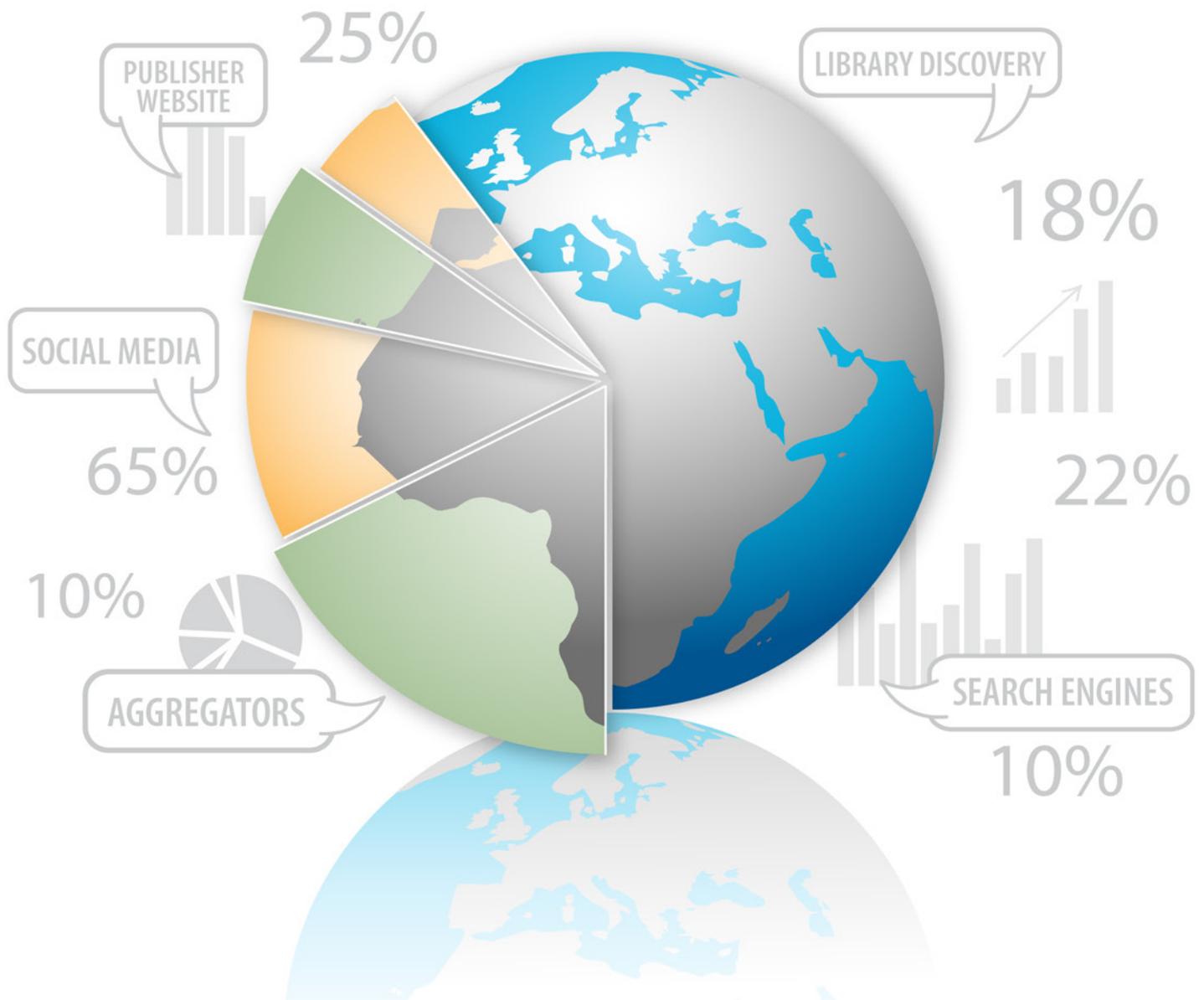


# How Readers Discover Content in Scholarly Publications

Trends in reader behaviour from 2005 to 2018



# *How Readers Discover Content in Scholarly Publications*

*By Tracy Gardner and Simon Inger*

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

Published August 2018

The full dataset can be found at <https://doi.org/10.6084/m9.figshare.7016735.v1>

The report can be downloaded at [renew.pub/discovery2018](https://renew.pub/discovery2018)

[www.renewpublishingconsultants.com](http://www.renewpublishingconsultants.com)



## TABLE OF CONTENTS

---

1.	Introduction .....	6
2.	Headline Findings.....	7
3.	Search .....	9
4.	Browse.....	36
5.	Delivery Sites for Journal Articles .....	37
6.	Mobile Devices.....	41
7.	Journal Web Site Features.....	43
8.	Methodology .....	49
9.	Demographics .....	55
10.	Conclusions.....	59
11.	About the Authors.....	60

## FIGURES

---

Figure 1 - Search, long-term demographic mix, see Methodology .....	9
Figure 2 - Search, academic sector life scientists, high-income 2012-2018.....	11
Figure 3 - Search, academic sector humanities, high-income 2012-2018 .....	11
Figure 4 - Library search, academic sector by subject, 2012-2018.....	12
Figure 5 - A&I search, by subject, 2012-2018.....	13
Figure 6 - Social Media search, academic sector, by subject, 2012-2018.....	14
Figure 7 - Search, academic researcher, 2012-2018.....	15
Figure 8 - Search, by lecturers, 2012-2018.....	15
Figure 9 - Search, variations by broad subject in high-income countries, 2018 – part a .....	16
Figure 10 - Search, variations by subject in high income countries, 2015 – part b .....	16
Figure 11 - Search by sector, 2018.....	17
Figure 12 - Search, by job role, high-income 2018 .....	18
Figure 13 - Search, by region, 2018 .....	18
Figure 14 - Search, by income classification, 2018 .....	19
Figure 15 - Search engines, high-income countries, by sector, 2018 .....	20
Figure 16 - Search engines, academic sector, by country, 2018 .....	21
Figure 17 - Google vs Google Scholar by subject, 2018 .....	22
Figure 18 - Google vs Google Scholar by subject, 2015 and 2012 .....	22
Figure 19 - Search engines, China, 2012-2018 .....	23
Figure 20 - Last article accessed, 2012-2015.....	24
Figure 21 - Last article accessed, chemistry, 2012-2018.....	24
Figure 22 - Last article accessed, by income, 2018.....	25

Figure 23 - Last article accessed, by broad subject, 2018 .....	25
Figure 24 - Last article accessed, by sector, 2018.....	26
Figure 25 - Book search by sector, 2018 .....	28
Figure 26 - Book search, academic sector, job role, 2018 .....	29
Figure 27 – Last book accessed, sector, 2018 .....	30
Figure 28 - Searching for video, high income by broad subject area, 2018 .....	31
Figure 29 - searching or video by sector, 2018.....	32
Figure 30 - last video accessed, sector, 2018 .....	33
Figure 31 - last video accessed, by broad subject area, 2018 .....	34
Figure 32 - Browse, 2005-2018.....	36
Figure 33 - Delivery sites, academic by income, 2018 .....	37
Figure 34 - Delivery, free resources vs publisher-controlled sites, 2018 .....	38
Figure 35 - Delivery site, high income countries, by sector, 2018.....	39
Figure 36 - Delivery site, high income, broad subject areas, 2018.....	39
Figure 37 - Delivery site, high income, by subject, 2018.....	40
Figure 38 - Devices, low income, 2012-2018.....	41
Figure 39 - Devices by income band, 2018.....	41
Figure 40 - Devices, medical sector, 2012-2018 .....	42
Figure 41 - Devices, job role, 2018.....	42
Figure 42 - Journal web site features, 2005-2018 .....	43
Figure 43 – Journal web site features, academic, 2012-2018 .....	45
Figure 44 - Features, income, 2018 .....	46
Figure 45 - Features, subject, 2018.....	48
Figure 46 – Screenshot of analytical tool .....	51
Figure 47 – Respondents by region, 2018.....	55
Figure 48 – Respondents by country of origin, 2018.....	55
Figure 49 – Respondents by sector, 2018 .....	56
Figure 50 – Respondents by job role, 2018.....	56
Figure 51 – Respondents by subject area and region, 2018 .....	57
Figure 52 – Respondents by World Bank income categories and subject (2018) .....	58



## 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: American Institute of Aeronautics and Astronautics (AIAA), American Psychological Association, Annual Reviews, Cambridge University Press, Institution of Engineering and Technology (IET), JAMA Network/American Medical Association (AMA), OECD, Public Library of Science (PLOS), Royal Society of Chemistry, SAGE Publishing, The Company of Biologists, TrendMD, Wiley.



# 1. INTRODUCTION

This report is the output of a large-scale survey of readers of scholarly publications (n=10977) and their behaviour in the discovery of journal articles, online books, and video. The survey was conducted during February, March and April of 2018.

As a consultancy, we work in all areas of content delivery and discovery, and it is through working with libraries that we see how obfuscated user navigation can be and why publisher analytics need to be treated with some caution. Many libraries around the world use proxies to access scholarly content, and they do this to provide a seamless navigational experience for on-site and off-site workers, perhaps to overcome limitations in their national network configuration, or just to outsource authentication management. With OCLC (with its EZProxy software and service), LibLynx, RemoteXS, JISC, Eduserv and others all providing this kind of technology, often highly integrated with the institutional single sign-on (SSO) solution, publisher analytics tools don't necessarily know even the country of origin of the reader, even though their access control system (ACS) probably will. Where publishers integrate their ACS into analytics they will have a good sense of the geographical origins of their readers, but if not, it is more of a lottery!

Reader navigation is also complicated. There are simple, straightforward journeys of course, perhaps from Google to content, or social media posting to content. But where library technology becomes involved, link servers hide from the target website (i.e. the publisher) the origin of the reader – did they start in the library technology, or perhaps within an A&I service or Google Scholar, or be working in an environment where a DOI reference still re-routes them via a library link server?

For these two reasons we have found that surveying is a useful technique to get to what at least the reader believes they are doing, even though we know surveys have their own problems, especially around perception. However, this brings with it another significant benefit of surveying: the reader perception of what they do will define how they relate the publications landscape to others, and those in a senior position will strongly influence expenditure and strategy in their institutions. Perception of reality is often more important than reality itself.

This survey builds upon previous surveys conducted by the authors in 2005, 2008, 2012 and 2015. For four key questions in the survey, it allows for longitudinal analysis over the thirteen-year span, and for a further three questions, allows for trend analyses between 2012, 2015 and 2018. 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 2018 survey includes new questions regarding discovery of video content. 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 which is why the bulk of the questions are focussed on journal content 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, Mendeley, and of course SciHub. 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.



## 2. HEADLINE FINDINGS

- In the academic sector as a whole, abstracting and indexing databases (A&Is) still appear to be the most important starting point in search. Academic search engines (such as Google Scholar) are more important than general search engines (such as Google). Figure 1, p9.
- Library discovery seems to have peaked in its importance-rating and is only holding a strong position in Humanities, Education and Social Sciences. Figure 4, p12.
- In 2012 we noted the particularly low traction of social media as a discovery tool in Humanities and Social Science, and it is in these two areas where we have seen some of the greatest increases. Some other subjects are showing a real decline indicating social media is less important as a discovery tool in some subjects. Figure 6, p14.
- Academic researchers (therefore excluding students etc.) rate academic search engines as the most important discovery resource when searching for journal articles. Figure 7, p15.
- A&Is are significantly more important in the medical and academic sector than for all other sectors, and for the medical sector, of more importance than academic search engines. Figure 11, p17.
- Librarians behave quite differently to everyone else in search, preferring professional search databases and library-acquired resources. This may point to a continued significant gap between what librarians recommend, and how their patrons behave. Figure 12, p18.
- There are significant differences in search behaviour from high to low-income countries, with people in lower incomes countries using a wider range of search resources than those in wealthier countries who seem to settle on just a few methods. Figure 14, p19.
- Only the academic sector uses Google Scholar more than Google. The corporate sector continues makes the least use of Google Scholar. Figure 15, p20.
- Google Scholar is the dominant search engine used for journal discovery in China, although the popularity of Baidu continues to grow. Figure 16, p21.
- Journal alerts have lost traction as a discovery method. Search alerts and bookmarks are used, but to the lowest extent of those resources tested. Figure 20, p24.
- Whilst searching as a discovery method dominates, it still only accounts for around 45% of people's behaviour. Around 55% of the time people found the article they needed via non-search behaviour.
- In comparison with journals, library web pages are seen as a very important discovery resource for books, now normally eclipsing A&Is, and in some sectors just as important as Google. Figure 25, p28.
- Discovery of video appears to be dominated by Google and YouTube, but medics in particular are finding relevant video content within the journals they read. Figure 29, p32.
- 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. Figure 34, p38.
- People in the medical sector are accessing journal articles from a free subject repository 25% of the time. This is significantly higher than all other sectors in high income countries, and presumably points to the success of PubMedCentral. Figure 35, p39.
- 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 the world's institutional repositories from these regions. Figure 36, p39.

- PubMedCentral is likely to account for the popularity of subject repositories in Medicine. Scientific social networking sites are not responsible for a very large proportion of the delivery of free content to readers in high income areas. Figure 36, p39.
- The growth in the use of mobile devices to access academic content has levelled off since 2012. Figure 38, p41.
- Related-articles remains the most popular feature of publisher websites. Content generated by the publisher such as news and editor's picks became more popular between 2005 and 2008, but are 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. Author-facing features have grown in popularity. Figure 42, p43.
- The features people find useful on a publisher website varies significantly across demographics. Related articles and reference linking are generally the most useful features. Social media sharing and article level metrics are not particularly popular in any demographic studied. Figure 43, p45.



### 3. SEARCH

During the lifetime of this series of reports, search has evolved greatly, and in many ways content discovery outside of the publisher silos has become quite a battleground for competing services and approaches. This survey looks at search in a number of ways. It looks at how people rate in importance a number of search options, looks at the competition between mainstream search engines, and also seeks to find out what proportion of the time people are actually doing something they would consider ‘search’ in the first place.

Content owners want to know where people start their journey of discovery so that they can optimise their working relationships with key discovery services and tailor their regional and sectoral marketing messages to suit the behaviour of the individuals in those markets.

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 librarians will 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 thirteen years, we are able to show how search behaviour has changed over time.

It is important to note that the results from 2012, 2015 and 2018 have been normalised to match the demographic sample of 2005 and 2008 where comparisons have been made over the full thirteen years (see section on Methodology). Therefore, these 13-year results show the responses biased towards (but not exclusively) researchers working within STM subjects, in academia, in the US and Europe.

### TREND ANALYSIS: JOURNAL ARTICLE SEARCH

#### TREND FROM 2005 – 2018

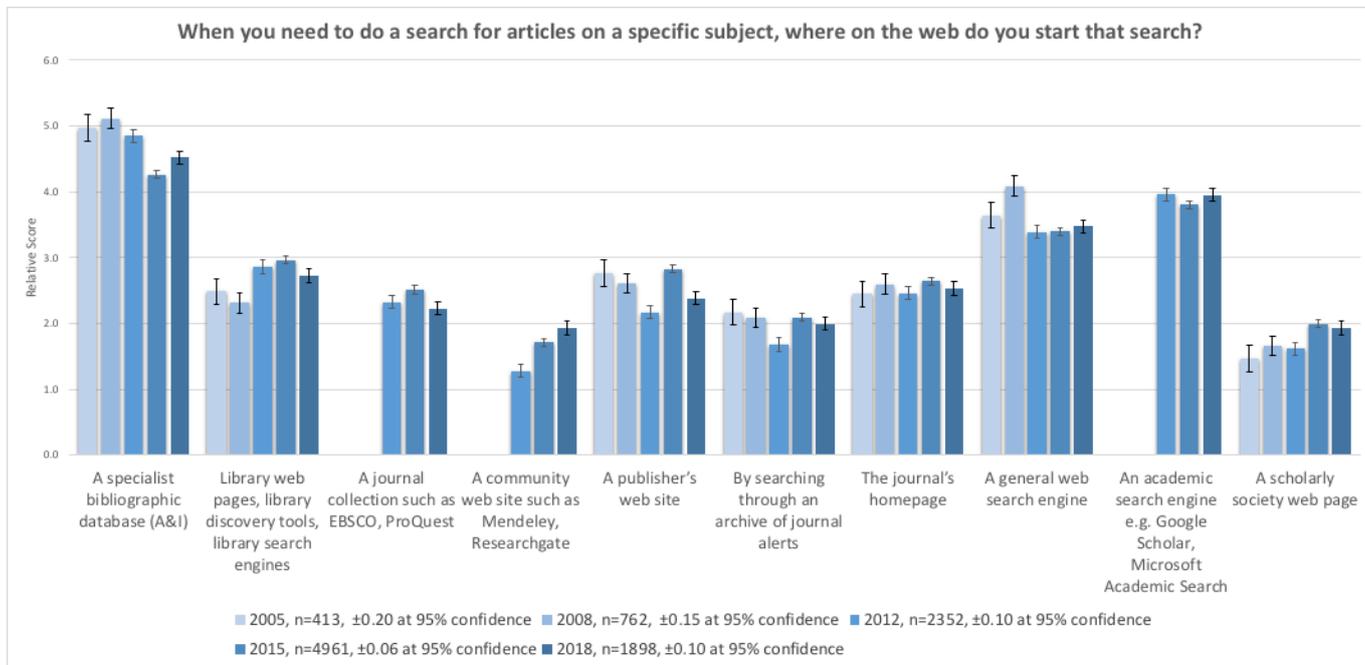


Figure 1 - Search, long-term demographic mix, see Methodology

Our methodology section describes the demographic for our 2005-2018 comparisons, which are biased to STM over HSS, and within STM biased more toward life sciences and medicine.

A key result is that A&Is are still marginally seen as the most important search resource, their importance had consistently dropped since 2008 but we see a growth in the importance of it as a discovery resource in 2018. General search engines are seen as important, but not as important as academic search engines. However, as we will see later, academic search engines are seen as the most important resource when only considering the academic researcher demographic.

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. But more detailed breakdowns discussed later in this report show that library web pages are gaining ground in some, but certainly not all, subject areas.

Community web sites, including Mendeley and ResearchGate, are acknowledged as increasingly important starting points for search, but remain in the bottom two in this survey. They show greater gains in other discovery behaviours, discussed later.

With the exception of community web sites, the perceived importance of the starting points for search seems to have stabilised and not subject to significant change.

A key message from this research is that all of the categories have their followers. Perhaps one of the oddest behaviours from the perspective of the authors of this report is the practice of some academics of archiving table of contents alerts and then subsequently searching them to reveal quickly articles that they had previously skimmed, a practice recently corroborated anecdotally by a librarian observing researcher behaviour.



*A&Is 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 in this demographic.*

In the past, academic researchers working in Life Sciences and Medicine 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 2 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 - 2018

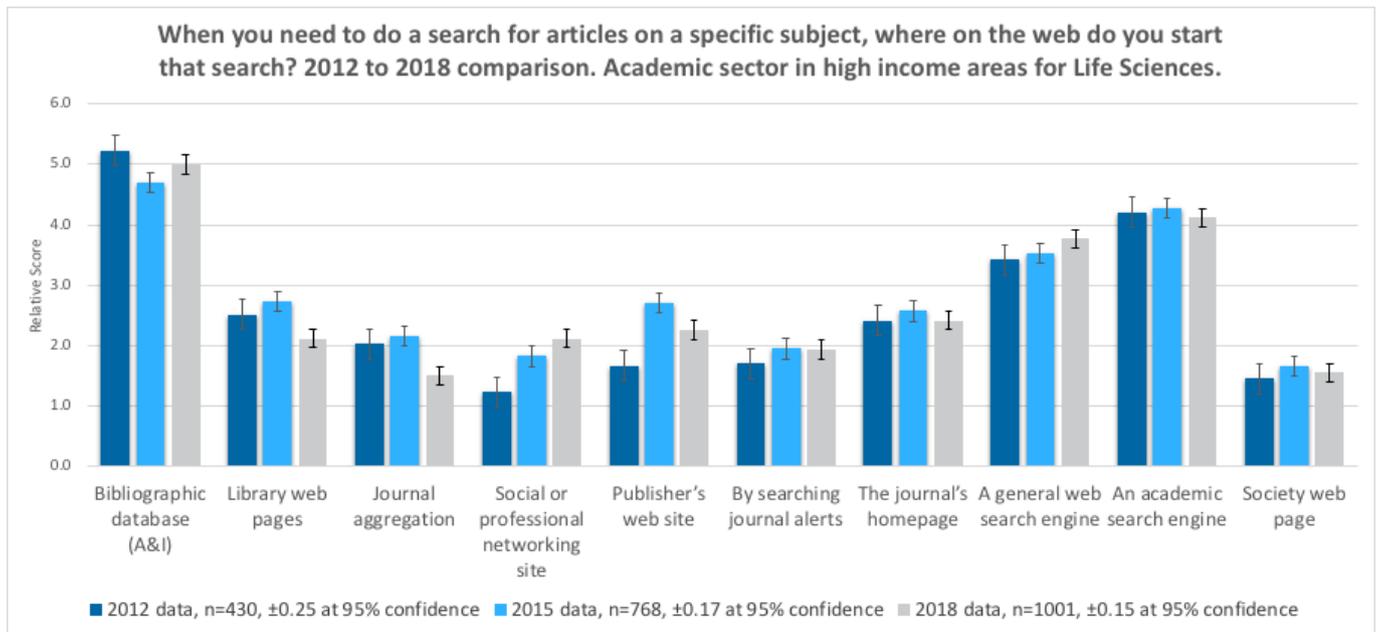


Figure 2 - Search, academic sector life scientists, high-income 2012-2018

A&Is are showing resilience and remain the most important resource. There have been significant gains in popularity of social and professional networking sites. Whilst the chart above only illustrates the results from high-income countries, we have found that the same general 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 3, we can see that the most significant change is that social media has grown as a search resource for people working in Humanities, and has begun to overtake some of the other options. More interestingly perhaps for this group is that journal aggregations have become less important and have now fallen below the level of library web pages and academic search engines.

