

How Readers Discover Content in Scholarly Publications

Trends in reader behaviour from 2005 to 2015



By Tracy Gardner and Simon Inger

How Readers Discover Content in Scholarly Publications

By Tracy Gardner and Simon Inger

Published by Renew Training, a trading name of Simon Inger Consulting Ltd

Published March 2016

ISBN 978-0-9573920-4-5

Renew Training

Fernhill, Church Lane, Drayton

Abingdon, OX14 4JS, United Kingdom

The entire survey results data set upon which this report is based and the analytical framework developed to interpret these data are available for purchase at

<http://sic.pub/discover>

TABLE OF CONTENTS

1. Introduction.....	6
2. Headline Findings	9
3. Search	11
4. Other Methods of Discovery	36
5. Delivery Sites for Journal Articles	38
6. Mobile Devices and Apps.....	43
7. Journal Web Site Features.....	48
8. Methodology	54
9. Demographics	60
10. Conclusions	64
11. About the Authors.....	65

FIGURES

Figure 1 - Some of the paths in reader navigation and discovery.....	6
Figure 2 - What the publisher can potentially measure	7
Figure 3 - What the library can potentially measure	7
Figure 4 - Search, high-income areas, STM, 2005-2015 trend	11
Figure 5 - Search, academic sector life scientists, high-income 2012-2015	12
Figure 6 - Search, academic sector humanities, high-income 2012-2015.....	13
Figure 7 - Library search, academic sector by subject, 2012-2015	14
Figure 8 - A&I search, by subject, 2012-2015	15
Figure 9 - Social Media search, academic sector, by subject, 2012-2015	16
Figure 10 - Publisher website in search, academic, 2012-2015	17
Figure 11 - Search, academic researcher, 2012-2015.....	18
Figure 12 - Search, by lecturers, 2012-2015.....	19
Figure 13 - Search, variations by subject in high-income countries, 2015 – part a	19
Figure 14 - Search, variations by subject in high income countries, 2015 – part b	20
Figure 15 - Search by sector, 2015.....	21
Figure 16 - Search, by job role, high-income 2015.....	22
Figure 17 - Search, by region, 2015	23
Figure 18 - Search, by income classification, 2015.....	24
Figure 19 - Library discovery by subject, 2015	25
Figure 20 - A&I by subject, 2015.....	25
Figure 21 - Search correlation by subject, 2015.....	26
Figure 22 - Search engines, high-income countries, by sector, 2015.....	27
Figure 23 - Search engines, academic sector, by country, 2015	28

Figure 24 - Google vs Google Scholar by subject, 2015	29
Figure 25 - Search engines, China, 2012-2015	30
Figure 26 - Last article accessed, 2012-2015	31
Figure 27 - Last article accessed, chemistry, 2012-2015	31
Figure 28 - Last article accessed, by income, 2015	32
Figure 29 - Last article accessed, by subject, 2015	33
Figure 30 - Last article accessed, by sector, 2015	33
Figure 31 - Book search by sector, 2015	34
Figure 32 - Book search, sector, educator/teacher, 2015	35
Figure 33 - Last book accessed, sector	35
Figure 34 - Citation search, 2005-2015	36
Figure 35 - Browse, 2005-2015	37
Figure 36 - Delivery sites, academic by income, 2015	38
Figure 37 - Delivery, free resources vs publisher-controlled sites, 2015	39
Figure 38 - Delivery site, high income countries, by sector, 2015	40
Figure 39 - Delivery site, high income, STEM, 2015	41
Figure 40 - Delivery site, high income, STEM vs HSS, 2015	42
Figure 41 - Devices, low income, 2012-2015	43
Figure 42 - Devices by income, 2015	44
Figure 43 - Devices, medical sector, 2012-2015	44
Figure 44 - Devices, job role, 2015	45
Figure 45 - App use, 2012-2015	45
Figure 46 - App use, by income, 2015	46
Figure 47 - Devices, books, low income by sector, 2015	47
Figure 48 - Devices, books, job role, 2015	47
Figure 49 - Journal web site features, 2005-2015	48
Figure 50 - Journal web site features, academic, 2012-2015	50
Figure 51 - Features, income, academic, 2015	51
Figure 52 - Features, subject, 2015	53
Figure 53 - Screenshot of analytical tool	56
Figure 54 - Respondents by region	60
Figure 55 - Respondents by country of origin	60
Figure 56 - Respondents by sector	61
Figure 57 - Respondents by job role	61
Figure 58 - Respondents by subject area and region	62
Figure 59 - Respondents by World Bank income categories and subject (2011)	63

ACKNOWLEDGEMENTS

This work would not have been possible without the support of the following organisations, who between them executed hundreds of thousands of invitations to readers of scholarly and professional publications asking them to participate in this survey. Our heartfelt thanks go out to them all: Annual Reviews, American Theological Library Association (ATLA), The Bone & Joint Journal, Brill, Cambridge University Press, The IET, ingentaconnect, The JAMA Network/American Medical Association (AMA), OECD, SAGE, and Wiley.



BRILL



WILEY

1. INTRODUCTION

This report is the output of a large-scale survey of readers of scholarly publications (n=40439) and their behaviour in the discovery of journal articles and online books. The survey was conducted during October, November, and December of 2015. While usage statistics and analytics gathered by publishers, libraries and intermediaries can give us a partial view of discovery behaviour, there are many gaps in the knowledge that these can provide which we have endeavoured to fill by asking readers what tools they use in discovery.

This survey builds upon previous surveys conducted by the authors in 2005, 2008 and 2012. For four key questions in the survey, it allows for longitudinal analysis over the ten-year span, and for a further three questions allows for trend analyses between 2012 and 2015. The subtle shifts over time in reader preferences provide a valuable insight into reader navigation, the features that they find useful in publisher web sites, and the role and effectiveness of library technologies. For the first time, the 2015 survey includes three new questions regarding discovery of online books. Please refer to section 8 Methodology and section 9 Demographics, for a full discussion of the survey methodology and the demographics of those responding to the survey.

The discovery of journal content is certainly more refined than for online books, with a range of discovery methods available for most resources. Historically, journal articles tended to be available on a limited number of platforms, usually the publisher's official web site and any sanctioned aggregation of its content, such as EBSCO and ProQuest collections. However, further incarnations of articles are increasingly discoverable in institutional repositories, subject repositories (especially PubMedCentral), as well as a range of other sites including ResearchGate, Academia.edu, and Mendeley. This has further complicated discovery since (at the time of writing) no single discovery service indexes all of these incarnations, and no single discovery service seems to index at least one incarnation of all of the content available. This limitation is partly one of business relationships, partly a lack of understanding of metadata distribution, and sometimes political constraints.

The diagram below shows some of the paths open to a selection of reader types in discovering journal content online, and demonstrates the complexity of reader navigation.

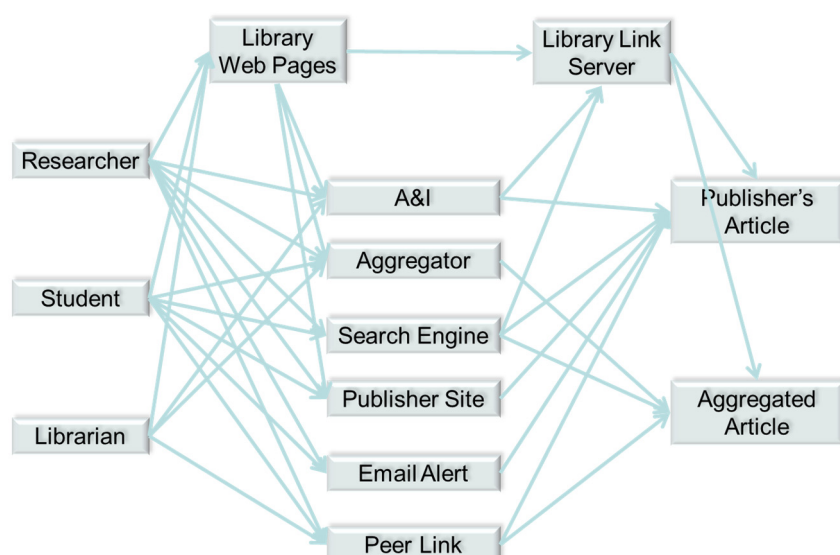


Figure 1 - Some of the paths in reader navigation and discovery

Libraries provide two layers of navigation, although these layers are increasingly indistinguishable to the reader. Library web pages are a discovery tool in their own right and range from simple catalogue listings of titles right through to advanced resource discovery solutions (RDS). In addition, though, there is the library link server (or link resolver), which is often configured to intermediate traffic from many other discovery resources and route the reader through to the most appropriate incarnation of the content (usually the subscribed-to

incarnation) for the reader. For more advanced libraries, this can be remarkably powerful, intermediating traffic not only from library web pages and abstracting & indexing services

(A&Is), but also from mainstream search engines, a feat sometimes achieved through careful programming of the institutional proxy or gateway. Google Scholar, as a specialist scholarly search engine, allows individuals to save institutional link server details as part of their preferences, thus revealing additional, library-intermediated pathways to content directly from the search results page (see configuration of library link server in Google Scholar preferences at https://scholar.google.co.uk/scholar_settings?sciifh=1&as_sdt=0,5#2).

However, library discovery, in fact most discovery channels, have failed to keep pace with the rise of mobile devices in journal discovery, especially smartphones. Publishers continue to invest in mobile delivery (and fully responsive web sites) but the discovery layers are failing to keep pace. This may drive more readers to use publisher web sites directly for discovery, something librarians generally discourage, as by definition this restricts the breadth of search undertaken.

Of the great variety of pathways (examples of which are illustrated in Figure 1 above) that readers utilise to discover content, usage statistics and web analytics can reveal some of this navigation. From the publisher's point of view, it will know how much use was made of its own content by each institution, but most likely not know the reader demographic within that institution e.g. job role. The publisher will also know something of which of the discovery platforms the reader has arrived from, and also the institutional breakdown of usage of any aggregated databases. Currently, the publisher uses this limited knowledge to help show value to its clients and also to inform the design of its web site. Given the needs of its paying clients (the libraries), publishers need to know more about this navigational behaviour.



The library knows more of the individual's identity, and has the potential to know a lot about discovery platforms as long as the reader navigates to the content via a link resolver, but the library doesn't know about the complexities of navigation for those who operate outside of the library-intermediated environment; however, the library will get usage data from publishers, but this is not married up with a reader profile. There is an increasing need for libraries to be able to demonstrate value of e-resources. In its simplest form this is usage, but more advanced analysis relies on understanding which job roles were responsible for each type of usage, and in the final analysis, libraries need to be able to show evidence of a positive outcome in return for acquisition of information products.

Figure 2 - What the publisher can potentially measure

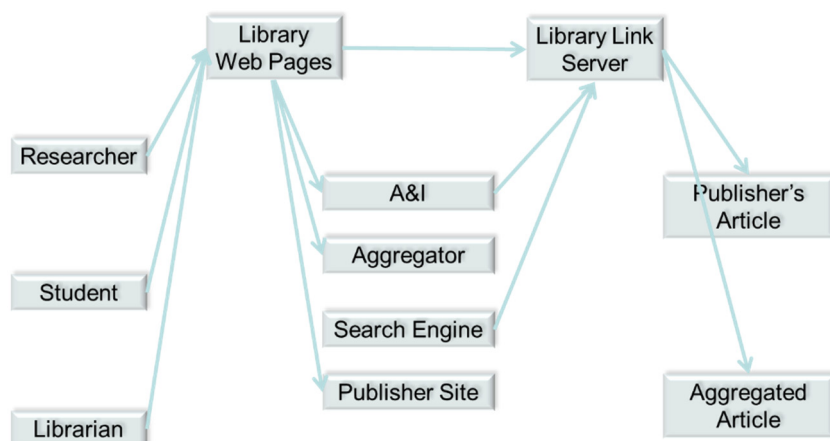


Figure 3 - What the library can potentially measure

This research aims to fill some of these knowledge gaps by surveying what readers believe their actions in discovery to be.

Our previous reports showed that readers are more likely to arrive within a journal web site at the article or abstract level than anywhere else and since then publishers have responded by changing their web sites so that more of the features and functions are visible from that landing page.

Online book discovery is far less mature than that of journals. This is partly due to the fact that most online books are available in multiple silos, each with their own specific terms of use, rights management, and interface. Each silo comes with its own search engine, and the metadata that would allow for external discovery is not extensively shared with discovery partners. Most online book platforms seem to be seeking to own the discovery as well as the delivery which is in stark contrast to how journal publishers have grown to behave, where maximised external discovery is seen as key to their success.

Another factor may simply be that there is less money to be made in books publishing than journals publishing, and by extension in books discovery versus journals discovery. Major academic libraries spend much more money on journals than books, and have only recently started to focus on significantly improving the findability of books in their online book collections.

2. HEADLINE FINDINGS

- As a starting point for search, A&Is seem to be in a slight decline when looked at in aggregate across all regions and sectors, but remain the most important. Figure 4, p11
- Academic researchers in high income countries now rate library discovery as highly as A&Is, and rate academic search engines as the most important discovery resource when searching for journal articles. Figure 11, p18
- Library discovery services have made significant advances in importance in search for academic researchers, and for all roles in hard sciences in the academic sector. As an average across all subjects and sectors, however, they have not grown in importance in since 2012. Figure 4, p11; Figure 7, p14; Figure 11, p18
- More than half of all journal content delivery appears to be from free incarnations of articles. There appears to be a clear PubMedCentral effect in the medical sector. Social media sites appear to be a significant source of free articles in lower income countries. Figure 37, p39
- In academic STM in higher income countries, academic search engines are now more important than general search engines. Figure 4, p11
- Table of Contents alerts have reduced in popularity in all measures across the survey. Figure 26, p31; Figure 35, p37
- There appears to be an increased role for social media in discovery. Figure 9, p16
- Online book discovery varies significantly by sector, with academics preferring library web pages marginally over general web search engines, the medical sector preferring A&I services and library web over search engines, but all other sectors preferring search engines over other forms of discovery. Figure 31, p34
- Publisher web sites are becoming more popular as a search resource, although this is less true for people in wealthier countries. Figure 10, p17; Figure 18, p24
- Google Scholar is used more than Google in the academic sector, but less than Google in all other sectors. Figure 22, p27
- A perceived lack of awareness of Google Scholar in poorer nations appears to be leading to a reduced use of free incarnations of content in institutional repositories from these regions. Page 40
- Readers in low income countries use their mobiles to access journals more than their counterparts in richer countries. However, access by phone still accounts for only about 10% of the use. Figure 42, p44
- A&Is continue to be the most important search method in the medical sector. Figure 15, p21
- The primary method of journals discovery is search, but even more so for online books. Figure 33, p35
- App use for journal discovery is still low. Figure 45, p45
- The most highly sought-after features of journal web sites are changing. Figure 49, p48

- Access to supporting data from a publisher website is more important to people in high income countries than people in lower income countries. Figure 50, p50
- Across all demographics there is no significant appreciation of the availability of social media sharing or article-level metrics, even though most publishers feel that these are essential features. Figure 52, p53
- Librarians behave quite differently to everyone else in search, preferring professional search databases and library-acquired resources. Figure 16, p22

3. SEARCH

One of the big questions it is very hard to answer with just analytics is where do people start their search when they are looking for journal articles – analytics will only tell you which resource people came to your site from and not where they started on their journey. Only by understanding where people start will publishers be able to ensure their content distribution priorities are in order, and will librarians be able to ensure they are helping their patrons on their journey of discovery. There are many variances in the search behaviour of people by sector, region, job type and subject discipline and to date the only practical way to get a true understanding of their chosen discovery resources is to ask them. Given that we have been asking this question of researchers, students, information professionals, lecturers and others over the last ten years, we are able to show how search behaviour has changed over time.

It is important to note that the results from 2012 and 2015 have been normalised to match the demographic sample of 2005 and 2008 (see section on Methodology). Therefore, these 10-year results show the responses predominantly from researchers working within STM subjects, in academia, in the US and Europe.

TREND ANALYSIS: JOURNAL ARTICLE SEARCH

TREND FROM 2005 - 2015

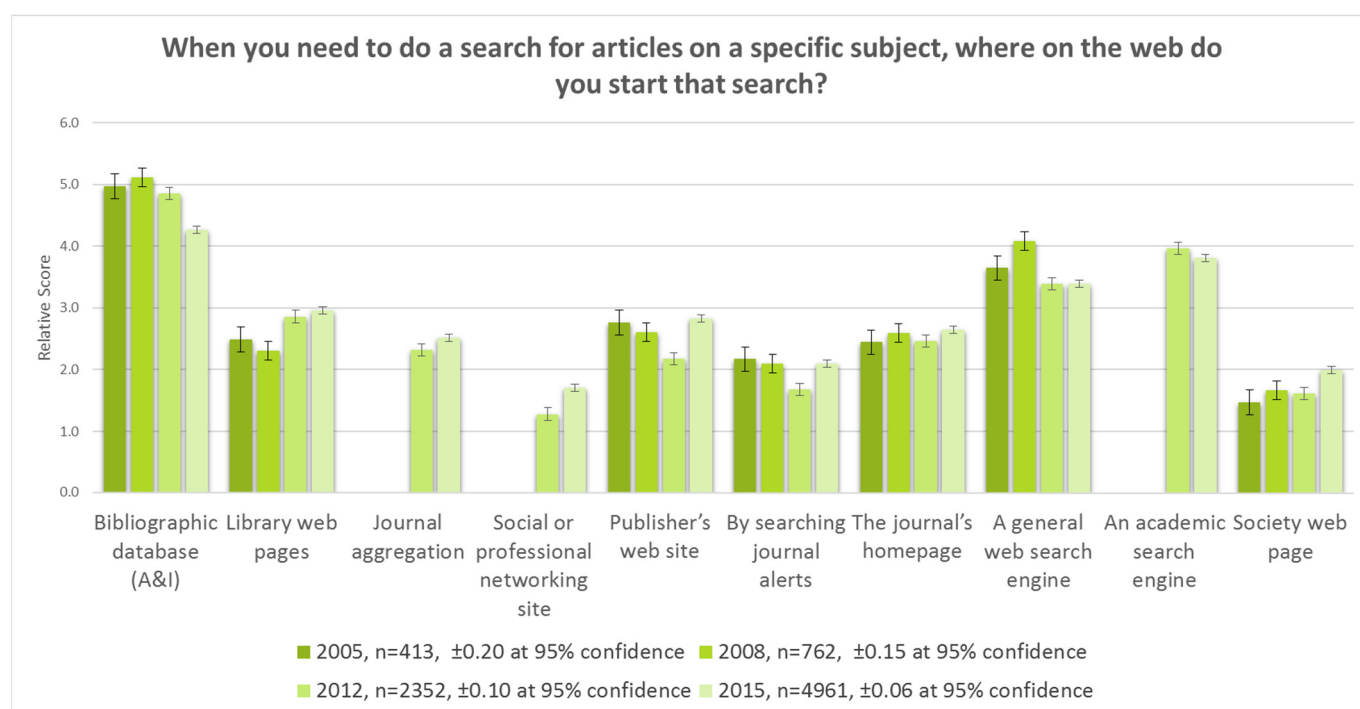


Figure 4 - Search, high-income areas, STM, 2005-2015 trend

A key result is that whilst A&Is are still marginally the most important search resource, their importance has consistently dropped since 2008. General search engines have lost some ground here to academic search engines.

Library discovery became more important for search in 2012 and has stayed steady since then. This may be seen as a disappointing result for libraries, who continue to invest in their resource discovery technologies.

There has been growth in the importance of journal aggregations and social media for search.

All search resources that are under publisher control – publisher website, journal alerts, journal homepage and society webpage - have made gains. This may go contrary to what we believe to be logical behaviour, since no one publisher site provides an index to all the relevant content in a subject area. Librarians themselves avoid this behaviour, see Figure 16 page 22, and presumably advise their patrons against it too. However, these results seem to show that more people are willing to use search on a publisher site and then presumably repeat that on another publisher site, and again for a number of relevant publishers.



A&Is are showing some decline in importance, but still appear to be the most important starting point. Academic search engines are more important than general search engines. Library discovery services have not grown in importance since 2012. Publisher controlled search resources are less important overall, but are growing.

In the past, academic researchers working in Life Sciences have told us that A&Is, particularly PubMed, are their key discovery resource type. We wanted to see if this had changed since 2012. Figure 5 shows the results for search behaviour for life scientists in academia working in high income countries. We used a sample from high income countries only to reduce the chance of funding limitations biasing the results.

TREND BY SUBJECT, 2012 - 2015

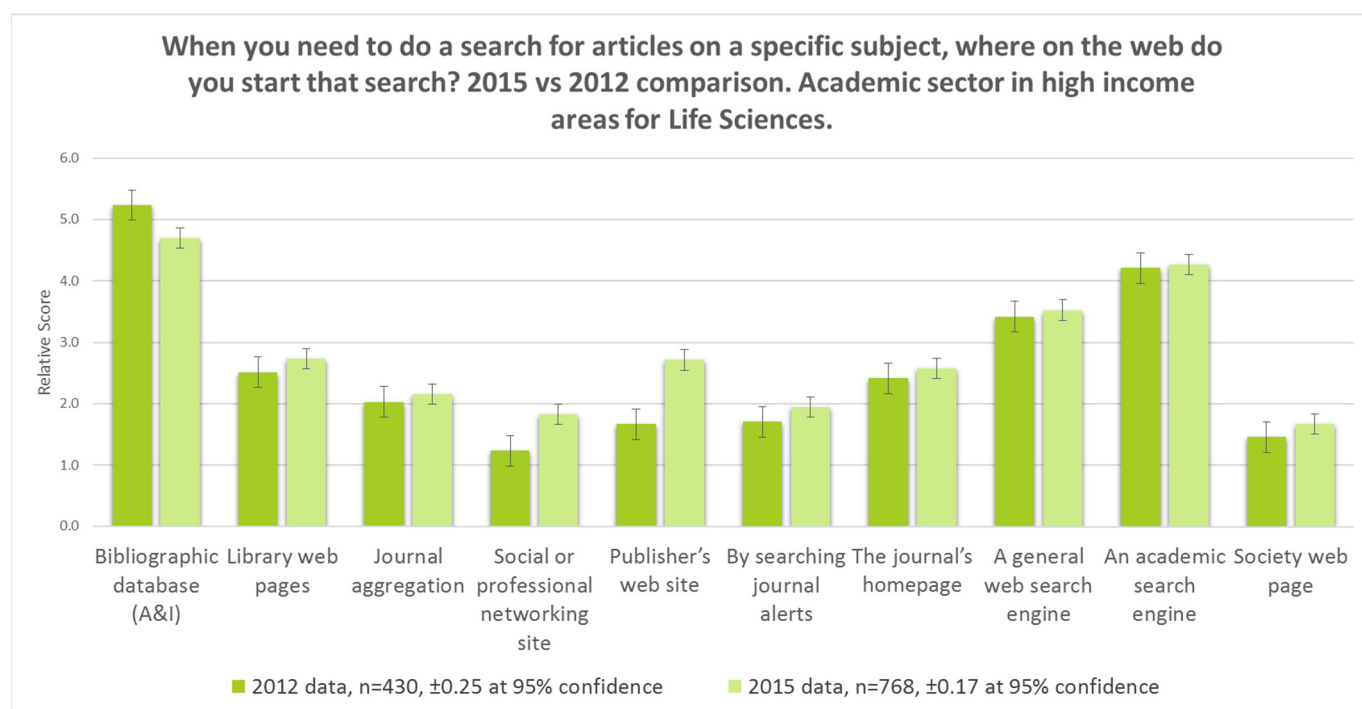


Figure 5 - Search, academic sector life scientists, high-income 2012-2015

A&Is show a slight decline, but remain the most important resource. There have been significant gains in popularity of social and professional networking sites and the publisher's web site. A&Is have possibly lost out to social media and the publisher websites. Whilst the chart above only illustrates the results from high-income countries, we have found that the same picture emerges regardless of income classification.

If we look at exactly the same comparison in Humanities, we see quite a different picture. We can look in more detail at the significant difference in the search behaviour of people working in Life Sciences and Humanities in a moment but if we first look at the changes since

2012, illustrated in Figure 6, we can see that the only significant change is that social media has grown as a search resource for people working in Humanities, although overall it is still less important than everything else. More interestingly perhaps for this group is that journal aggregations have become less important and are at the same level as library web pages and academic search engines.

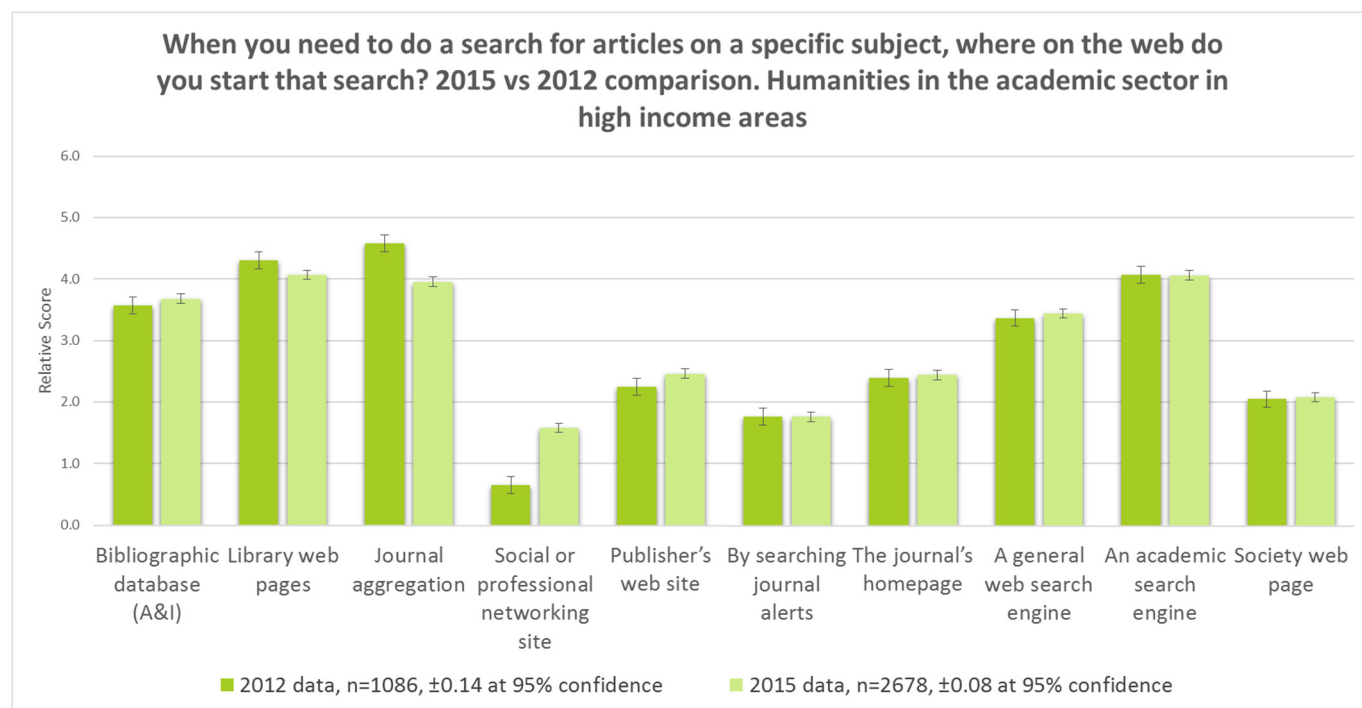


Figure 6 - Search, academic sector humanities, high-income 2012-2015

There has been much debate in recent years about the importance of library discovery technology. In Figure 7, below, we look at the importance of library discovery over time in the academic sector by subject. This uses the subject classifications deployed in 2012, not the larger, expanded set from 2015.

By finding, filtering, shaping, curating, certifying, editing, promoting, disseminating and rewarding ideas, a scholarly publisher helps convert them into cultural products that enable their transmission from the minds of writers and editors into the minds of readers.

Ziyad Marar, Global Publishing Director



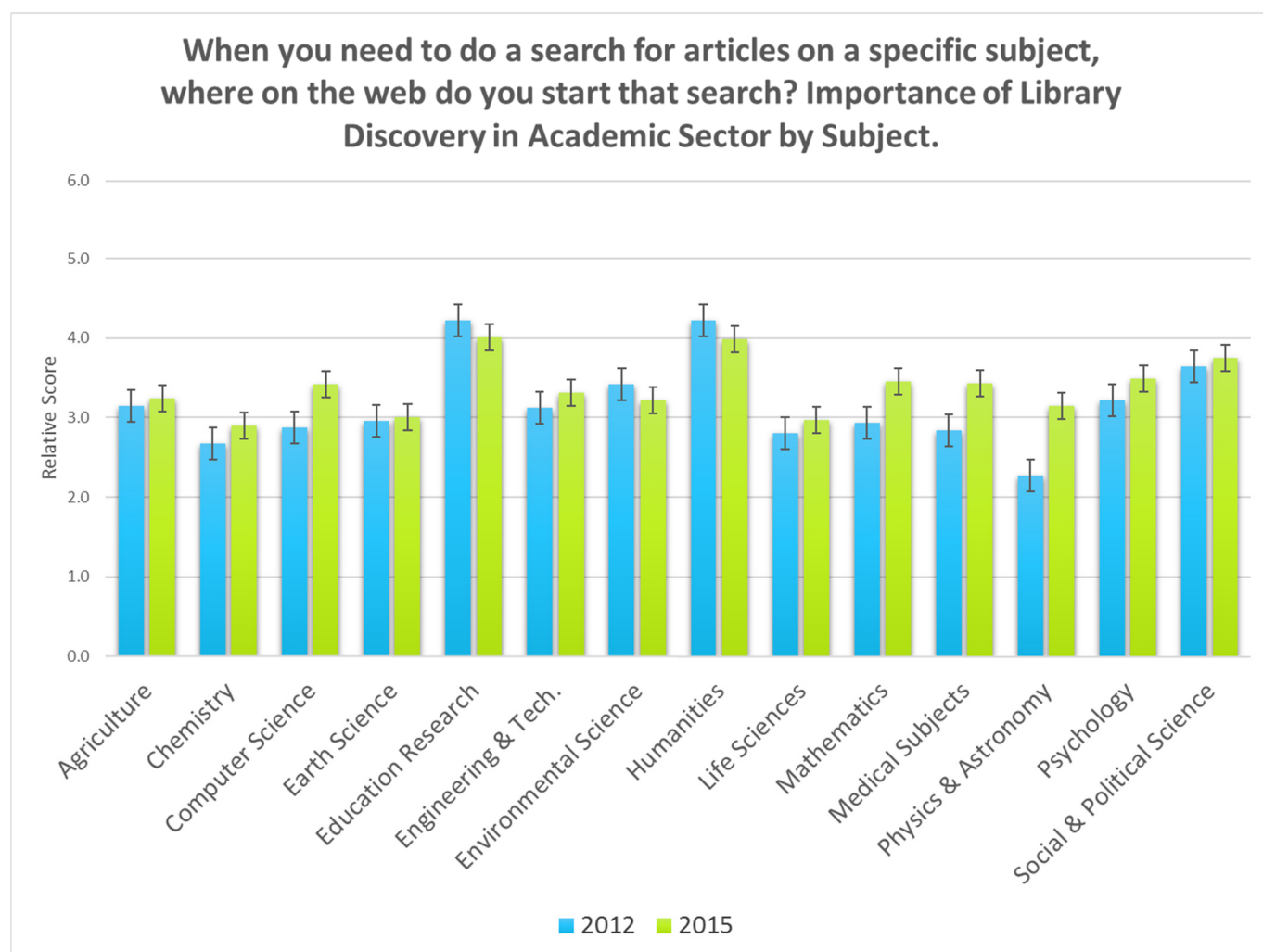


Figure 7 - Library search, academic sector by subject, 2012-2015

This shows an indicative decline in the use of library discovery for both Education Research and Humanities, perhaps indicating early-adopting subject areas have subsequently re-evaluated its utility. However, these are still the subject areas where library discovery is strongest. Maths, Computer Science, Medical Subjects and Physics, have all shown significant rises in the importance of this search resource.