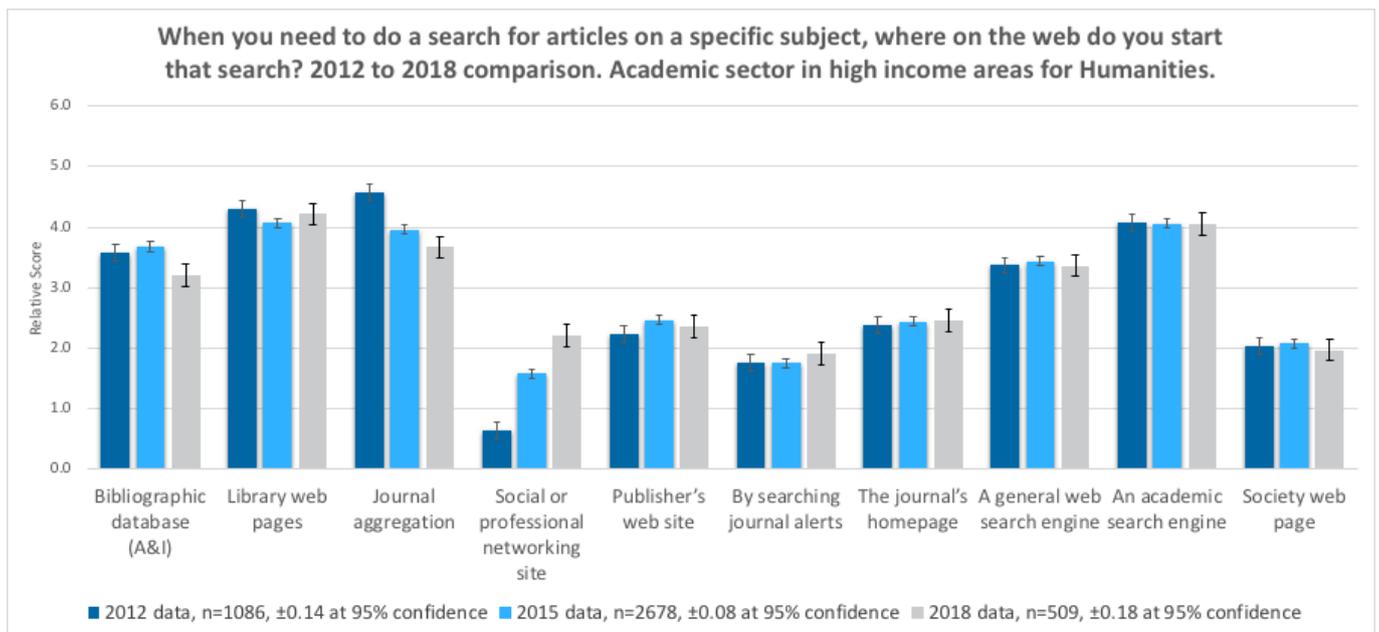


Figure 3 - Search, academic sector humanities, high-income 2012-2018

There has been much debate in recent years about the importance of library discovery technology. In Figure 4, 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 and 2018.

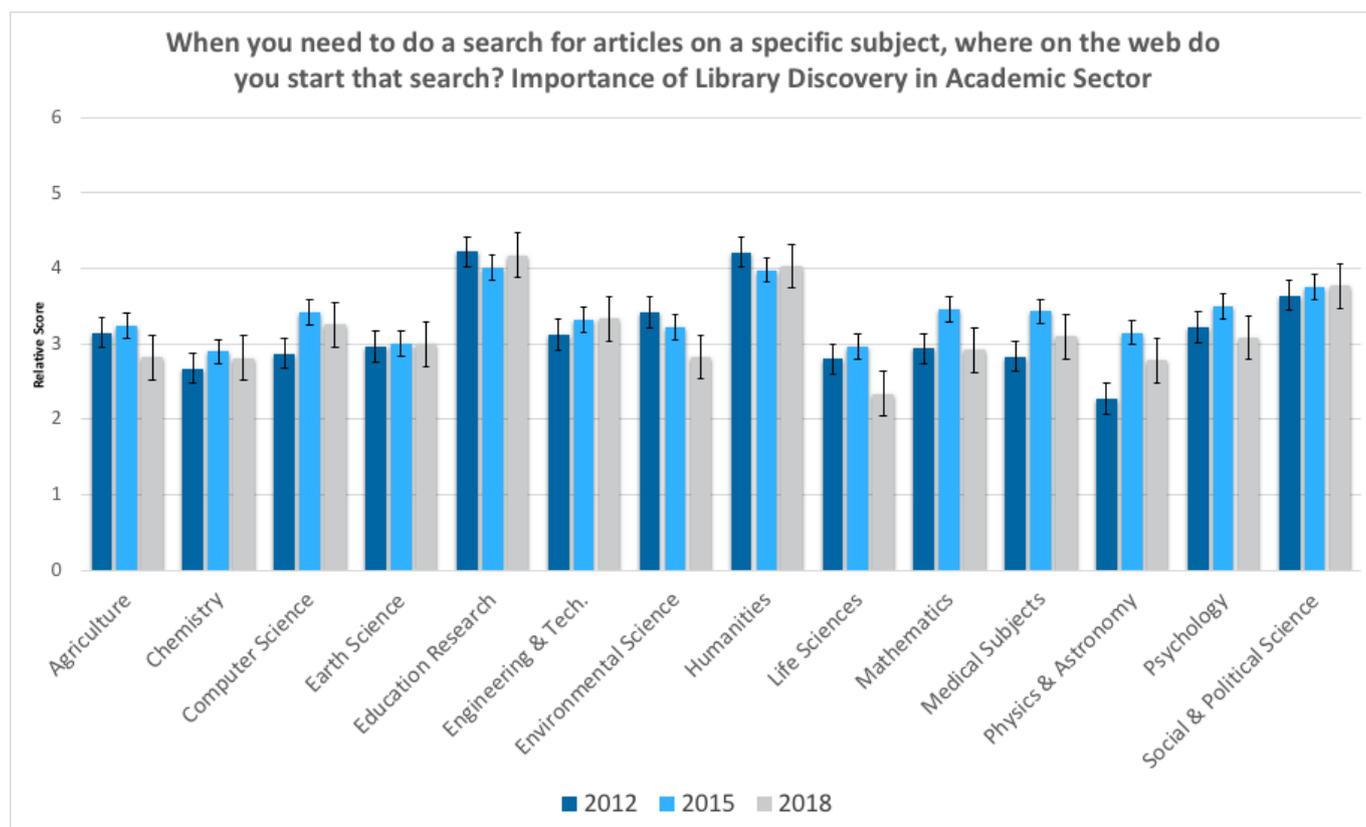


Figure 4 - Library search, academic sector by subject, 2012-2018

In absolute terms we can see the significance of library web pages in discovery in Education Research, Humanities and to a lesser extent, Social and Political Science. Many of the other subject areas have shown an indicative decline (or in the cases of Life Sciences and Maths a real decline) in the importance of library web pages in discovery. The results don't show any growth in any subject area since 2015.



*Library discovery seems to have peaked in its importance rating, and is only holding a strong position in Humanities, Education and Social Sciences.*

If we look at the importance of A&Is over time by subject in the academic sector we see that there has been a downward trend in Earth Science and Environmental Science, but one presumes the presence of SciFinder and PubMed in Chemistry, Life Sciences and Medicine are major drivers in keeping A&Is as key in those subjects.

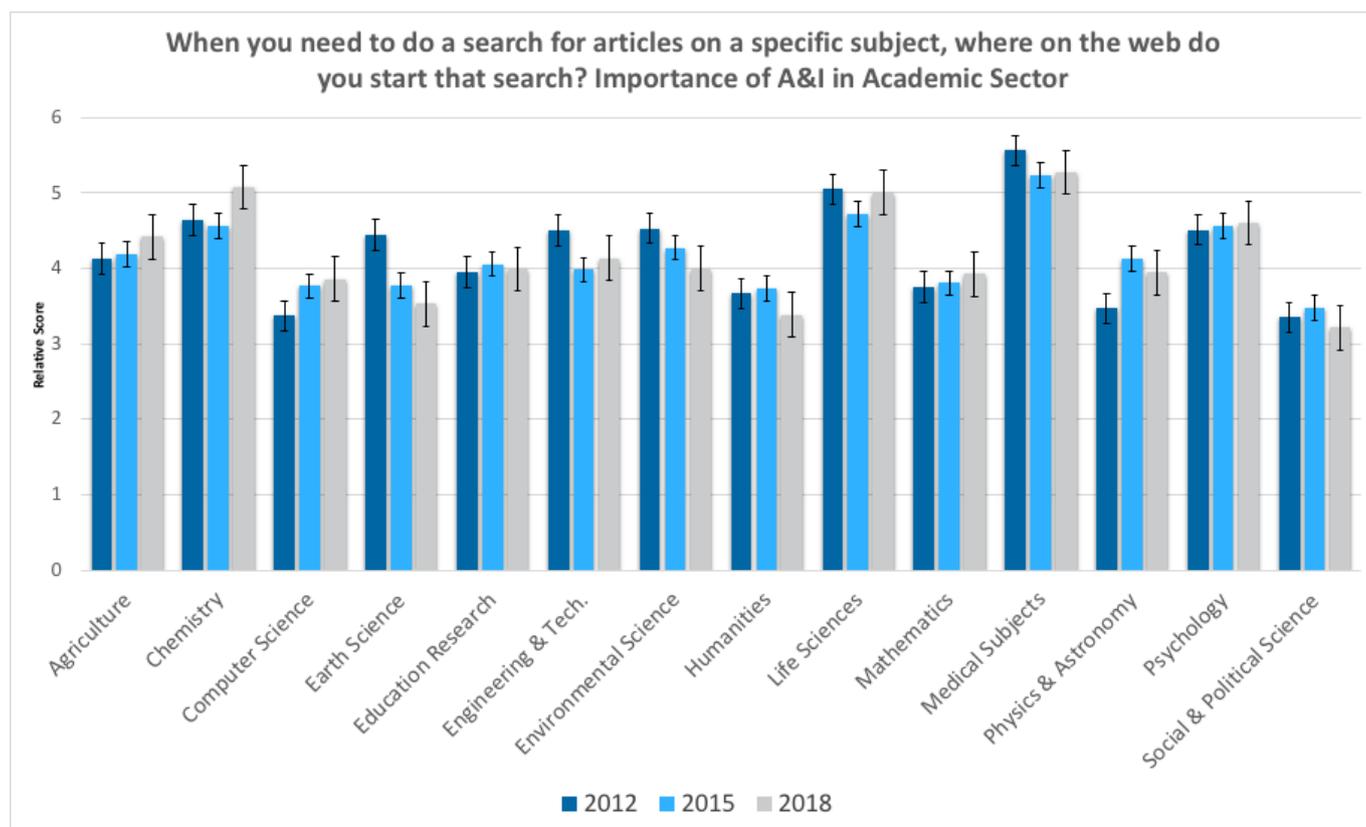
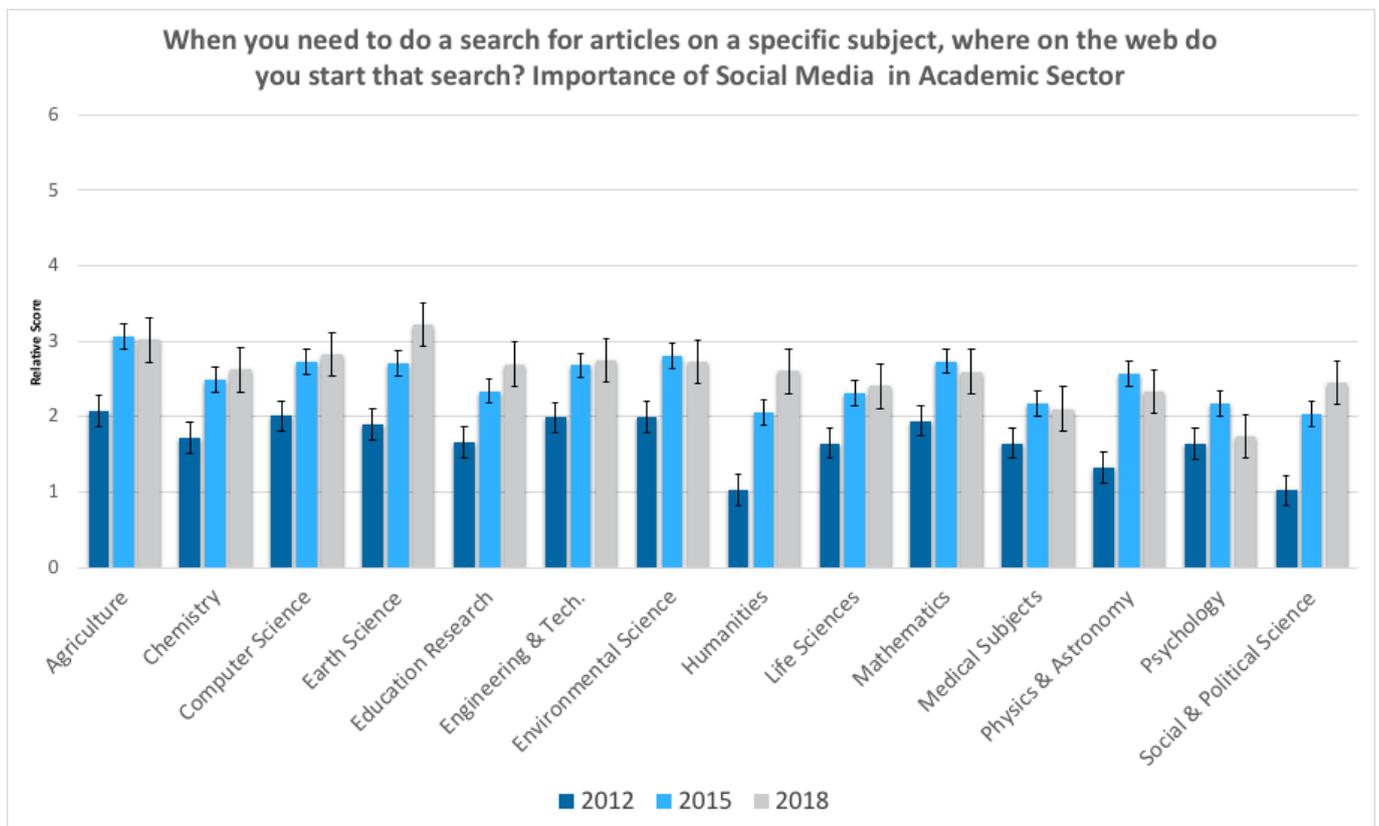


Figure 5 - A&I search, by subject, 2012-2018

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 since 2012, but for most subject areas that rating has now peaked. There is still real growth, however, in Education Research, Humanities and Social Science.





**Figure 6 - Social Media search, academic sector, by subject, 2012-2018**



*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. Some other subjects are showing real decline.*

## ANNUAL REVIEWS

SERVING THE RESEARCH COMMUNITY FOR MORE THAN 85 YEARS

**Annual Reviews articles save readers time by synthesizing primary research literature and identifying the principal contributions to their field.**

Founded in 1932, Annual Reviews provides researchers, professors, and scientific professionals with highly cited review journals across 50 scholarly disciplines. The articles each contain an average of 150 citations to key primary research contributions, serving as a high-quality filter for the research community. They hold a unique place in the scholarly communication ecosystem by enabling the transfer of expert knowledge to other scholars and society as a whole.

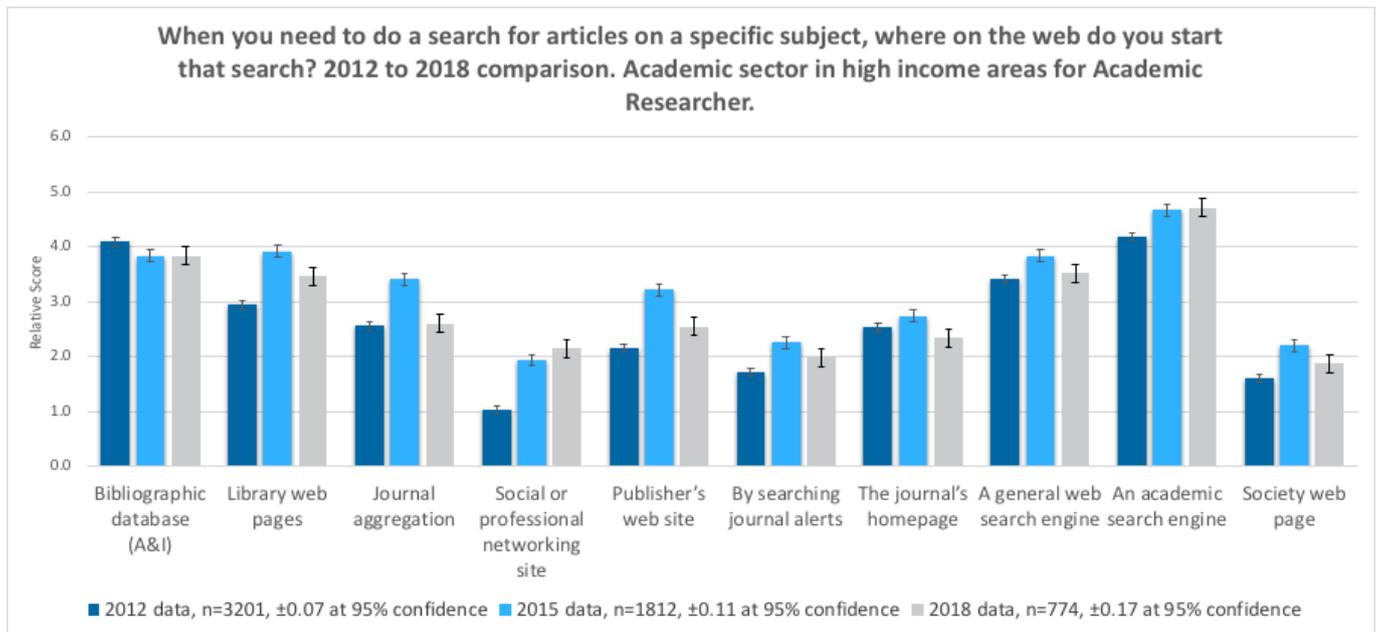
To create this impactful content, Annual Reviews relies on expert Editorial Committees in each discipline and facilitates meetings at which members discuss trends and select relevant topics for review. Authoritative authors are then invited to submit comprehensive review articles, which are often described as "required reading" for the research community.

Annual Reviews journals are among the most highly ranked by impact factor according to the Journal Citation Reports® (Clarivate Analytics).

**CONNECT WITH OUR EXPERTS | [www.annualreviews.org](http://www.annualreviews.org)**

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

## TREND BY JOB ROLE, 2012 - 2018



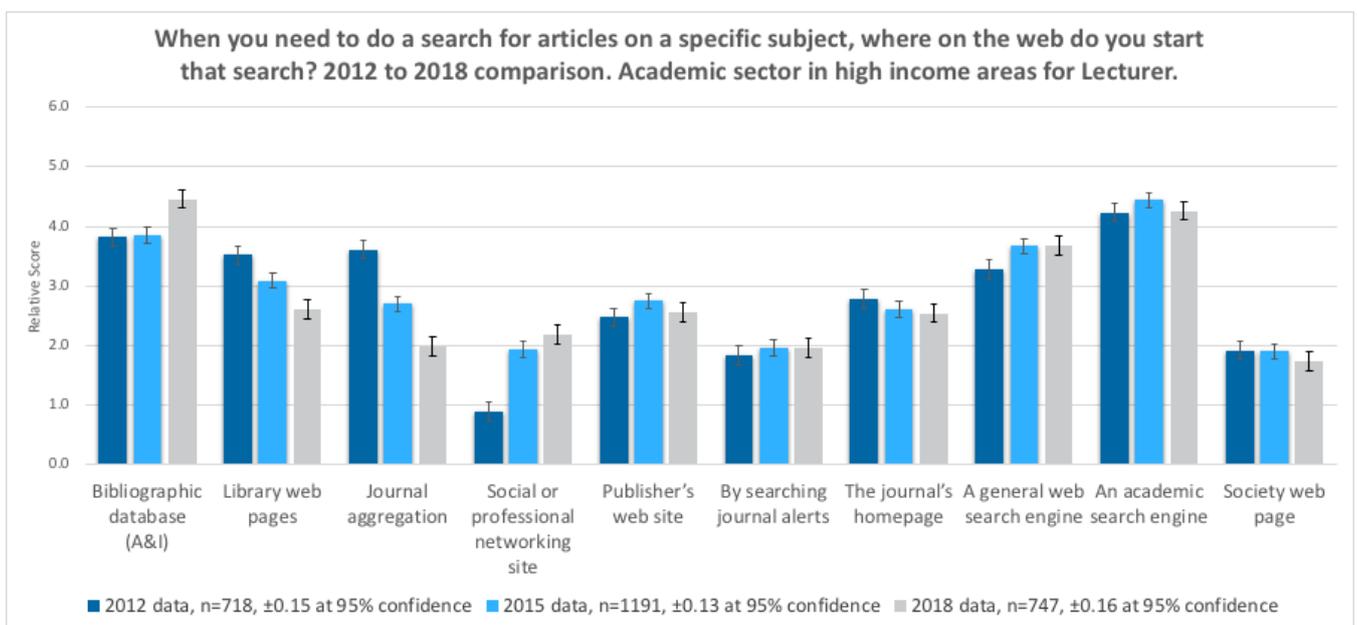
**Figure 7 - Search, academic researcher, 2012-2018**

Figure 7 shows that whilst A&Is are still important to academic researchers, they were overtaken by academic search engines in 2015, and that change in ranking has been consolidated in our 2018 results. In 2015, library web pages had become as important as A&Is to this group, but since then it has dropped away somewhat on average, although as we have seen there are some subject areas for which this isn't the case.



*Academic researchers, taken as a whole, rate academic search engines as the most important discovery resource when searching for journal articles.*

This picture changes dramatically for lecturers, in Figure 8. For this group, journal aggregations and library web pages have become significantly less important and less important than their researcher counterparts find them. Academic search engines and A&Is remain the most important.



**Figure 8 - Search, by lecturers, 2012-2018**

SEARCH BY SUBJECT

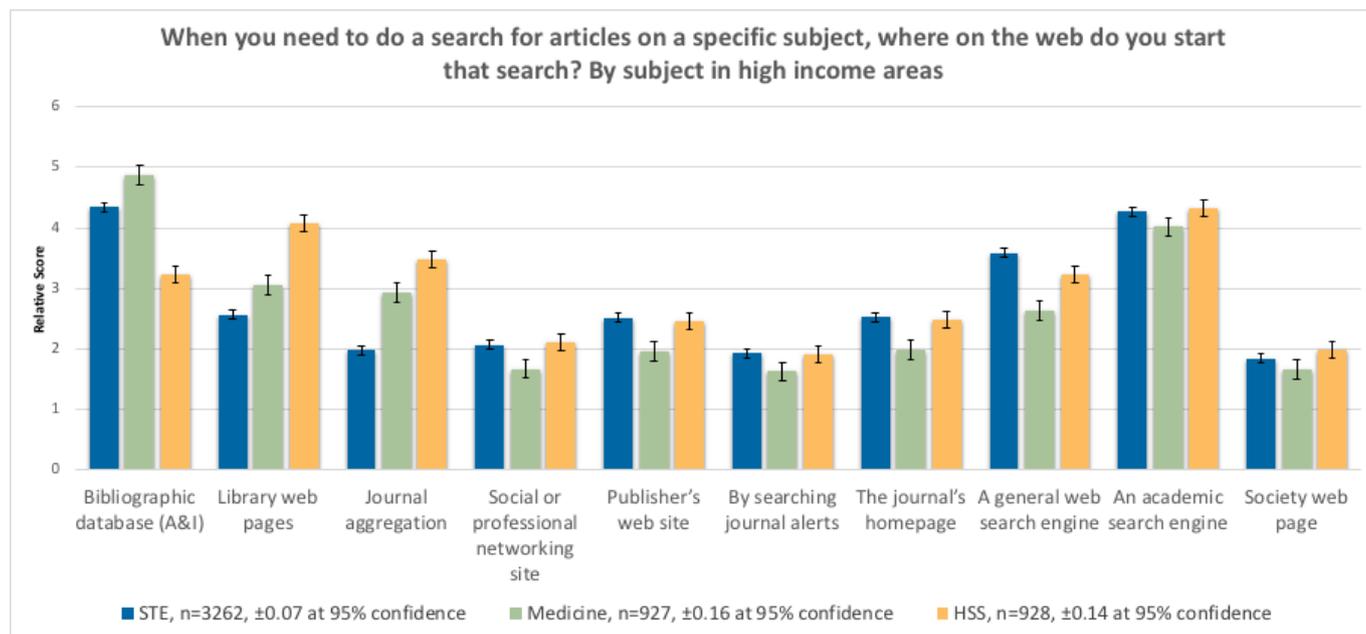


Figure 9 - Search, variations by broad subject in high-income countries, 2018 – part a

Figure 9 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, library web pages are almost on a par with academic search engines in Humanities & Social Science, whereas for scientific, technical and engineering subjects, A&Is are on a par with academic search engines. The relatively poor showing for A&Is in HSS is compatible with a changing focus for libraries in discovery for these subjects, where it appears easier to satisfy demand for discovery through broad library discovery rather than subscribing to specialist databases.

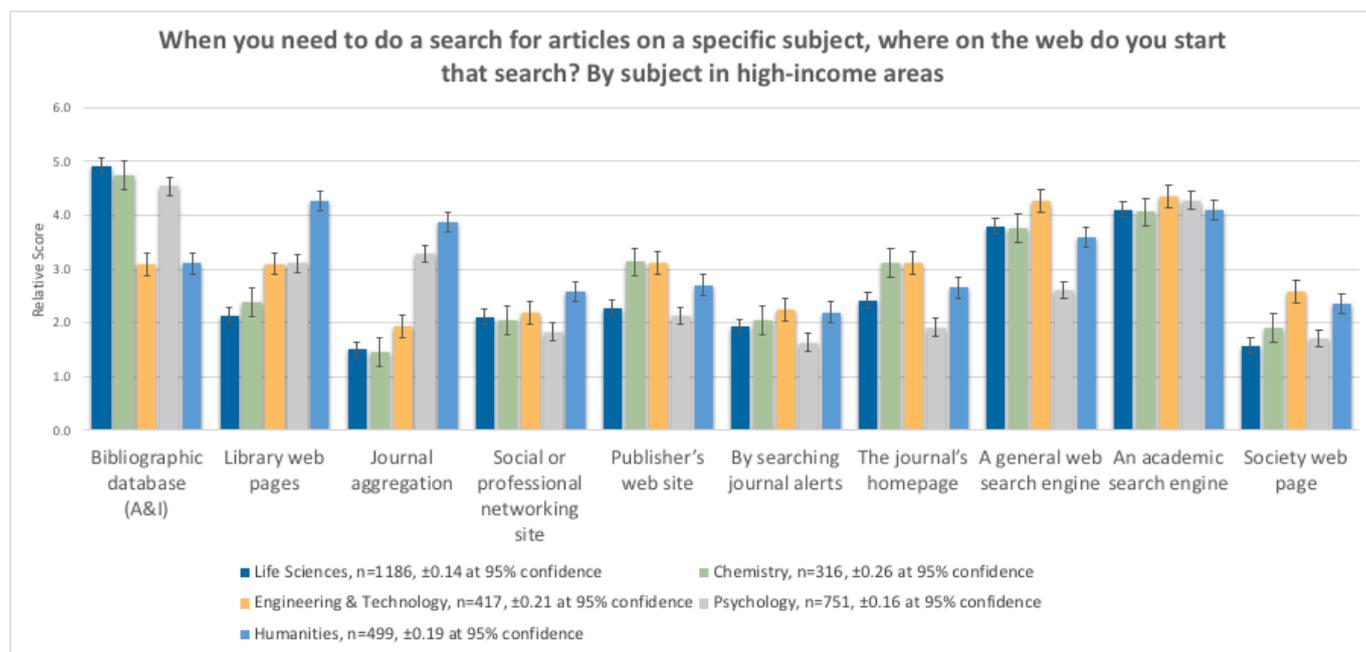


Figure 10 - Search, variations by subject in high income countries, 2015 – part b



As we dig further down into the subject results, Figure 10 shows us how little engineers seem to value bibliographic databases in comparison to academic search engines. Whilst every demographic values the academic search engine, the other discovery methods all have significant variations between subject areas, presumably reflecting factors such as the market penetration of the relevant A&I, what content is included in aggregations, how libraries make up for shortfall in discovery in some subjects with their own resource discovery services, and so on.

### SEARCH BY SECTOR

Figure 11 shows us the sectoral differences in search behaviour. The most significant difference in behaviour is in the medical sector where search importance is dominated by the A&I, with a resultant decrease in importance of academic search engines in that space.

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.

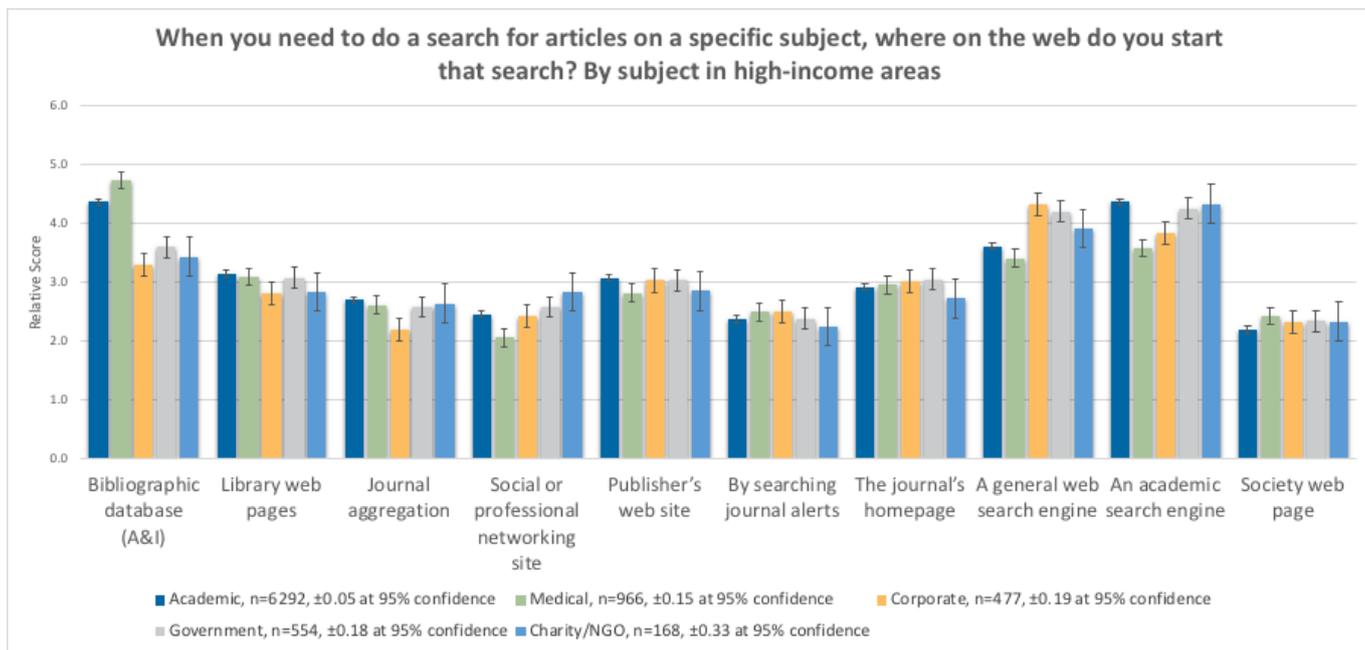


Figure 11 - Search by sector, 2018



*A&Is are significantly more important in the medical and academic sector than for all other sectors, and for the medical sector, of more importance than academic search engines.*

### SEARCH BY JOB ROLE

Figure 12 shows how search behaviour differs by job role across all sectors and regions – the differences are significant.

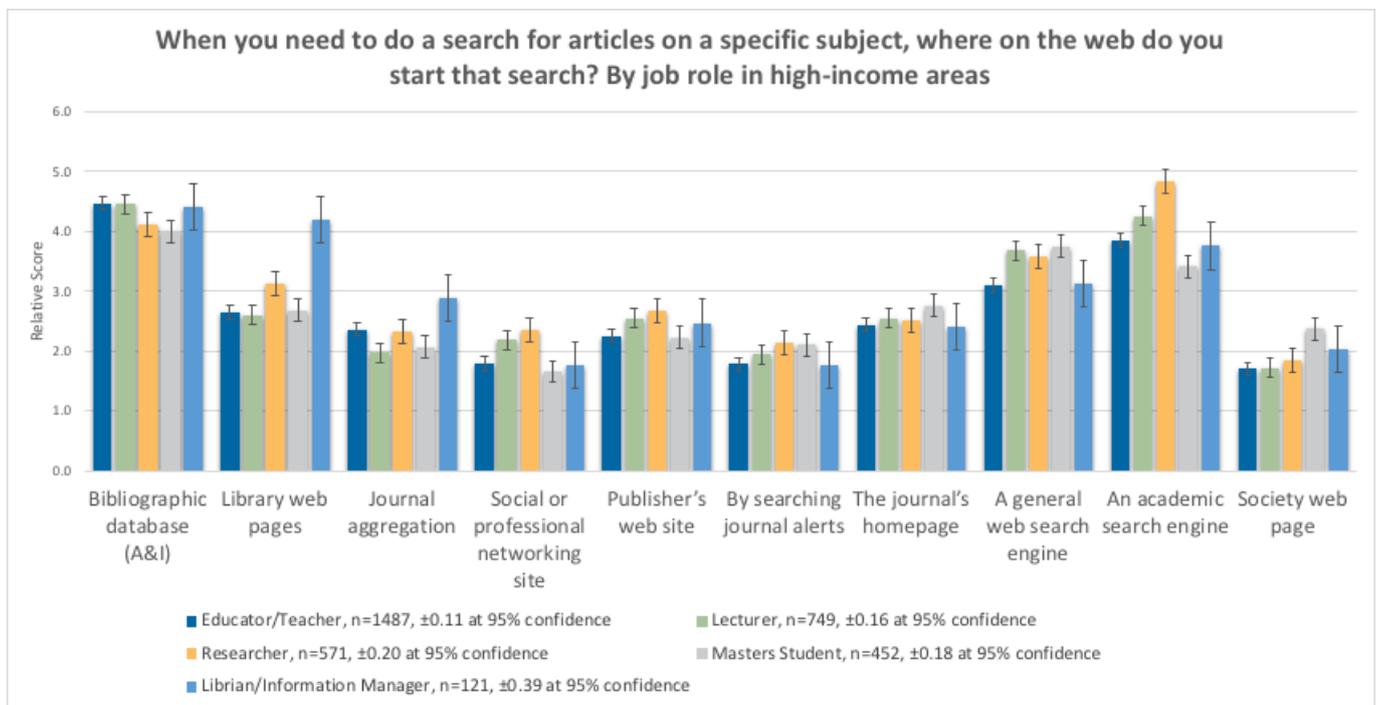


Figure 12 - Search, by job role, high-income 2018



Librarians behave quite differently to everyone else in search, preferring professional search databases and library-acquired resources. This may point to a continued 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 favour academic search engines less than everyone else (just as they did in the 2015 results).

### SEARCH BY REGION

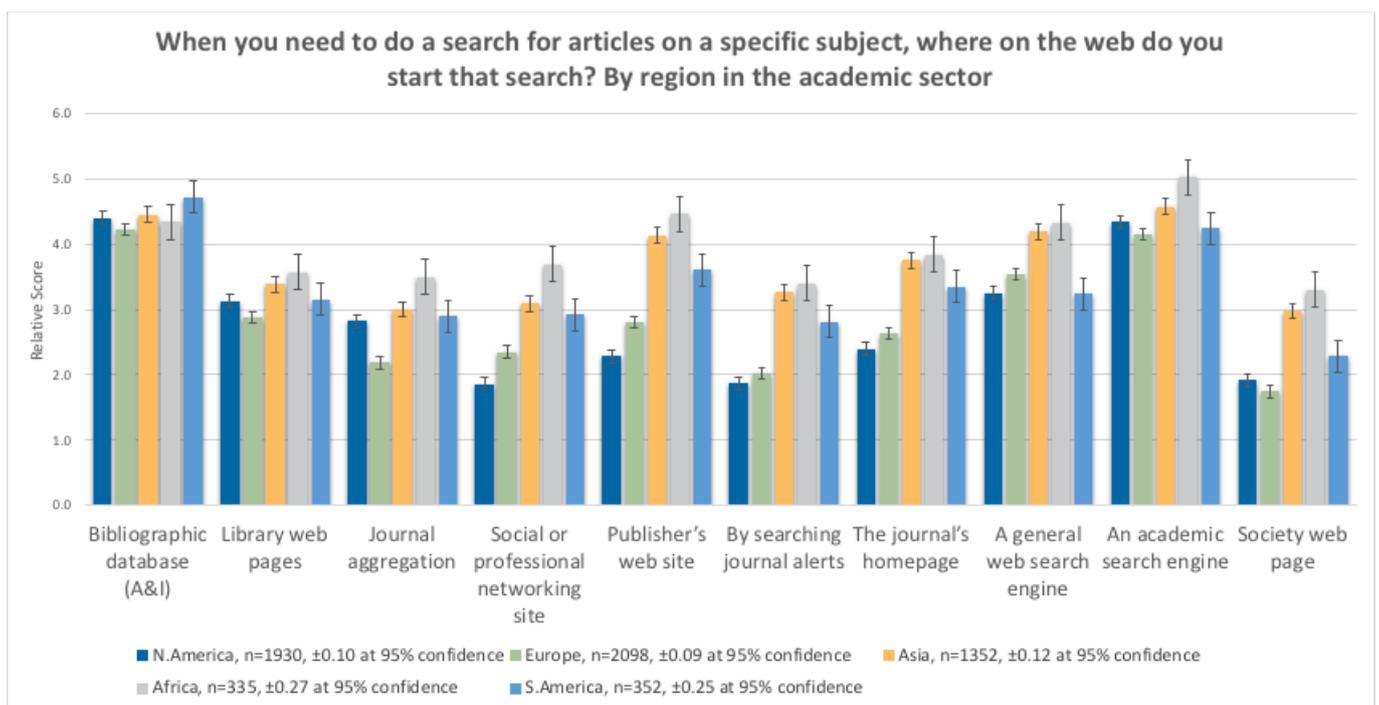
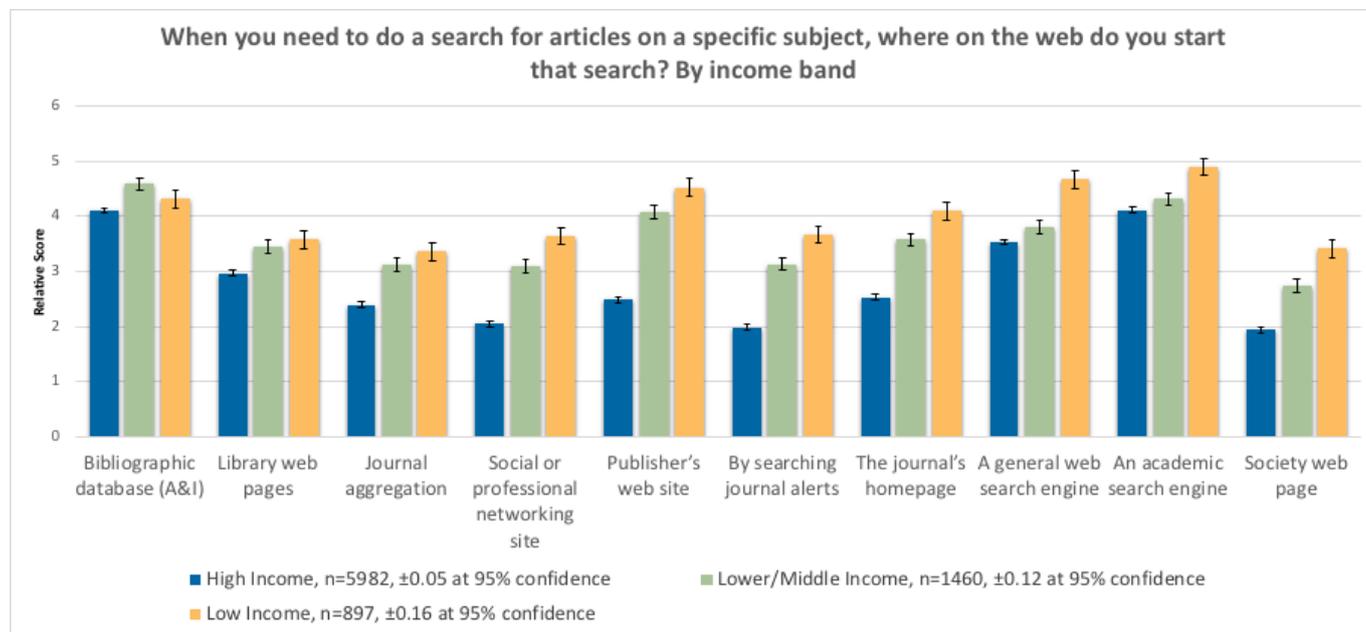


Figure 13 - Search, by region, 2018



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. This is consistent with the 2015 findings. Likewise, people in Asia, Africa and South America appear to value ToC alerts as a search resource.

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.



**Figure 14 - Search, by income classification, 2018**

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, but also a tendency in lower income countries to say everything is of high importance. Or perhaps readers in low income countries have developed a wider range of search strategies in order to locate the content of relevance to them, and are less reliant on premium services and library infrastructure.