Library discovery has shown some significant gains in rating in some subject areas, whilst perhaps declining in those areas that were early adopters. These gains could be because library discovery has got better at indexing certain subjects, or perhaps librarians have been more successful at marketing their resources.

If we look at the importance of A&Is over time by subject in the academic sector we see that there has been a slight indicative drop in their importance to Life Sciences and Medicine, however, they are still very important to people working in these areas. A&Is have become significantly less important in Earth Sciences and Engineering. There has been a growth in importance of A&Is in Physics and an indicatively positive change in Computer Science.

When you need to do a search for articles on a specific subject, where on the web do you start that search? Importance of A&Is in Academic Sector by Subject

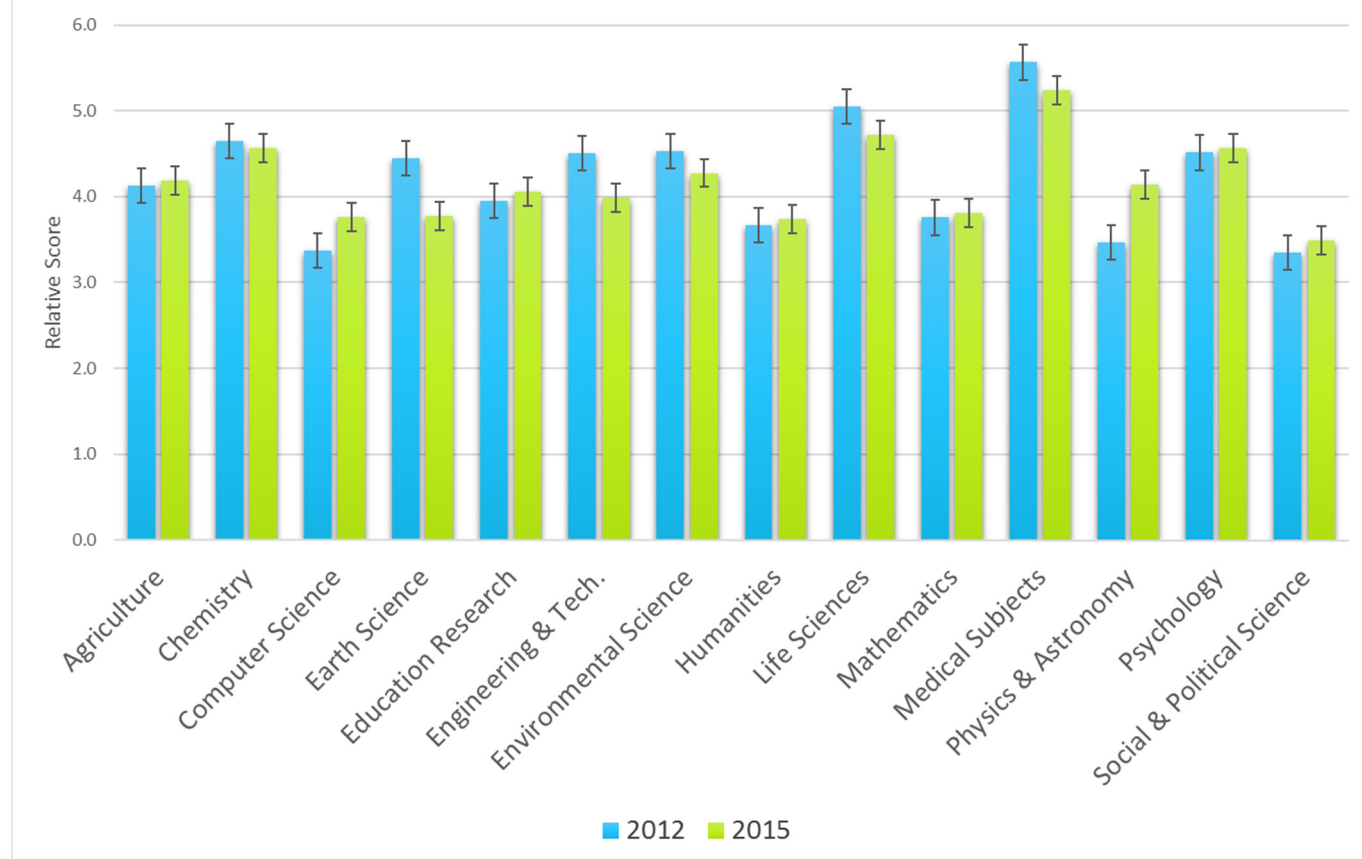


Figure 8 - A&I search, by subject, 2012-2015

If we look at one of the less important discovery routes, albeit one that many people are discussing, we can see that Social Media (and in our question we specifically included Facebook, Twitter, Mendeley and ResearchGate as examples) has become significantly more important in all subject areas.

ANNUAL REVIEWS CONNECT WITH OUR EXPERTS

SPARK YOUR RESEARCH CONNECTION

Annual Reviews publishes 46 relevant, impactful, and highly cited review journals in the Biomedical, Life, Physical and Social Sciences. Researchers have relied on Annual Reviews journals since 1932 to intelligently synthesize the overwhelming volume of scientific literature and deliver the ideas and citations that will advance research further, faster.

Our invited experts cut out the noise and provide reliable entry points into the scholarship of a given discipline. In today's environment of rapid information proliferation, review articles are more useful than ever before and save scholars valuable research time.

Visit www.annualreviews.org for:

- Timely; Relevant Reviews
- Expert Figures and Illustrations
- Content Alerts
- Advanced Search



ANNUAL REVIEWS

T: 650.493.4400/800.523.8635 (US/CAN) | E: service@annualreviews.org

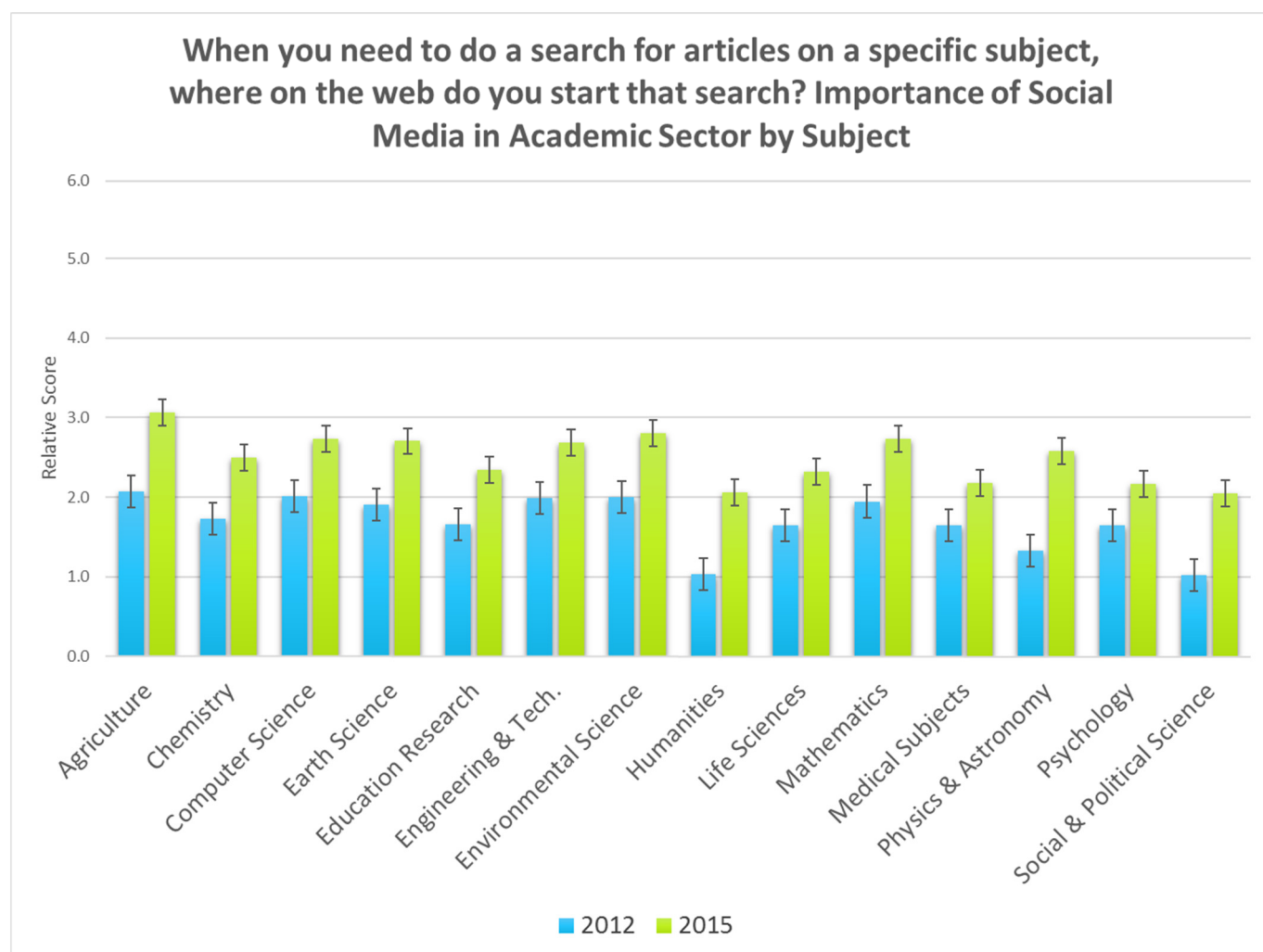


Figure 9 - Social Media search, academic sector, by subject, 2012-2015



In our 2012 report, we noted the particularly low traction of social media in Humanities and Social Science, and it is in these two areas where we have seen some of the greatest increases, along with Physics and Agriculture. It is in Medical Subjects and Psychology where the growth is smallest.

Perhaps the most significant jump in importance across all subject areas is that of the publisher website which has grown most significantly in Agriculture, Earth Science, Engineering, Environmental Sciences, Life Sciences, Medicine, and Psychology.

When you need to do a search for articles on a specific subject, where on the web do you start that search? Importance of the Publisher Website in Academic Sector by Subject

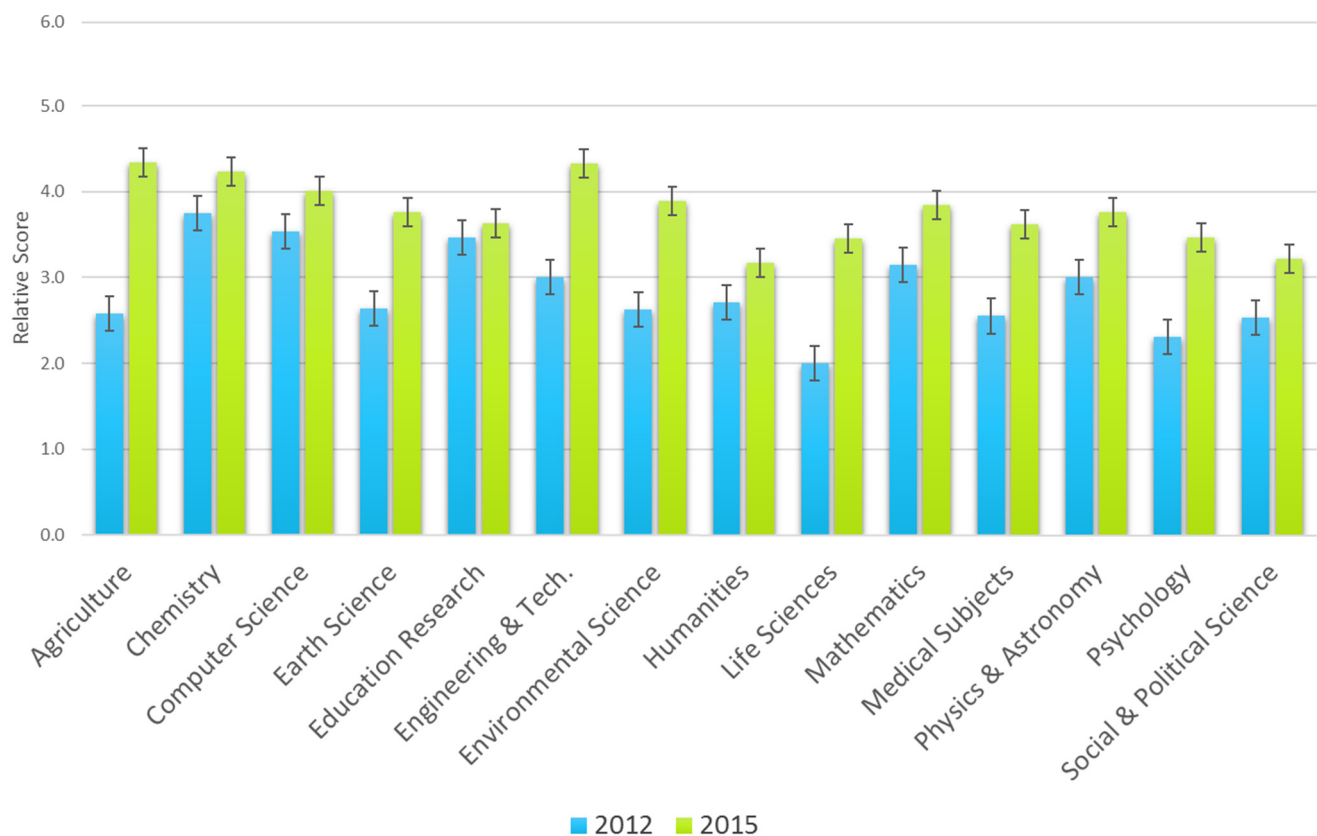


Figure 10 - Publisher website in search, academic, 2012-2015



This growth in importance of publisher web sites in search could be because publisher websites have improved, because publisher marketing has improved, the relationship with the end user has improved (including engagement through social media) or because people have become frustrated with other discovery resources.



World-Class Digital Engineering Research

- 450+ DRM free eBooks
- 18,000+ research articles
- 15 Million + A&I abstracts
- 7,000+ video presentations
- 70,000+ archive articles
- 1,400+ conference proceedings



To request a free trial or for more information, please contact sales@theiet.org

TREND BY JOB ROLE, 2012 - 2015

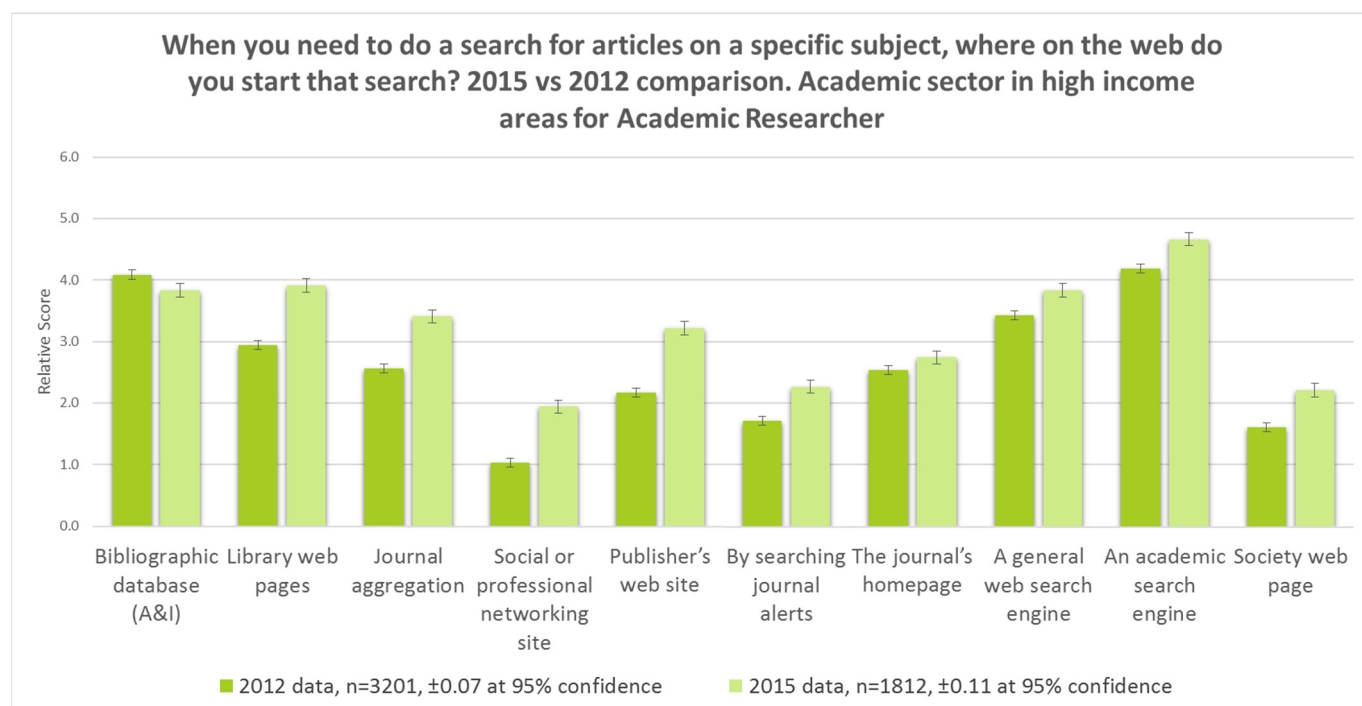


Figure 11 - Search, academic researcher, 2012-2015

Figure 11 shows that whilst A&Is are still important to academic researchers, all other discovery resources have seen growth, indicating this group is branching out across alternative search resources. Academic search engines have now overtaken A&Is, and library web pages are now as important as A&Is to this group which shows a significant trend upwards in the importance of library technology. Is this because libraries are doing a much better job of encouraging their more senior patrons to use the library website? Likewise, journal aggregations, publisher websites and social media sites have grown significantly in importance for academic researchers – perhaps this group have become more aware of resources in general over the last 3 years.



Academic researchers now rate library discovery as highly as A&Is, and rate academic search engines as the most important discovery resource when searching for journal articles.

This picture changes dramatically for lecturers, in Figure 12. For this group, journal aggregations and library web pages have become less important and social media has become more so. Academic search engines and A&Is remain the most important. Although a much smaller sample, we have also found indications in the data that students behave analogously to lecturers.

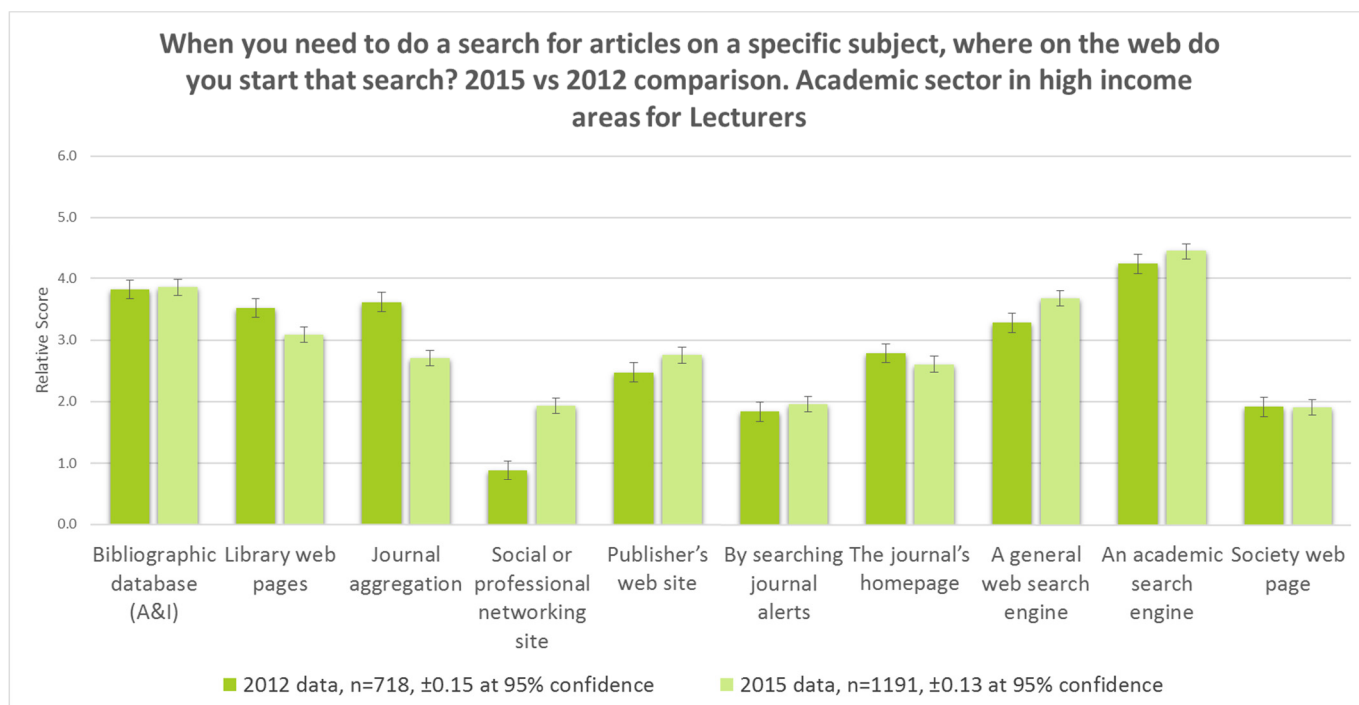


Figure 12 - Search, by lecturers, 2012-2015

JOURNAL ARTICLE SEARCH, 2015

SEARCH BY SUBJECT

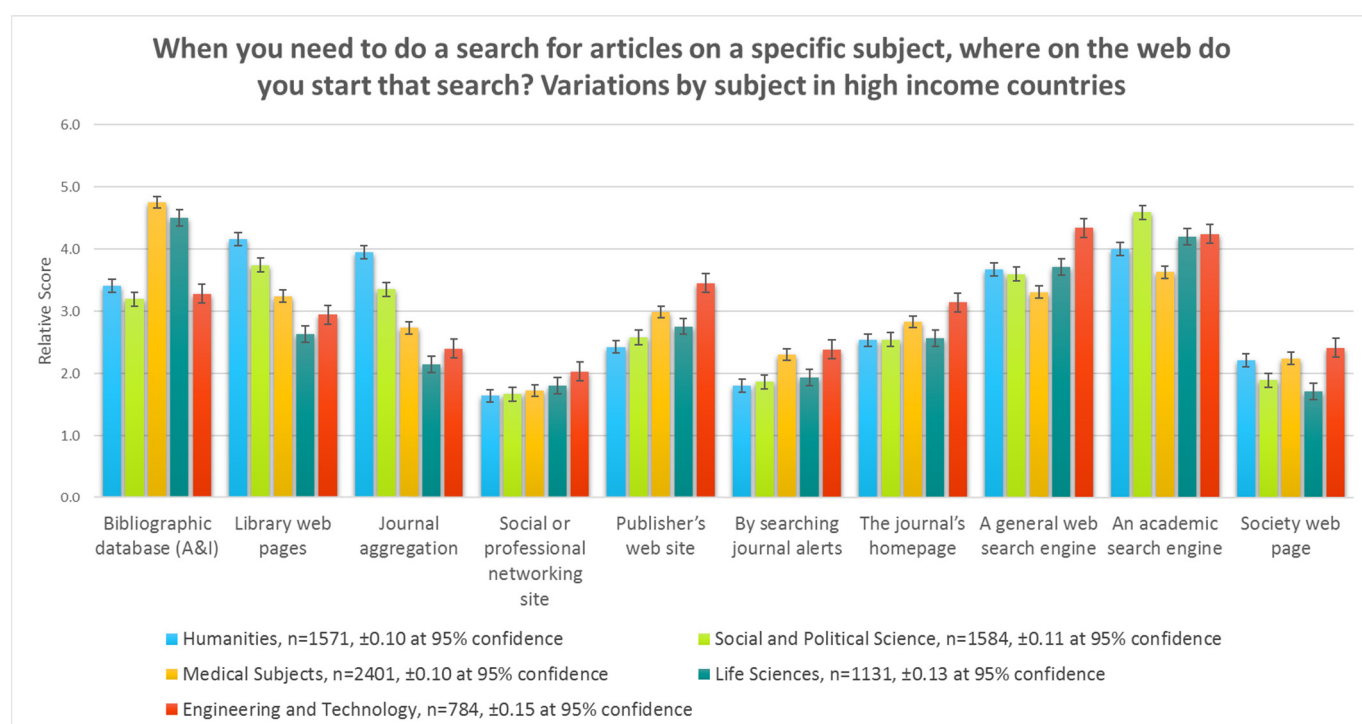


Figure 13 - Search, variations by subject in high-income countries, 2015 – part a

Figure 13 shows that there are significant variations in the search behaviour of people by subject area. A&Is are by far the most important resource for people working and studying within Medical Subjects. This is true to a slightly lesser extent for people working and studying in Life Sciences (and we have already discussed the slight decline in importance of A&Is in Life Sciences over time on page 12).

The results by subject for library web pages and journal aggregation map each other closely. People working in most subjects say that academic search engines such as Google Scholar are more important than general search engines for their subjects. People working in Engineering & Technology rate general web search engines, such as Google, as slightly more important. For people working in Social and Political Science, academic search engines are the most important resource by some margin. It should be noted that in this question we asked the respondents to state how important the resource is to them. In a later question (see Figure 24, page 29) we ask respondents about what proportion of their time is spent on each search engine type. Most of the results are consistent, but some provide anomalies, notably in Religion.

The other significant result is that the publisher website is as important, if not slightly more important, than A&Is for people working in Engineering and Technology.

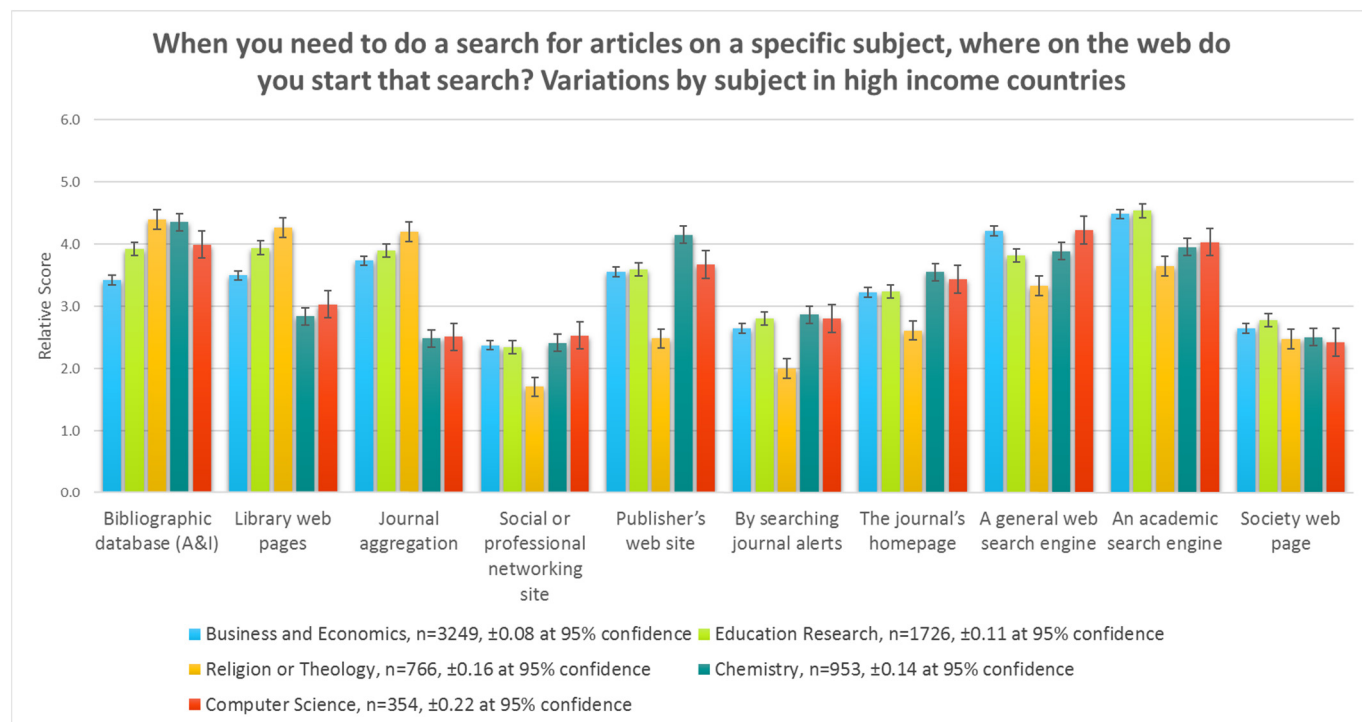


Figure 14 - Search, variations by subject in high income countries, 2015 – part b

A&Is, library discovery and journal aggregations are of nearly equal importance to people working in Religion & Theology. All other discovery resources are far less important to them, even eclipsing academic search engines.

People working in Chemistry rated A&Is as the most important discovery resource, closely followed by the publisher website. Their use of general search engines and academic search engines is comparable. Library web pages and journal aggregations are far less important to people in Chemistry and Computer Science. People working in Business and Economics favour search engines significantly more than all other resources.

SEARCH BY SECTOR

Figure 15 shows us the sectoral differences in search behaviour. The most significant difference in behaviour is in the medical sector where search is dominated by the A&I – library web pages, the publisher website, and search engines are all runners up with a similar level of importance to each other.

People working in the corporate sector rely on free search engines more than anything else, and rate general search engines higher than academic search engines. We see the reverse for people working in academia, and for all other sectors their use of general and academic search engines is comparable.

People in the corporate sector have slightly less use for library web pages.

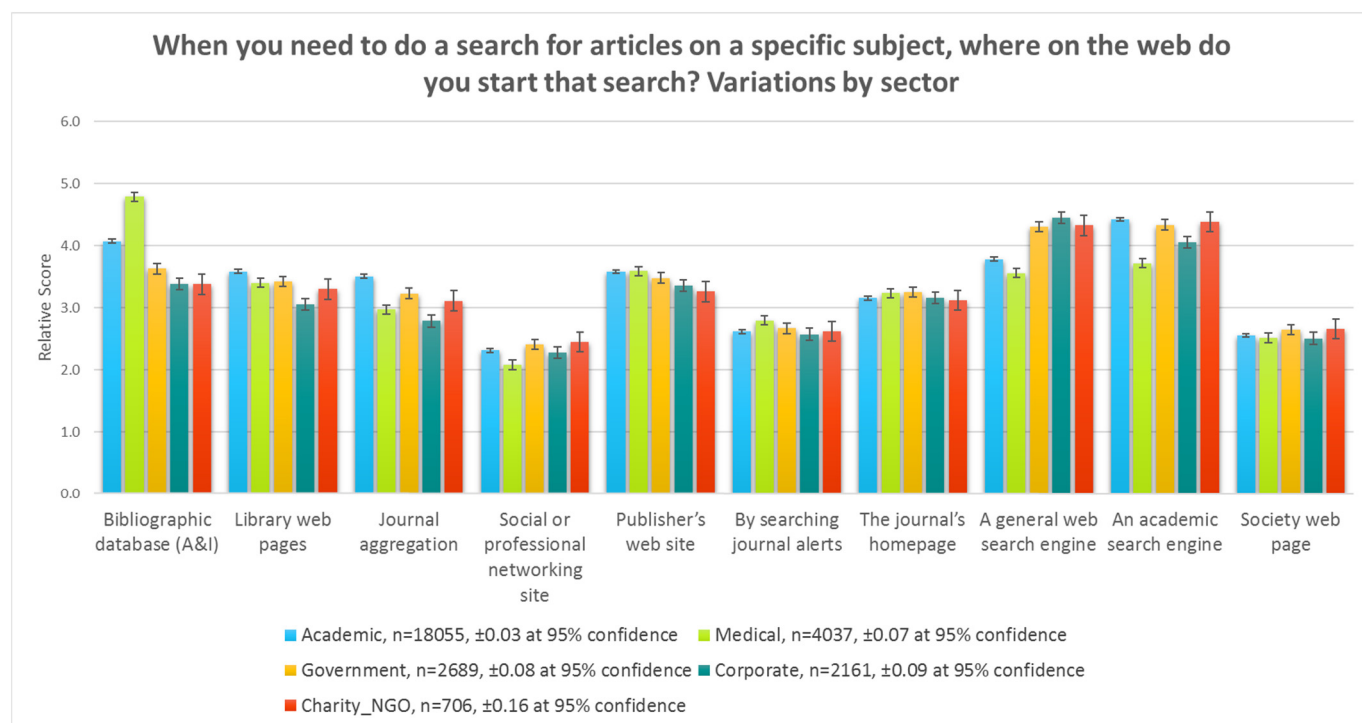


Figure 15 - Search by sector, 2015

Celebrating
70
Years

Established in 1946, the American Theological Library Association (ATLA) is a professional association of over 800 individual, institutional, and affiliate members providing programs, products and services for theological and religious studies libraries and librarians. ATLA offers a prestigious product line of electronic resources to support the scholarly study of religion and theology, including the *ATLA Religion Database®* (ATLA RDB®) and *ATLASerials®* (ATLAS®).

Proud sponsor of
*How Readers Discover Content in
Scholarly Publications*

ATLA
www.atla.com



A&Is are still significantly more important in the medical sector than for all other sectors, and is significantly more important than all other resources.

SEARCH BY JOB ROLE

Figure 16 shows how search behaviour differs by job role across all sectors and regions – the differences are significant.

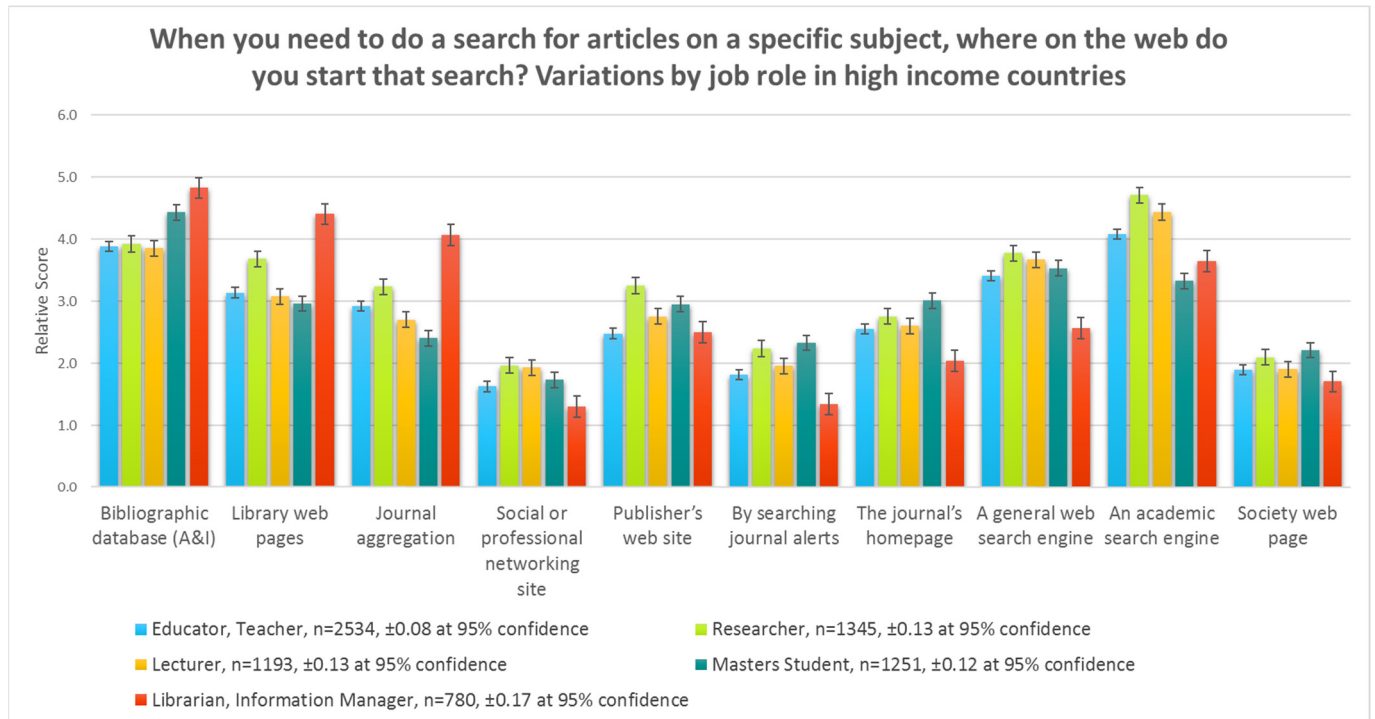


Figure 16 - Search, by job role, high-income 2015



Librarians behave quite differently to everyone else in search, preferring professional search databases and library-acquired resources. This may point to a significant gap between what librarians recommend, and how their patrons behave.

Perhaps as expected, lecturers and teachers behave in a similar manner. Somewhat surprisingly, masters students use academic search engines less than everyone else and they favour the A&I, even more so than researchers (all sectors). When we repeat this comparison, looking only at researchers from the academic sector, it still holds true.

“I maintain a series of Google Scholar Alerts to stay informed. I frequently use library staff for a particular literature search. There is not yet a single destination that works adequately for interdisciplinary content search. Google Scholar, which I use most, is not that great.”

“For subject searches, I would advise any user to take full advantage of the bells & whistles available within each specialized database, esp looking at the expertise of the user who needs the info. We do have faculty who tell students to "go to JSTOR" and find articles, which is really limiting and inappropriate for many searches, esp. where currency is important. Regular Google is okay for some searches but brings back too much add'l. garbage; Google Scholar is better in part because the library's links have been added (and, often, even work, as do the direct publisher links that have been added, making the direct linking easier.”

“I use google images to find articles”

“If I have a process I don't know what it is. I just try everything I can think of!”

SEARCH BY REGION

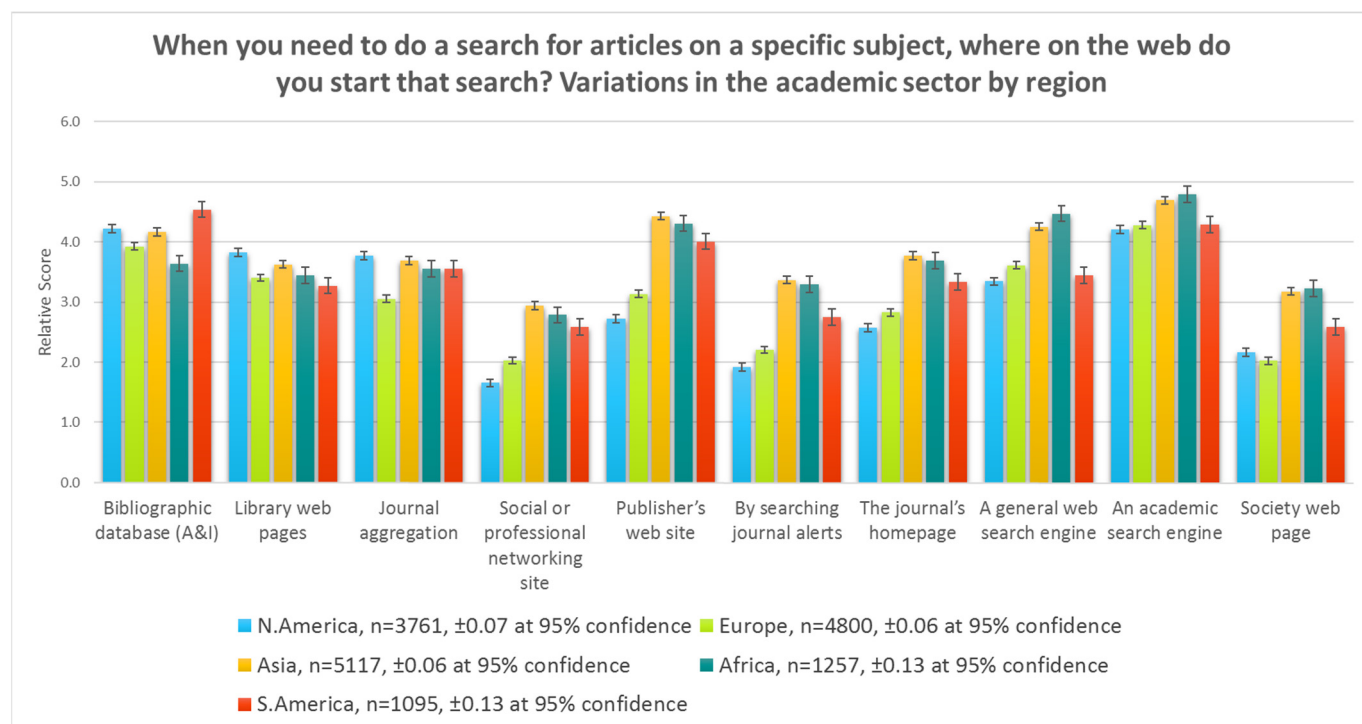


Figure 17 - Search, by region, 2015

There are significant regional differences in the importance of search resources. People in Asia, Africa and South America think publisher websites have a similar level of importance to search engines and we can see that people in Europe and North America think they are far less important. People in Africa and Asia still appear to value ToC alerts as a search resource.

A&Is are the most important search resource to people in South America.

Journal aggregations are far less important in Europe than they are in North America and social media is less important in North America and Europe than any other region. This may be due to the amount of free material available in some of the social media sites.

The home of **scholarly research**

ingenta**connect** >



More than
5 million articles
 From over **11,000**
 publications

Start your search at: **ingentaconnect.com** @ingentaconnect | help@ingentaconnect.com

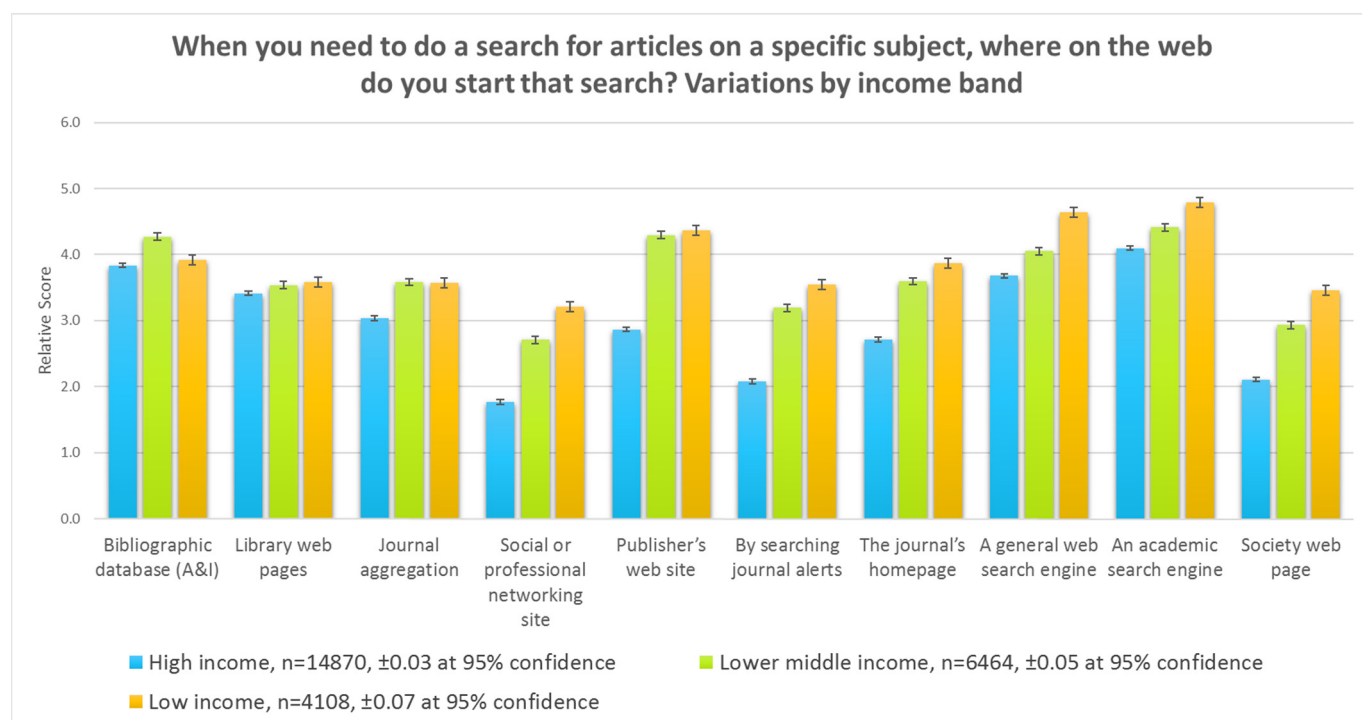


Figure 18 - Search, by income classification, 2015

A&Is are slightly more important to people in lower/middle income countries. Poorer countries are more reliant on free search resources. There is a significant difference in the ranking of social media sites as a search resource between high and low income countries.



Publisher website, journal alerts, society web pages and journal alerts (i.e., all resources under publisher control) are far less important to people in wealthier countries

IMPORTANCE OF DISCOVERY RESOURCES IN SEARCH BY SUBJECT AREA

Figure 19, below, shows us that people in HSS (Humanities and Social Sciences) subjects still favour library discovery over their STEM (Scientific, Technical, Engineering and Medical) counterparts.

“Sometimes a library discovery tool is okay, but not for serious research. The search results aren't accurate enough and the tool often doesn't respond well to specific queries.”

“Library and information service in my employing organisation undertakes literature searches on my behalf”

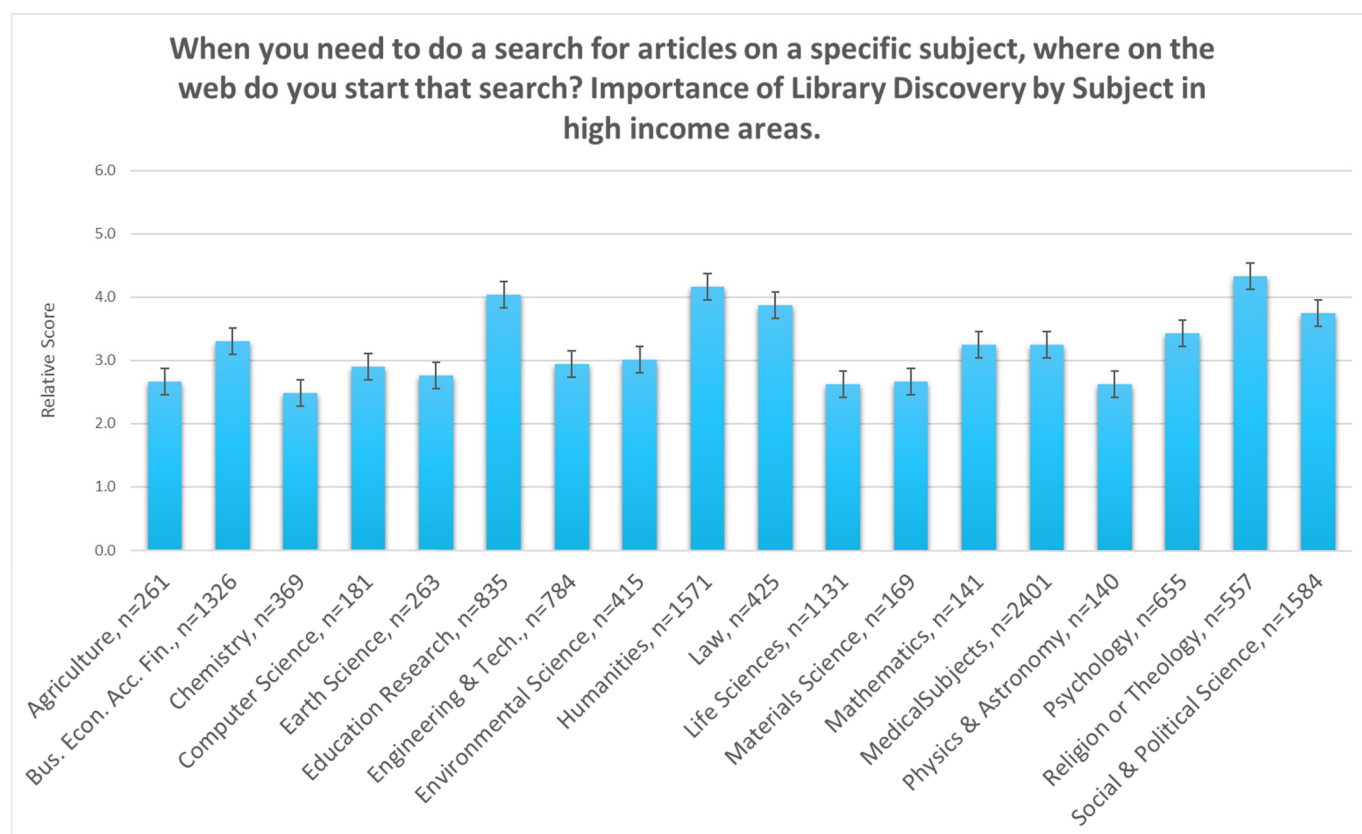


Figure 19 - Library discovery by subject, 2015

For A&Is the picture is quite different – A&Is are still important in many subjects, particularly in Medicine, Life Sciences, Psychology, Chemistry, and Religion.

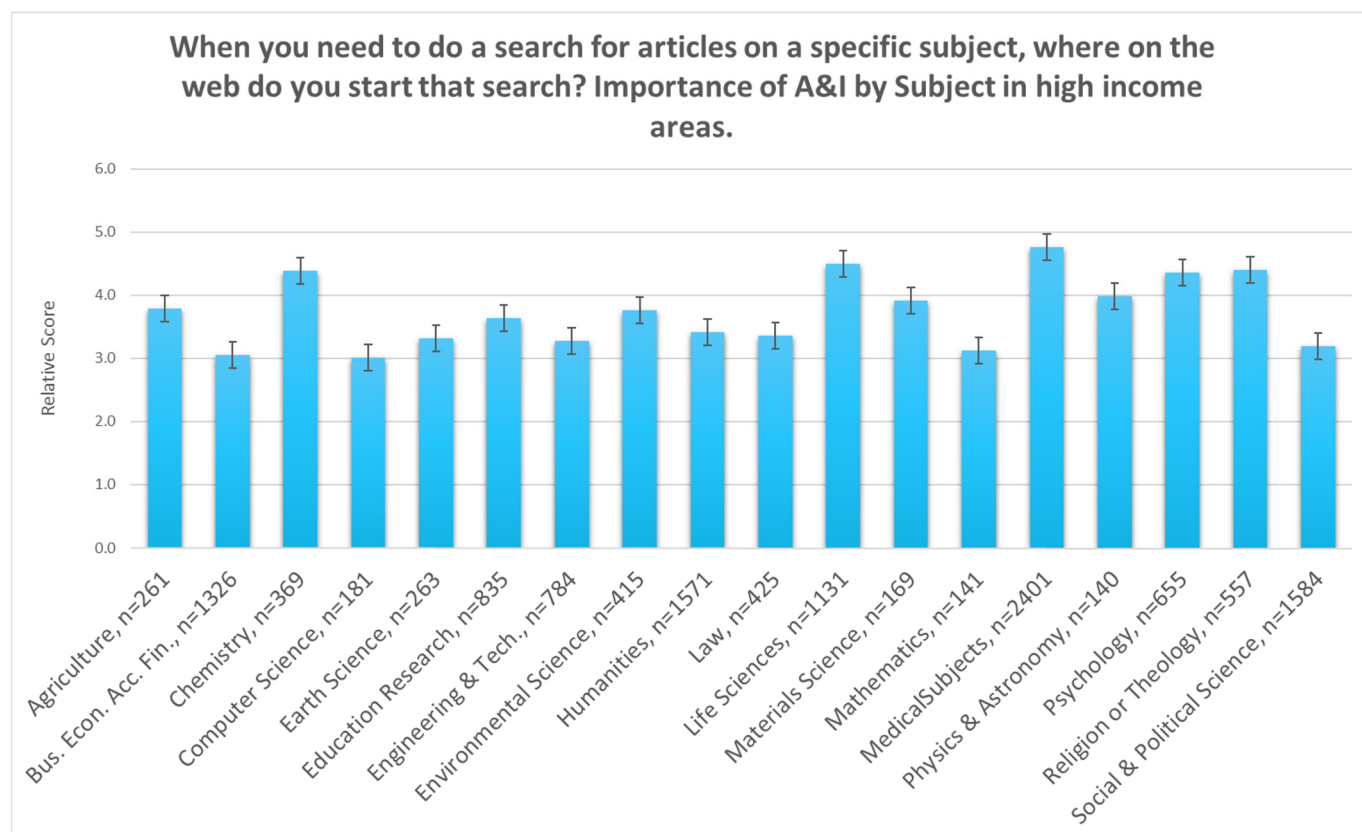


Figure 20 - A&I by subject, 2015

We did the same analysis for Social Media (which included Mendeley and ResearchGate as examples in the questions) and there were some differences between HSS and STEM – people studying and working in STEM subjects (apart from Physics) were slightly more inclined to use social media for search than people working and studying in HSS subjects.

CORRELATION IN SEARCH BEHAVIOUR

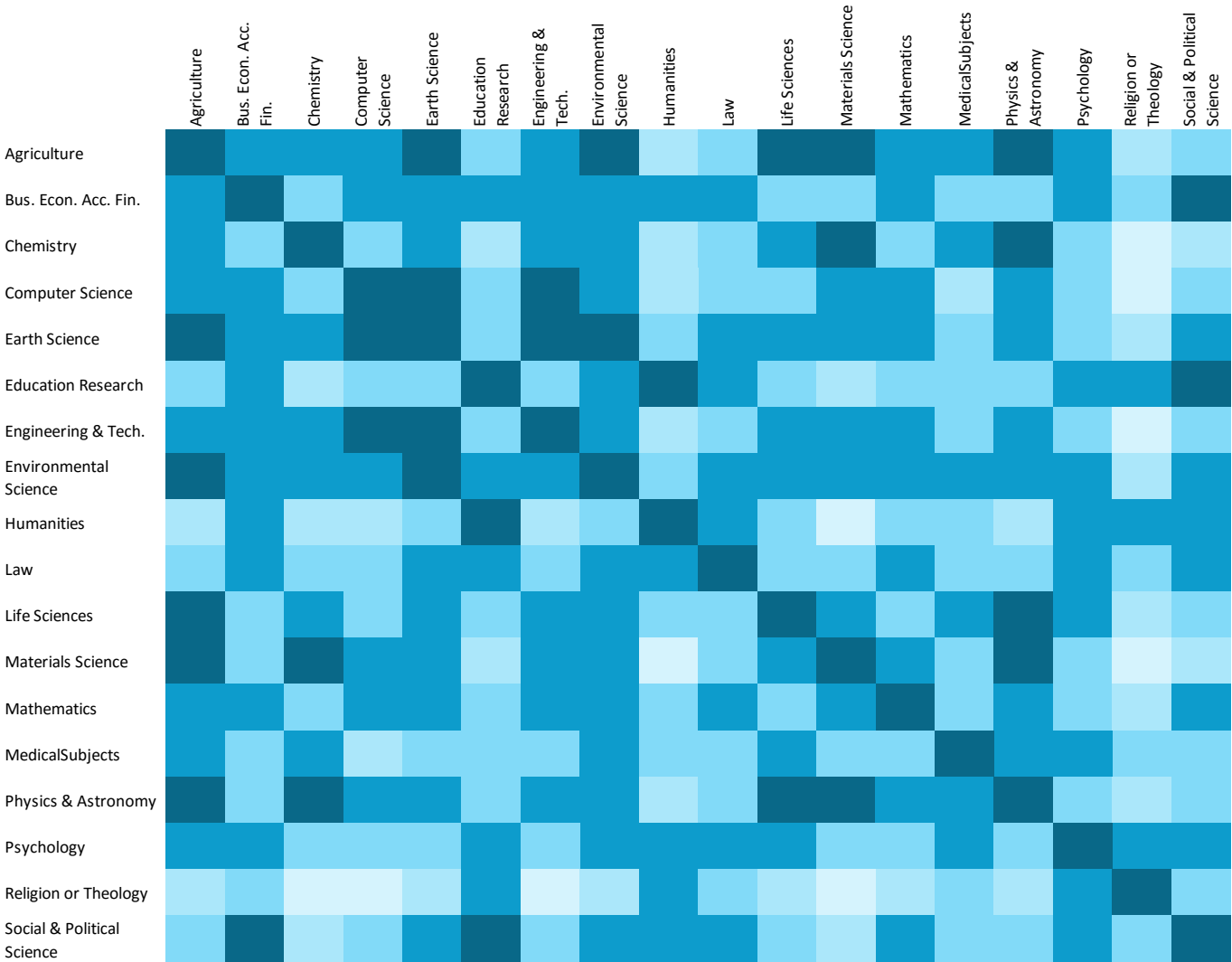


Figure 21 - Search correlation by subject, 2015

Correlation Key		
Low	High	Key
0.95	1.00	
0.80	0.95	
0.60	0.80	
0.40	0.60	
0.20	0.40	
0.00	0.20	

The heat map above shows the correlation of search behaviour by subject, created by looking at search behaviour for each subject and each category. Some correlations are somewhat obvious, such as Education Research and Humanities, and Chemistry and Physics but others are more surprising. For example, the search behaviour of people in Psychology is closely correlated with many other areas such as Religion, Agriculture and Social Science. People working in Agriculture have similar behaviour to people working in Physics and Astronomy.

SEARCH ENGINE PREFERENCE

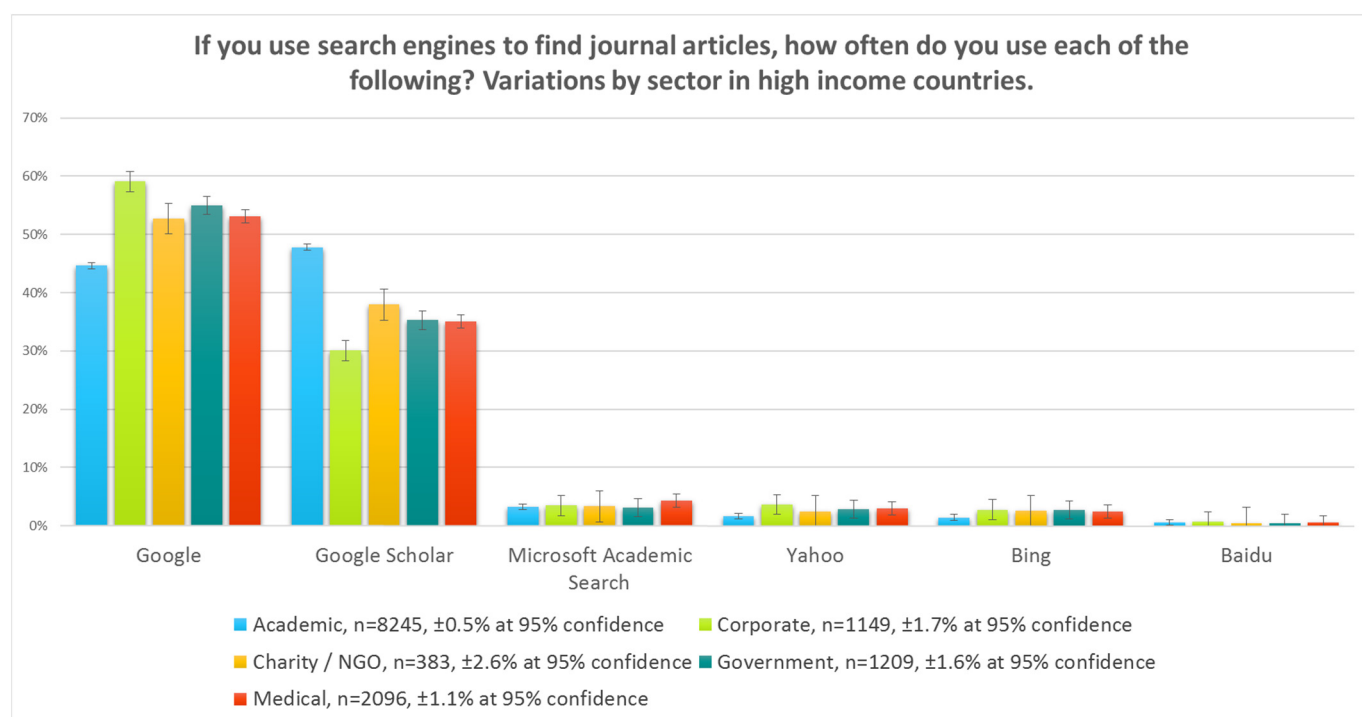


Figure 22 - Search engines, high-income countries, by sector, 2015



Figure 22 shows us that it is only in the academic sector where journal readers use Google Scholar more than they do Google. It should be noted that in previous questions we asked respondents to indicate how important each starting point was to them, whereas here we have asked them about the frequency of use. As a consequence, the relative differences between Google and Google Scholar may be slightly different to the earlier findings, although the earlier questions will have been effectively a sum of the academic versus the non-academic search engines above. The corporate sector makes the least use of Google Scholar, and this may be an awareness issue.