*There are significant differences in search behaviour from high to low-income countries, with people in lower incomes countries using perhaps a wider range of search resources than those in wealthier countries who seem to settle on just a few methods.*

SEARCH ENGINE PREFERENCE - 2018

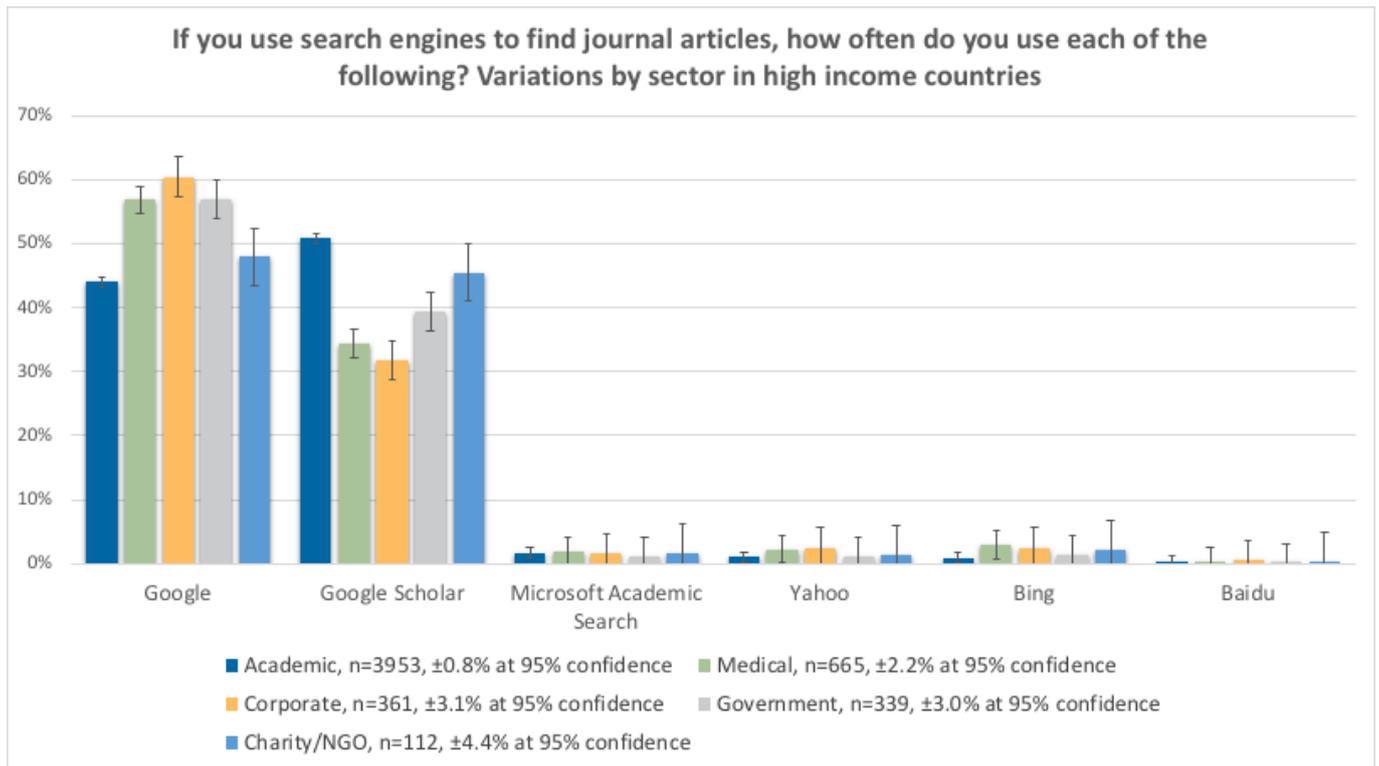


Figure 15 - Search engines, high-income countries, by sector, 2018

Figure 15 shows us that it is still 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 continues to make the least use of Google Scholar, and this may be an awareness issue.





Only the academic sector uses Google Scholar more than Google. The corporate sector continues makes the least use of Google Scholar, and this may be an awareness issue.

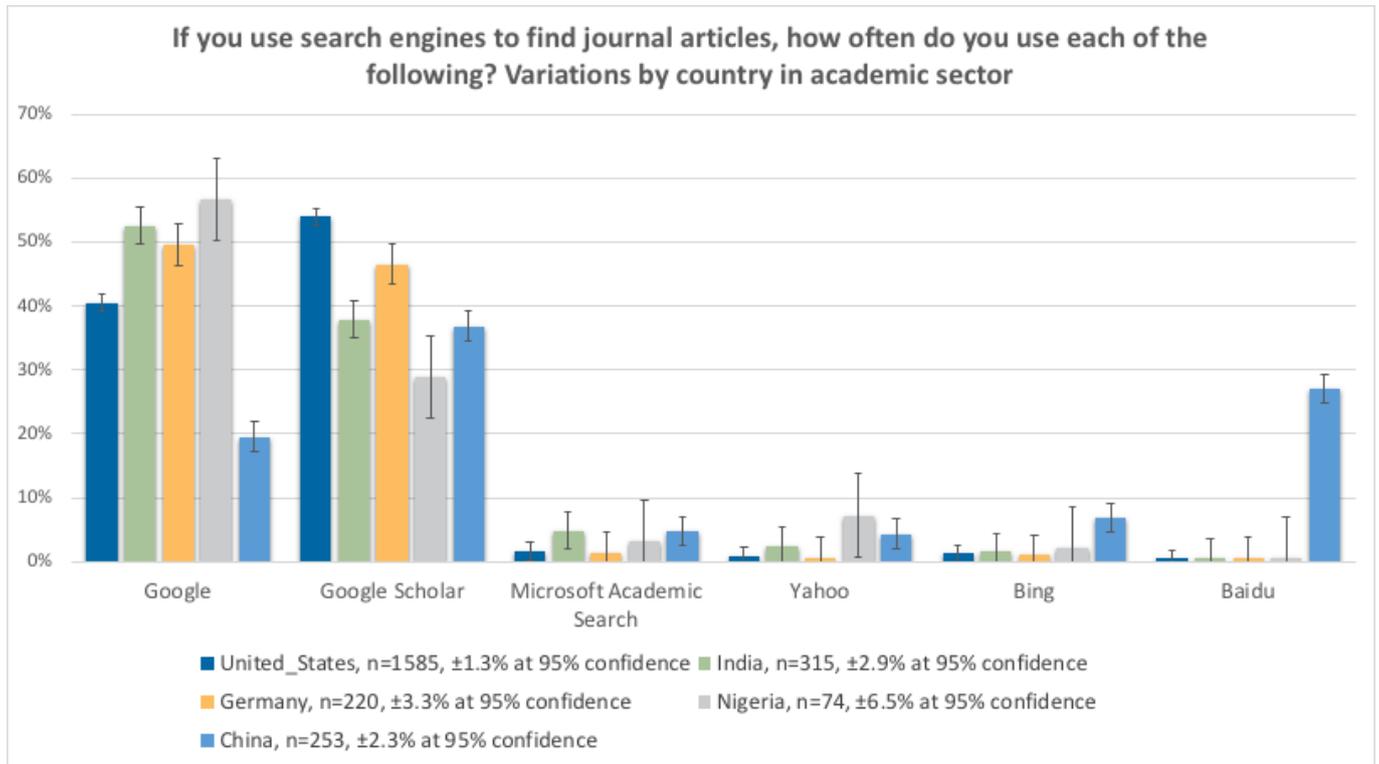


Figure 16 - Search engines, academic sector, by country, 2018

There are significant regional variations in the adoption of Google Scholar over Google within all sectors. Figure 16 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 (since not only is Google Scholar free, but it also exposes many free versions of the content itself). Use of Google in China is known to be restricted to a degree, and it is not surprising that Baidu has a greater presence than Google in China.

**OECD iLibrary**  
[www.oecd-ilibrary.org](http://www.oecd-ilibrary.org)  
Explore over 200,000 books, papers and datasets, covering 17 different themes



Google Scholar is the dominant search engine used for journal discovery in China, although the popularity of Baidu continues to grow.

Further analysis of the relative use of Google Scholar versus Google in the academic sector in high-income countries is shown in Figure 17.

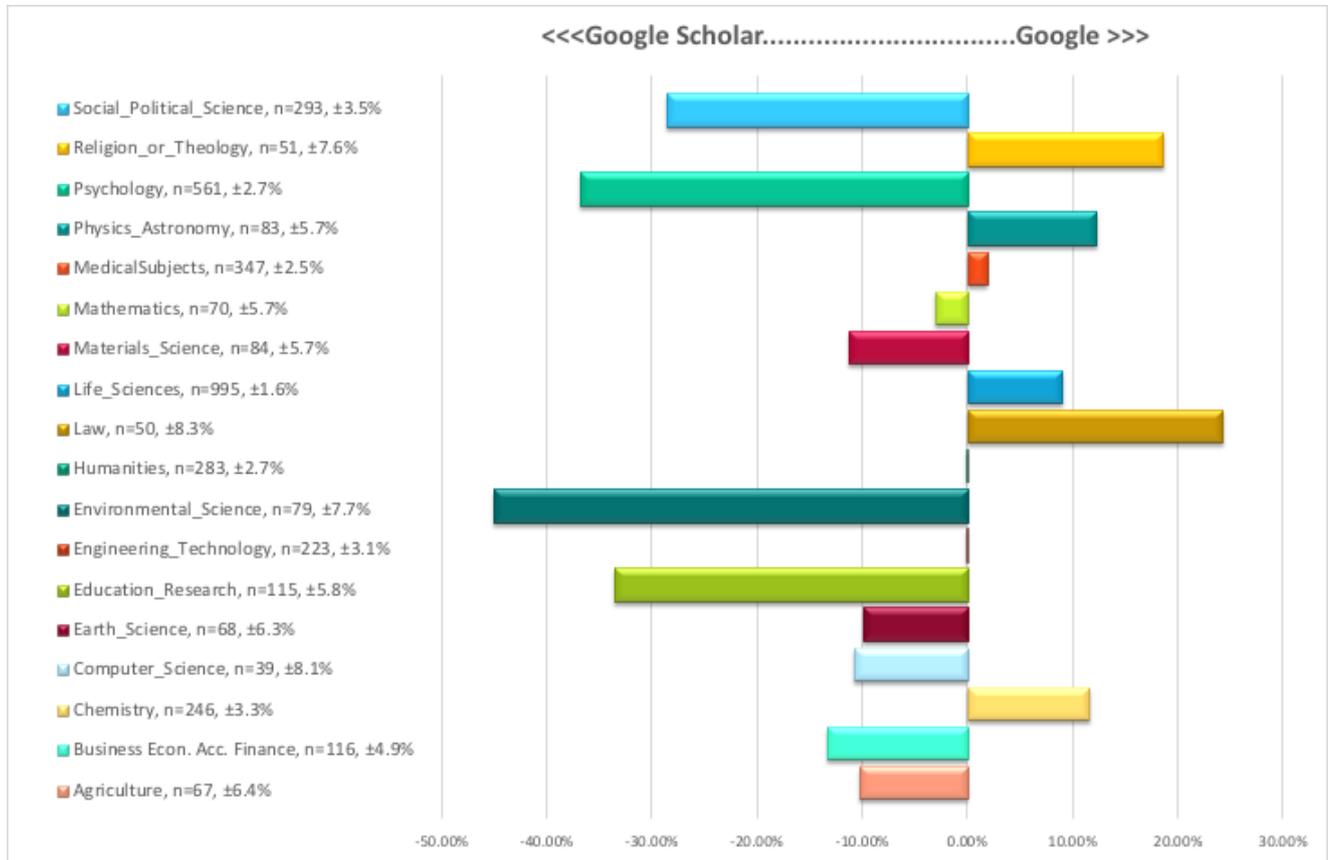


Figure 17 - Google vs Google Scholar by subject, 2018

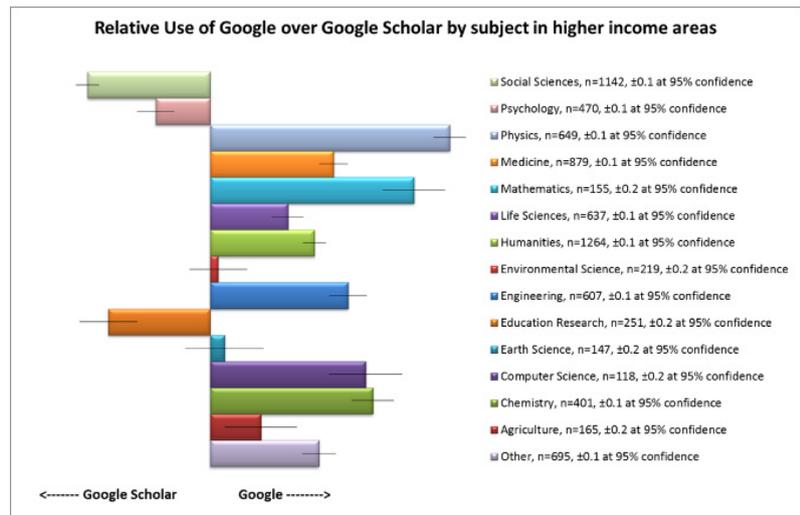
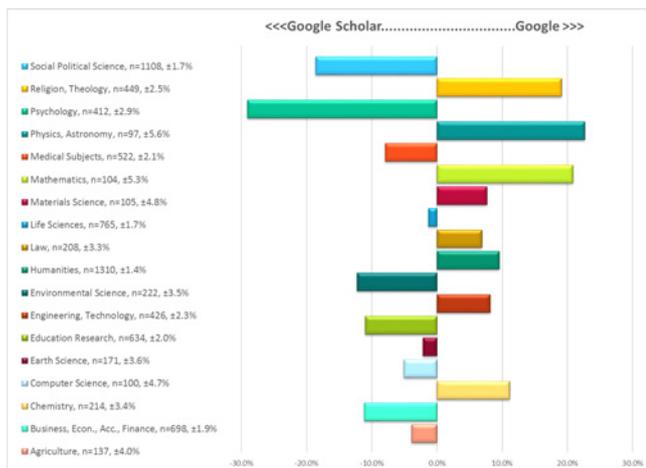


Figure 18 - Google vs Google Scholar by subject, 2015 and 2012

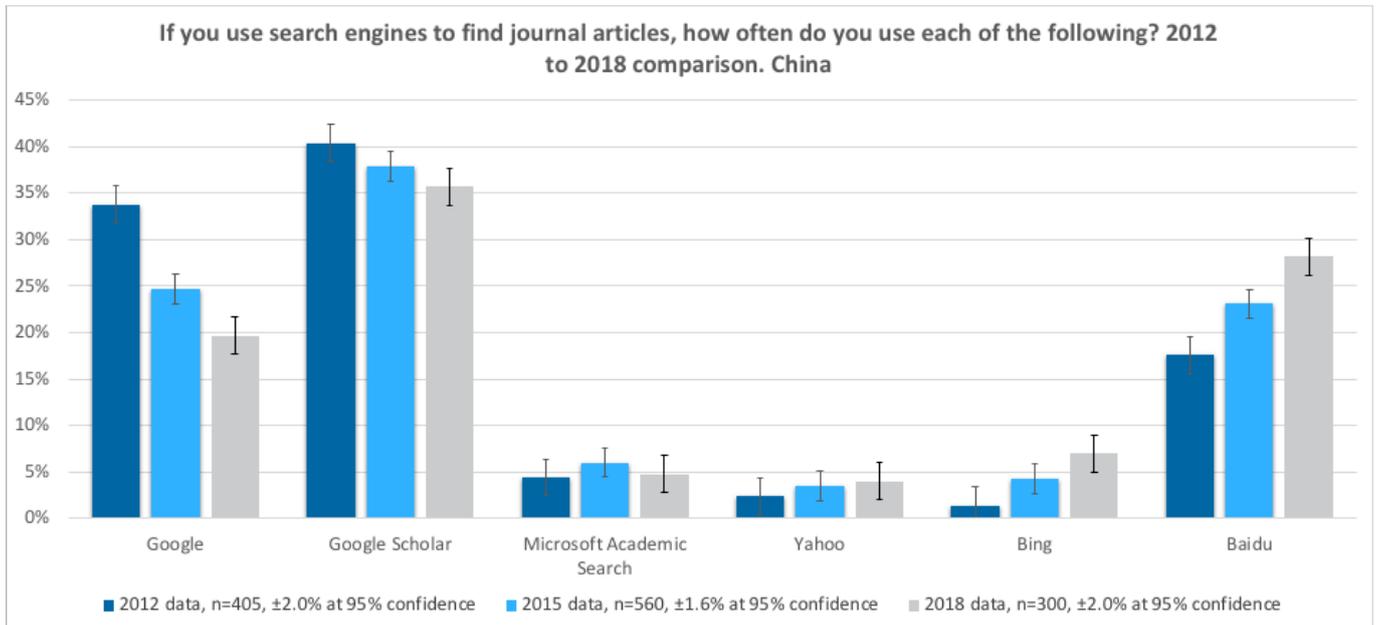
Figure 17 shows a large variance by subject area in people's preference for Google vs Google Scholar. People in Social Sciences, Education, Psychology all use Google Scholar more to find journal articles. However, people working in Law and Religion & Theology prefer to use Google – this is surprising as you might assume that there is more similarity in the behaviour of people working and studying in HSS fields. It is also a shift from the data from the 2015 and 2012 surveys, shown in



Figure 18. Whilst the overall shift is away from Google towards Google Scholar, this is not universally true.



*There is an increasing reliance on Google Scholar by people working and studying in the academic sector.*



**Figure 19 - Search engines, China, 2012-2018**

Figure 19 shows us that the shift in usage in China has been significant, as those leaving Google seem to have shifted predominantly to Baidu. Looking at the entire 6-year span, there is a real and significant decrease in the use of Google Scholar.

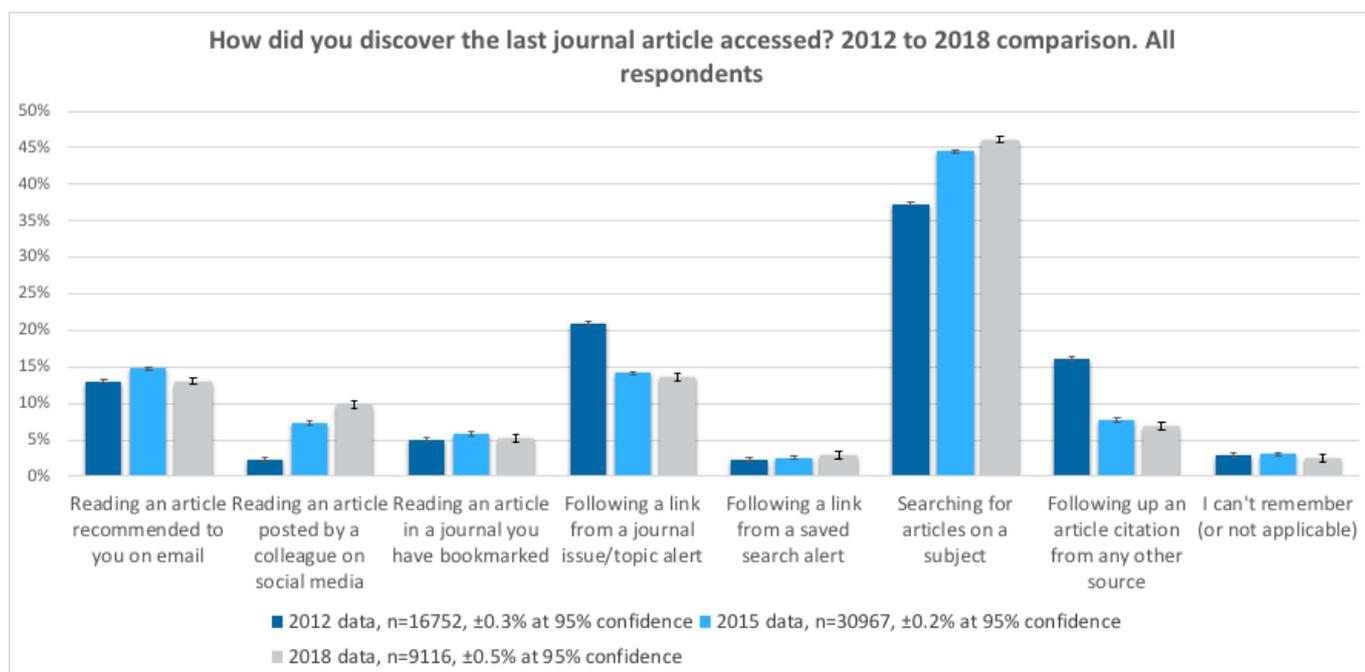
In other countries, just as in China, studies have shown us that the bulk of the shift in favour of Google Scholar over Google is as a consequence of the decline of use of Google more than it is an increase in the use of Google Scholar. The difference appears to have been made up in the use of less popular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

## LAST ARTICLE ACCESSED

### LAST ARTICLE ACCESSED: 2012-2015

In the past three surveys 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 20 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.