OECDiLibrary
www.oecd-ilibrary.org
 Explore over 200,000 books, papers and datasets, covering 17 different themes

OECD

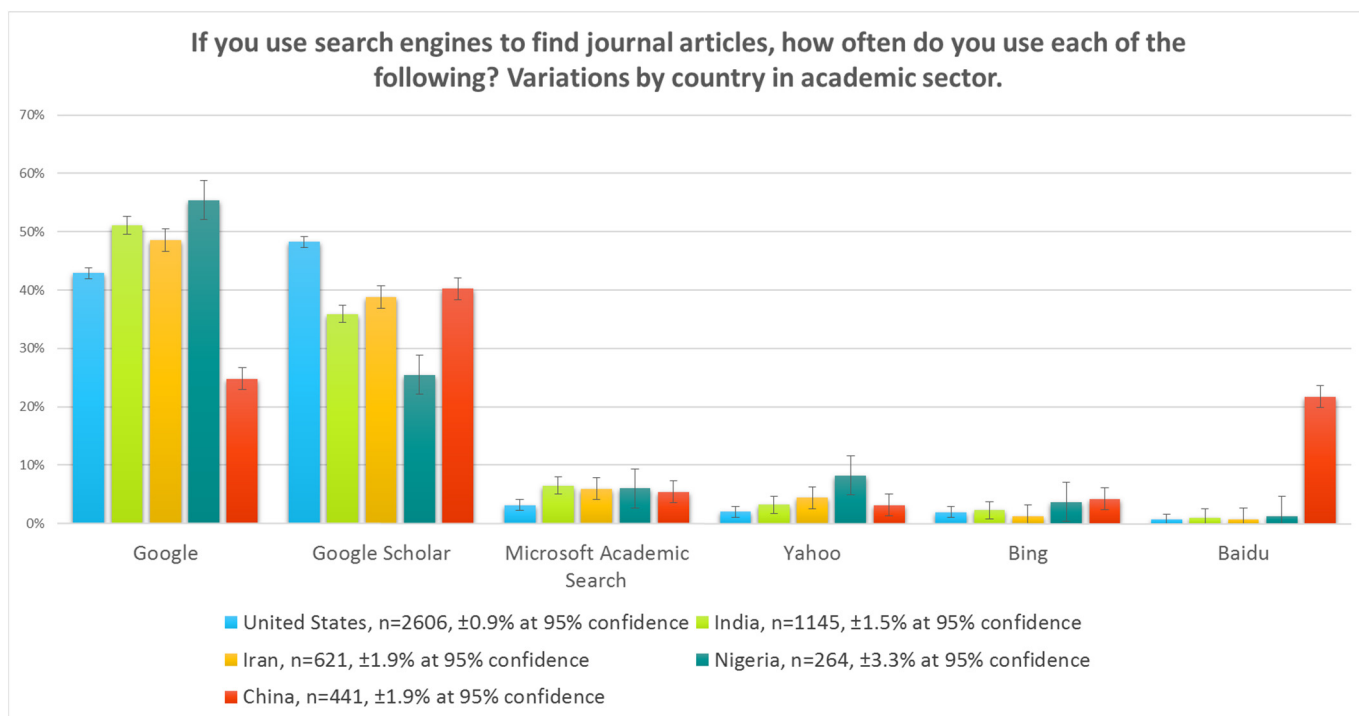


Figure 23 - Search engines, academic sector, by country, 2015

There are significant regional variations in the adoption of Google Scholar over Google within all sectors. Figure 23 shows the differences for the academic sector only. Google Scholar is used more than Google in the USA, a behaviour that we found mirrored in a large number of countries, such as United Kingdom, Netherlands, Germany, Brazil. The use of Google Scholar seems to be reduced in many African and Asian countries. We presume this is a problem of awareness, since it cannot be a problem of price. Use of Google in China is known to be restricted to a degree, and it is not surprising that Baidu has as strong a presence as Google in China.

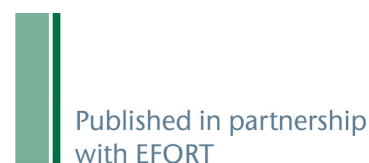


Google Scholar is the dominant search engine used for journal discovery in China, even though there is an indicative decrease in its use since 2012.

Further analysis of the relative use of Google Scholar versus Google in the academic sector in high-income countries is shown in Figure 24.



www.boneandjoint.org.uk



www.efort.org/openreviews

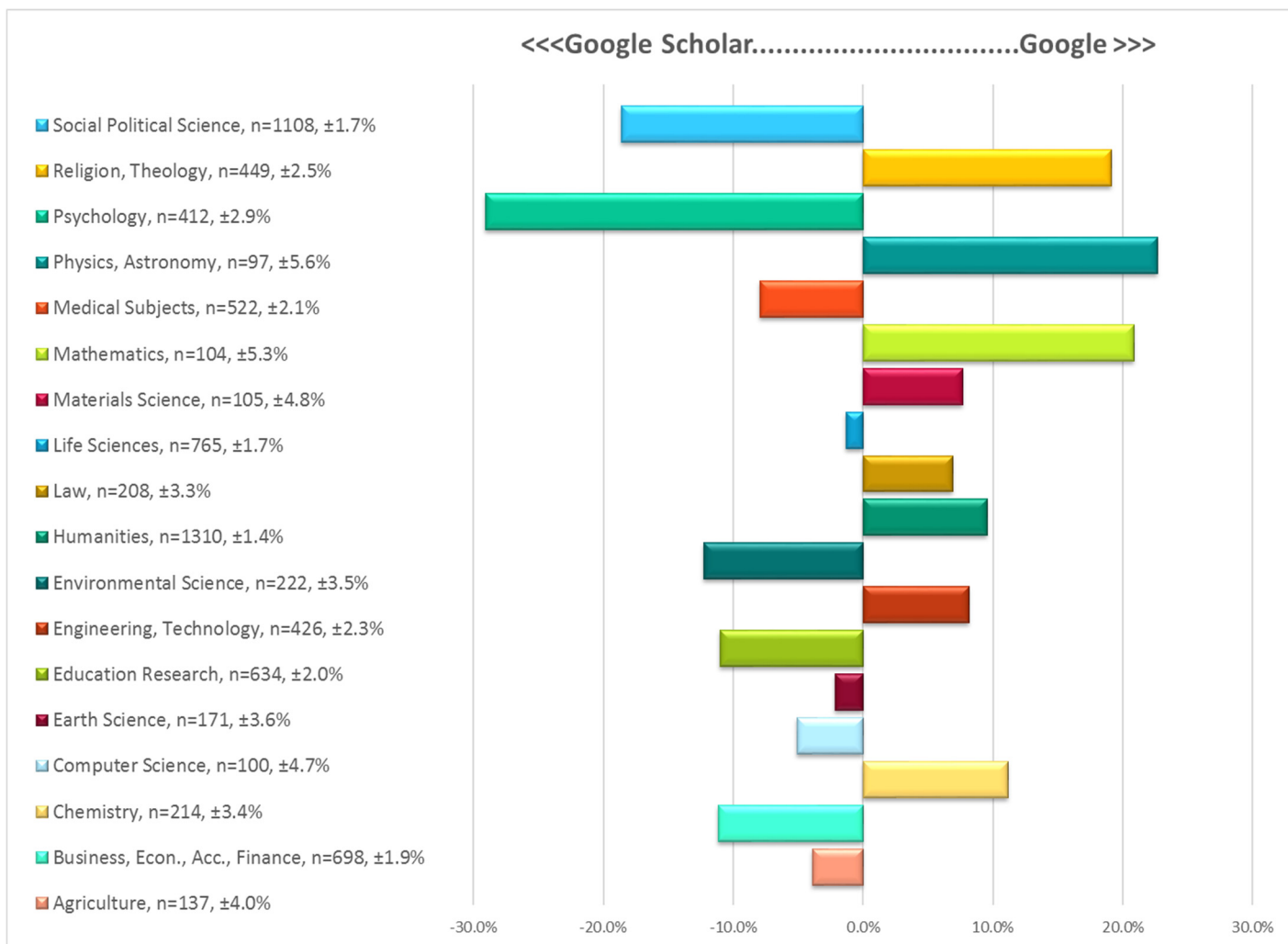
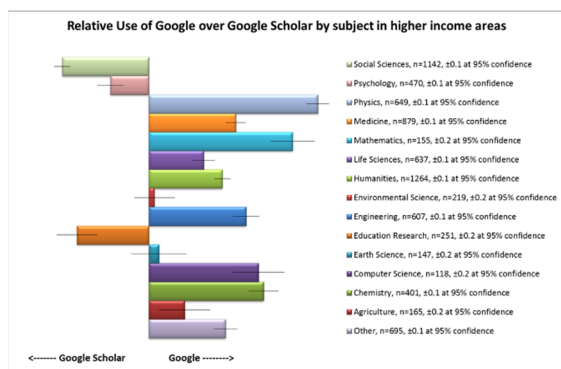


Figure 24 - Google vs Google Scholar by subject, 2015

There is quite a large variance by subject area in people's preference for Google vs Google Scholar. People in Social Sciences, Education, Law, and Business use Google Scholar more to find journal articles. However, people working in Humanities and Religion & Theology prefer to use Google – this is surprising as you might assume that there is more similarity in the behaviour of people working and studying in HSS fields. This echoes the similarities shown in the correlation heat map in Figure 21.

People working in Psychology have a strong preference for Google Scholar, people working in Physics and Mathematics have a strong preference for Google.



TREND FROM 2012 - 2015

A brief visual inspection of the thumbnail of the 2012 results shown (left) clearly shows the shift from 2012 to 2015 as nearly all the data points have moved starkly in the favour of Google Scholar.

During the period since the 2012 study, Google changed the way it indexed scholarly articles behind a paywall, making it much harder for individuals to see search results in many journals, especially the ones to which

they do not subscribe. It is not unreasonable, therefore, to expect a shift to Google Scholar which does not have similar restrictions on what it will display to the user.



There is an increasing reliance on Google Scholar by people working and studying in the academic sector. This is largely due to people leaving Google more than it is to people joining Google Scholar.

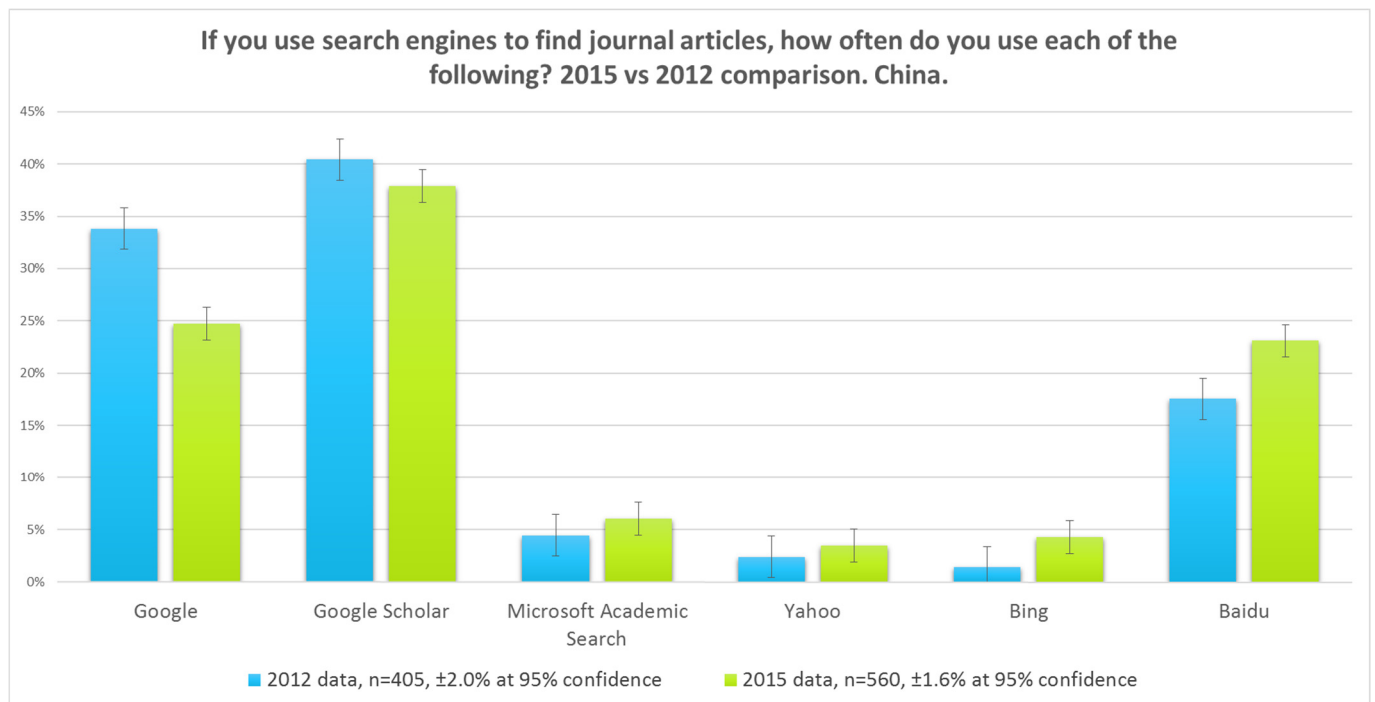


Figure 25 - Search engines, China, 2012-2015

Figure 25 shows us that the shift in usage in China has been significant, as those leaving Google seem to have shifted predominantly to Baidu. There is perhaps an indicative decrease in the use of Google Scholar.

In other countries, just as in China, studies have shown us that the bulk of the shift in favour of Google Scholar over Google is as a consequence of the decline of use of Google more than it is an increase in the use of Google Scholar. The difference appears to have been made up in the use of less popular search engines such as Bing.

LAST ARTICLE ACCESSED

LAST ARTICLE ACCESSED: 2012-2015

In 2012 and 2015 we asked how people had found the last article they accessed – via a search; clicking on links in social media; following a recommendation in an email; via a journal alert; or by any other means.

The results show that search is dominant – people are actively searching more than they are discovering articles any other way. Figure 26 shows us the data from all respondents who answered this question but the same results were shown in most other sectors, subjects, regions and job roles.

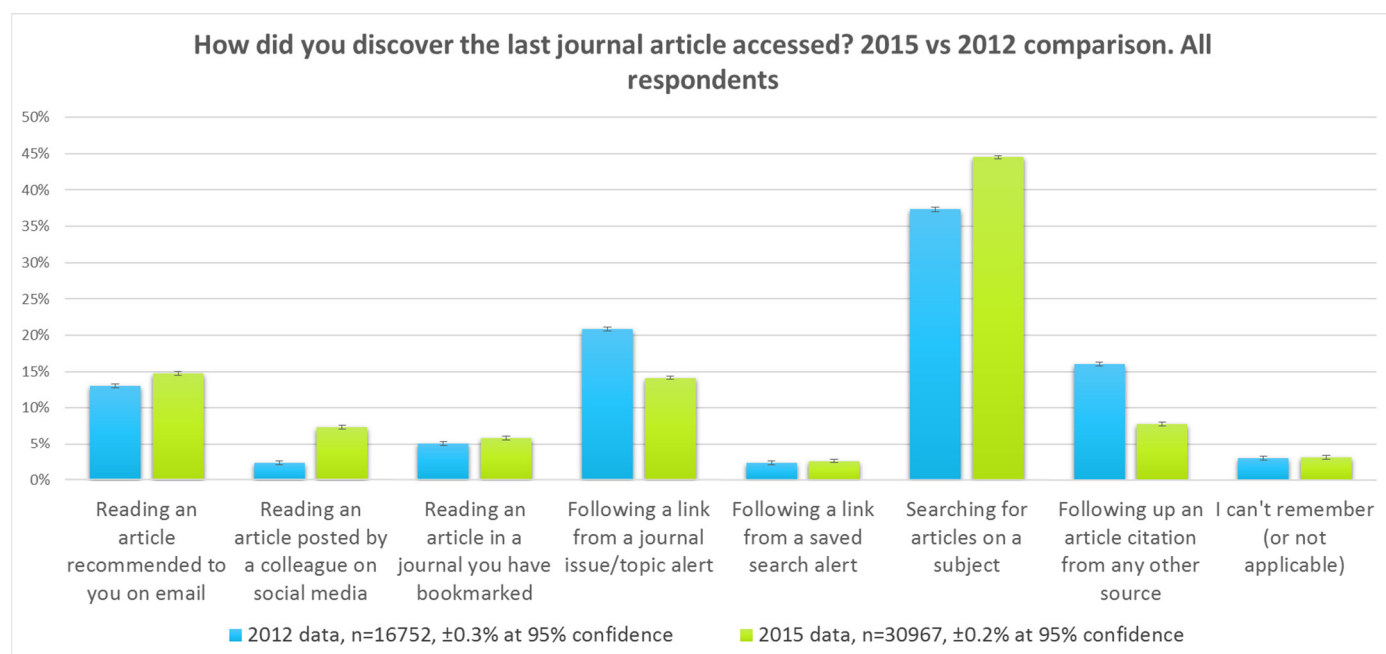


Figure 26 - Last article accessed, 2012-2015



These results show us that journal alerts have lost traction as a discovery method but social media has become slightly more popular. Search alerts and bookmarks were used to access the last article by only 2% and 5% of people respectively.

Perhaps one of the most marked changes is in Chemistry (Figure 27) where journal alerts were once as popular as search in the context of the last article that was accessed. Have people in Chemistry become bored of ToC Alerts or has searching improved for them?

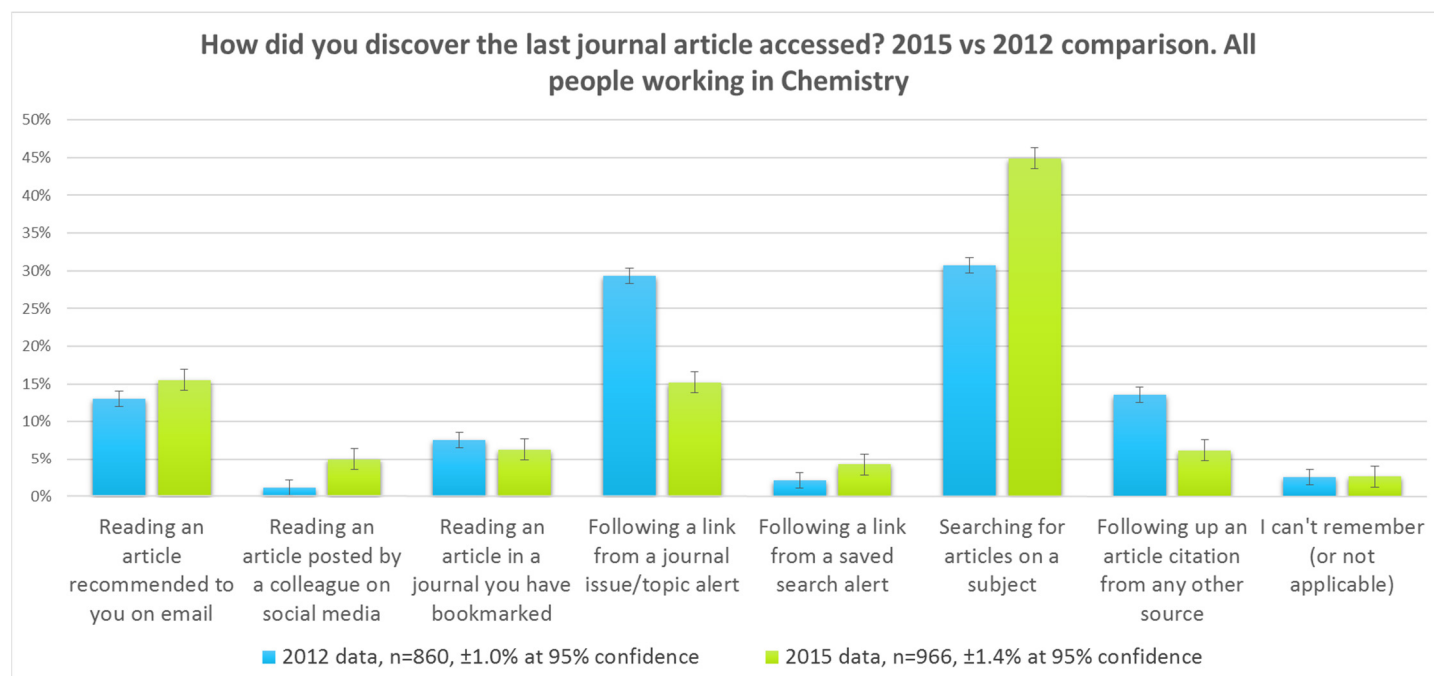


Figure 27 - Last article accessed, chemistry, 2012-2015

LAST ARTICLE ACCESSED: 2015

We tested to see if there was any difference in the responses to this question by wealth of the country. People in high income countries were slightly more likely to use journal alerts but search is still the dominant behaviour.

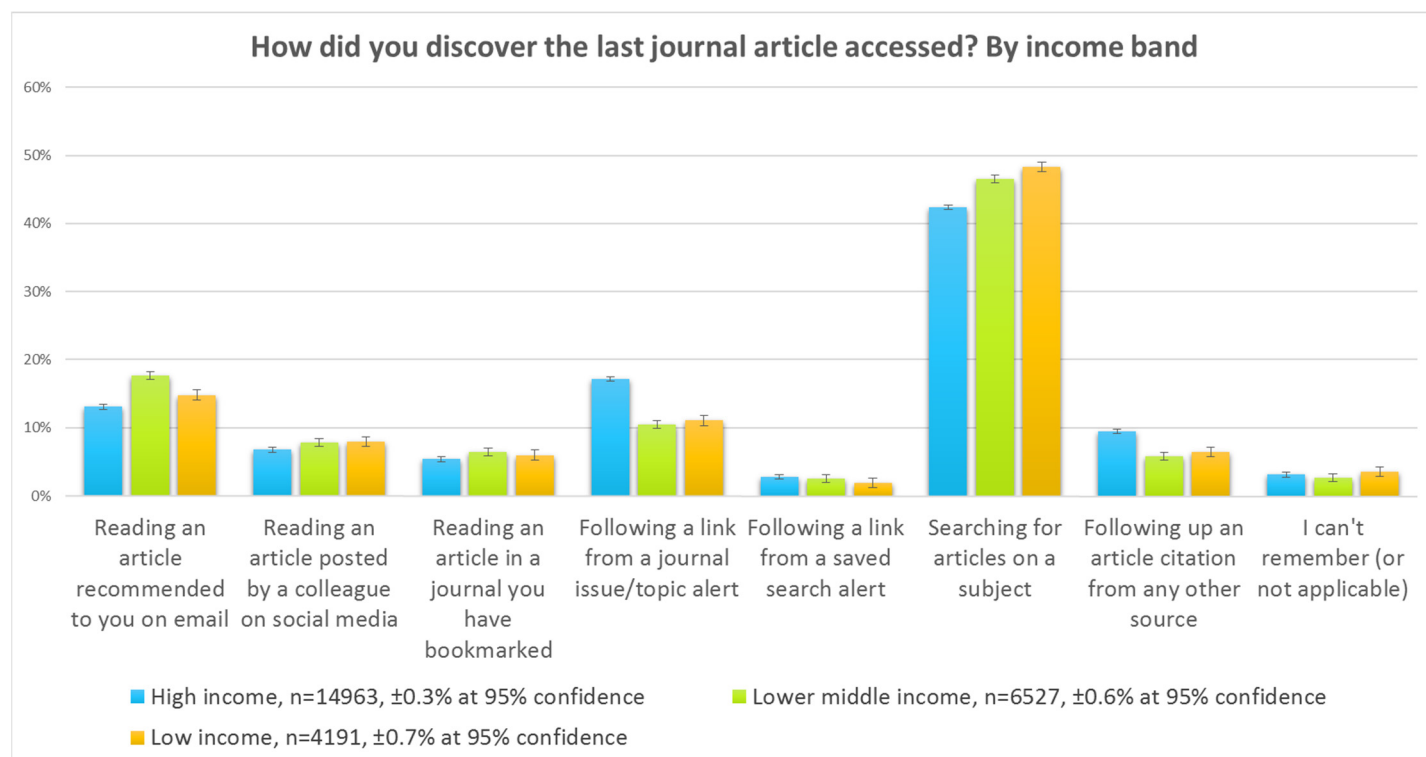


Figure 28 - Last article accessed, by income, 2015

Figure 29 shows that people in Humanities are much less likely to use ToC alerts and have “other sources” they may use. In the context of this question, “other sources” included Wikipedia.

“I often start the search as the result of a newspaper article or review online or from a physical book or newspaper I am reading”

“A very important resource for me for areas that are peripheral to my areas of expertise is Wikipedia. It is usually a good source of 1) basic information and 2) leads to research material. I also use the Encyclopaedia Britannica for introductory information about more traditional subjects.”

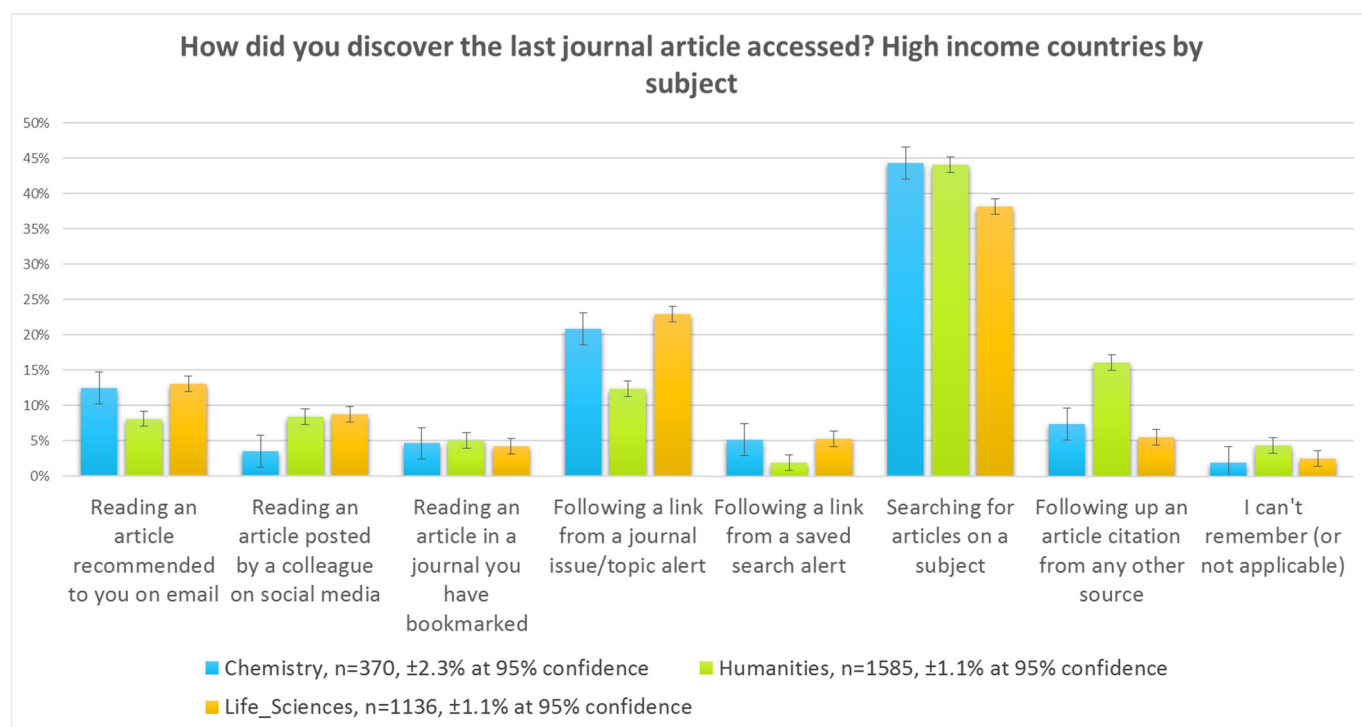


Figure 29 - Last article accessed, by subject, 2015

Figure 30 shows there is not so much variance by sector: academics are slightly more likely to search than people in other sectors. However, people in Medicine click on links in emails from colleagues more often than they follow journal alerts but the differences are small.

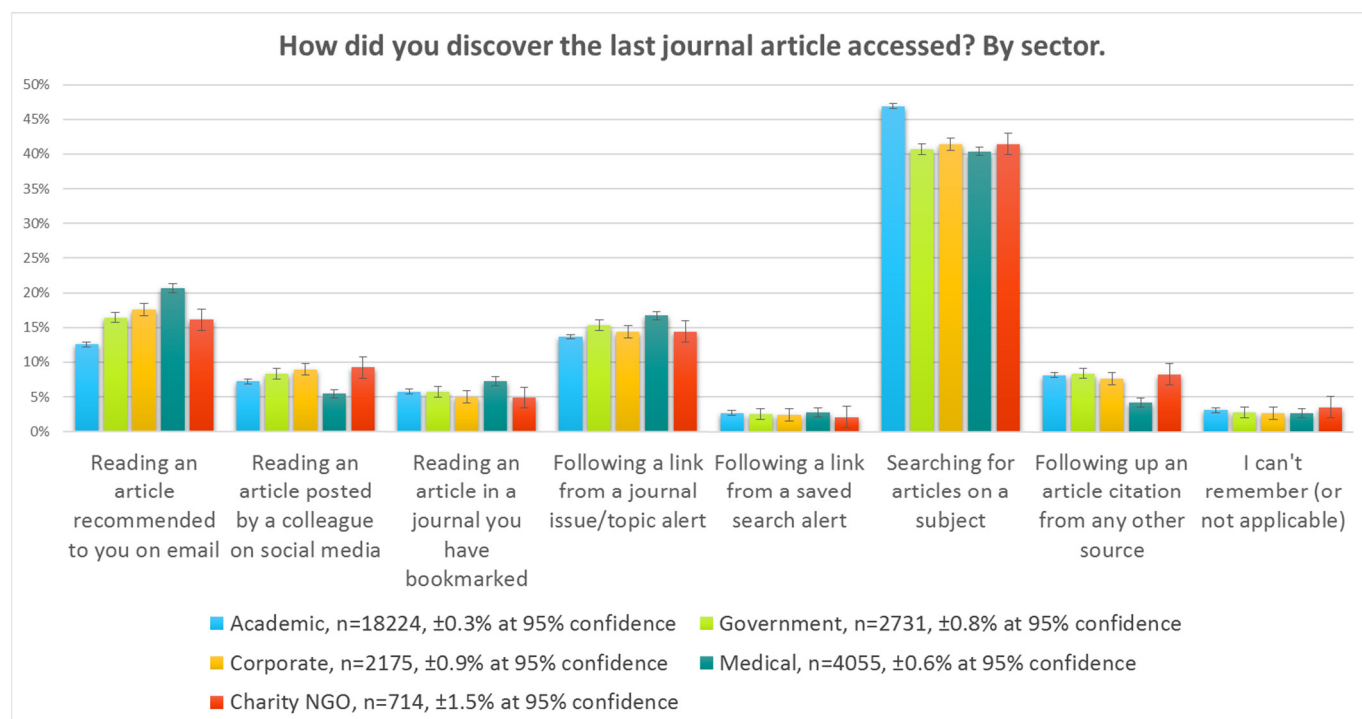


Figure 30 - Last article accessed, by sector, 2015

SEARCHING FOR BOOKS

The 2015 research for the first time included a section on book discovery. We asked where people started when they were looking for books, and asked what they were doing to find the last online book they read.