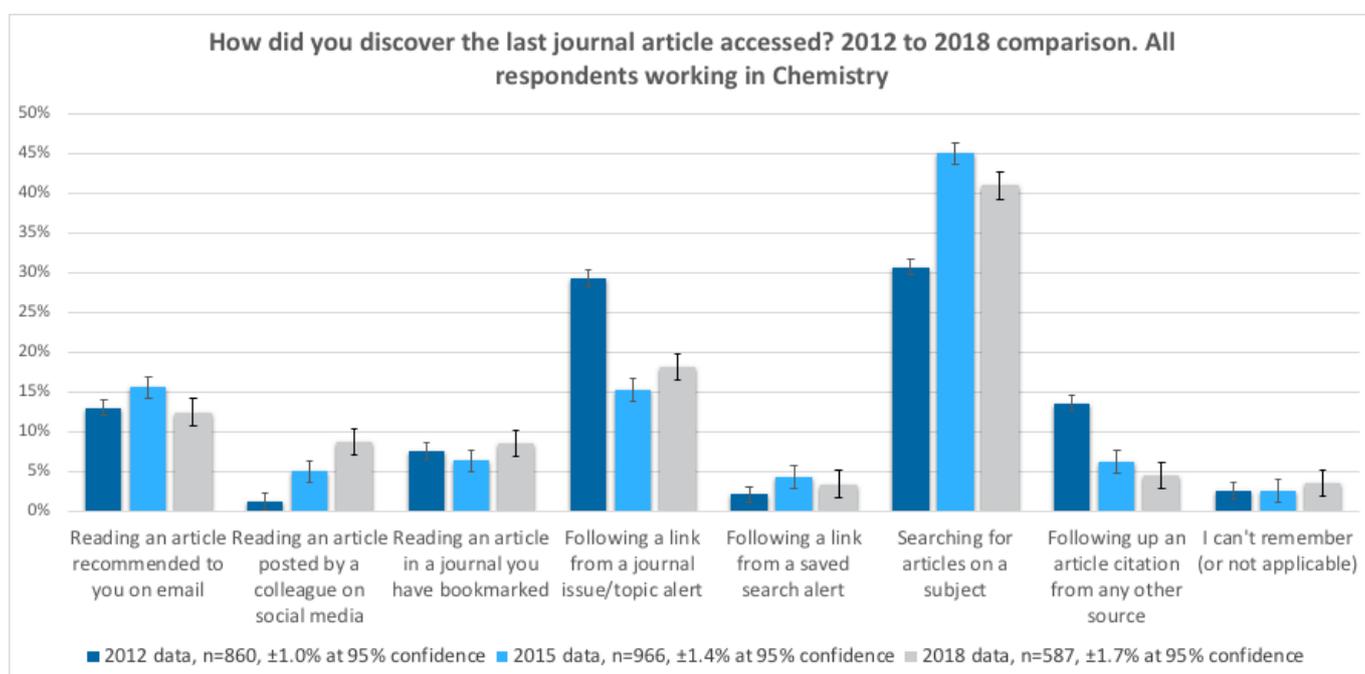


**Figure 20 - Last article accessed, 2012-2015**



*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 are used, but to the lowest extent of those resources tested.*

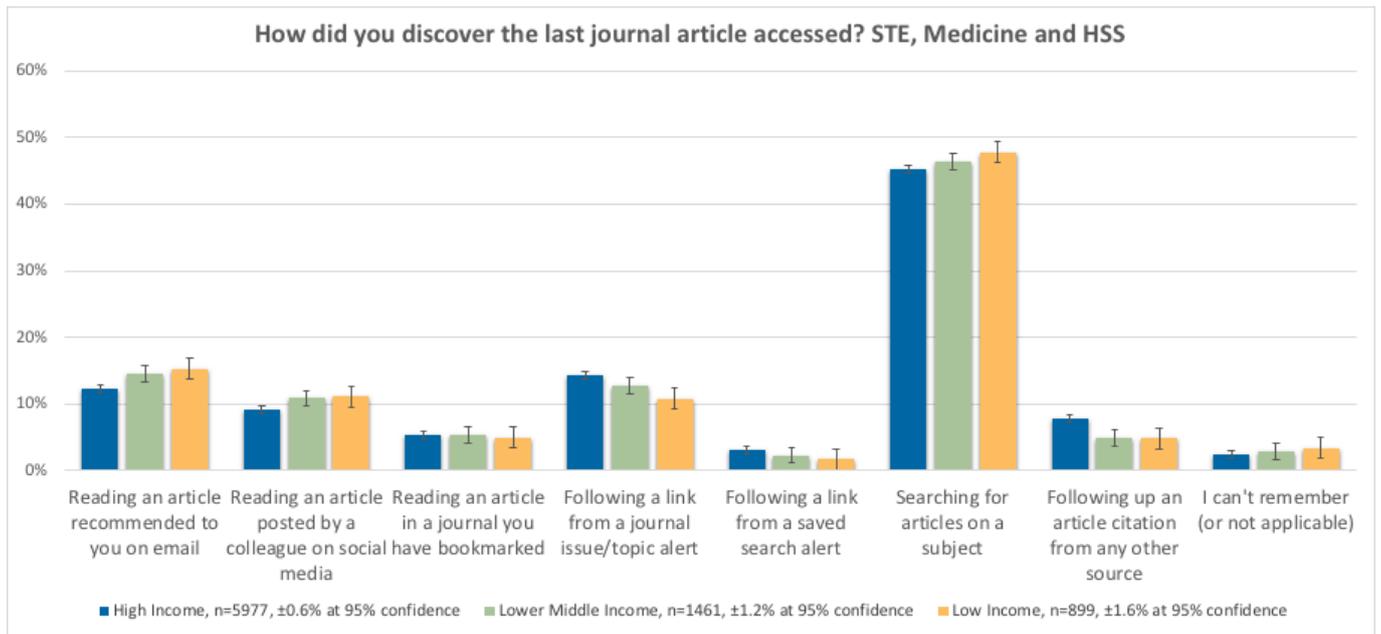
Perhaps one of the most marked changes is in Chemistry (Figure 21) where journal alerts were once as popular as search in the context of the last article that was accessed. Have people in Chemistry become bored of ToC Alerts or has searching improved for them?



**Figure 21 - Last article accessed, chemistry, 2012-2018**

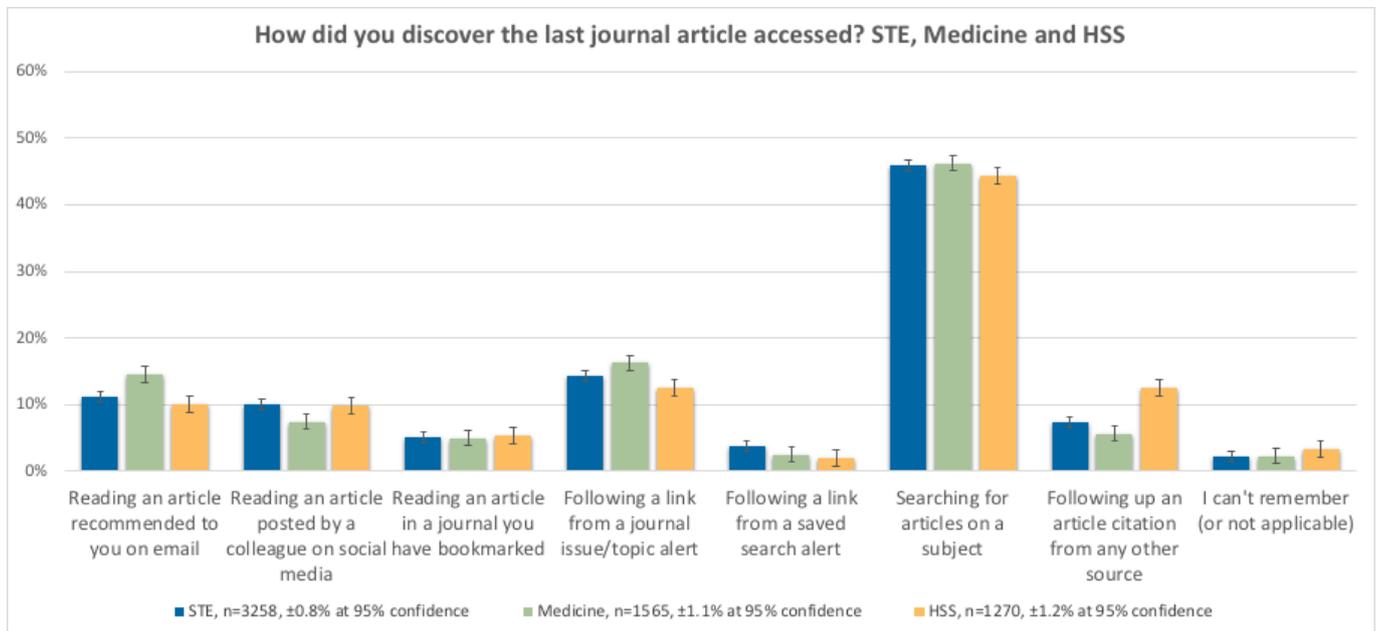
**LAST ARTICLE ACCESSED: 2018**

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.



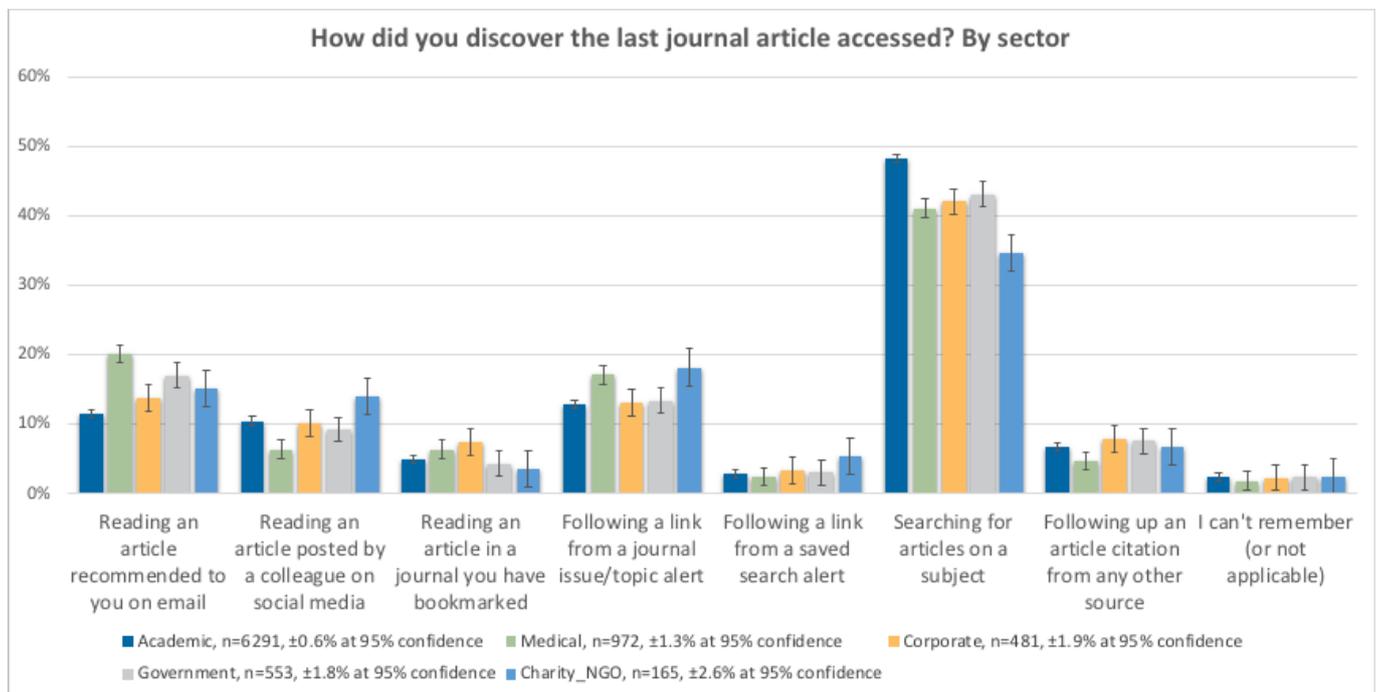
**Figure 22 - Last article accessed, by income, 2018**

Figure 23 looks at the variation by high level subject area (STE – Science Technical Engineering, Medicine, HSS – Humanities and Social Science). There are subtle differences, but mostly shows that people in HSS have “other sources” they may use. In the context of this question, “other sources” included Wikipedia.



**Figure 23 - Last article accessed, by broad subject, 2018**

Figure 24 shows there is some 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. People working in the charity and NGO sector seem to make the least use of search.



**Figure 24 - Last article accessed, by sector, 2018**

### Selected Journal Search Comments

- ❖ *I look at the academic web page of an expert in that field--to look at their teaching resources, syllabi, CV, working papers. This helps me find not only links to articles but also figure out additional search terms.*
- ❖ *Ask the librarian who completes the search for me*
- ❖ *Blog posts from research news digests*
- ❖ *Every search begins with the WVU Library home page, then click databases, then enter Pubmed, and that finds everything I need most of the time. Occasionally I will sign on to the website of some of the subspecialty journals to which I subscribe.*
- ❖ *Frankly, printed paper still is the best starting point...*
- ❖ *Google Images*
- ❖ *Google the subject and review, find the most recent review on the subject then follow the literature references therein*
- ❖ *I follow researchers and the references in their published articles. The first problem is finding the researchers. I do not do general topic searches until I already have a good mental map of a research area. Then it is a secondary activity to scan for articles and topic areas that I have overlooked.*
- ❖ *I may look on the websites of people that I know work in the area to see their full list of publications.*
- ❖ *I often start from the reference list of the paper that I'm currently reading, and then jump from reference to reference.*
- ❖ *I prefer to avoid Google if I can because of their practice of tracking and datamining. I prefer to use DuckDuckGo because of their policy of protecting the privacy of those who use their service.*
- ❖ *I try and exhaust free sources first. As I study, I catalog items I want to retrieve from paid sites. Idea being to subscribe for a single session and download as many resources as I can to save money. The*

*reason I put neutral for Journal publishers web sites is that I don't ever go there first since I have to pay to see anything. I do, however, review these pages at some point to get scope of publisher's coverage of scholarship in a given area so that I can learn where what is published."*

- ❖ *"Professional network sites like academia and personal faculty websites are useful for avoiding silly paywalls for Journals."*
- ❖ *"1. Google 2. ADS 3. arXiv 4. Sci-hub."*
- ❖ *"A large proportion of searching is looking through the latest issues of a set list of journals and downloading the ones that are interesting. I still place a high value on owning my own copy of books and articles, in pdf format or on Kindle."*
- ❖ *"A lot of the time I search materials - books and journals - using the key subject titles to locate them. From there I either try downloading them from journal websites, access them from libraries or from researchgate/SSRN sites."*
- ❖ *"A mix of Google Scholar, including alerts, Mendeley, and the library website is the recipe."*
- ❖ *"Alerts are the first motivation."*
- ❖ *"Also through websites and social media posts from vendors, museums, research centres, and funding agencies."*
- ❖ *"Basically, I open a bunch of internet windows and look through all of the databases I can think of until my searches becomes redundant and my research starts overlapping. The particular research question I'm trying to answer changes where I look first and the success I tend to have with either."*
- ❖ *"I usually begin a global Google Scholar search, and when I find a good articles, I read up on the authors and then search more by the same people and review the references in the article. I then also ask leaders in the field who they recommend and read those suggestions as well. I then read critiques of whatever "top" articles there are and compare the scholarly backgrounds of both sides. I then try and pinpoint key ideas in the articles on the topic, and then summarize the concepts. Once summarized, I do a new search to be sure I covered all the angles."*
- ❖ *"I use ResearchGate frequently because this is probably the only social network platform I see any utility in and I check it everyday. I also use Google Scholar to search for papers based on keywords/subjects because the search engine is free, accessible, and easy to use."*

# SEARCHING FOR BOOKS 2018

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. There has been no significant change in the results since then.

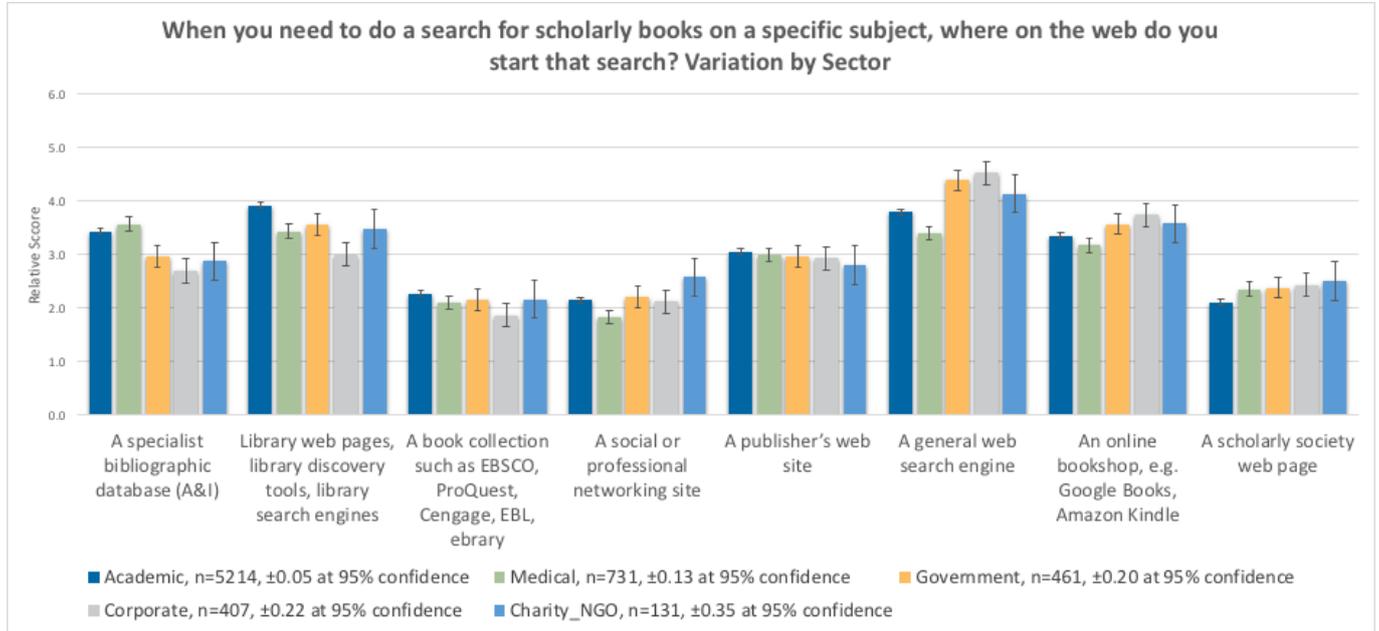


Figure 25 - Book search by sector, 2018

Figure 25 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 Google, A&Is and their library as equally important for book search. This is a shift from 2015 since when Google has made some gains. People in the Academic sector think their library website is the most important resource for book discovery, followed very closely by Google.



*In comparison with journals, library web pages are seen as a very important discovery resource for books, now normally eclipsing A&Is, and in some sectors just as important as Google.*

**TREND MD** Grow your readership

### Why do scholarly publishers use TrendMD?

- Drive article usage and citations
- Increase author submissions
- Grow readership of new journals new subject areas
- Boost usage at specific locations and/or institutions
- Increase subscription revenue

support@trendmd.com

1-888-589-5561

www.trendmd.com

### JOURNAL OF CLINICAL ONCOLOGY

Official Journal of the American Society of Clinical Oncology

Enter words / phrases / DOI / ISBN / authors / keywords / etc.

Home | New Articles | Issues | Browse By Topic | Special Content | Authors | Subscribers | About ASCO

#### REVIEW ARTICLES

#### Systematic Review of the Impact of Cancer Survivorship Care Plans on Health Outcomes and Health Care Delivery

Paul B. Jacobsen, Antonio P. Dellosa, Tara O. Henderson, Deborah K. Mayer, Cheryl S. Moskowitz, Elnora D. Paskett, Show More

Abstract | Full Text | PDF | Figures and Tables

**Purpose**  
Numerous organizations recommend that patients with cancer receive a survivorship care plan (SCP) comprising treatment summary and follow-up care plans. Among current barriers to implementation are providers' concern about the strength of evidence that SCPs improve outcomes. This systematic review evaluates whether delivery of SCPs has a positive impact on health outcomes and health care delivery for cancer survivors.

**Methods**  
Randomized and nonrandomized studies evaluating patient-reported outcomes, health care use, and disease outcomes after delivery of SCPs were identified by searching MEDLINE, Embase, PsycINFO, Cumulative Index to Nursing and Allied Health Literature, and Cochrane Library. Data extracted by independent raters were summarized on the basis of qualitative synthesis.

**Results**  
Eleven nonrandomized and 13 randomized studies met inclusion criteria. Variability was evident in the cancer types, SCP delivery timing and method, SCP recipients and content, SCP-related outcomes, and disease outcomes assessed. Nonrandomized study findings yielded descriptive information on satisfaction, adherence, and related SCPs. Randomized study findings were generally negative for the most commonly assessed outcomes (ie, QOL, functional, and psychosocial well-being). Findings were positive for long-term adherence to treatment recommendations and patient-reported health status.

#### We recommend

ASCO Clinical Practice Guideline: Optimizing Anticancer Therapy in Metastatic Non-Castrate Prostate Cancer

By Matthew Stenger, The ASCO Post

Timing of Chemotherapy, Ra 223 in Advanced Prostate Cancer  
ASCO Daily News

Metastatic Castration-Sensitive Prostate Cancer: Optimizing Patient Selection and Treatment  
Andrew W. Hahn et al., ASCO - New EdBook

No Survival Benefit Reported With Docetaxel Added to Hormone Therapy in Metastatic Prostate Cancer  
Amanda L. Graham et al., J Med Internet Res, 2013

Deep Androgen Suppression Plus Abiraterone and Prednisone: Effective Strategy for Hormone-Naive Prostate Cancer  
Namkee G Choi et al., J Med Internet Res, 2013

Recommended by TREND MD

We place your articles here

If we look at different job roles in the Academic sector, we can see how much researchers now value library web pages for books. Teachers in academia are somewhat less reliant on Google for book searches than other sectors.