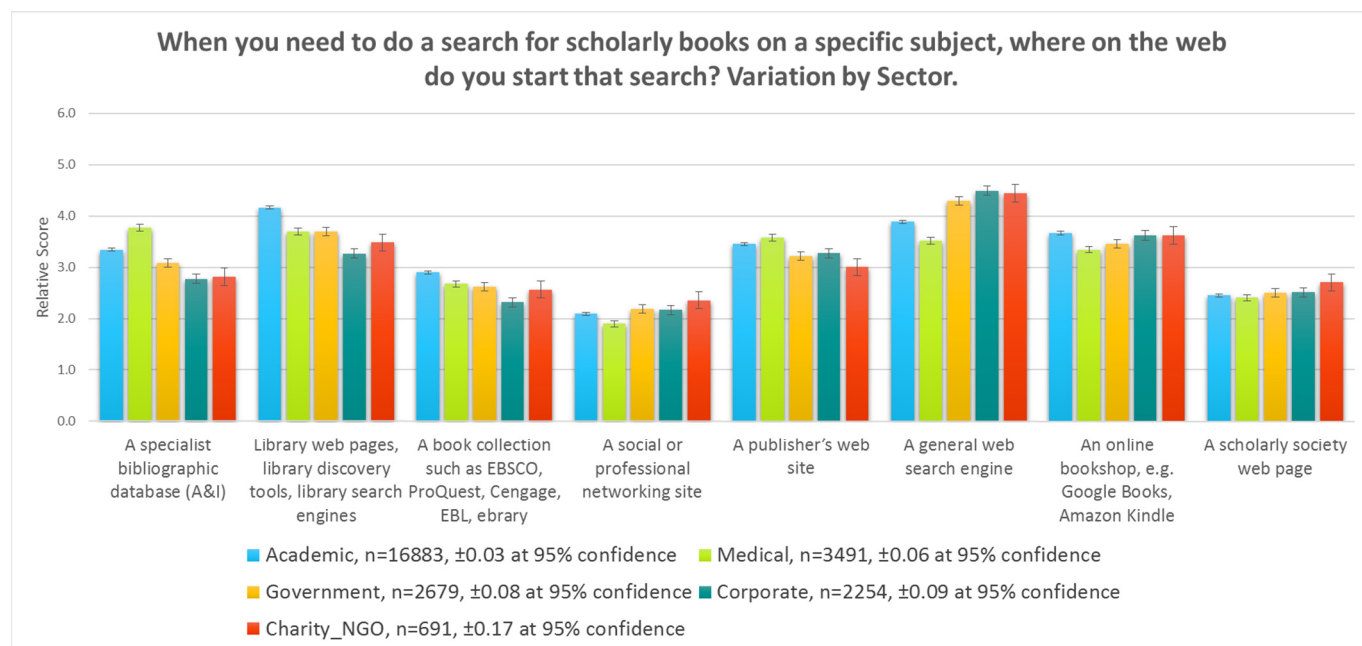


Figure 31 - Book search by sector, 2015

Figure 31 shows that people working in the Government, Corporate and Charity sectors think Google is the most important discovery resource for books. People working in the Medical sector rate both A&Is and their library as equally important for book search and rate Google as much less important than all the other sectors. People in the Academic sector think their library website is the most important resource for book discovery.



Just as for journals, it is people working and studying in the medical sector who rate A&Is as the most important resources for book discovery.

CAMBRIDGE

28,000 ebooks

Available to institutions via **Evidence Based Acquisition** –
for data driven decisions

ebooks.cambridge.org



CAMBRIDGE
UNIVERSITY PRESS

If we look at one particular job role in three sectors, we can see how educators/teachers rate different search resources for books. Teachers in academia are significantly less reliant on Google for book searches than other sectors.

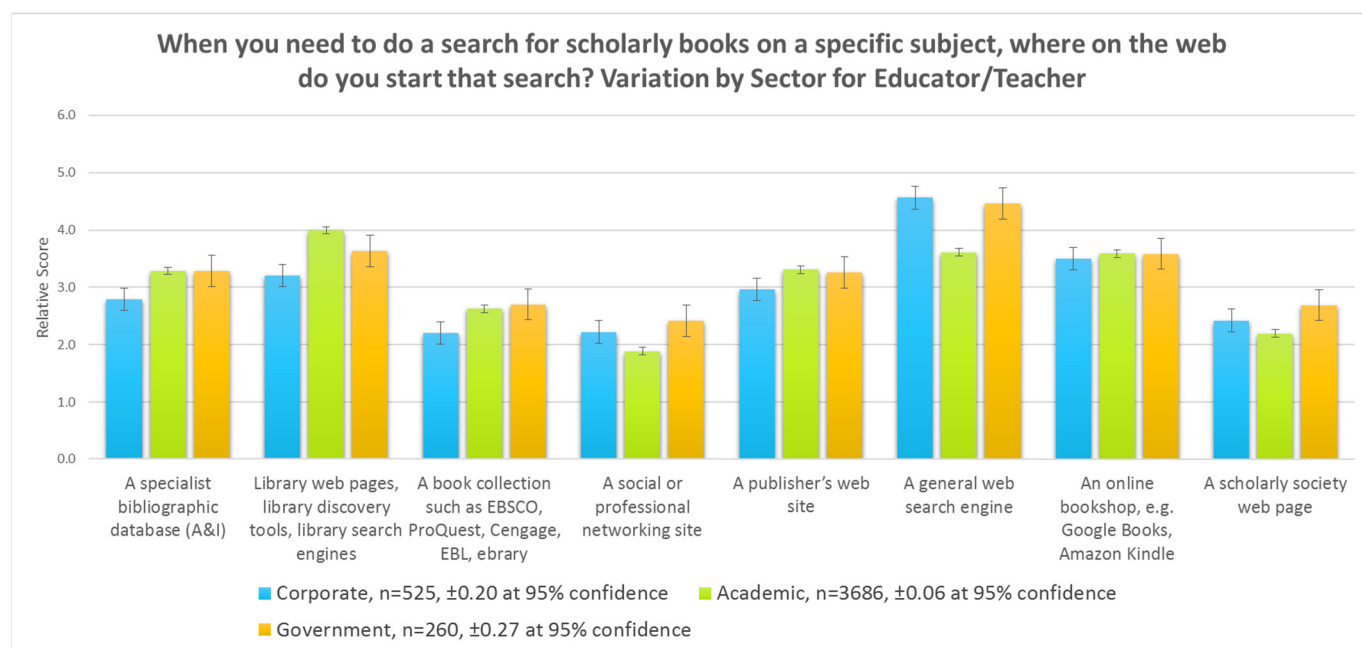


Figure 32 - Book search, sector, educator/teacher, 2015

LAST BOOK ACCESSED

Search is the single dominant method of online book discovery. Figure 33 is typical of the results no matter what demographics are studied. However, it still only makes up 45% or so of the discovery, with all other methods of discovery, such as emailed recommendations making up another 45%, with 10% unable to recall.

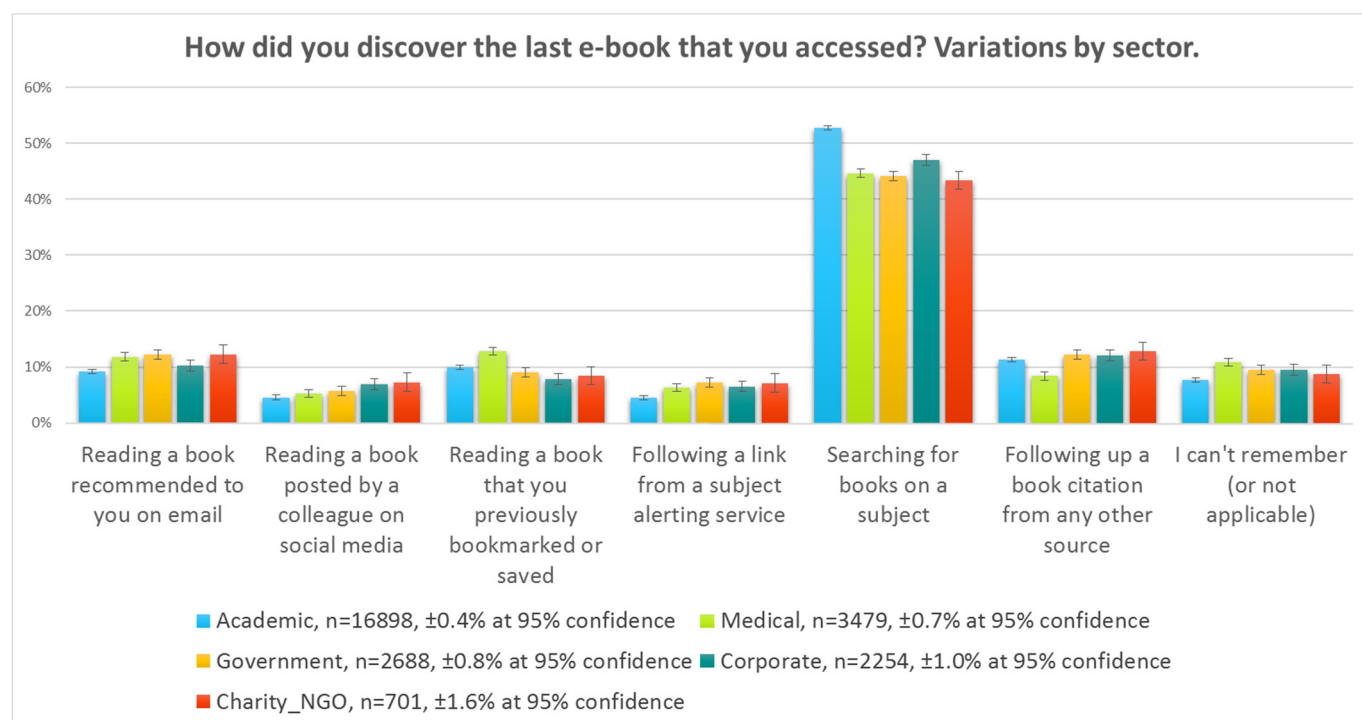


Figure 33 – Last book accessed, sector

4. OTHER METHODS OF DISCOVERY

FINDING SPECIFIC JOURNAL ARTICLES

Whilst search is arguably the most important type of researcher behaviour, it is interesting to study other reader activities and use cases.

This section looks at the resources people use when they already have a citation. The following graphs show where people start their search when they know exactly what article they are looking for.

The results for the trend analyses are based once again on a sample which is predominantly academic researchers working in STM, in North America and Europe.

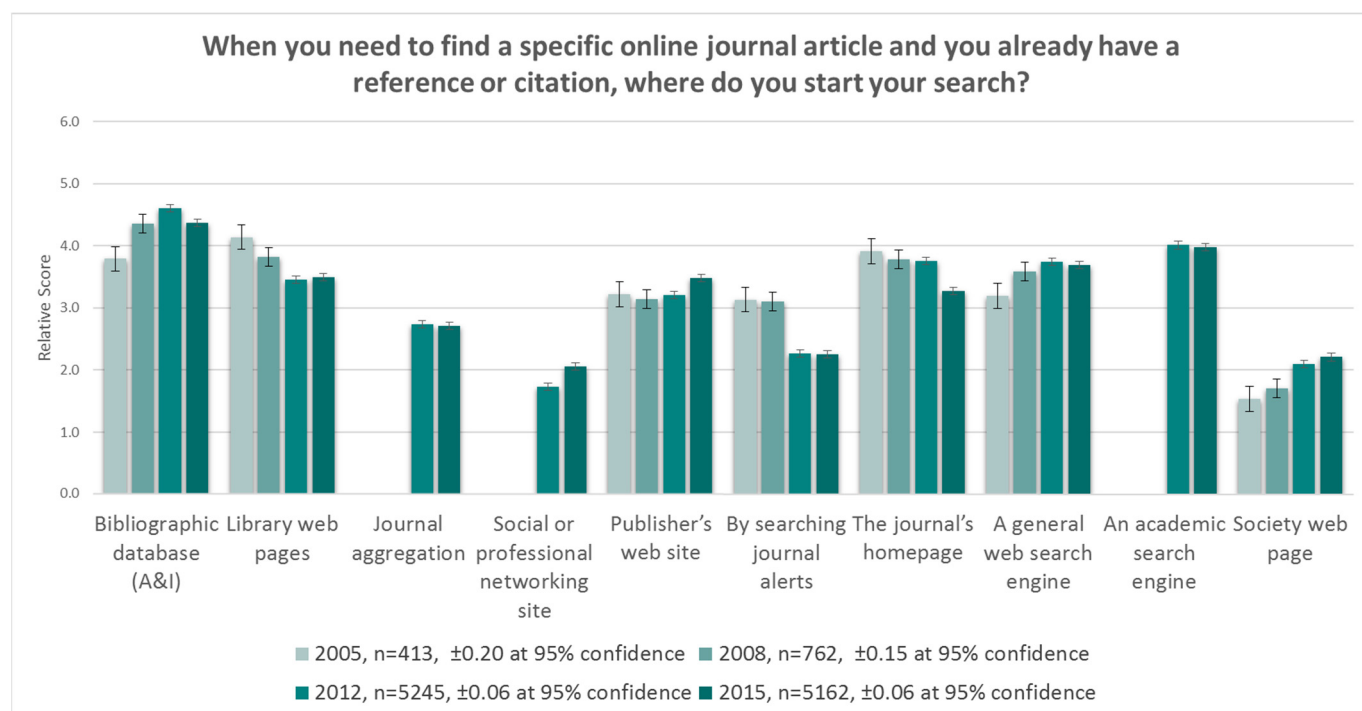


Figure 34 - Citation search, 2005-2015

It is perhaps counter-intuitive that readers use an A&I to look up a citation. Perhaps it is because they know they will get straight to the abstract (and possibly see citation statistics) which may help to determine whether they need to read the full text or not. However, there has been a slight decline in the importance of this resource for this use case since 2012. Google Scholar is now nearly as important for looking up a citation. Library Discovery is less popular for this activity, as is the journal homepage – perhaps fewer people are bookmarking their favourite journal for this purpose. However, the publisher (and society) website has seen a slight increase since 2012 – interesting as it has been a commonly held belief that readers are not interested or even aware of publisher brands and often do not know who publishes their favourite journals. Searching through an archive of journal alerts has declined in popularity for this use case.

BROWSE

A further use case tested by the survey, was the popularity of discovery resources used when readers want to stay up to date with the latest research in their area, and within their favourite journals. At this point the reader is perhaps in browse, rather than search, mode.

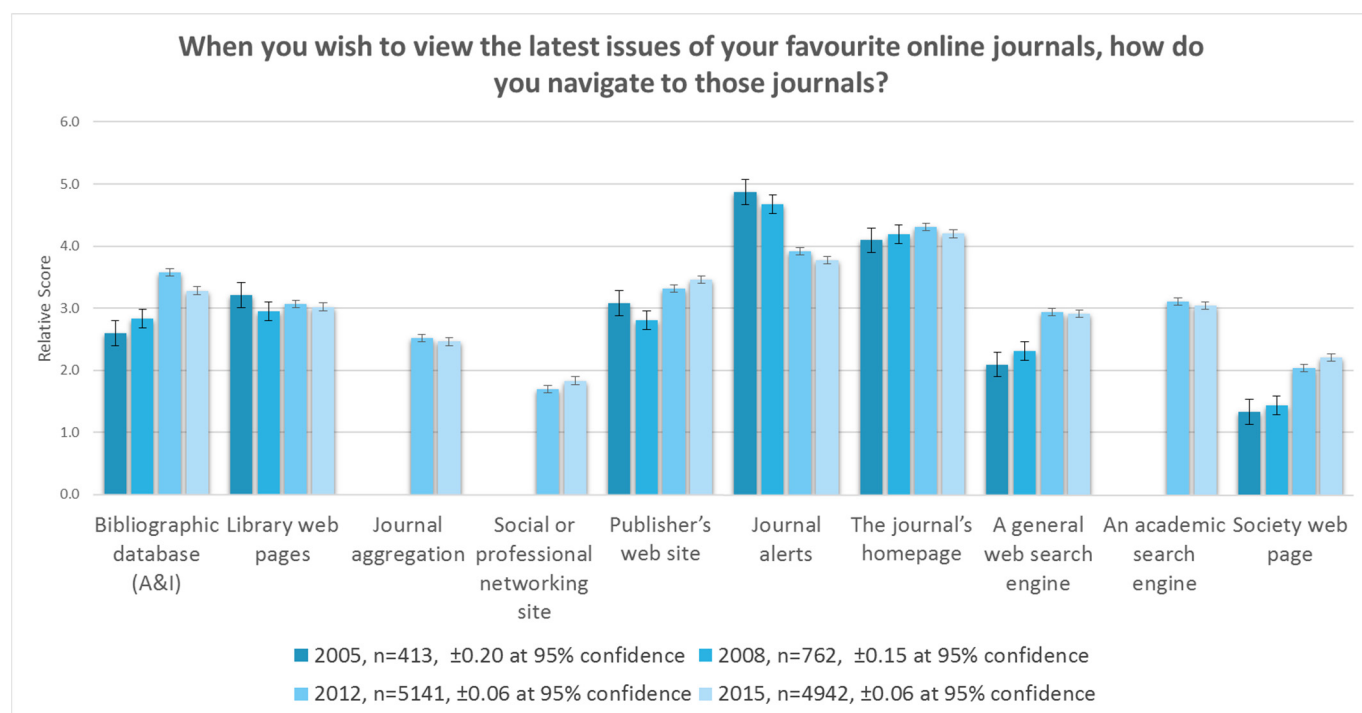


Figure 35 - Browse, 2005-2015

Journal alerts were once the most popular way of staying up to date with the latest research – users would receive an email alerting them to a new issue and either act on it straight away or perhaps save it for reference. As Figure 35 shows, this is certainly no longer the most popular method for browsing the latest research. Perhaps the most likely cause of this change is the move from issue-based online publishing to a more continuous publishing model, where articles are available online prior to the creation of an issue. This would clearly reduce the utility of the alert since by definition it is no longer a timely alert for materials that may have been available online for some time.

The journal homepage is the most popular method of browsing latest issues of key journals, with publisher and society web sites seeing some growth.

As would be expected, search engines are not as popular for this use case.



These results give a very strong indication that journal alerts are losing popularity as a method for staying up to date with the latest research.

“Email alerts with links can be handy & are pushed to me/the user - very convenient (unless the user is off-campus and must authenticate first). If just looking for the ToCs, then going straight to the journal site is the most efficient - works for owned as well as unsubscribed content in the majority of cases (assuming that the user has already used the library site to determine what is actually available directly). Some sites do not permit the user to access at all unless the "favourites" are owned, of course, so users who aren't sure of access may want to start with the library tools unless they are already aware of which journals are accessible to them.”

“MOST IMPORTANT: 3rd-party services such as BrowZine”

5. DELIVERY SITES FOR JOURNAL ARTICLES

There has much debate about the number of journal articles retrieved from resources other than the publisher website. For example, librarians and publishers are interested in how often a reader retrieves the article from an institutional repository, a subject repository or a professional/social networking site such as ResearchGate or Mendeley.

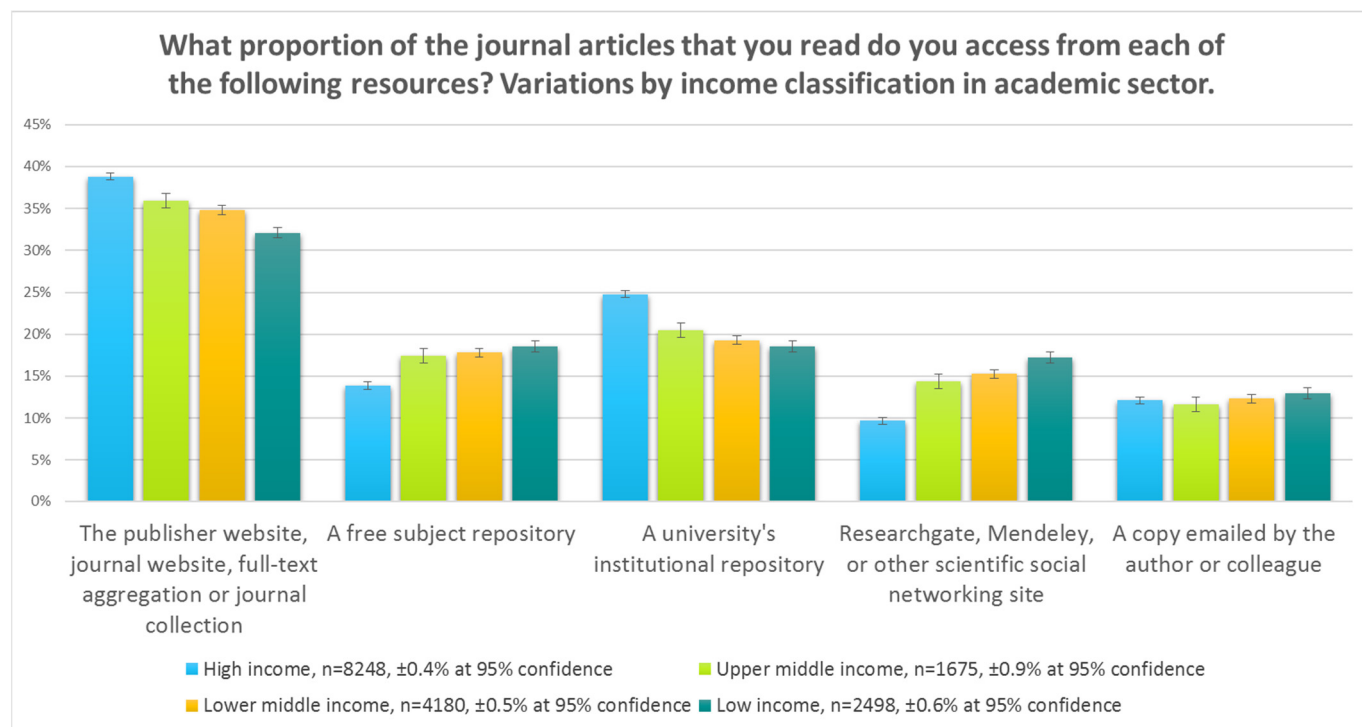


Figure 36 - Delivery sites, academic by income, 2015

Figure 36 illustrates delivery site options for readers from the academic sector across the world. (Please read 'Delivery Sites' in section 8, Methodology, for a discussion on the calculations used in this section.)

The use of the publisher web site for delivery varies simply by income bracket, with most use being made of them by people from higher income countries who one presumes are more likely to be in institutions with sufficient funding to purchase these resources. However, with projects like Hinari and Agora in place, one might have expected an improved position for low income countries.

The use of institutional repositories seems to mirror the likelihood of an institution having installed one – i.e. in higher income countries. This also tends to indicate that institutional repositories are used more by people who are familiar with them in their own institutions, rather than being seen as a global resource for delivery. After all, in aggregate, people in lower income countries use delivery sites that typically house free versions of articles the most (i.e. the sum of the use of sites other than the publisher or aggregator - Figure 37), and yet don't use institutional repositories as much as part of the mix.

The reach of the subject repository, however, seems to be much more global, and is used proportionately more in poorer countries, as are sites such as ResearchGate and Mendeley.

Figure 37 shows the same data, replotted as a sum of publisher-controlled delivery sites, versus those that it doesn't control. All of the resources listed, apart from the publisher and aggregator category, represent an opportunity to download the article for free. We can see that over half of the downloads are from these sites. In addition, it should be remembered

that many publisher sites contain open access journals, open access articles within subscription journals, free access to older material from many subscription journals, or gratis access as part of developing world initiatives. It follows, therefore, that the actual proportion of delivery that is free, is far greater than the amount shown here.

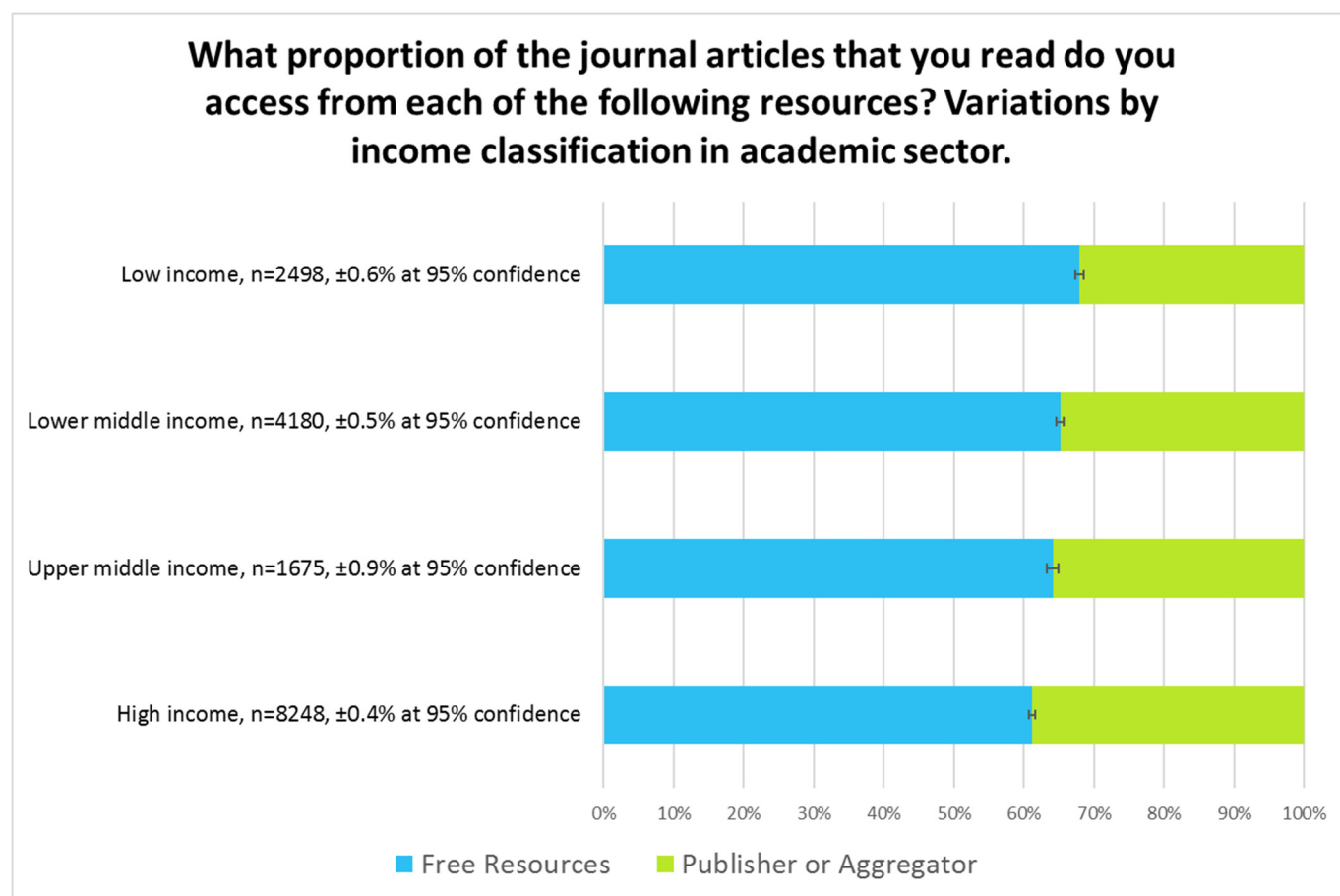


Figure 37 - Delivery, free resources vs publisher-controlled sites, 2015



For approximately 60% of the time, readers in high income countries in the academic sector are accessing articles from a free resource. This means that they are 1.5 times as likely to be reading an article from a free resource. In lower income countries this rises to over 2 times as likely.

If we now look at this by sector within high income countries, we can see there are some significant variances in the type of free resource the article is being retrieved from, although not so much in the overall “publisher site vs. free site” ratio.

“If I can't access an article of interest on the Web, I do a search for the email address of one of the article's authors and then send an email directly to the author explaining that I wish to read the entire article for a paper I'm writing (which I note I will then cite), and then request a copy of the article. I express much gratitude for the author's consideration of my request. The company I work for is not affiliated with a university, therefore I am unable to access to many articles through an academic database, as I did in previous jobs.”