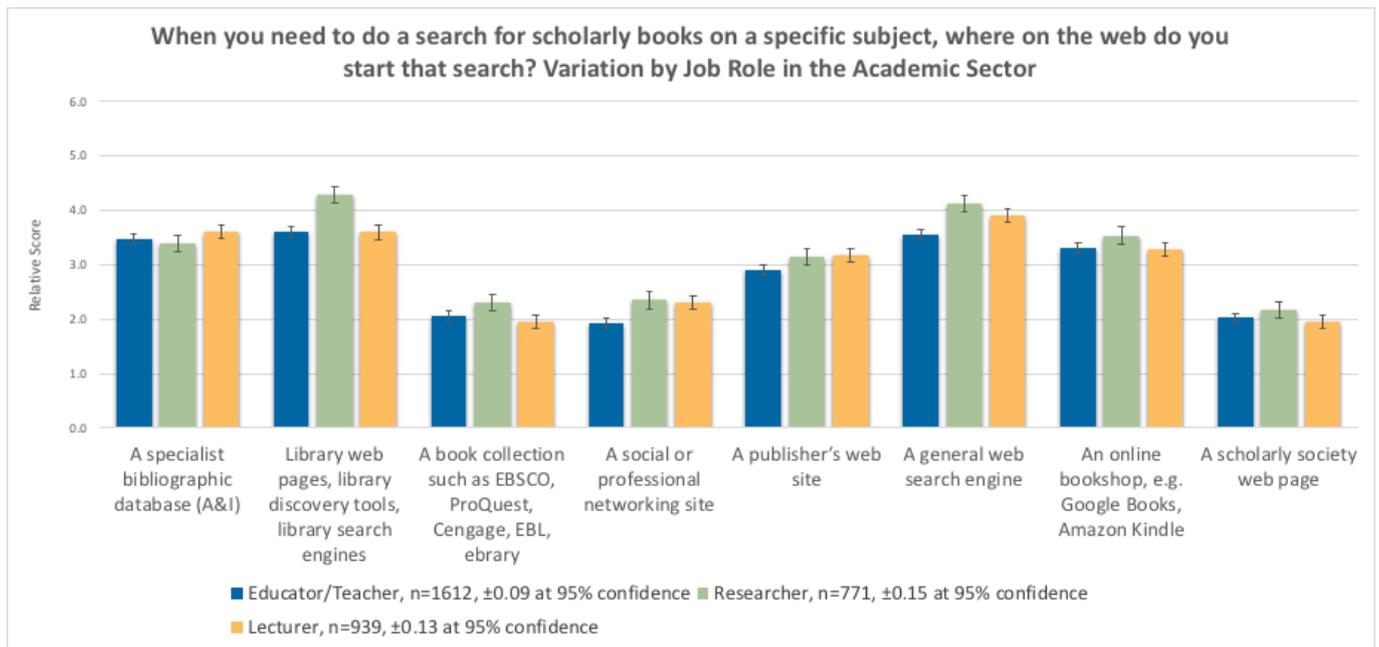


Figure 26 - Book search, academic sector, job role, 2018

### LAST BOOK ACCESSED

Search is the single dominant method of online book discovery, just as with journals. Figure 27 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.

# Cambridge Core

Bringing together over 32,000 ebooks  
and 380 journals

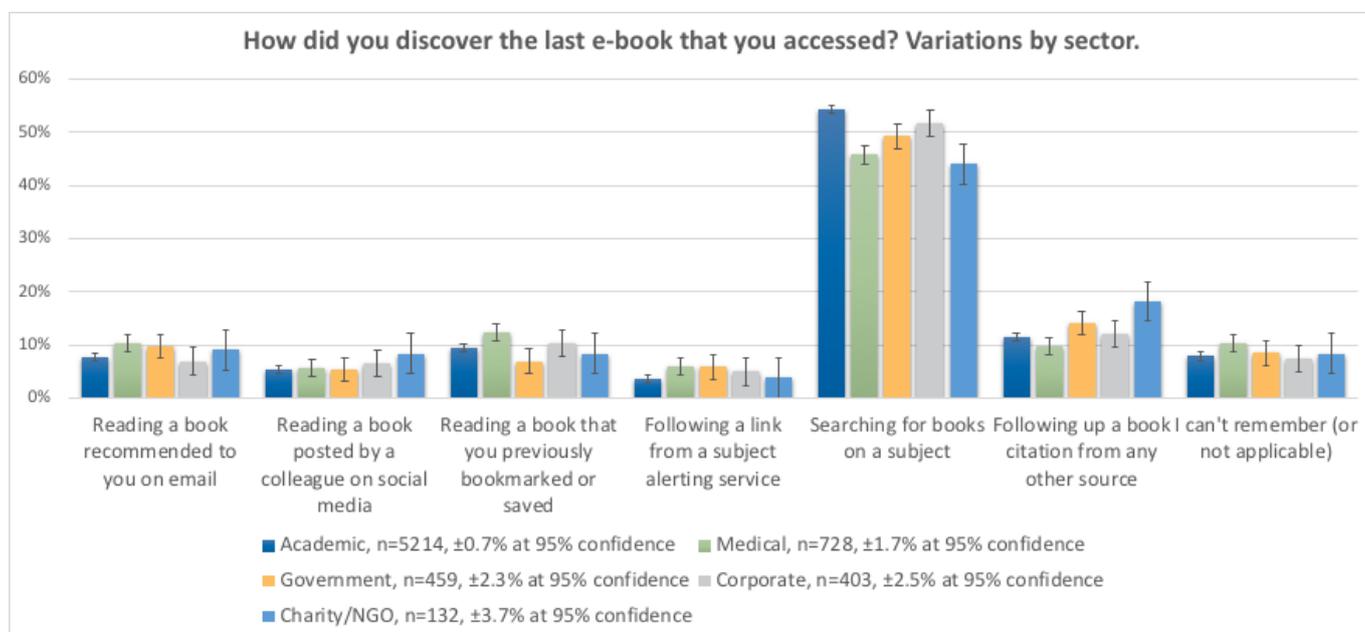
Powerful search  
functionality

Send content to  
Dropbox, Google Drive  
and Kindle

Easy citation with  
multiple styles

[cambridge.org/core](http://cambridge.org/core)

CAMBRIDGE  
UNIVERSITY PRESS



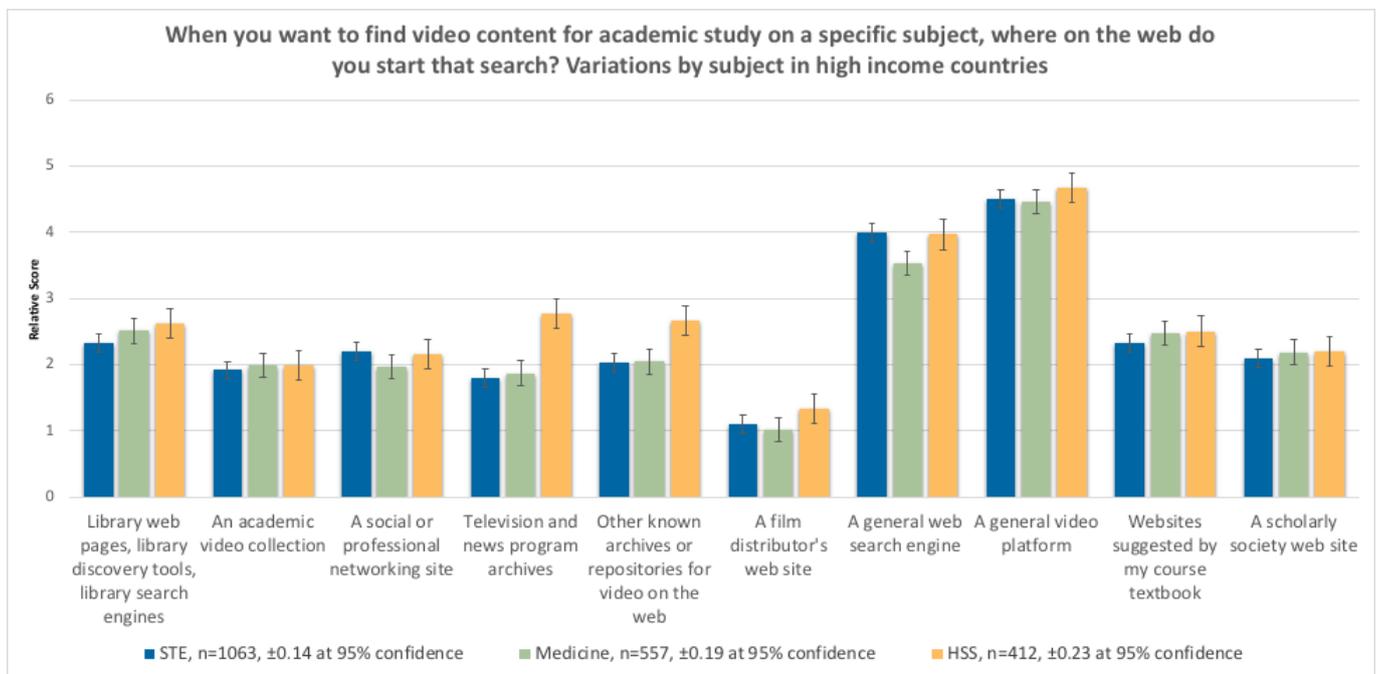
**Figure 27 – Last book accessed, sector, 2018**

### Selected Book Search Comments

- ❖ *“I seldom \*start\* with Ebsco/PQ unless it a specific format needed (eg, Thesis, newspaper item); but I may end up there if that is how our library accesses that particular book/resource.”*
- ❖ *“I don't search for books per se, I find citations of book chapters and then hunt down the book.”*
- ❖ *“I don't search for scholarly books - I only use them if I already know of them/used them during my studies.”*
- ❖ *“I follow the work of a particular medical expert.”*
- ❖ *“Professional associations which I am a member of, such as PMI, AS, IIBA, YES I allow access to webinars, books, handbooks, and articles as well as forums, which I use in my work to enhance my knowledge and learn about best practices.”*
- ❖ *“Amazon is usually where I start”*

## SEARCHING FOR VIDEO

For the first time in 2018, we have included survey questions on video discovery. Video resources already seem to be the default destination for the young for all kinds of instructional needs, and libraries appear to be rapidly increasing their holdings of educational content.

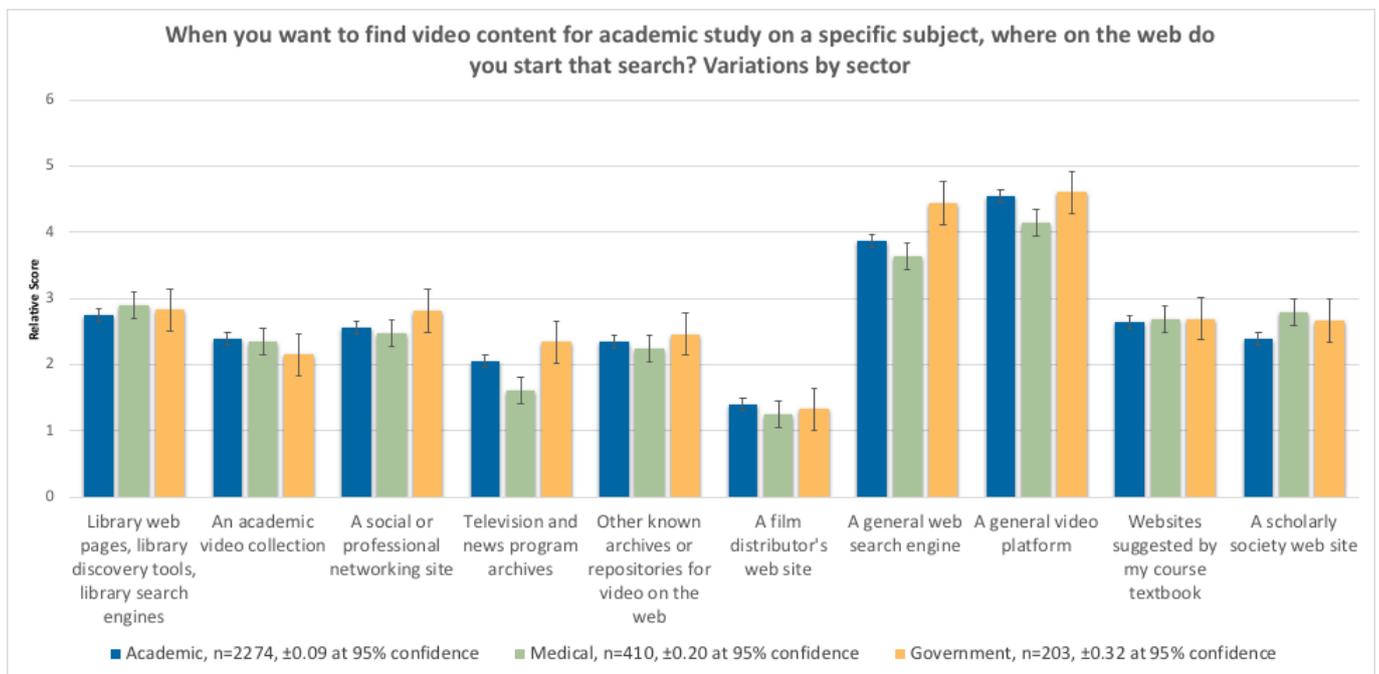


**Figure 28 - Searching for video, high income by broad subject area, 2018**

Figure 28 shows how respondents rate the importance of a range of discovery resources in video, organised by broad subject area. It appears a general video platform, like YouTube, is seen as more important than a general web search engine, such as Google, even though the general search engine will index content from YouTube and other video platforms, making it a broader search. This is very different to what we see in journals, where the dominant discovery platforms are not also the content host, but is more similar to how the e-books market has evolved.

There are not many significant differences in the ratings from one discipline to the next, except in HSS, where TV and News archives are perhaps unsurprisingly seen as a good starting point, and so are other specialist archives and repositories.

Beyond the Google and YouTube dominance, we do see libraries and other resources suggested by text books as being slightly more important than the others, which perhaps shows some effective level of institutional organisation evolving in video.



**Figure 29 - searching or video by sector, 2018**

Figure 29 shows us the same question, with the answers split out by sector. Of interest here perhaps is that the medical sector places less importance on the general web and general video platforms than other sectors, but without having a strong preference for any of the other categories.

## LAST VIDEO ACCESSED

Just as with other content types, we took the opportunity to ask the question about how the respondents found their most recent video that they watched. The results are shown in Figure 30. Whilst people from all sectors used search as their discovery method the most, there are significant differences in the magnitudes of their responses. Academics were noticeably more often searching, but the medics are finding a lot of their video content in journal articles, indicating just how much this particular content type is taking off as additional or primary content in medical journals.



*Discovery of video appears to be dominated by Google and YouTube, but medics in particular are finding relevant video content within the journals they read.*

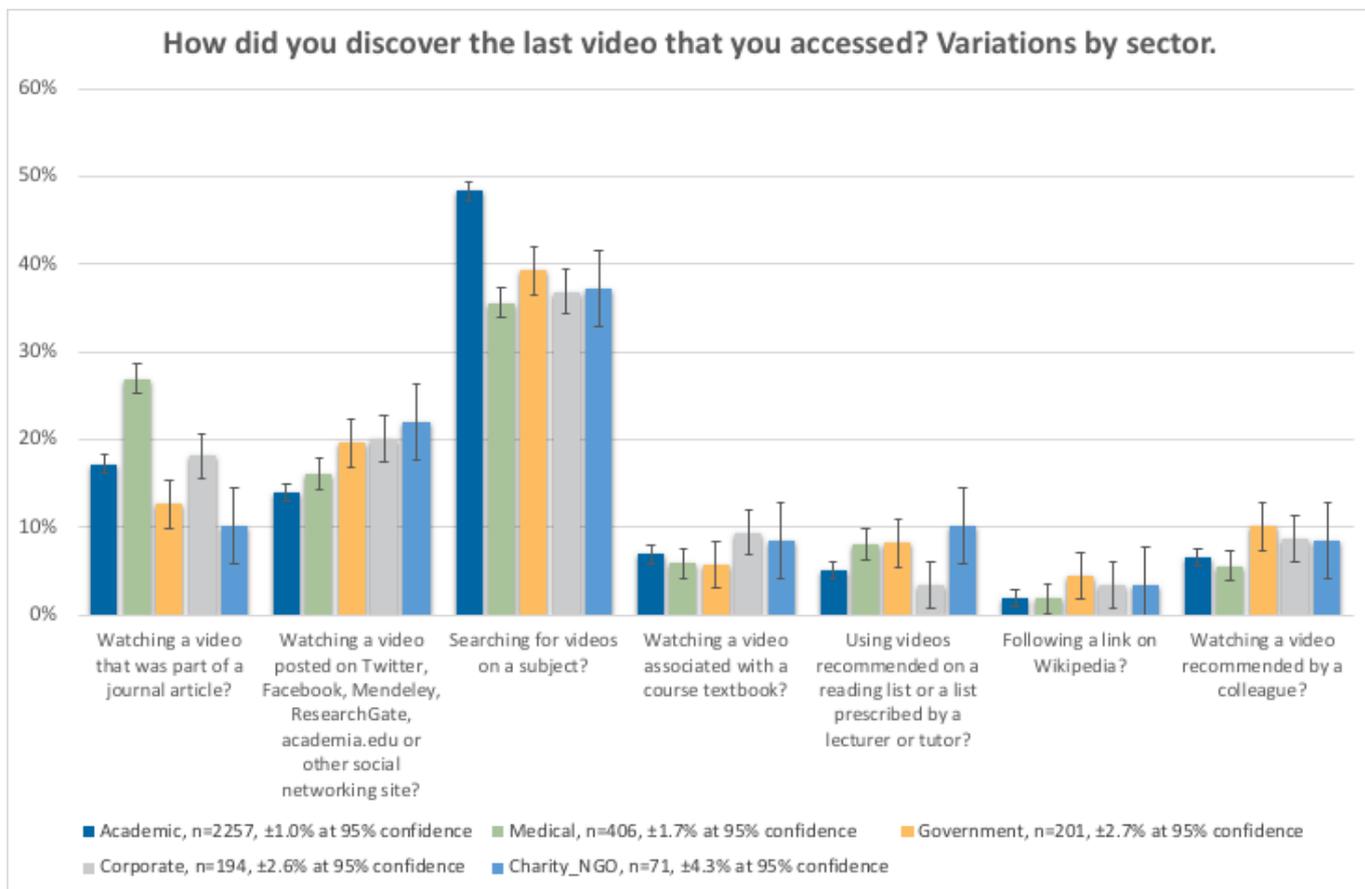


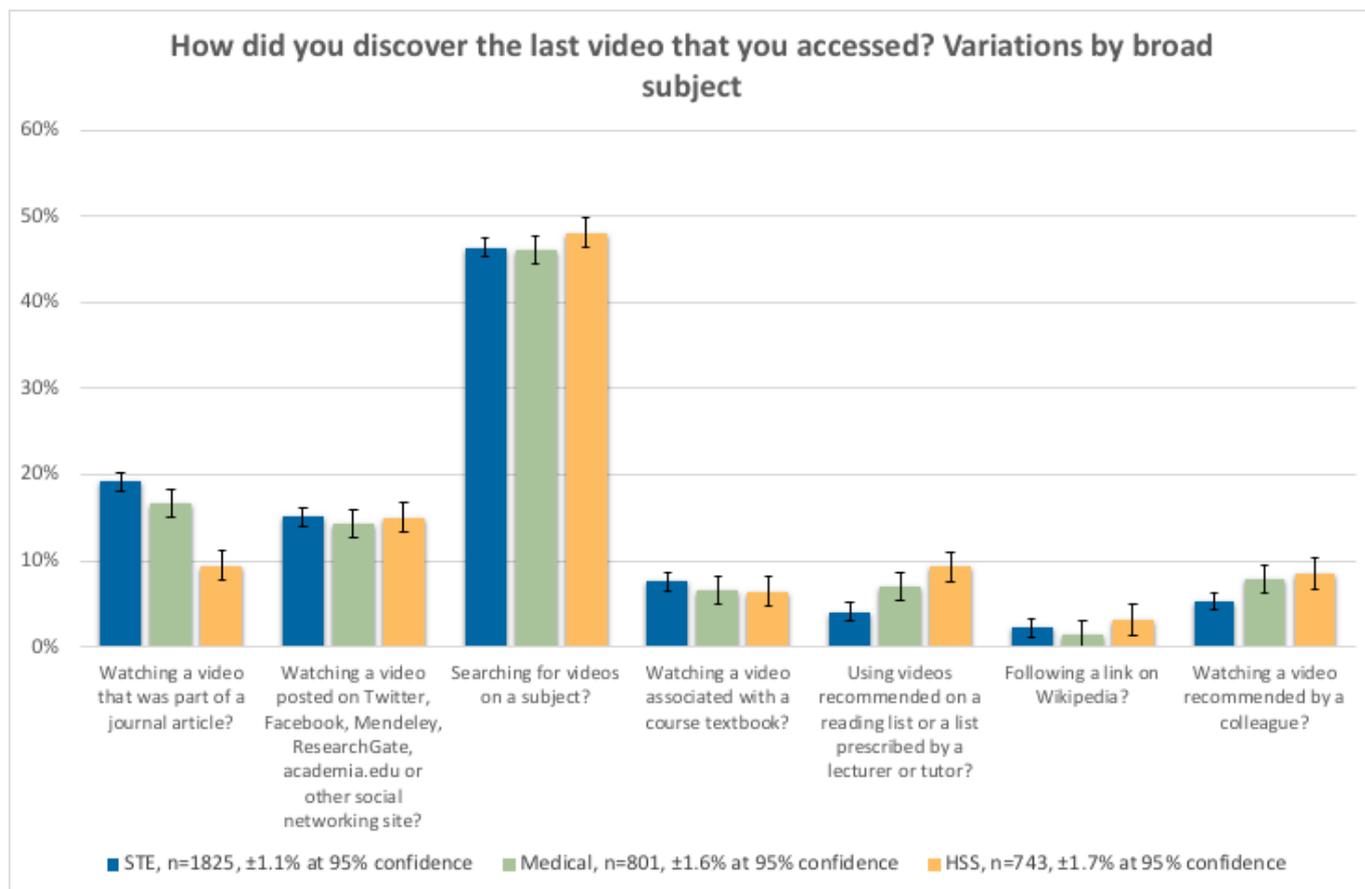
Figure 30 - last video accessed, sector, 2018



By guaranteeing our independence indefinitely, Sara has given SAGE an exceptional gift. It allows us to plan for the long term, ensuring that we continue to support the dissemination of usable knowledge to educate a global community well into the future.