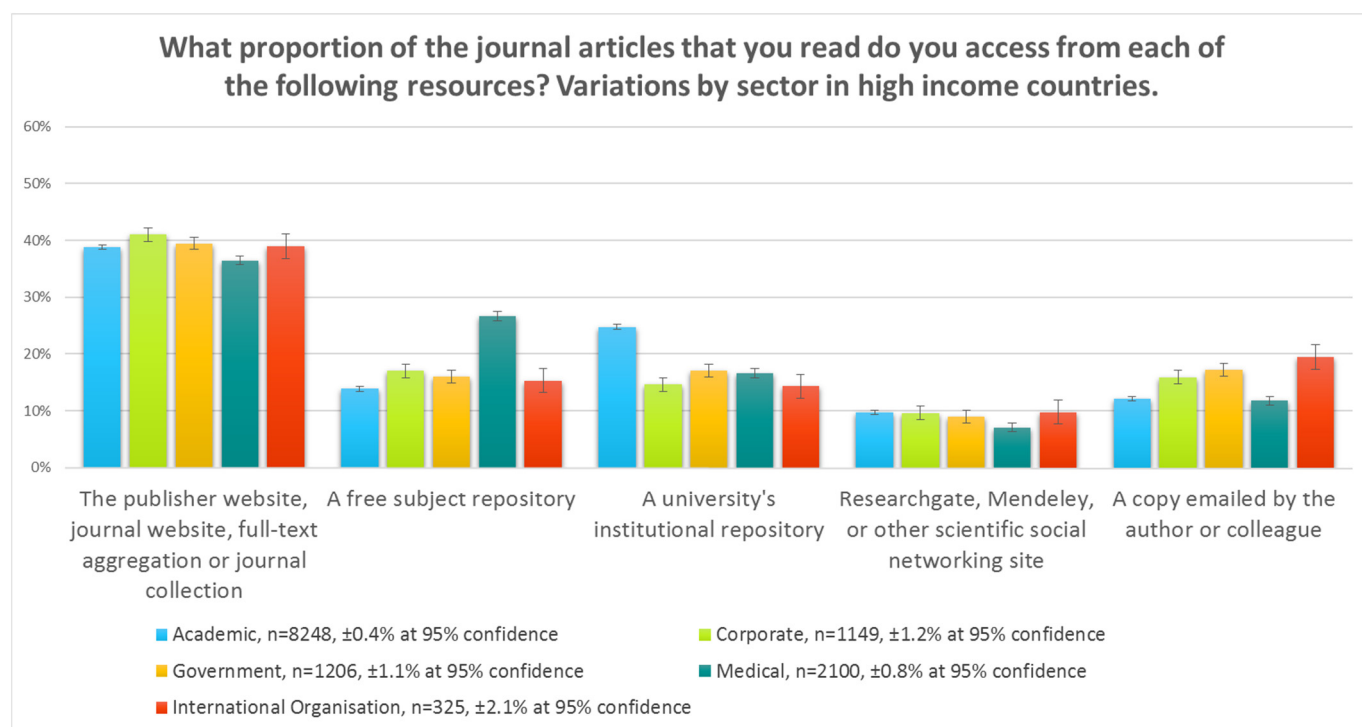


Figure 38 - Delivery site, high income countries, by sector, 2015

Articles come from sites such as ResearchGate and Mendeley less than 10% of the time in high income countries. Institutional repositories are significantly more important in the academic sector potentially indicating people in this sector are more aware of institutional repositories. It follows from the findings on page 27 where we showed that people in academia, in higher income countries are more likely to use Google Scholar for discovery than others, that, since Google Scholar will highlight results from institutional repositories, then academics will retrieve articles from repositories more, and that people in poorer countries will do so less.



People in the medical sector are accessing journal articles from a free subject resource 25% of the time. This is significantly higher than all other sectors in high income countries.



It seems probable that a lack of awareness of Google Scholar in poorer nations is leading to a reduced use of free incarnations of content in institutional repositories from these regions.

Figure 39 shows that people working in Chemistry are more likely to access their journal article from the publisher website – over 50% of articles are accessed this way compared to 37% of articles in the Life Sciences.

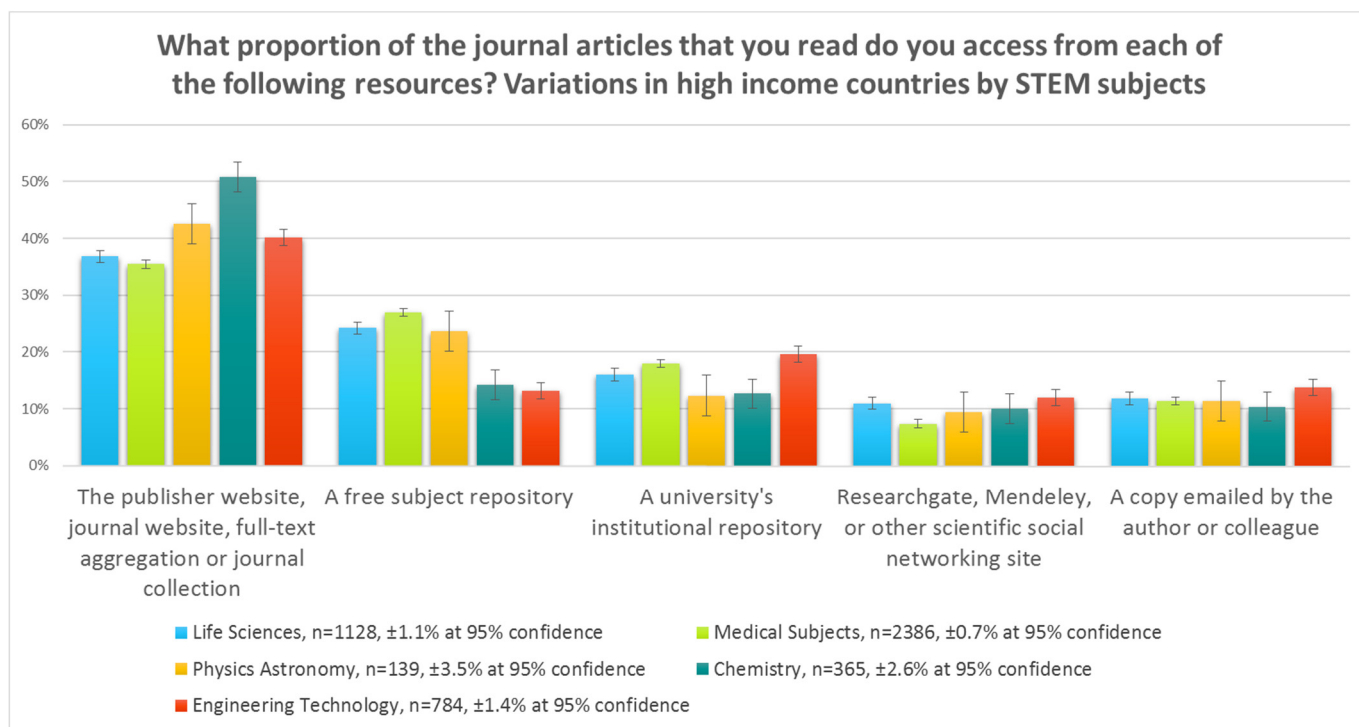


Figure 39 - Delivery site, high income, STEM, 2015

We saw little variation in delivery sites across the HSS subjects, however in comparison with STEM we can see some significant differences. The absence of high-profile repositories in HSS is well-evidenced in Figure 40.

Discover BrillOnline.com

The majority of Brill's Online Resources are available on five state-of-the-art platforms, each of which are home to one specific content type:



- booksandjournals.brillonline.com
- referenceworks.brillonline.com
- bibliographies.brillonline.com
- primarysources.brillonline.com
- dictionaries.brillonline.com

Interested?

Visit brillonline.com for more information.

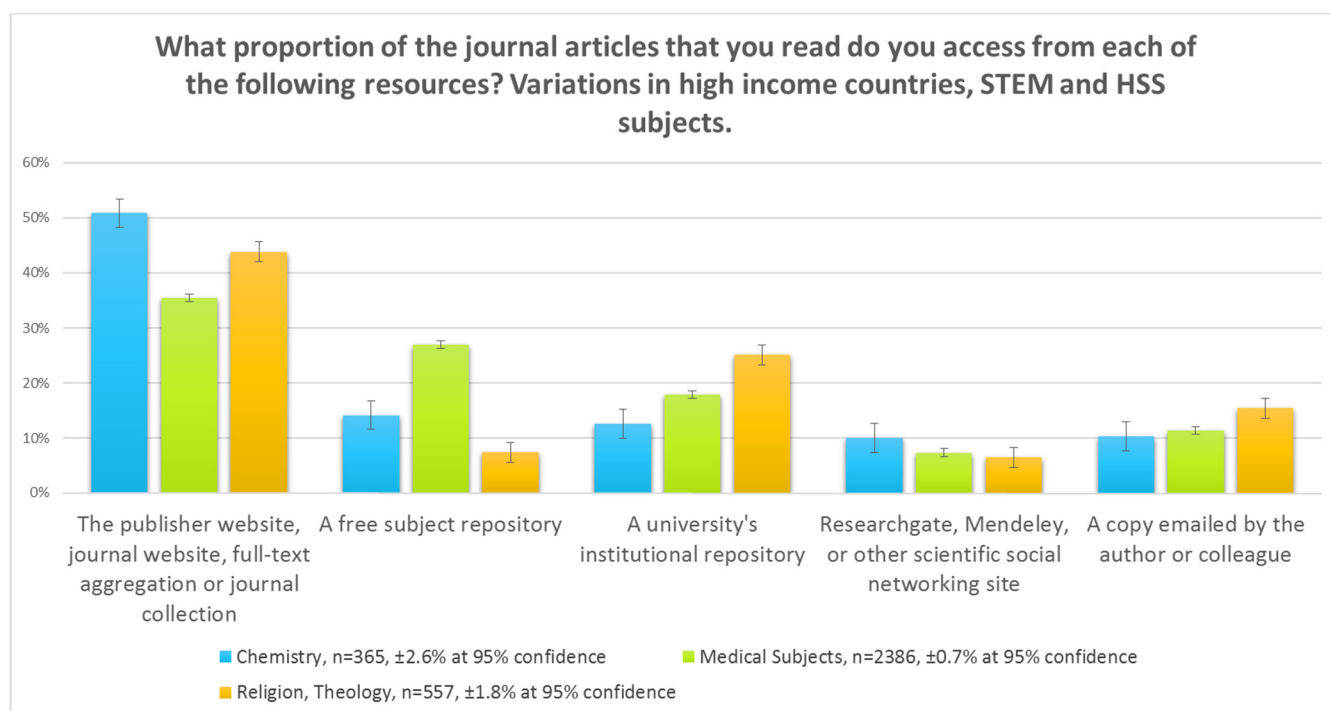


Figure 40 - Delivery site, high income, STEM vs HSS, 2015



PubMedCentral is likely to account for the popularity of subject repositories in Medicine. Scientific social networking sites are not responsible for a large proportion of the delivery of free content to readers in high income areas.

6. MOBILE DEVICES AND APPS

DEVICES USED TO FIND AND ACCESS JOURNAL ARTICLES

In 2012 we started to study the device types that readers use to access journal articles. We can see that, in low income countries, there has been a shift from desktop computers to mobile access using phones and tablets since 2012 – the use of laptops has stayed steady at around 50%.

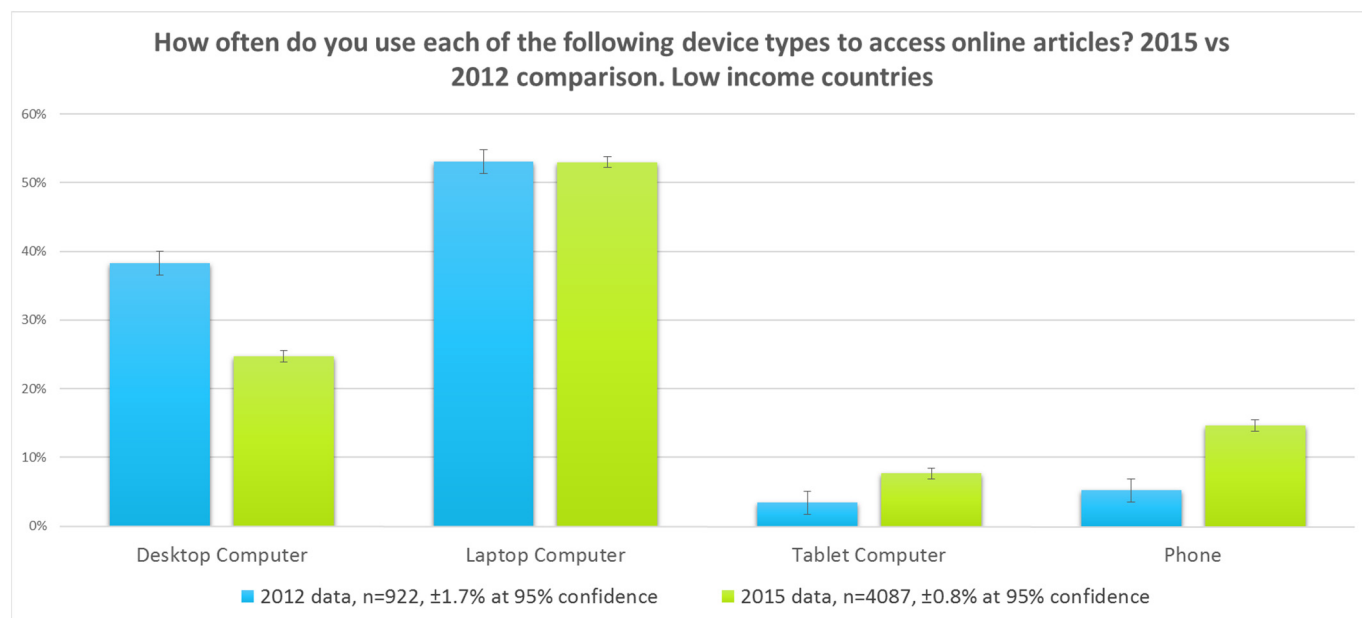


Figure 41 - Devices, low income, 2012-2015

We looked at this comparison for high income countries and the drop in use of desktop computers was only around 4% in favour of phones and tablets.

Looking at the 2015 results only, in Figure 42, we can see there is some variance by income classification.

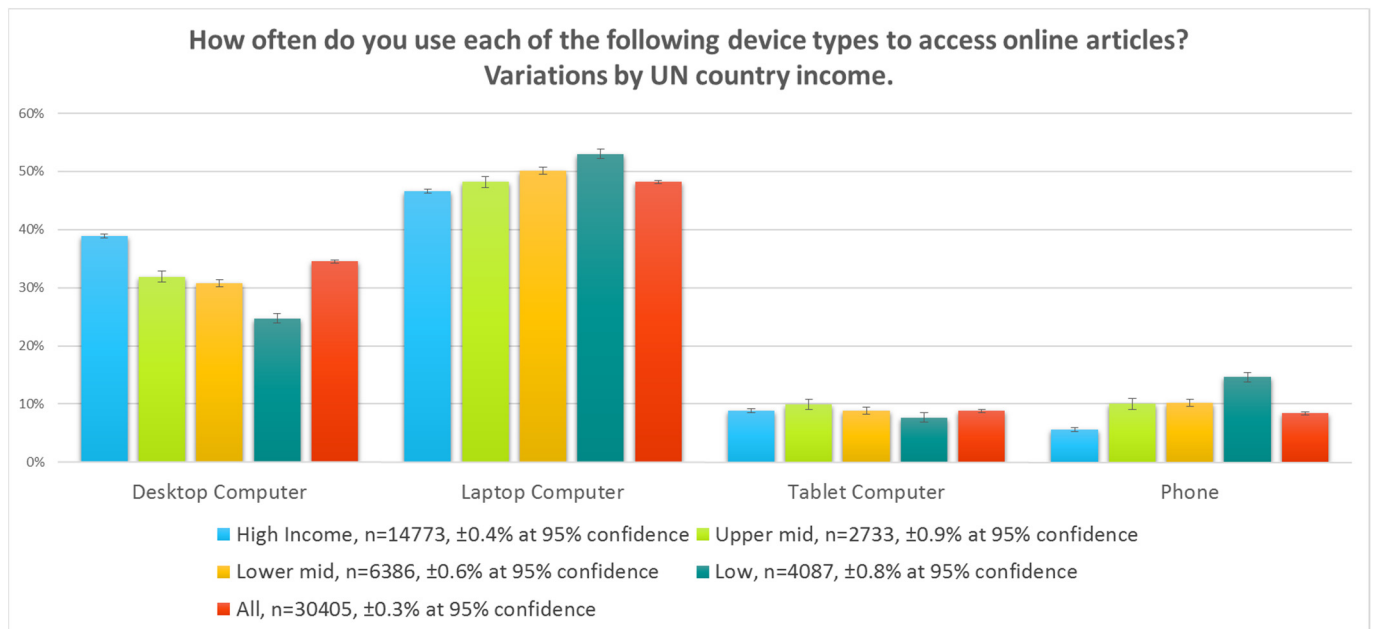


Figure 42 - Devices by income, 2015

It seems that desktop computers are legacy devices mostly used in the developed world, and that more mobile forms of computing are favoured in poorer countries.

Figure 43 shows that in the medical sector across the world, there is growth in the use of tablets and mobile phones to access journal articles and quite a dramatic drop in the use of desktops.

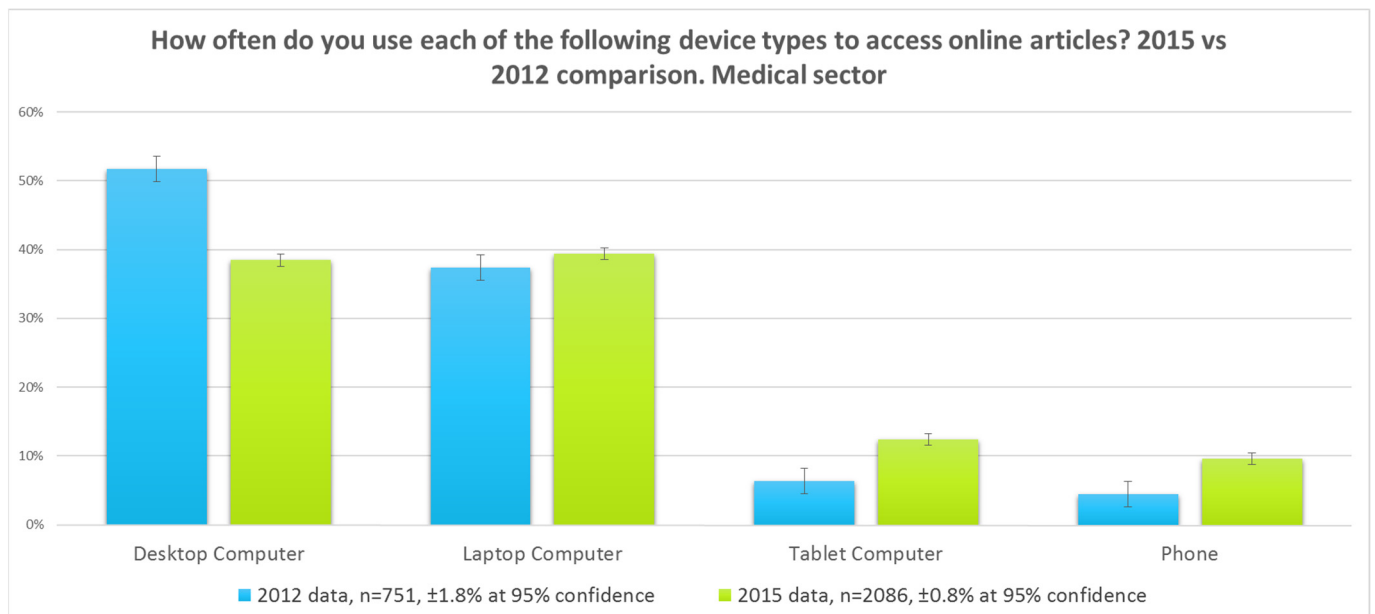


Figure 43 - Devices, medical sector, 2012-2015



Given that the user experience does not vary a great deal if they are using a desktop, laptop or a tablet, it's really the adoption of smartphones for finding and accessing content which has the greatest impact on service providers. Overall we see that whilst there is growth in the use of phones to access journal articles, especially in low income countries, it still accounts for only about 10% of the use. Many discovery services, however, are not mobile-friendly. Publisher sites are often mobile-friendly and this may account in part for the rise in their popularity.

There are also some variances by job role. Librarians are still much more likely to use desktop computers than anyone else. Researchers have very little use of phones and tablets to access journal articles and this has not changed since 2012.

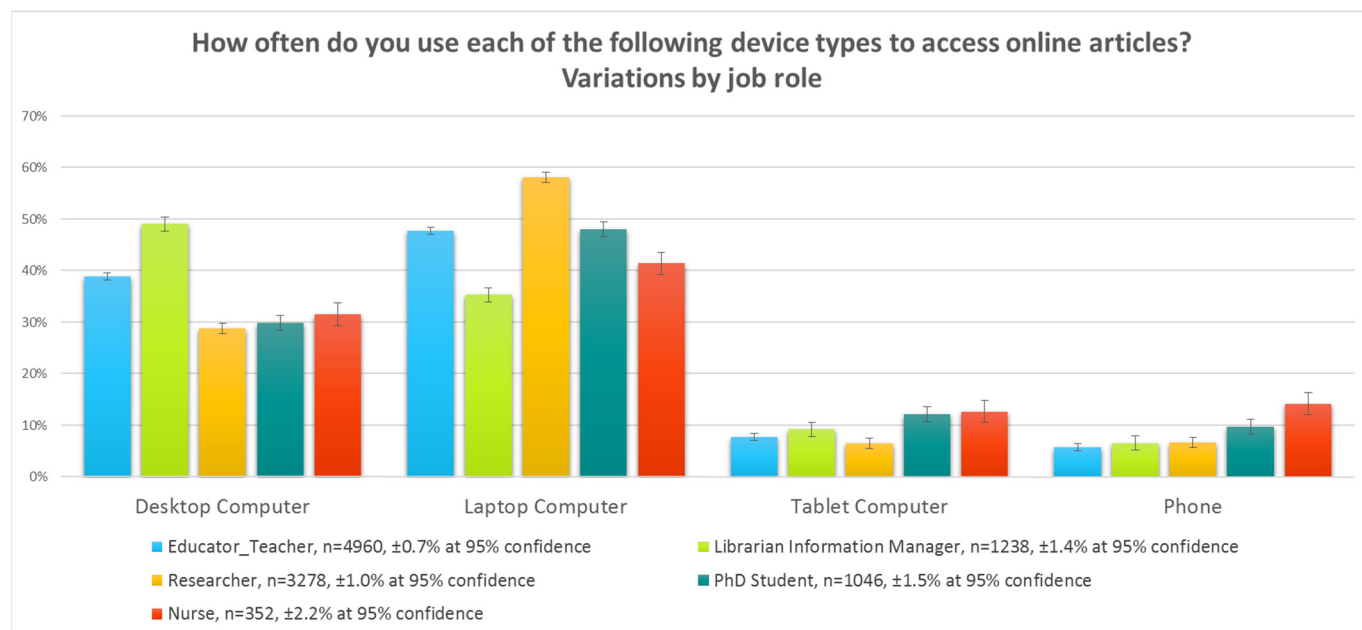


Figure 44 - Devices, job role, 2015

APPS USED TO DISCOVER AND READ JOURNAL ARTICLES

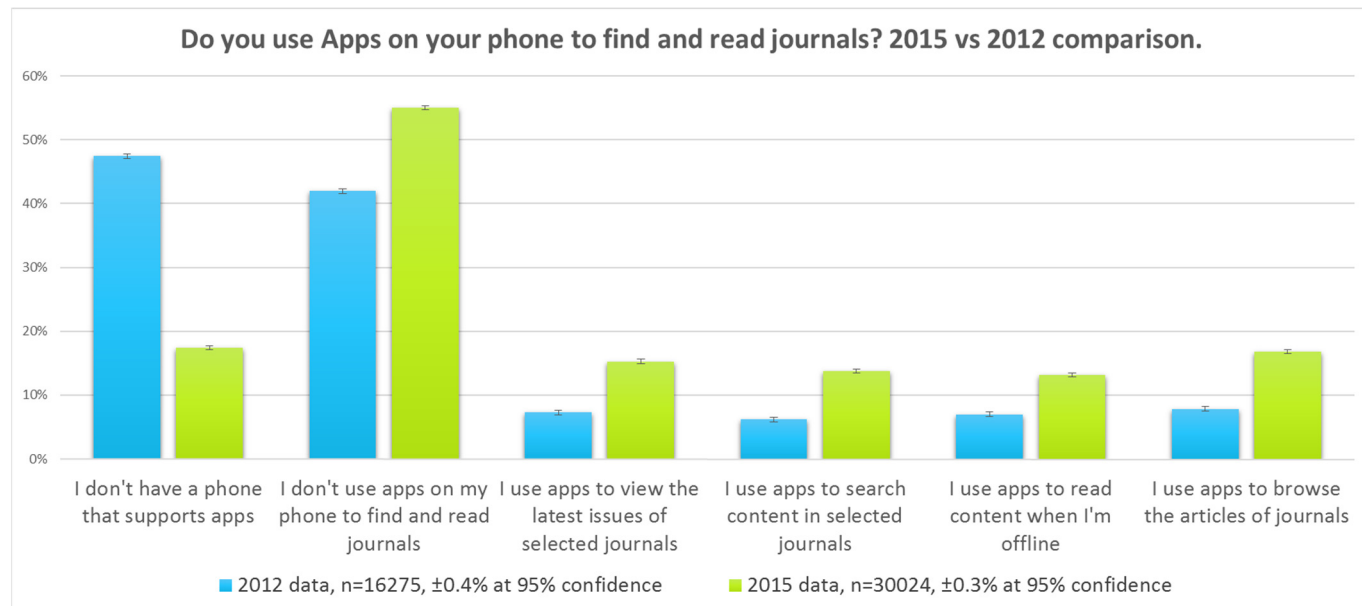


Figure 45 - App use, 2012-2015

The number of people who do not have a phone which supports apps has decreased dramatically, only 18% of people now don't have a smart phone compared with nearly 50% of people in 2012. However, it does not appear that an increase in the ownership of smartphones has translated into an increase in app use for journal article discovery and consumption. In fact, a higher proportion of people don't use apps on their phones than in 2012. For those who do use apps on their phones, the range of applications used have

increased. (Note: In this question, respondents were able to choose multiple answers – all those that apply. This means that the percentage responses does not need to add to 100%. This also means that each of the four measures on the right can increase even against a backdrop of a reduction in the overall proportion of people using smartphones for journal discovery. It simply means that those that use smartphones for journal discovery now use them in more ways than they did previously.)

If we look at the results from 2015 (Figure 46), we can see that people in low and low/middle income countries are more likely to use apps to find and discover journal articles than those in high income countries. However, overall use of apps is still low.

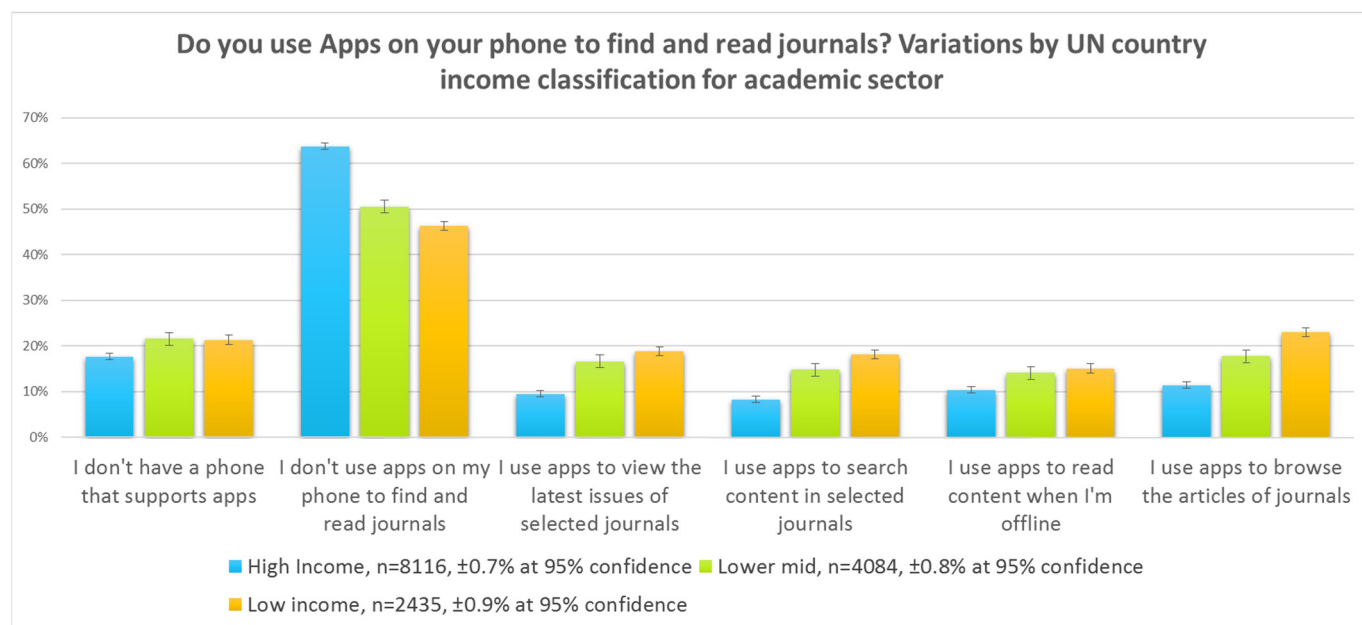


Figure 46 - App use, by income, 2015

When we looked at these results by subject in high income countries, we noted some variance from subject to subject in app use – for example, people in Medicine are more likely to use apps than people in Chemistry.

DEVICES USED TO FIND AND ACCESS BOOKS

People in low income countries use phones and tablets more so than people in richer countries to access books as well as journals. We looked to see if there was any variance by sector in low income countries and appears there is some, although important to note that access by laptop still dominates and phone and tablet use for online book consumption is still relatively low (see Figure 47).

People in the academic sector are more likely to use laptops than in other sectors and their take up of phones and tablets is quite low. Accessing books online via a phone is more likely in the medical, corporate and charity sectors.

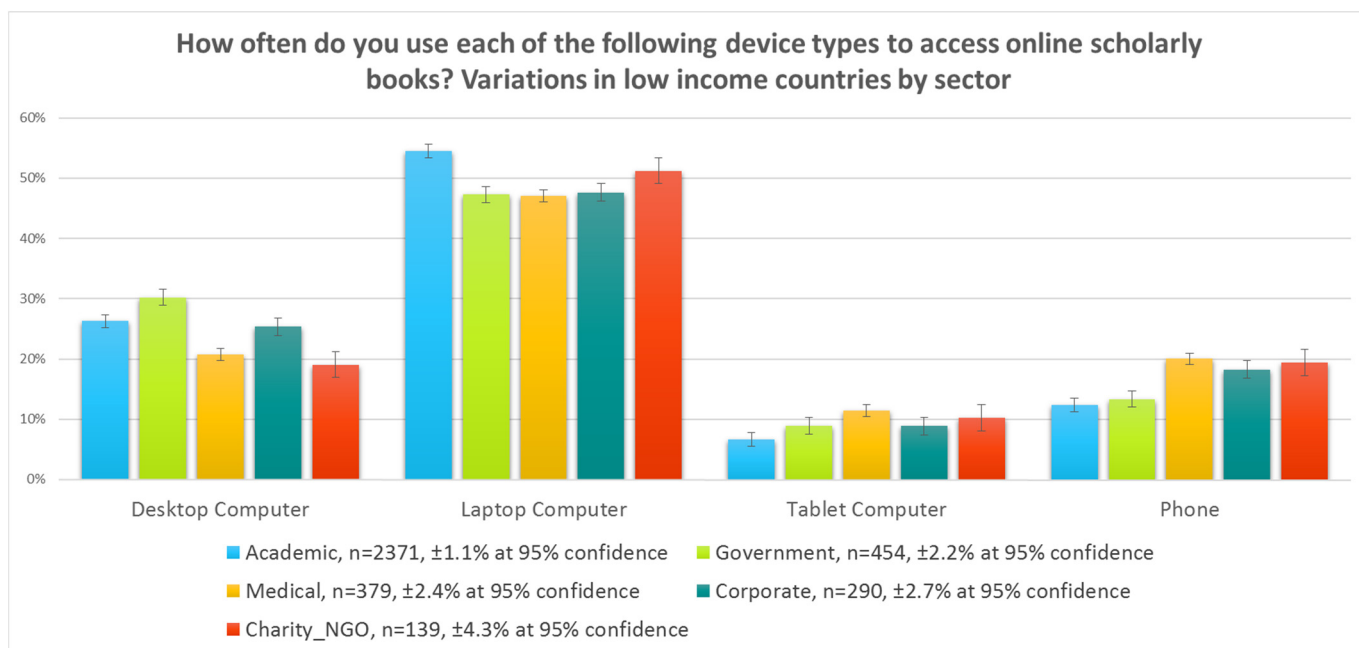


Figure 47 - Devices, books, low income by sector, 2015

If we look in aggregate across all sectors and income classifications, we can see there are also some variances in job role (Figure 48). People in Engineering seem much more tied to their desktop than anyone else – even their use of laptops is low.