Find out more about SAGE's unique vision  
[www.sagepublishing.com](http://www.sagepublishing.com)





**Figure 31 - last video accessed, by broad subject area, 2018**

Looking at the same data by broad subject area, we see almost exactly the same behaviour for each subject area in search, but significant differences elsewhere. STEM subjects are much more likely to find video content in their journals, accounting for nearly 20% of discovery in STE subjects. Just as we have seen in journals, people working in medicine are more likely to have content recommended to them by a colleague than those in other STE subjects.

In all cases search accounts for less than half of the usage. Greater than half of the usage comes from what is essentially browse, or recommendations, or required ‘reading’.

#### Selected Video Search Comments

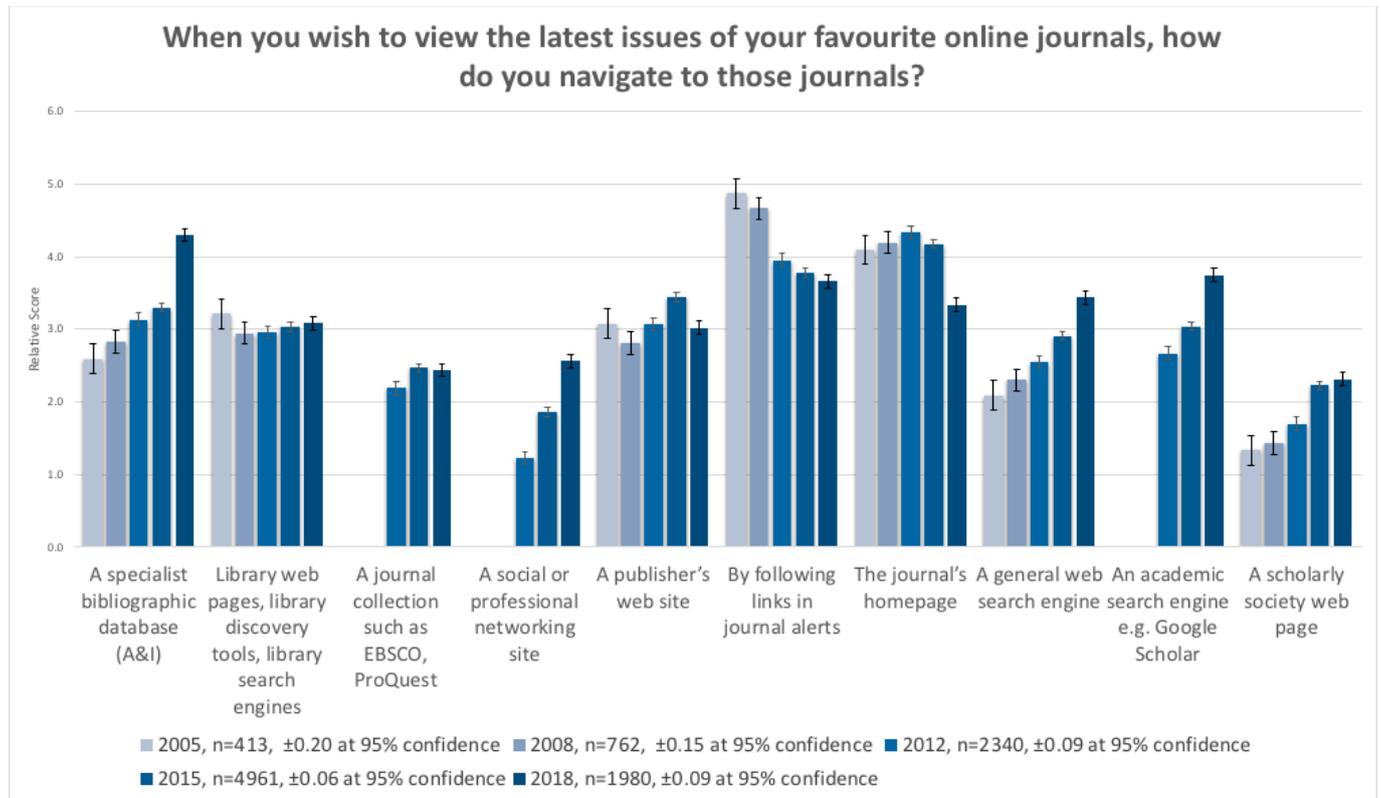
- ❖ *“Google’s specialized video search tool.”*
- ❖ *“I didn’t know these (Academic Video Online (Alexander Street Press), SAGE Video, Kanopy) existed! Cool!”*
- ❖ *“I mostly use videos that are included as SI in articles -- I find them when I read the article.”*
- ❖ *“Corporate website with expert interviews”*
- ❖ *“Course sites such as edX, Coursera, PluraSight”*
- ❖ *“Direct links from journal articles”*
- ❖ *“I watch a lot of webinars, whose sponsors send me email announcements. Some of the webinar providers have accessible lists of them on their web sites”*



- ❖ *“JOVE”*
- ❖ *“Mostly I use youtube for videos, or I generally google and then move forward from site to site, accessing articles/ sites following links from there or googling stuff/terms/people/works/words I find in the articles, and so on and on - and so sometimes I WIND UP finding videos in arcane archives of old interviews, scientific hoards, whatever, which are vastly interesting, but had not originally set out, to find them or find any video.”*
- ❖ *“My librarian knows where to find things.”*
- ❖ *“Professional society and newsletter linking”*
- ❖ *“RSC experiments teaching video.”*
- ❖ *“Some academics maintain blogs and regularly alert readers to videos that stand out in terms of quality of instruction, unique perspective, or thoroughness in coverage of a given topic.”*
- ❖ *“I use TED quite often and mostly law enforcement you tubes”*
- ❖ *“Academic video producers”*

## 4. 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 32 - Browse, 2005-2018**

Survey on survey we have seen a steady rise in the importance of abstracting and indexing databases in browse, but this time around we have seen a very large leap. The same has become true of search engines. Many A&Is have some faceted search functions which can effectively be used as browse, and so this may be accounting for the behaviour. In addition, perhaps readers are using Google Scholar to find a title or a few titles, then limit by year, which would effectively be more browse-like than search-like. Perhaps a take-home message here is that browse and search are becoming more blurred and more similar.

People are using social or professional networking sites for staying up to date, perhaps knowing that if they follow certain pages or people, the latest research will be presented to them in an easy-to-consume manner.

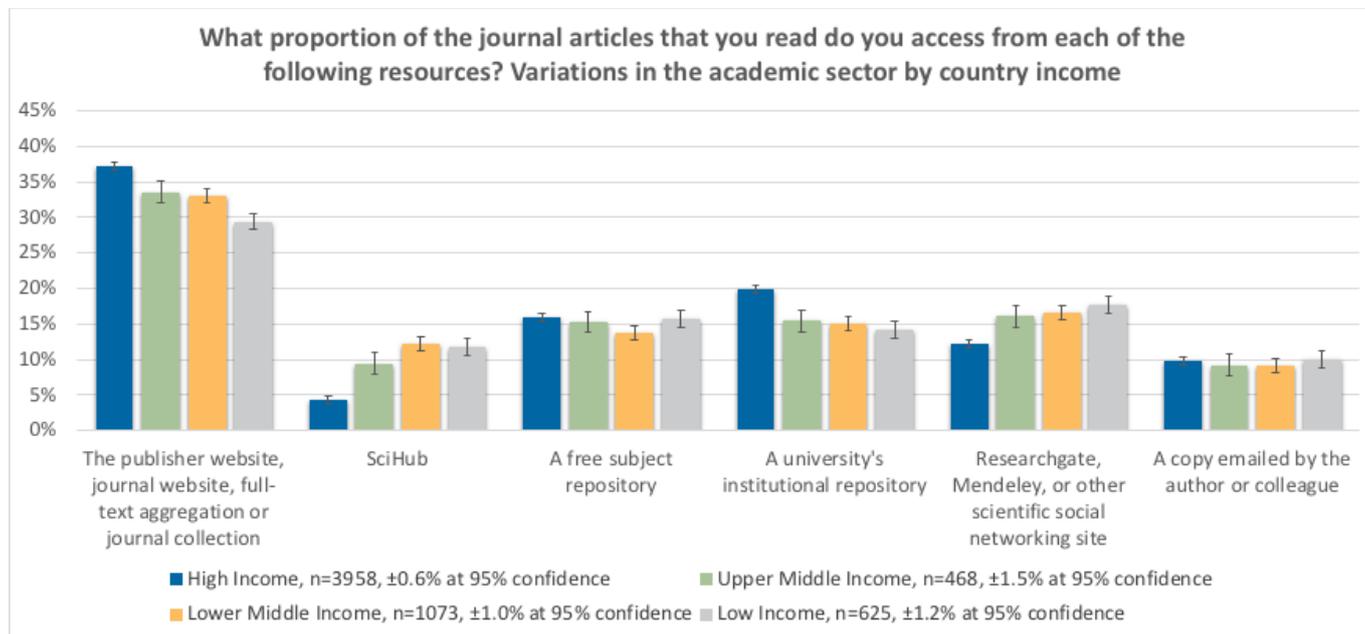
Journal alerts were once the most popular way of staying up to date with the latest research – users would receive an email alerting them to a new issue and either act on it straight away or perhaps save it for reference. As Figure 32 shows, this is certainly no longer the most popular method for browsing 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.



*These results give a very strong indication that journal alerts have lost their popularity as a method for staying up to date with the latest research, and have been replaced by discovery resources which are more normally seen as search resources.*

## 5. DELIVERY SITES FOR JOURNAL ARTICLES

The rise in popularity of sites like ResearchGate coupled with the increase in the articles being shared in subject and institutional repositories, and the expansion of Sci-Hub has led publishers and librarians to question how often users are actually downloading the published version of record.



**Figure 33 - Delivery sites, academic by income, 2018**

Figure 33 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 34), 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.

This year's research looked at the impact on delivery of SciHub for the first time, and the results show a level of usage of up to 13% in some territories, and under 5% in higher income countries.

Figure 34 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.

### What proportion of the journal articles that you read do you access from each of the following resources? Variations in the academic sector by country income

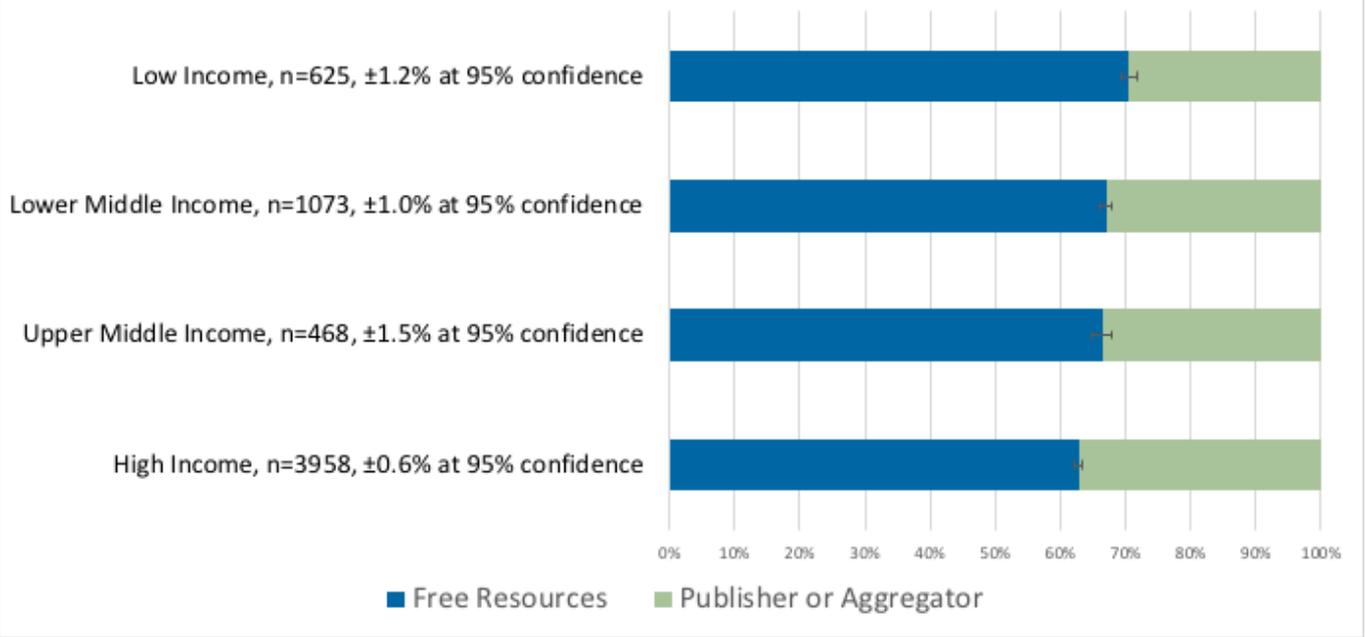


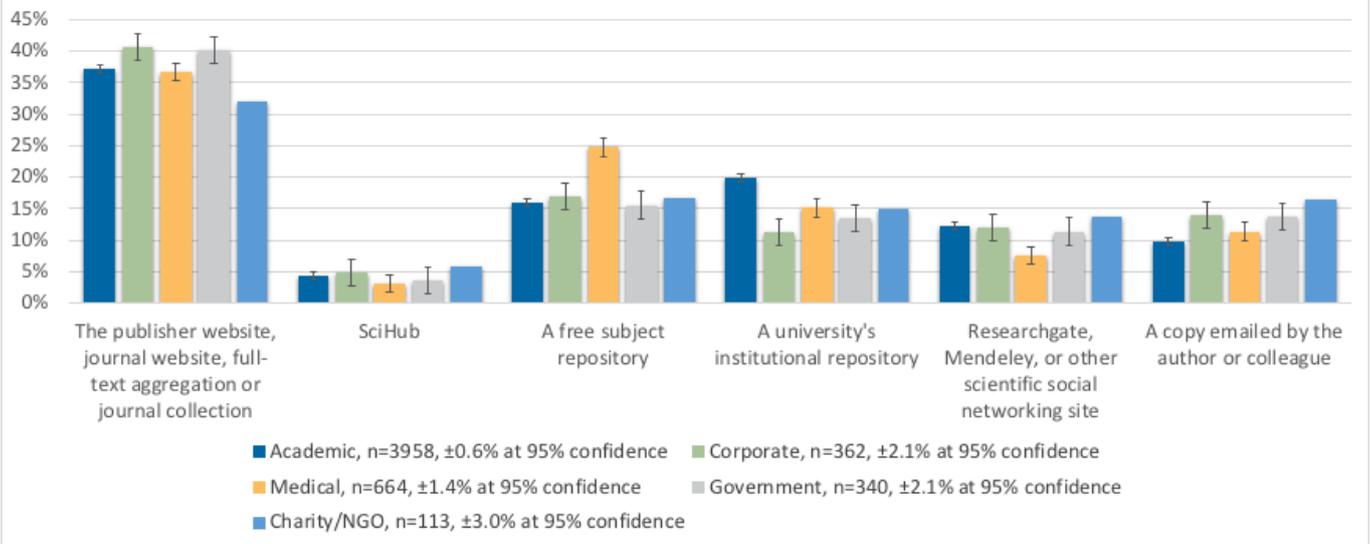
Figure 34 - Delivery, free resources vs publisher-controlled sites, 2018



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.

### What proportion of the journal articles that you read do you access from each of the following resources? Variations by sector in high income countries



**Figure 35 - Delivery site, high income countries, by sector, 2018**

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 20 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 use a subject repository significantly more than other sectors do, and that repository is presumably PubMedCentral.

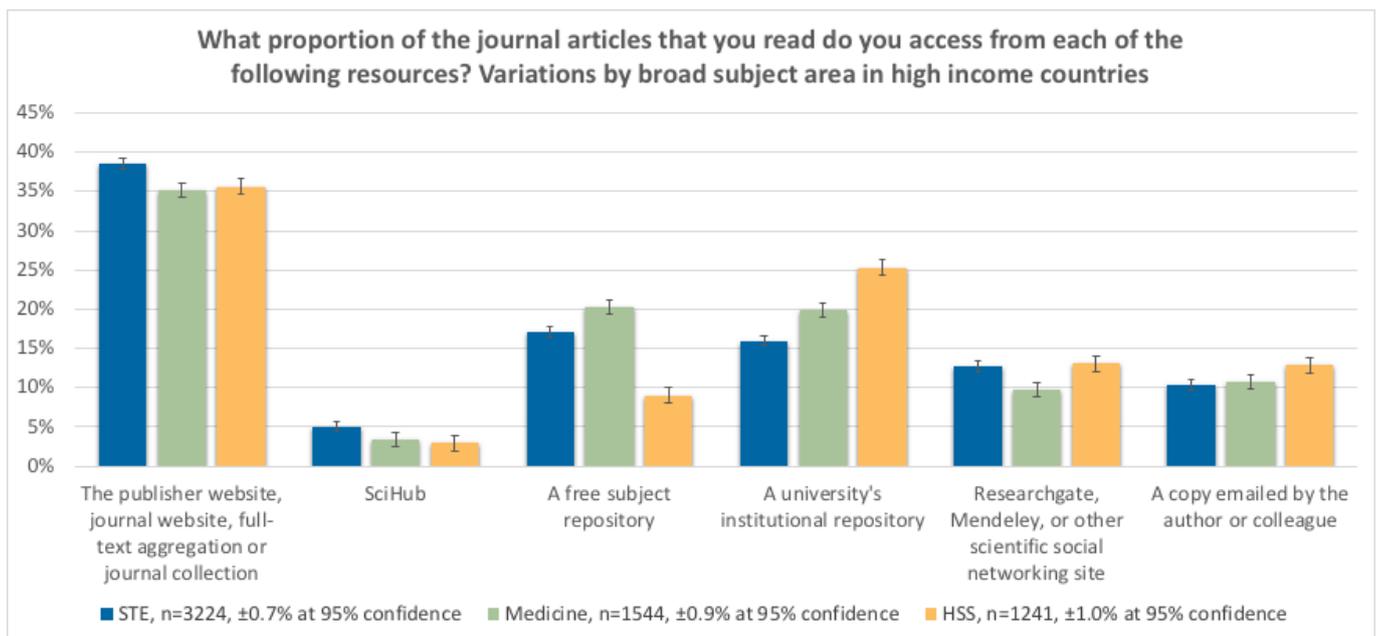


*People in the medical sector are accessing journal articles from a free subject repository 25% of the time. This is significantly higher than all other sectors in high income countries.*



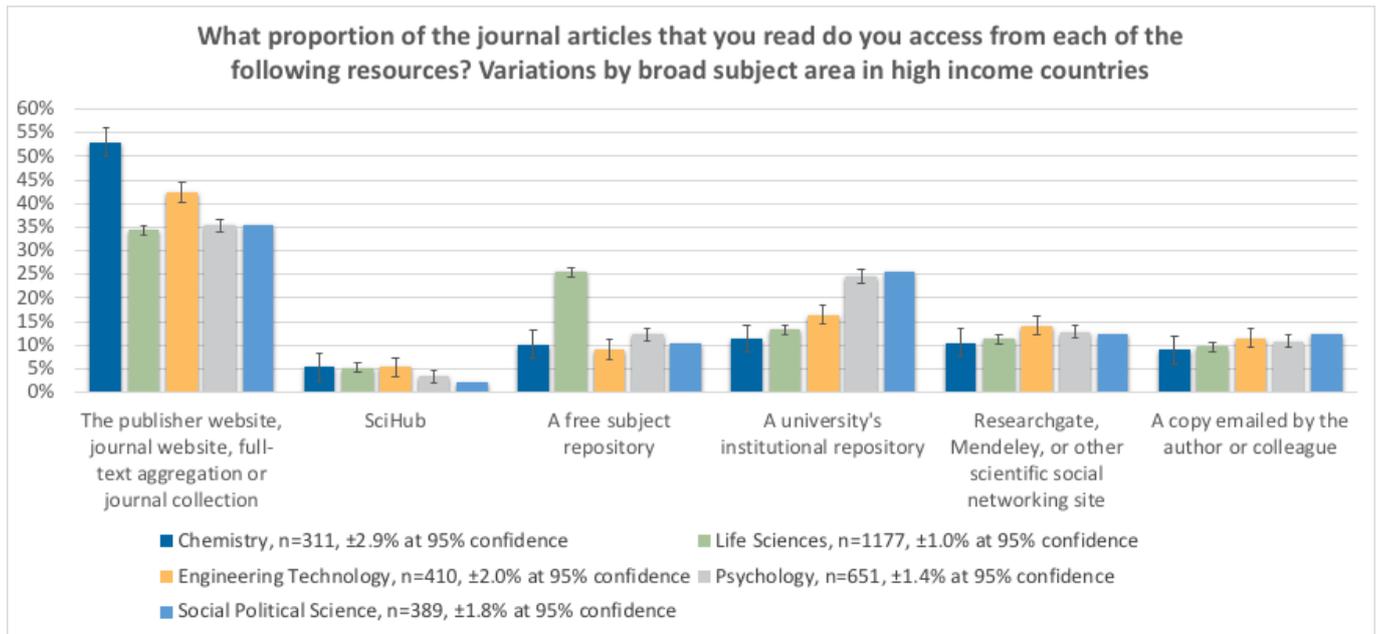
*It seems probable that a lack of awareness of Google Scholar in poorer nations is leading to a reduced use of free incarnations of content in institutional repositories from these regions.*

Figure 36 shows that the total percentage of downloads from publisher websites is very similar from broad subject to subject, but the differences in other resources is very significant. There are clearly fewer subject repositories in Humanities and Social Sciences, but workers in these areas make significant use of institutional repositories instead. ResearchGate, Mendeley and other similar networking sites are significantly more used than SciHub.