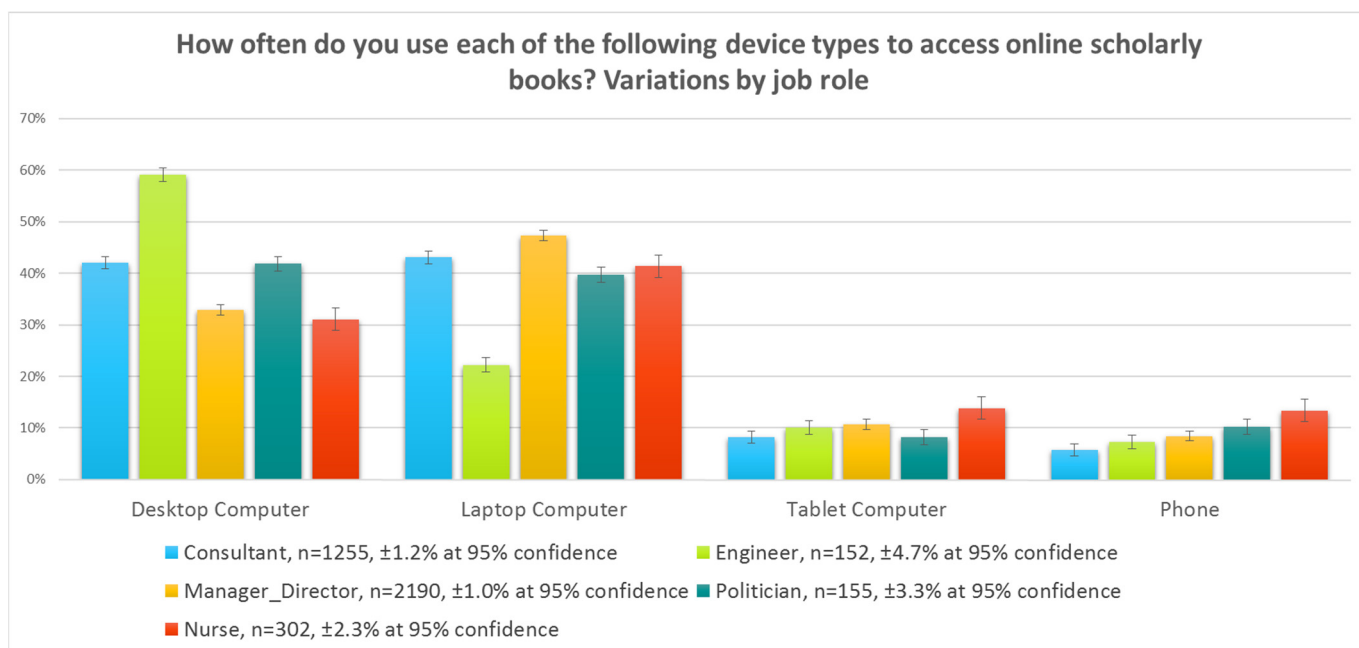


Figure 48 - Devices, books, job role, 2015

7. JOURNAL WEB SITE FEATURES

The survey asked people to state how useful they found a range of publisher website features. We also have the same data for comparison over a ten-year period. During that time, we have updated the feature list to reflect changes in web sites. (Respondents aren't asked to rank the items, merely rate their utility, so changing the number of options does not affect the importance of other features.)

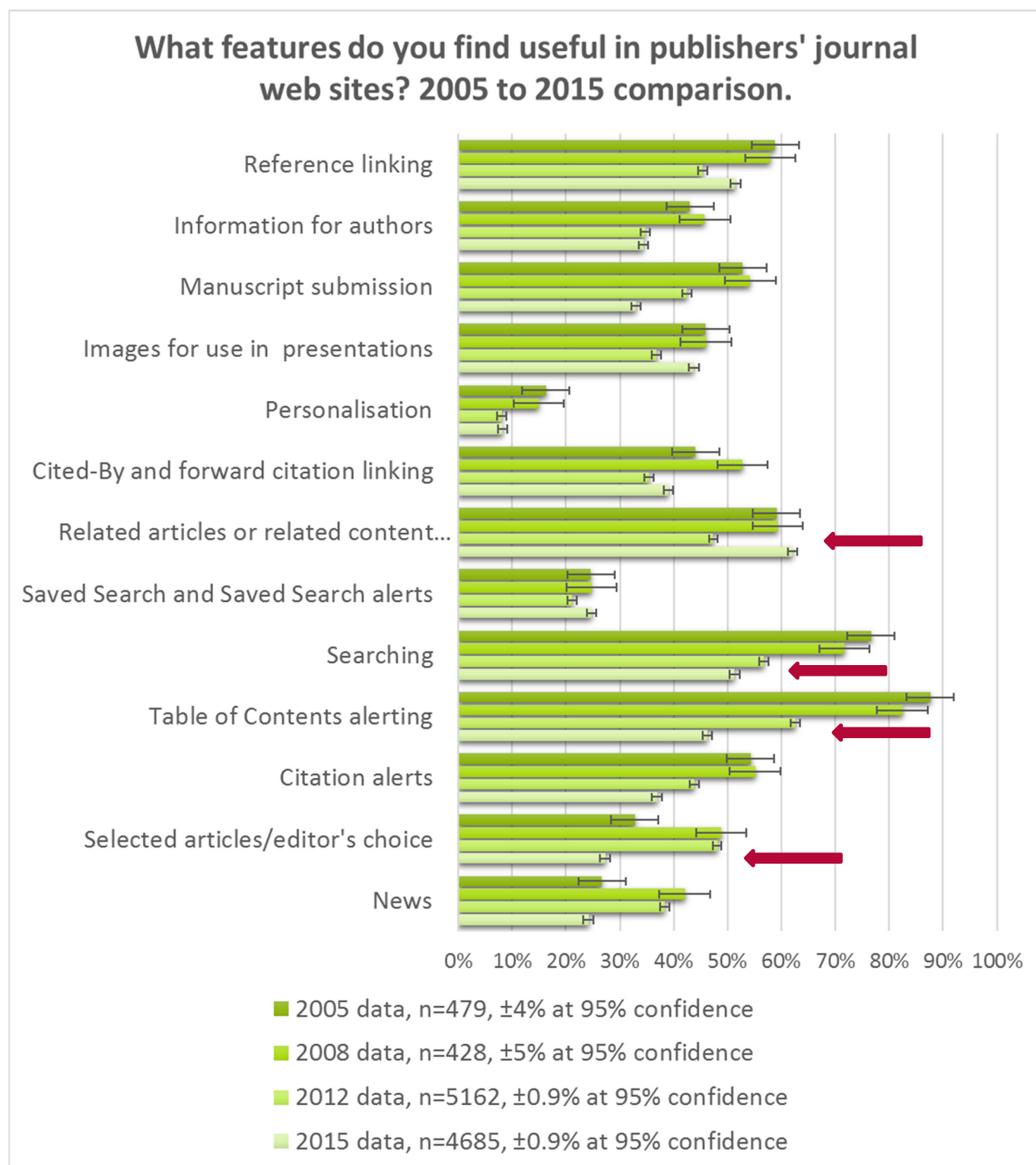


Figure 49 - Journal web site features, 2005-2015

The usefulness of Table of Contents Alerting on a publisher website has significantly decreased since 2005. This in line with our findings on search and browse reported earlier in this report.

Related Articles functionality is the only area that has become more useful since 2012 but it has only caught up with its popularity score from 2005 and 2008.

Whilst News and Editors Choice/Selected Articles type content saw an increase in popularity in 2008 and 2012, this type of feature seems to have peaked in its popularity; only around a quarter of respondents found either of these features useful – a significant drop since 2012.

It is interesting that whilst we have shown that publisher websites have experienced a growth in importance as a starting point in search, fewer people have rated the search functionality as useful – 50% of respondents said they found search features useful now compared with 75% of respondents 10 years ago. One theory is that respondents feel somewhat driven to search on publisher web sites because they are unaware of, or have no access to, other discovery resources. Alternatively, perhaps publishers have become much better at marketing their web sites.

Figure 50 looks at the 2012-2015 trend in the academic sector. We see that the ability to download images, reference linking, and related articles have all become more popular. In the academic sector, searching has not lost as much ground as other sectors (as shown in Figure 49).



Related-articles is now the most popular feature of a publisher website. Content generated by the publisher such as news and editor's picks became more popular between 2005 and 2008, but is now back at a similar level to 2005. This perhaps indicates the time-poor researcher does not want to be distracted by non-journal content.

What features do you find useful in publishers' journal web sites? 2015 vs 2012 comparison. Academic sector, 2015 versus 2012

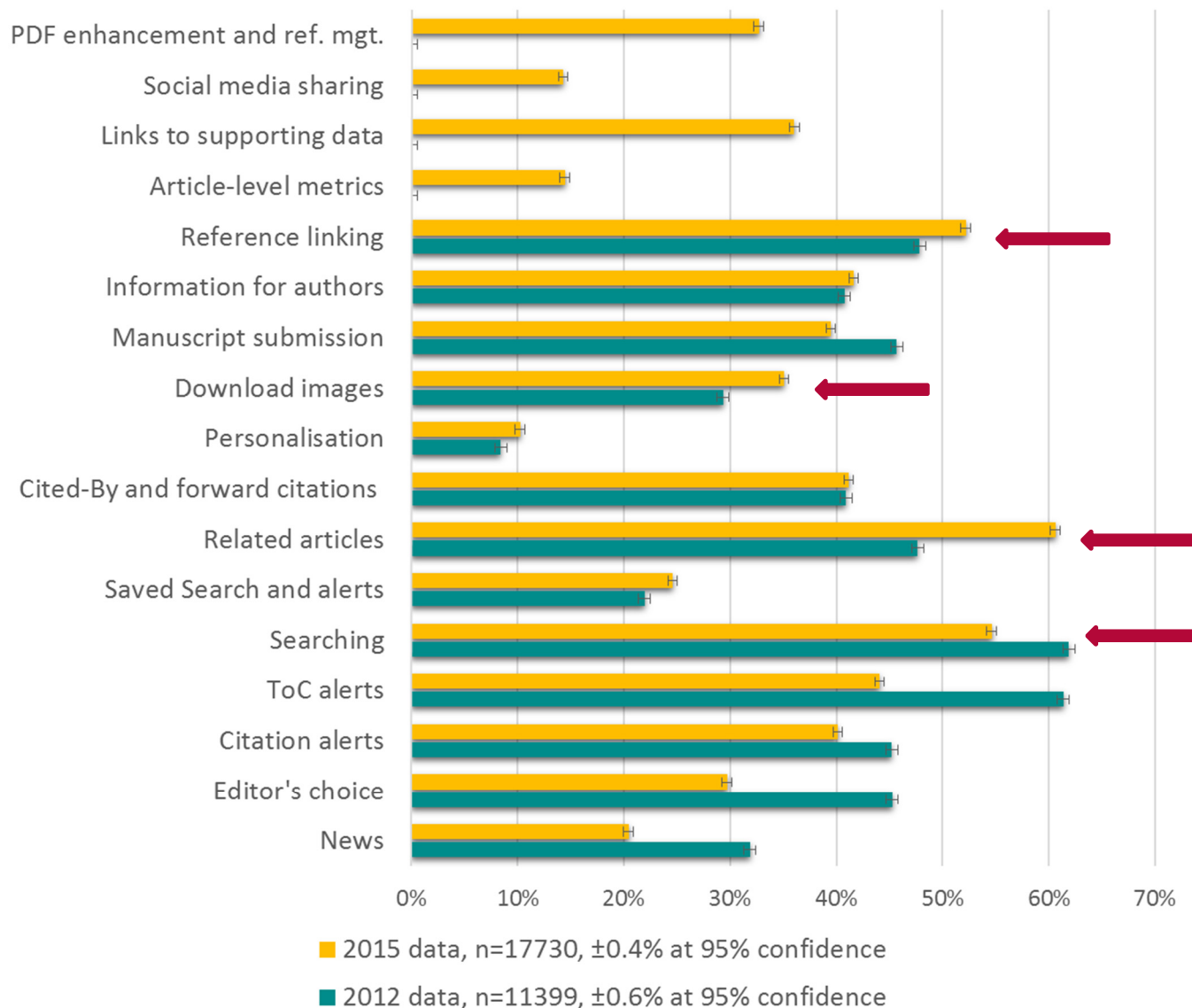


Figure 50 – Journal web site features, academic, 2012-2015

Looking at the data for just 2015 (Figure 51 and Figure 52) we can delve a little deeper into the differences in demographics.

People in lower income countries value PDF enhancement tools far more than people in in high income countries.

People in higher income countries value ToC alerts and Cited by information more than people in lower income countries – this could perhaps be because of a lack of understanding about what these features are.



People in higher income countries value links to supporting data more than those in lower income countries.

What features do you find useful in publishers' journal web sites? Variations by World Bank income classification for academic sector.

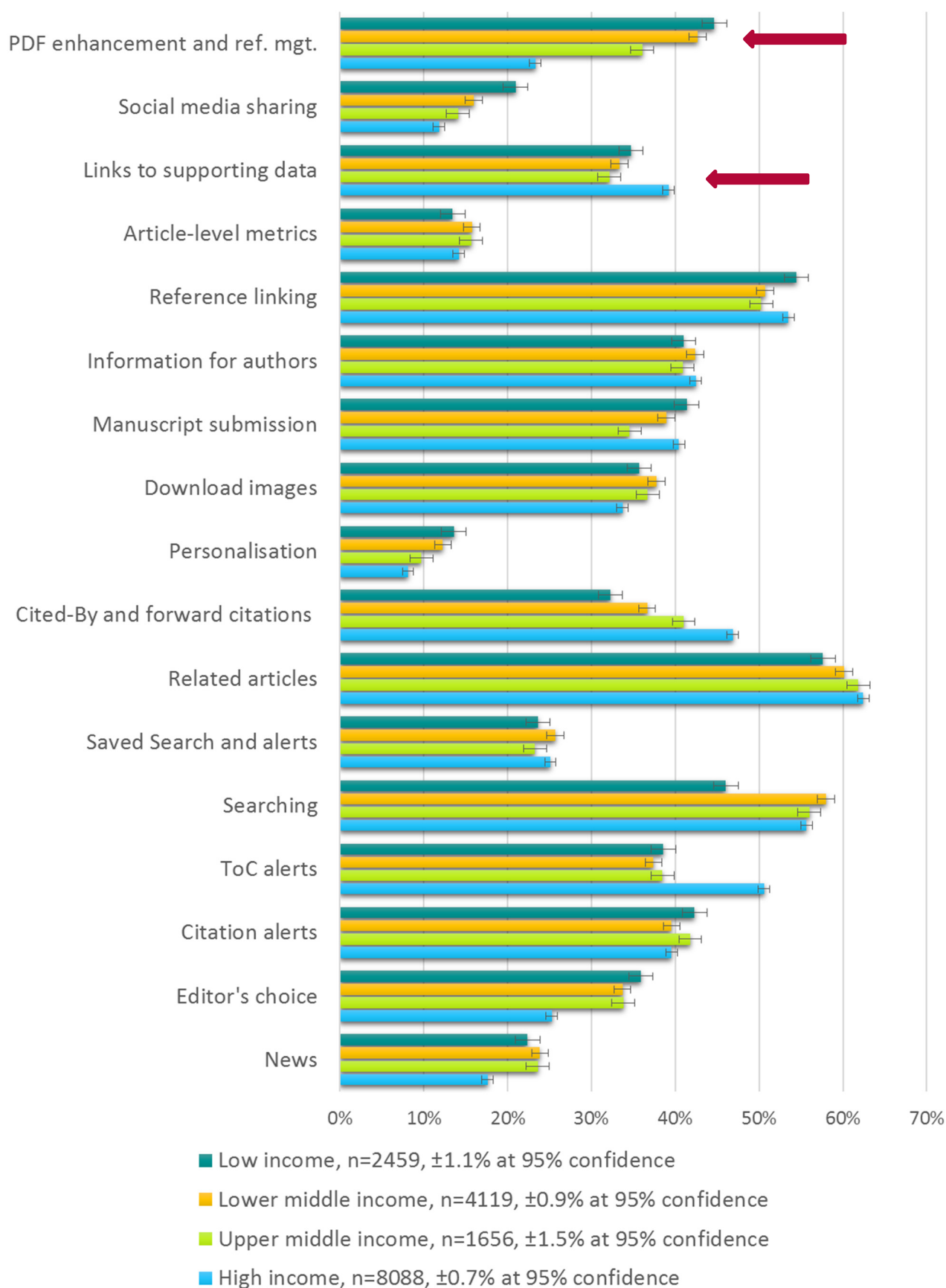


Figure 51 - Features, income, academic, 2015

Figure 52 shows the usefulness of features across subject areas. We can see there are many significant differences. People in Chemistry and Life Sciences are far more interested in links to supporting data than people working in Engineering and Humanities.

People in Humanities still value ToC alerts more than people in other subject areas, perhaps because continuous publishing is not as prevalent in Humanities as the sciences.



It is interesting that in all measures and demographics, we have not seen any significant appreciation of the availability of social media sharing or article-level metrics, even though we know that most publishers feel that these are essential features.

Ensuring Discoverability for Your Success

Librarians

Assured content
discoverability for
Patrons and Users



INASP

Research4life

EBSCOhost Discovery Services

Ex Libris Primo

WorldCat Local

Google Scholar

Anywhere Article

Content Alerts

CrossRef

Web of Science

PubMed

RSS Feed

Altmetrics

ReadCube

Authors & Researchers

Ensure discoverability
of your research



Wiley Author Toolkits

Kudos

ORCID

Publons

What features do you find useful in publishers' journal web sites? Variations by subject area.

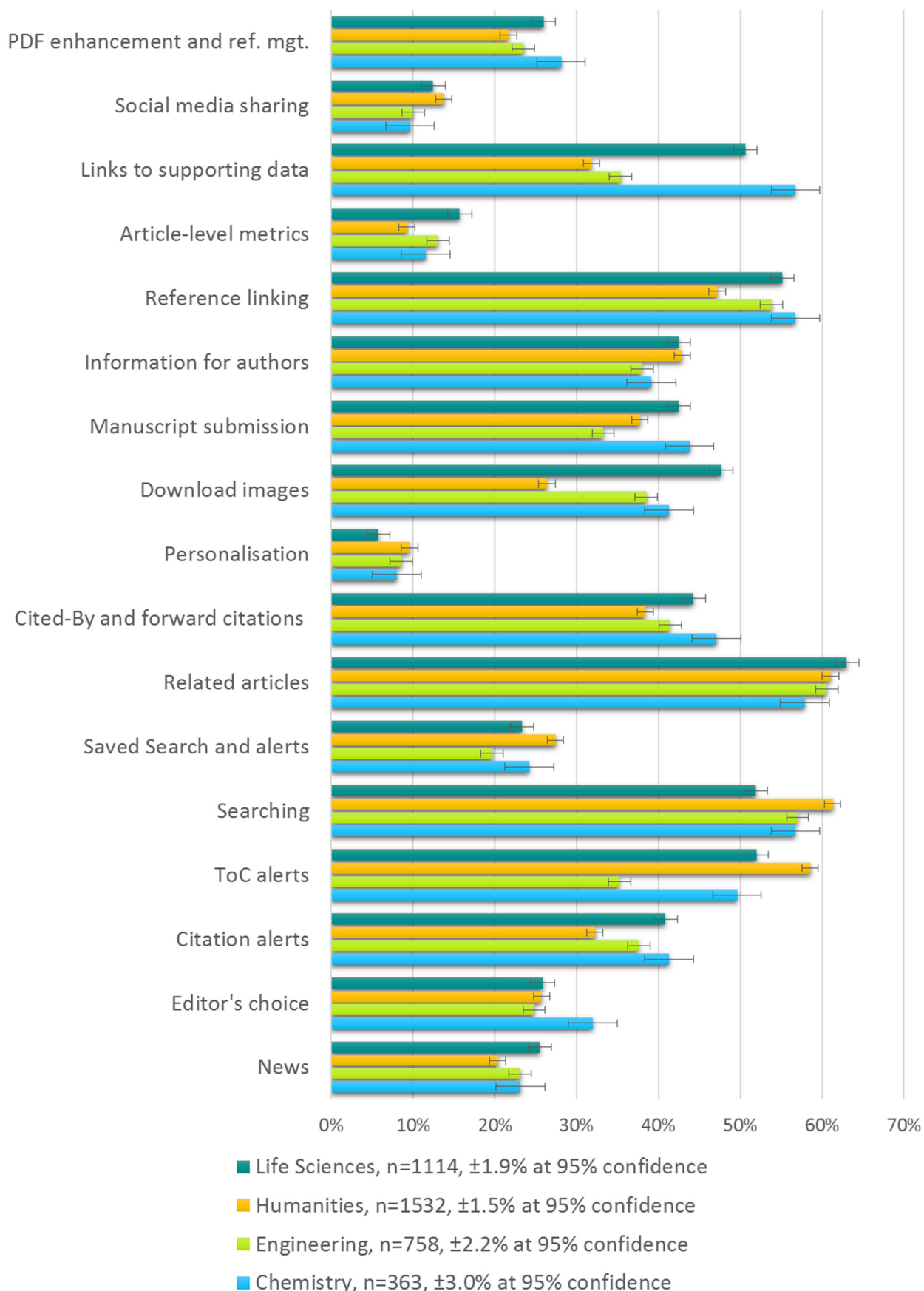


Figure 52 - Features, subject, 2015

8. METHODOLOGY

This research carries on from, and expands upon, previous research undertaken in 2005, 2008, and 2012 (also by Simon Inger and Tracy Gardner) and attempts to follow the trends in behaviour over that period of time. Naturally, each time the survey is repeated, the authors have sought to keep the questions as consistent as possible with the questions in earlier surveys whilst keeping terminology current and tracking new developments. For this reason, the three key questions on reader behaviour were modified a little, some options reclassified and additional options created. However, since those questions don't limit how many starting points the reader acknowledges as being important, this approach should have minimal impact on the results for any option present in the survey all the way from 2005 to 2015. No questions have been dropped since the research in 2012.

One of the key limitations of the older surveys was their reach. However, both the 2012 and 2015 surveys have reached very significant numbers of individuals. As a consequence, demographic breakdowns are possible, by region, by subject, by major countries, and by World Bank income classifications. Our primary goal was to give us sufficient responses within each subject category to make for meaningful comparisons. We identified the subject areas we needed for our study and created a list of publishers and intermediaries who had content in those subject areas. We approached them over a period of six months until we had sufficient organisations as supporters to give us the best chance of reaching the numbers we needed.

Of course, since the invitations to take the survey were sent out by many publishers, some individuals will have been invited to take the survey more than once. At no point did we have sight of any email addresses from any of the supporting organisations and so no de-duplication was possible. However, we attempted to mitigate the effects of individuals taking the survey more than once by careful wording in the survey invitation – the standard invitation pointed out that if the respondent had received the survey invitation more than once, it was not intended as a prompt to take the survey twice, nor an indication that any previous response had not been received. In addition, the survey was incentivised, with three prizes of \$100 Amazon vouchers or equivalent, but it was made clear that duplicate prize draw entries would be ignored.

SAMPLE

For the reasons outlined above the sample used for the survey was not a random sample – surveyees were selected by our supporters who themselves were selected by their likely subject coverage. In addition, the contacts that each publisher used for the survey will be quite highly engaged with the publisher or intermediary – all of the contacts used will have *opted-in* to receiving emails of this type. The supporting organisations are based in the US, UK, France and The Netherlands, but their clientele are truly global, and the regional demographic breakdowns are testament to this.

SURVEY DESIGN AND RESPONSE RATE

To ensure the highest response rate the survey was limited in size. We timed the survey to take about 9 minutes and we advertised it as such.

40443 people embarked on the survey, and 39378 of them completed the demographics stage. At the end of the demographics, 36562 people elected to take the journals survey, and 2816 elected to take the books survey directly. Of the 36562 people who took the journals survey, 29991 of them completed the section on journals. At the end of the journals

part, 27656 people elected to continue to take the books survey too. This gave us a total of 30472 starting the books survey, and 29110 of them completed it.

In summary, therefore, 29991 people completed the journals survey (74%), and 29110 people completed the books survey (72%).

Each survey partner used a similar wording for their email invitations. Some used additional methods and media to reach potential surveyees and used various samples of their contact databases. As such, response rates are not always directly comparable. However, the approximate response rates were between 1% and 3%. Generally speaking, publishers who sent their invites out later in the process received a lower response rate, probably because their target audience had already been invited to the survey by another supporter.

BIAS AND LIMITATIONS

Of course any survey is limited in what it can achieve. We tried not to be leading in our questioning but there are always limitations in the language used. In addition, although the audience was international, the survey was only conducted in English and so the interpretation of the language may be a factor in the accuracy of the responses. However, it should be noted that the surveyees will mostly consume journals and books in English.

The survey used invitations from our supporters, which were not necessarily a completely representative sample of the world of journal and online book readers. Furthermore, due to data privacy/data protection rules, all those invited to the survey via email will be quite highly engaged with the publisher and have *opted-in* to receiving emails like these. This may bias their attitude to publisher web site features, or indeed their likelihood of having bookmarked a particular journal page. It may be, therefore, that external discovery platforms are even more important than portrayed in these results.

No control sample was used in this survey. With so many variables, of subject, region, job role and sector, it was decided that creating a control sample would be beyond the scope of this research project.

ANALYSIS

Analysis was performed using Excel. Many of the results were reduced to binary arrays to make further comparisons easy to calculate. A tool was developed so that a number of comparisons could be made for each survey question easily. Figure 53 shows how the tool works, creating 5 series of data, each as a result of multiple demographics.

ERROR CALCULATIONS

Throughout this document we have tried to keep demographic breakdowns of the data to quite large samples so as to minimise errors. All the charts include error bars calculated at a 95% confidence interval, which of course in itself brings assumptions about the nature of the distribution of answers. The reader should be aware that with average luck, one in twenty of these calculations will be insufficient and the true reading will be out of the boundaries shown in the charts. They may not be far further adrift, but as with all survey results, the reader needs to exercise caution in interpreting any certainty in the outcomes shown.

For simplicity, the error bars shown for a given data series are the same. For example, if one measurement representing 30% of the responses carried with it an error of $\pm 5\%$, and another measurement of 20% of the responses carried with it an error of $\pm 4\%$, the chart would be plotted with both errors at $\pm 4.5\%$. Therefore, some of the error bars for “popular”

results will be slightly overstated. In reality, the plots in this report generally carry much smaller error ranges than in this example, and the differences between them would therefore be much smaller still.

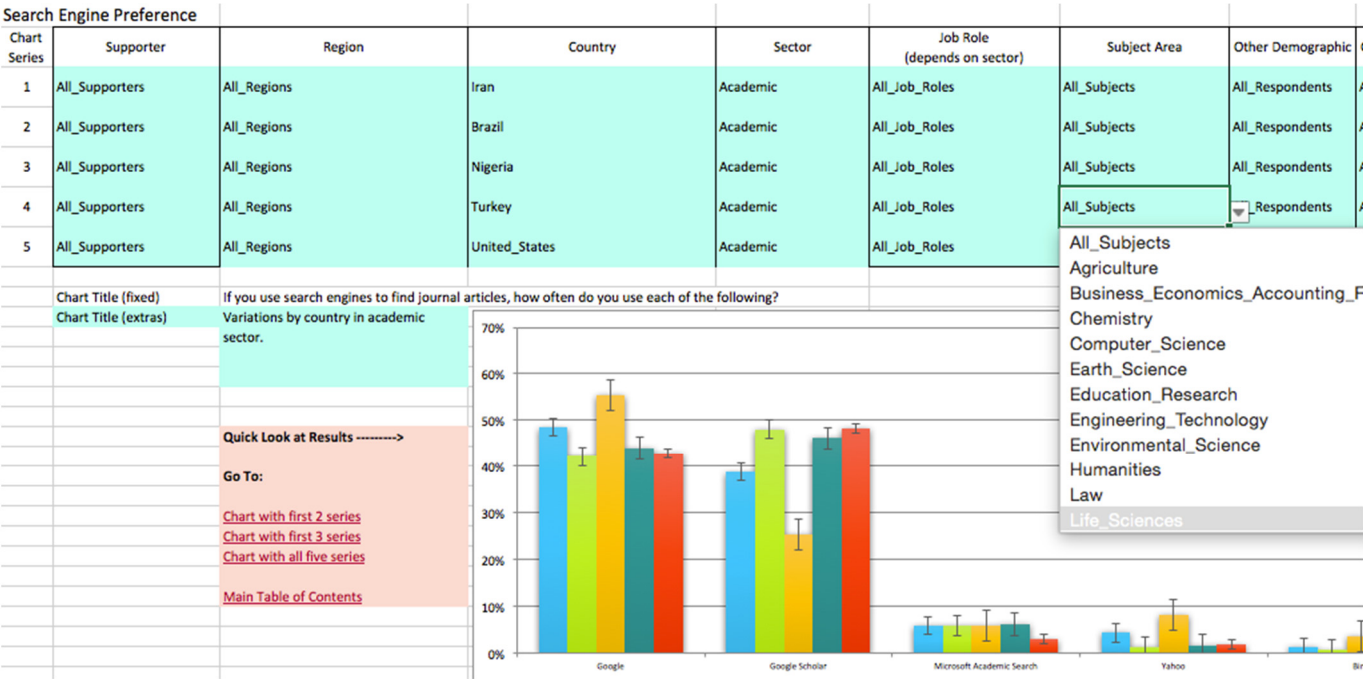


Figure 53 – Screenshot of analytical tool

In some of the charts, with multiple breakdowns of the data, the error bars shown are those for the series being analysed with the fewest responses. This means that the errors for the remaining breakdowns are overstated. The reader should make a mental note of the response numbers 'n' clearly shown in each chart in making a judgement about the accuracy of the data. In these charts the actual error bars will be smaller than those shown for the great majority of the data points – a very cautious view of error has been used in these charts.

COMPARISON WITH 2005 AND 2008

The survey, as conducted in 2008, was formed as similarly as possible to the original survey conducted in 2005. In both cases an invitation to take the survey was emailed to a large selection of readers of the supporters of each of those items of research. As already noted, it was imperative that the survey used language and terms as similar as possible to the prior research, that the medium of collection was the same (online survey) and the temptation to add greatly to the survey was resisted.

In 2005 and 2008, invitees to the survey were taken solely from those who had signed up for ToC alerts. It was noted in those surveys that this potentially affected the popularity of ToC alerts as a starting point in the findings. In 2012 and 2015, this bias is somewhat removed (although the exact extent we cannot know), and this seems to be apparent in some of the results.

Between 2005 and 2008, there was a shift in response demographic towards Life Scientists in North America, away from other demographics. In making those comparisons these shifts were taken into account. In comparing the data with the 2012 and 2015 data we have created a random sampling of the respondents to the 2012 and 2015 data down to the 2008 levels. This random sampling included a probability calculation which created broadly the same demographic breakdown in the 2012 and 2015 “down-sampled” set as in the 2008 set. For

example, approximately 1 in 30 of the responses of those in Humanities were used, and 40% of them would be in North America. Overall the method seems to have created a set of data with similar demographics as the 2008 set, but not identical (of course).

The trend information is not available for all of the questions asked in this survey, because a number of them are new for 2012 and 2015, and some of the discovery platform options were introduced for the first time in 2012 also (such as Mendeley). Also some of the questions asked in 2005 and 2008 are now considered redundant – the practices that they recommended are now part of received wisdom.

DISCOVERY RESOURCE CHARACTERISTICS

Readers have a wide choice of where to undertake content discovery. This survey asked readers to rate the importance of a range of starting points for discovery in for several different use cases. For the time being most of these starting points are reasonably distinct, but the authors recognise that some of the boundaries between different types of discovery resource are being blurred, and will continue to be so in the further. All of the potential starting points for discovery have clear differentiating USPs for the reader, as discussed below.

ABSTRACTING AND INDEXING SERVICES OR SPECIALIST BIBLIOGRAPHIC DATABASES

Major subject A&Is – e.g. Scopus, PubMed, Web of Science, focus on structured access to the highest quality information within a discipline. They typically cover all the key literature but not necessarily all the literature in a discipline. Their utility flows from the perceived certainty and reassurance that they offer to users in providing the authoritative source of search results within a discipline. However, they cannot boast universal coverage of the literature – they provide good coverage of a defined subject niche, but reduce the serendipitous discovery of peripheral material. Also, many A&Is are sold at a premium, which in itself is a barrier to their use. Examples from a wide range of subjects were given in the survey questions to help surveyees understand this classification.

LIBRARY WEB WAGES

Many libraries invest significantly in the development and utility of their web pages as a starting point for discovery. Library controlled web space usually has the advantage of linking only to content that has been paid for by the library and meets library selection criteria. The library's deployment of link resolver technology has further strengthened their importance. Not only are libraries now the primary purchasers of content for their staff, researchers and students, they are also, where link resolver and associated technology has been deployed, the main determinants of how different, relevant resources are presented and offered to end users; the way in which the user navigates to a publisher site; and also what part of the site the user is delivered to.

A JOURNAL COLLECTION, OR AGGREGATION

This survey studies the importance of aggregators in journal discovery, though recognising that this may be a closed environment – discovery in such a database can lead only to content in that database. Aggregated databases are normally a separate incarnation of a collection of journal content licensed to an aggregator and sold and delivered independently to a publisher's primary content incarnation.

A COMMUNITY WEB SITE, SOCIAL NETWORKING SITES

The market for community web sites has consolidated in recent years to a few standout sites, such as Mendeley, ResearchGate, academia.edu, which offer broader coverage than services that are no longer with us like UniPhy and Redcat. In addition, Twitter, Facebook and Wikipedia are places where links are shared and citations embedded for others to follow.

A PUBLISHER'S WEB SITE

Publisher web sites, of course, contain only a fraction of the available literature in a given subject area, unless that publisher has almost complete dominance of the subject area under consideration. Despite this clear limitation when searching for new content, the size of these publisher collections and the often superior interface design make these sites appealing to users, even though information specialists would advise against using them in lieu of larger search collections.

EMAIL BASED ALERTS

Email based alerts, whether table of contents alerts, or saved search alerts have an advantage because they are under user control, and most likely are set up for content that the user knows he has access rights to use. By definition, the resource has already gained the user's trust.

THE JOURNAL'S HOMEPAGE

In many ways, the journal's home page is not dissimilar in characteristics to email based alerts. A user has bookmarked a journal home page presumably because he has experience of the journal, has access to it (a subscriber) and has respect and trust for it.

GENERAL WEB SEARCH ENGINES

The main strengths of search engines such as Google are their simplicity, broad coverage and the fact they are free to use. Their speed allows for search to be refined and retrieved quickly and is a frequently cited reason for their popularity. Google changed its policy on the indexing of pay-walled content since our 2012 survey, so its apparent coverage to many users will have decreased.

ACADEMIC SEARCH ENGINES

In the earliest studies we did not separate out the use of academic search engines, such as Google Scholar, from general ones, like Google. But with the advent of a change in Google's policy of the indexing of pay-walled content the relative popularity of this class of engines deserves its own category. These search engines achieve some measure of quality by selection and the addition of citations to results is a clear differentiator over the general search engine.

A SCHOLARLY SOCIETY WEB PAGE

Society web pages have much the same appeal as a journal homepage. Society members usually have access privileges to the society journals through the site. One presumes, though, that the brand affinity for members with the society is even stronger than with the journal.

One of the survey questions tackles the question of content delivery. We asked 'What proportion of the journal articles that you read do you access from each of the following resources?' Surveyees were given a range of delivery site types and asked to simply say whether they used those sites more than 70% of the time, between 30% and 70% of the time, less than 30% of the time, or not at all. Naturally, some people voted for more than one item as more than 70%, or more than 4 items between 30% and 70%. To mitigate against this effect, we took the following approach.

For each response we assigned an approximate figure, of 75%, 35%, 15% and 0% to each of the four answers, added up the total of the percentages and created a correction factor for each person. For example, if a person's answers added up to 200%, we divided all their responses by 2, i.e. the correction factor. We calculated a standard deviation figure for the correction factors. Through an iterative method we altered the assumed percentages above to reduce the standard deviation of correction factors to its minimum value, and so created a best fit of how to interpret the range of percentages given. These are given below:

Survey Response	Used Figure
More than 70%	71.0%
Between 30% and 70%	38.1%
Less than 30%	9.9%
Not at all	0.0%

With each response given the percentage above, the correction factor was reapplied to each respondent's answers so that no-one could account for more than 100% of their time.

Answers from this question provide valuable insight into readers' delivery options and so we wanted to test the effect of different derivations of these figures on the final published results. We found that any reasonable combination of numbers in the above table consistently produced the headline results reported in this work, and although changing the sizes of each result in absolute terms, never affected the relative ranking of results, and critically, never placed the use of free resources at less than the publisher-controlled resources.

9. DEMOGRAPHICS

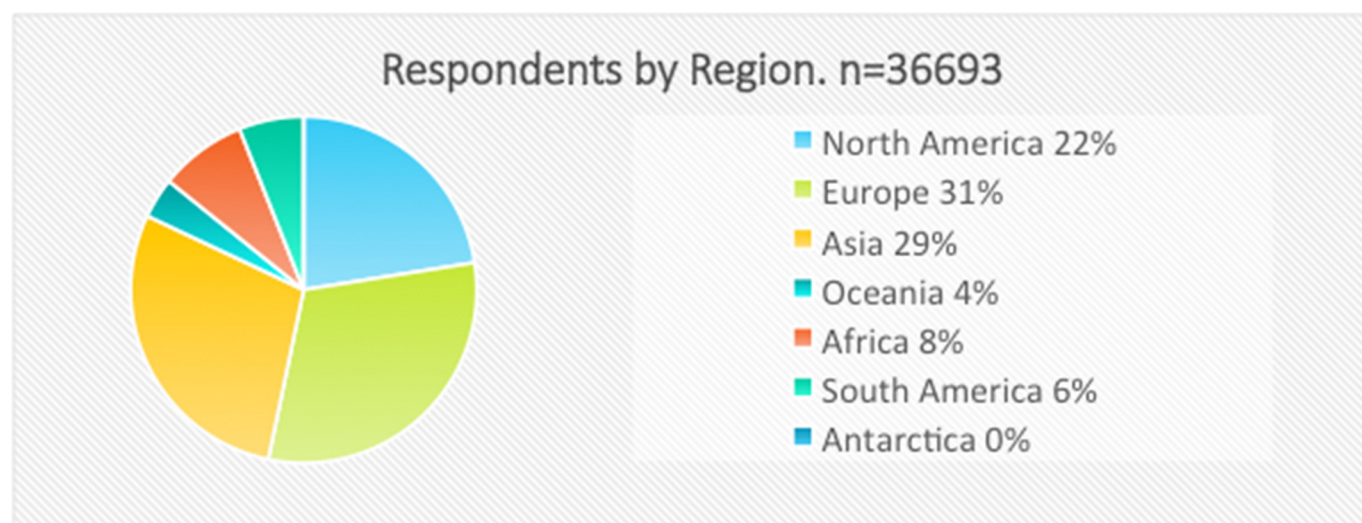


Figure 54 – Respondents by region

As shown in Figure 54, the respondents to the survey come from an excellent regional spread and this allows for significant regional (and in some cases country) breakdowns, see below. 36698 people told us their country (and hence region).

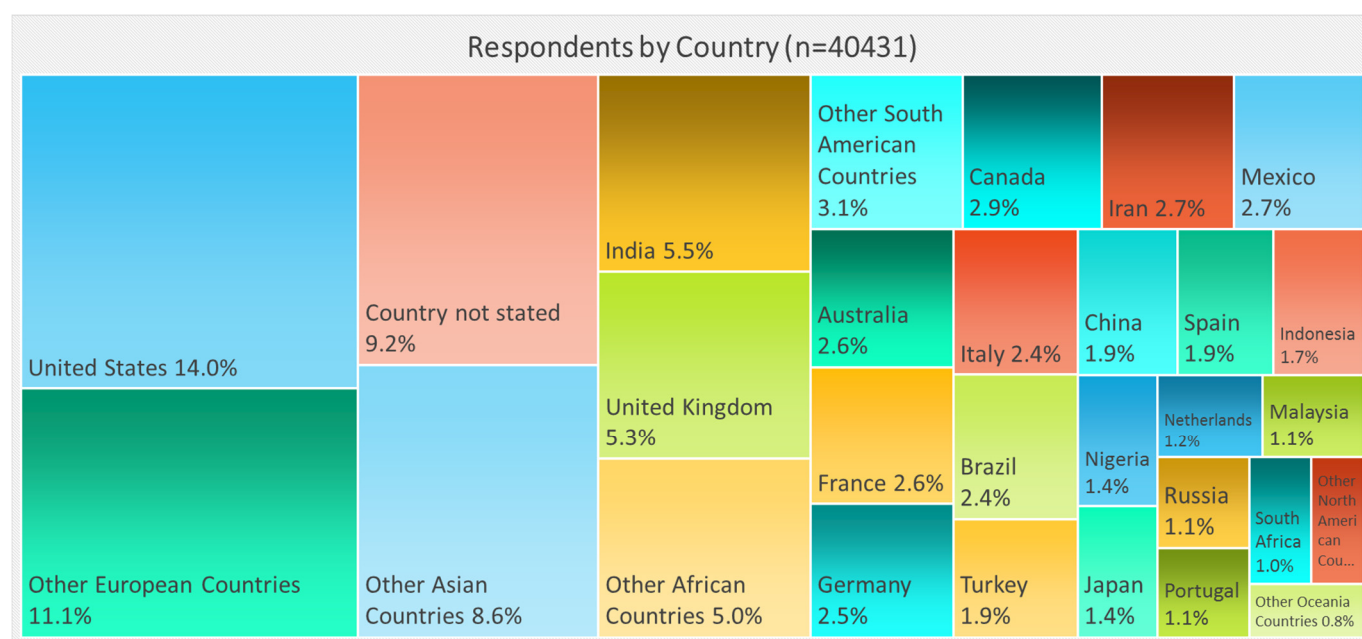


Figure 55 – Respondents by country of origin

For the purposes of being able to apply demographic breakdowns by country for all of the survey questions, we separately identified all countries with more than 400 respondents and then further clustered all other respondents within a region as one. In Europe this means we can analyse the behaviour of individuals from UK, Germany, Italy, France, Spain, Portugal, and The Netherlands and all of the other European respondents are grouped together as "All Other European Countries". The same approach was taken with all other Continents. This year's survey showed a reduced proportion of respondents from China than in the 2012 survey, but the absolute number of respondents from China is similar to the 2012 number.

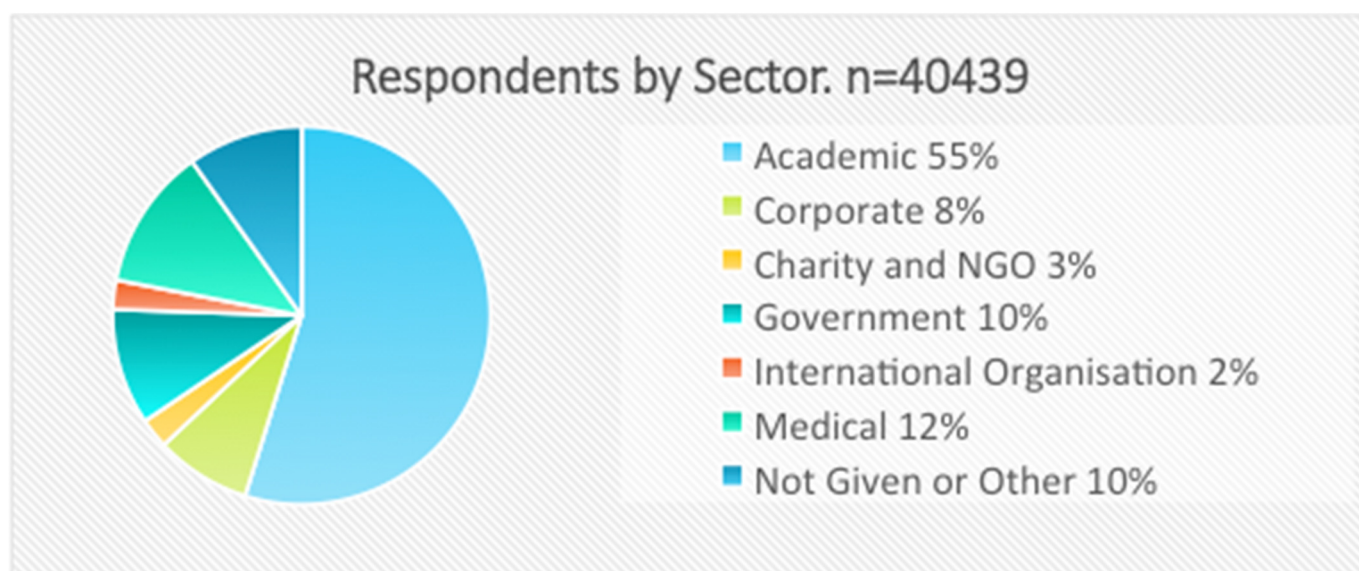


Figure 56 – Respondents by sector

The sector-breakdown is also very good, with sufficient numbers in all but Charity/NGO and International Organisation to allow for further breakdowns by subject, region, and income. 40439 people told us which sector they worked in. Although the academic sector dominates, even the 8% of respondents from the Corporate sector represents over 3000 people.

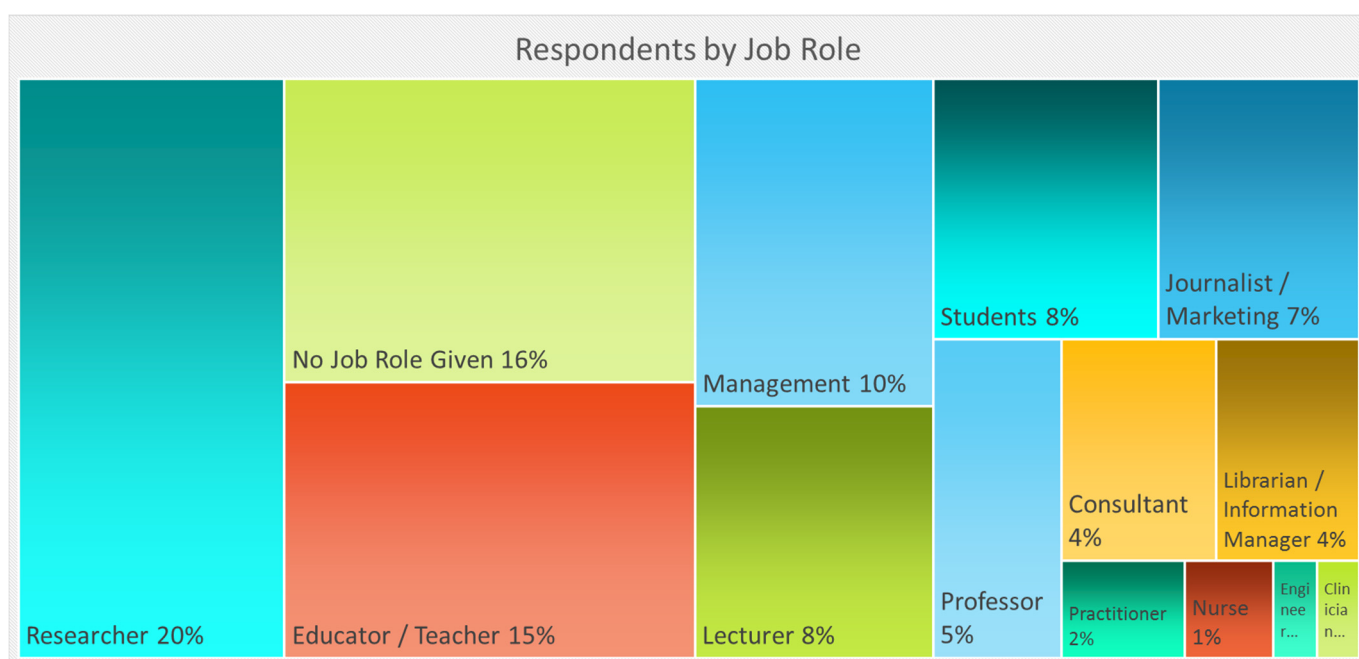


Figure 57 – Respondents by job role

The numbers of respondents by job role allows for meaningful comparisons to be made for almost all roles, and for most of the roles there are sufficient numbers to permit meaningful analyses in combination with other demographics such as region or country.

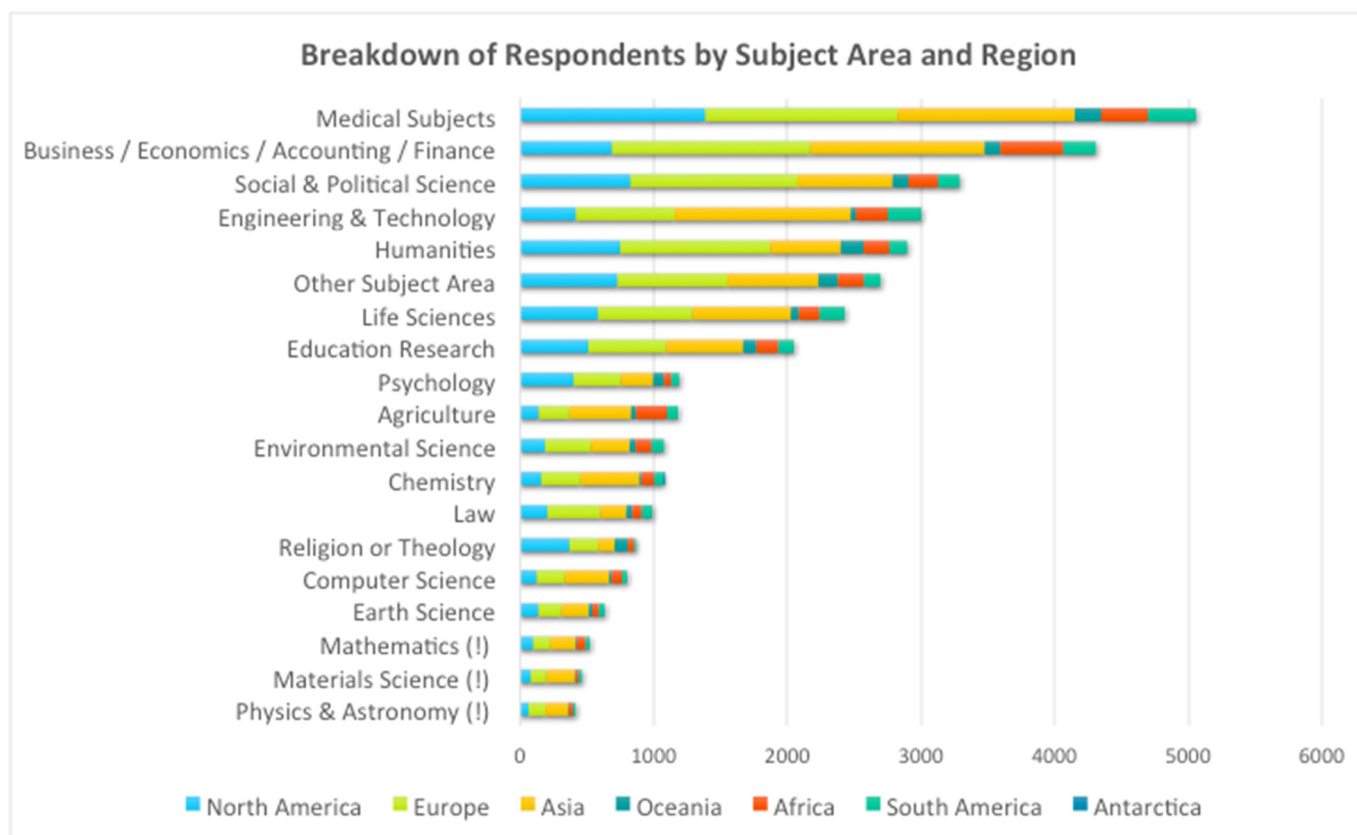


Figure 58 – Respondents by subject area and region

When embarking on this project, our aim was to get around 1000 responses in each subject area, so that detailed demographic analysis would be possible within subjects with minimal error-bars. This was achieved for twelve of the subject areas, and three more subject areas got over 600 responses, which still allows for some useful further sub-division. Materials Science, Physics and Mathematics (marked with '!') are the least well-represented in the data, although in absolute terms there are enough individuals in these areas to allow for a useful subject-based analysis, even if not when combined with a further demographic, other than perhaps the Academic sector demographic, or the Asia demographic.

Within these subject areas we have achieved a useful regional split, although we do not have any corroborating evidence to show whether or not these are in any way in proportion to the numbers of individuals in each subject area in each of the regions. The relative response by country income within a given subject is shown in the chart below. It shows that Agriculture has a strong presence in poorer nations, while Humanities, Psychology, Religion & Theology are more confined to wealthier nations.

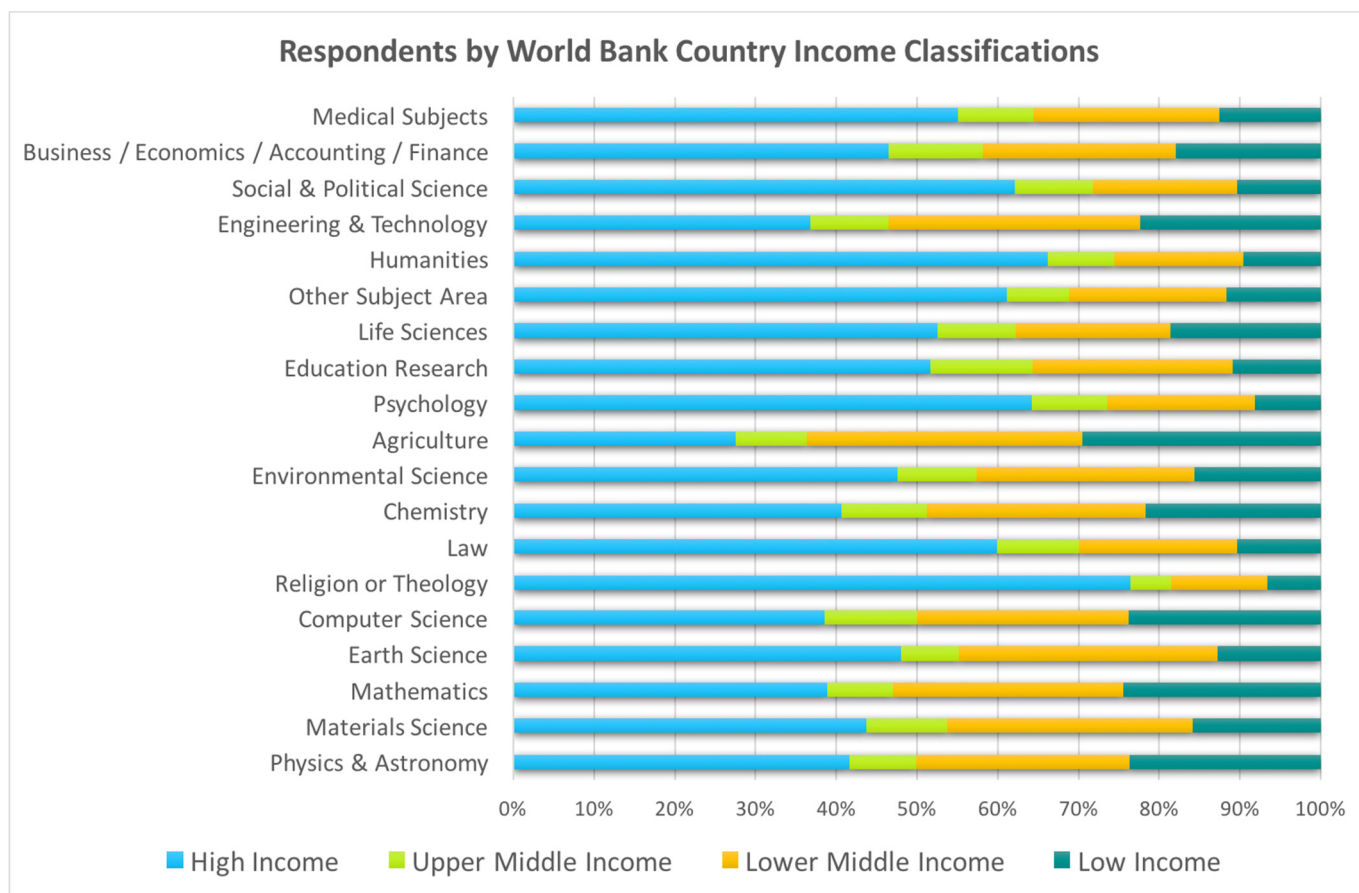


Figure 59 – Respondents by World Bank income categories and subject (2011)

The individuals' countries were mapped onto World Bank income categories which are themselves calculated as GNI per capita: low income, \$1,025 or less; lower middle income, \$1,026 - \$4,035; upper middle income, \$4,036 - \$12,475; and high income, \$12,476 or more, as measured in 2011. We used the 2011 classification rather than the updated ones to allow for comparisons across time – our current data structure does not allow for changes in income classification of countries in our longitudinal analyses. Of greatest significance to this survey is that Argentina and Hungary are now both classified as high income countries, but results from these countries will still be found in the survey in the upper-middle income classification.

10. CONCLUSIONS

Content discovery remains a highly competitive space with no clear winner – people use a wide range of discovery methods and tactics for perfectly good and logical reasons. There is no single “right” answer.

A key differentiator in behaviour appears to be the awareness of the range of resources available, evidenced in many comments. This awareness seems to be strongest in the academic sector in wealthier nations where presumably information literacy training is having a positive effect, coupled with the availability of a wide range of discovery options. In the corporate and charity sectors, even in the developed world, readers seem to show a lower level of search sophistication; for example, relying on general search engines, even in the face of free scholarly discovery resources. For some, of course, the limited availability of content resources dictates an alternative discovery and delivery strategy.

The investment and improvements in both library discovery services and publisher websites seem to be paying off. Both of these categories have seen rises in popularity, especially in the academic sector, since 2012. At the same time the Abstracting and Indexing databases (A&Is) seem to have lost some ground, whilst remaining, in many subjects and sectors, the most important discovery resource.

11. ABOUT THE AUTHORS

SIMON INGER

Simon Inger has been working in journals since 1987, when he joined B.H.Blackwell, the Oxford-based subscription agent. In late 1994 he founded CatchWord, the world's first journal platform service provider and ran that business until its acquisition by Ingenta in 2001 (now Publishing Technology). Since 2002, Simon has worked as a consultant specializing in journals, including: sales, marketing & pricing; journal delivery & platform selection; fulfilment & editorial systems selection; management; financial planning; product development; market research; content development; content discovery; library technology; standards; strategic business reviews; and open access publishing. In addition, Simon runs training courses for librarians for UKSG in Europe, as well as courses for publishers by Renew Training (a subsidiary of Simon Inger Consulting) in association with ALPSP. Simon co-founded Renew Consulting for Societies, which helps scholarly societies navigate the world of academic publishing and assists specifically with business reviews, contract renewals, writing tender documents and negotiating new partnership agreements. Simon works with not-for-profit publishing organisations in the UK and US, as well as commercial publishers, technology providers, intermediaries and service providers in journals publishing.

TRACY GARDNER

Tracy Gardner has worked in journals since 1997. She has a very broad view of publishing having worked for publishing technology companies (CatchWord and Ingenta), a not-for-profit publisher (CABI Publishing), and thereafter in consultancy, on various projects for publishers, intermediaries and libraries. Throughout her career she has been focused on improving the communication channels between publishers, intermediaries and librarians and understands the business of scholarly publishing from many different perspectives. Tracy has worked on a wide range of projects, including sales, marketing & pricing; journal delivery & platform selection; management; product development; market research; content discovery; library technology; strategic business reviews; and open access publishing. Tracy is a trainer on UKSG's E-Resources Technical Update course, and its Licence and Negotiation Skills course, as well as courses for publishers by Renew Training in association with ALPSP.

Read the Report? Want to know more?

Then unlock the detail with the full dataset and analytical tool.



Find out about the discovery habits of over 60,000 researchers; students; teachers; lecturers; professors; journalists; managers; clinicians; medics; librarians; government officials; engineers; and others, working across all sectors and in all regions, including high, middle, and low income countries.

And if you do not have time to analyse the data, but you do have questions that need answering, then commission your own bespoke report – we can cover a single sector, country, region or subject area, or compare a number of demographics.

Find out more at sic.pub/discover or email discover@simoningerconsulting.com for more information



@siox14 #howreadersdiscover