**Figure 36 - Delivery site, high income, broad subject areas, 2018**

Figure 37 shows how these differences break down to a selection of narrower subject areas. Chemists certainly prefer the publisher website and presumably have the funding that gives them access to these resources. Life scientists seem to be using PubMedCentral (subject repository), whereas the psychologists have greater reliance on institutional repositories.



**Figure 37 - Delivery site, high income, by subject, 2018**



*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.*

# PUBLICATIONS



## Looking for the Premier Publisher for the Aerospace Industry?

Visit the World's # 1 Source for Aerospace Information Today

[arc.aiaa.org](http://arc.aiaa.org)

**Books & eBooks:** The most comprehensive selection of aerospace titles

**Journals:** The leading repository of aerospace innovation

**Meeting Papers:** The most timely, cutting-edge aerospace research

**Standards:** The best practices guiding the aerospace community

**Aerospace America:** The magazine that shapes the future of aerospace

**Other Publications:** Momentum and Policy Papers

[aiaa.org/Publications](http://aiaa.org/Publications)



# 6. MOBILE DEVICES

## 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%.

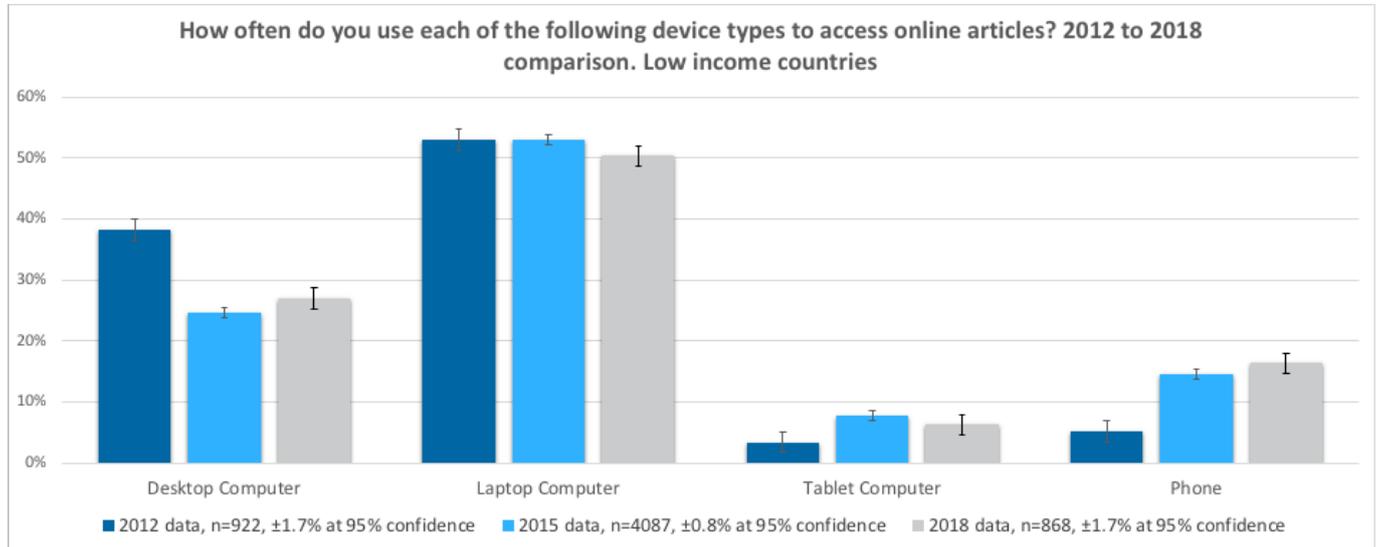


Figure 38 - Devices, low income, 2012-2018

We looked at this comparison for high income countries and the decline in use of desktop computers was only around 4% in favour of phones and tablets since 2012 and has stayed steady since then.

Looking at the 2018 results only, in Figure 39, we can see there is some variance by income classification.

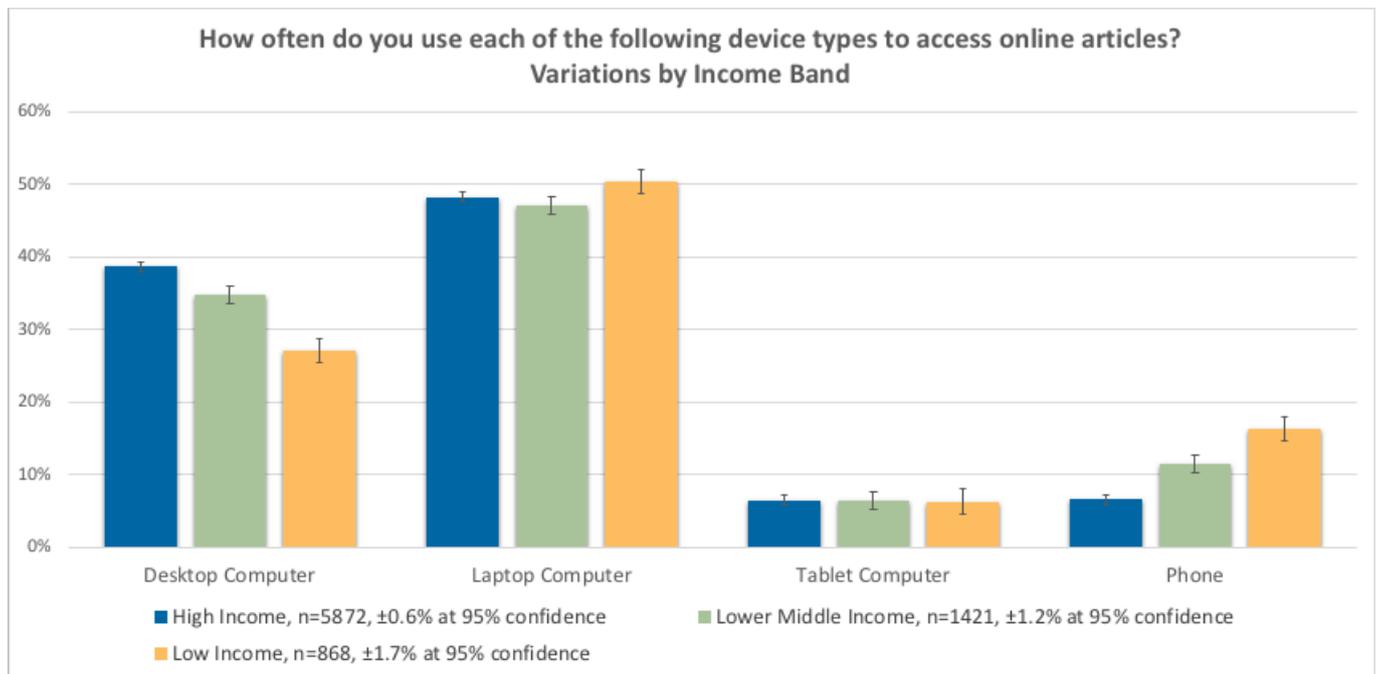


Figure 39 - Devices by income band, 2018

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 40 shows that in the medical sector across the world, there was growth in the use of tablets and mobile phones to access journal articles and but that has evened off since 2015.

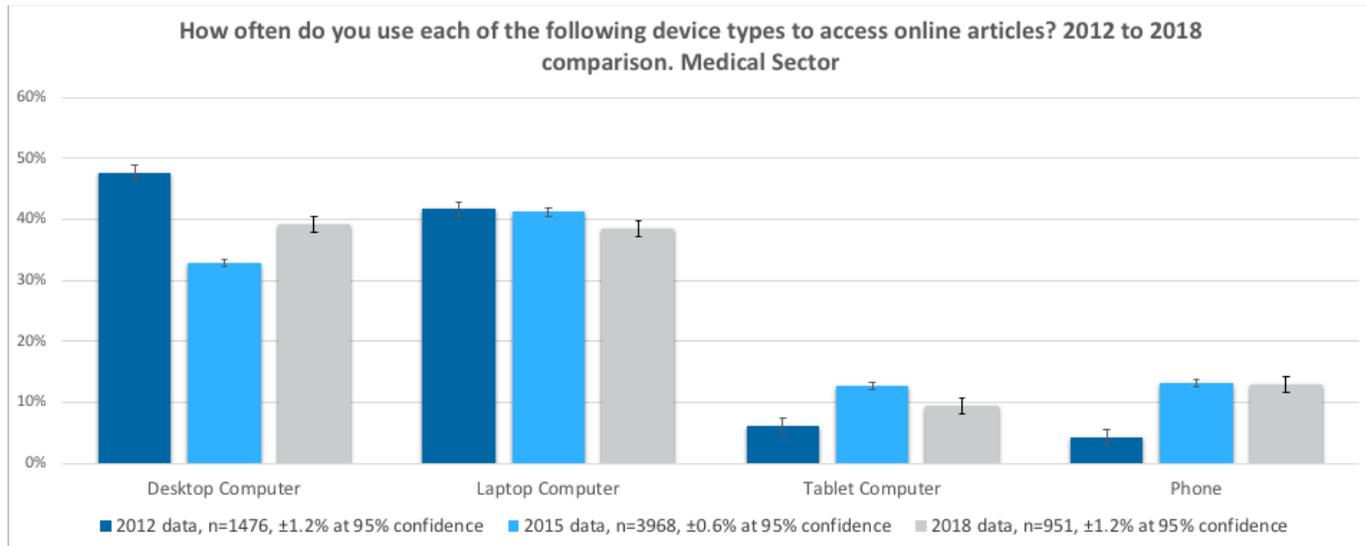


Figure 40 - Devices, medical sector, 2012-2018

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.

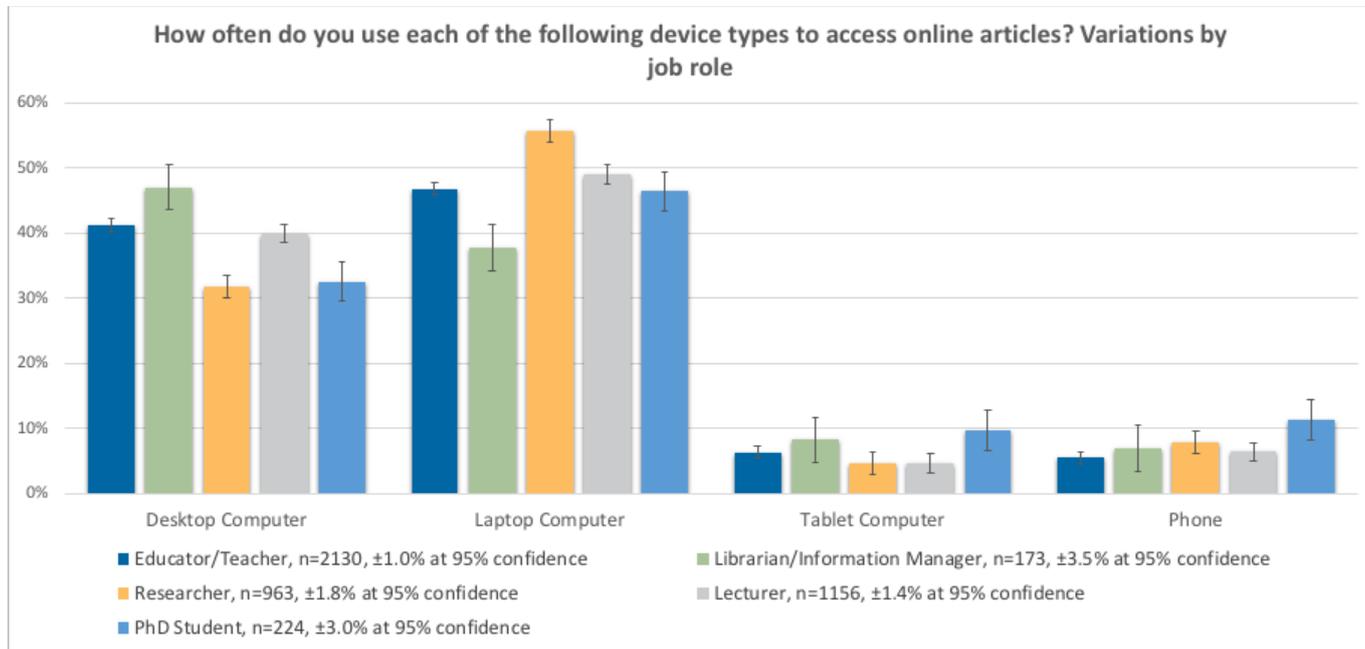


Figure 41 - Devices, job role, 2018



The growth in the use of mobile devices to access academic content has levelled off since 2012.



## 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.) Note that this chart uses a normalised population to match the 2005 demographic as noted earlier.

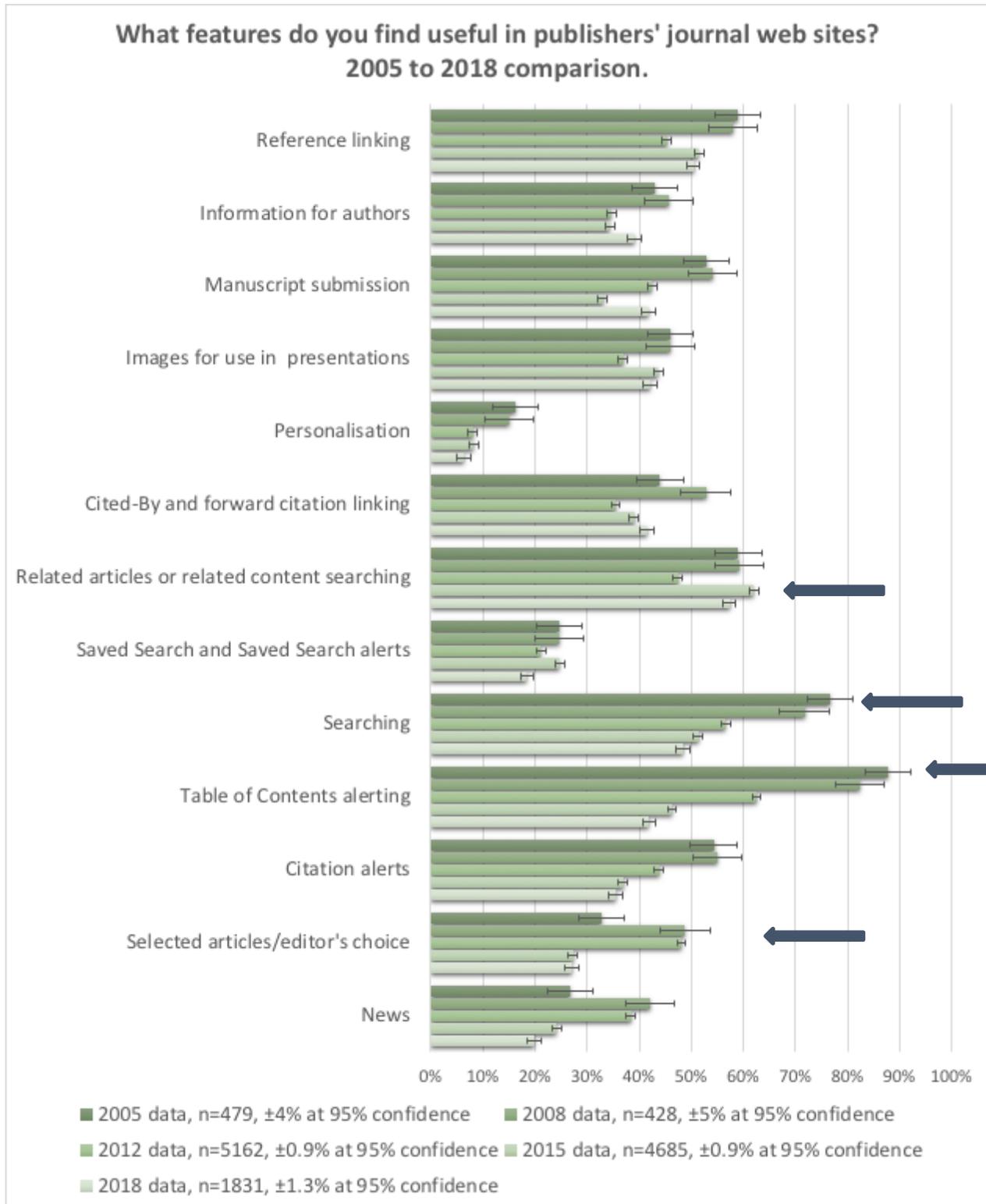


Figure 42 - Journal web site features, 2005-2018

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

Related Articles functionality is the most useful feature of those tested, and has maintained its position fairly consistently over the period of study.

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.

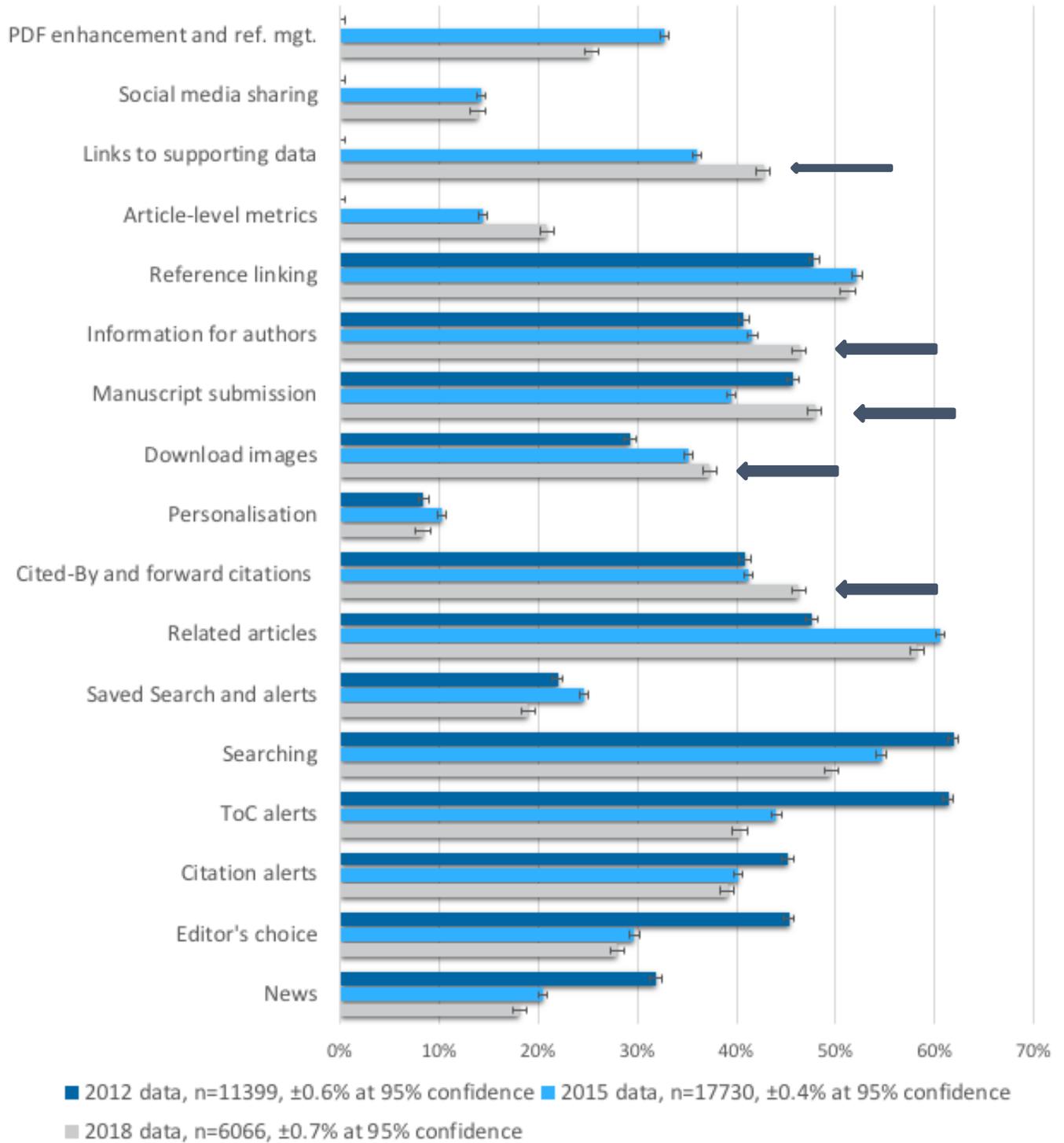
It is interesting that whilst we have shown that over an extended period of time that publisher websites have experienced a growth in importance as a starting point in search, fewer people have rated the search functionality as useful – under 50% of respondents said they found search features useful now compared with 75% of respondents 13 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. Or thirdly, it could be a matter of brand recognition and trust that means readers are more comfortable starting there.

Figure 43 looks at the 2012-2018 trend in the academic sector. We see that links to supporting data, information for authors, manuscript submission, the ability to download images and “cited by” have all become more popular. Related Articles is still the most popular feature.



*Related-articles remains 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. Author features have grown in popularity.*

## What features do you find useful in publishers' journal web sites? 2012 to 2018 comparison. Academic sector



**Figure 43 – Journal web site features, academic, 2012-2018**

Looking at the data for just 2018 (Figure 44 and Figure 45) 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 high income countries. They also value Editor's Choice type features far more than people in higher income countries – perhaps because these are often free and it's an easy way of getting up to speed with the important research in their area.

People in higher income countries value links to supporting data and “cited-by” features more than those in lower income countries.

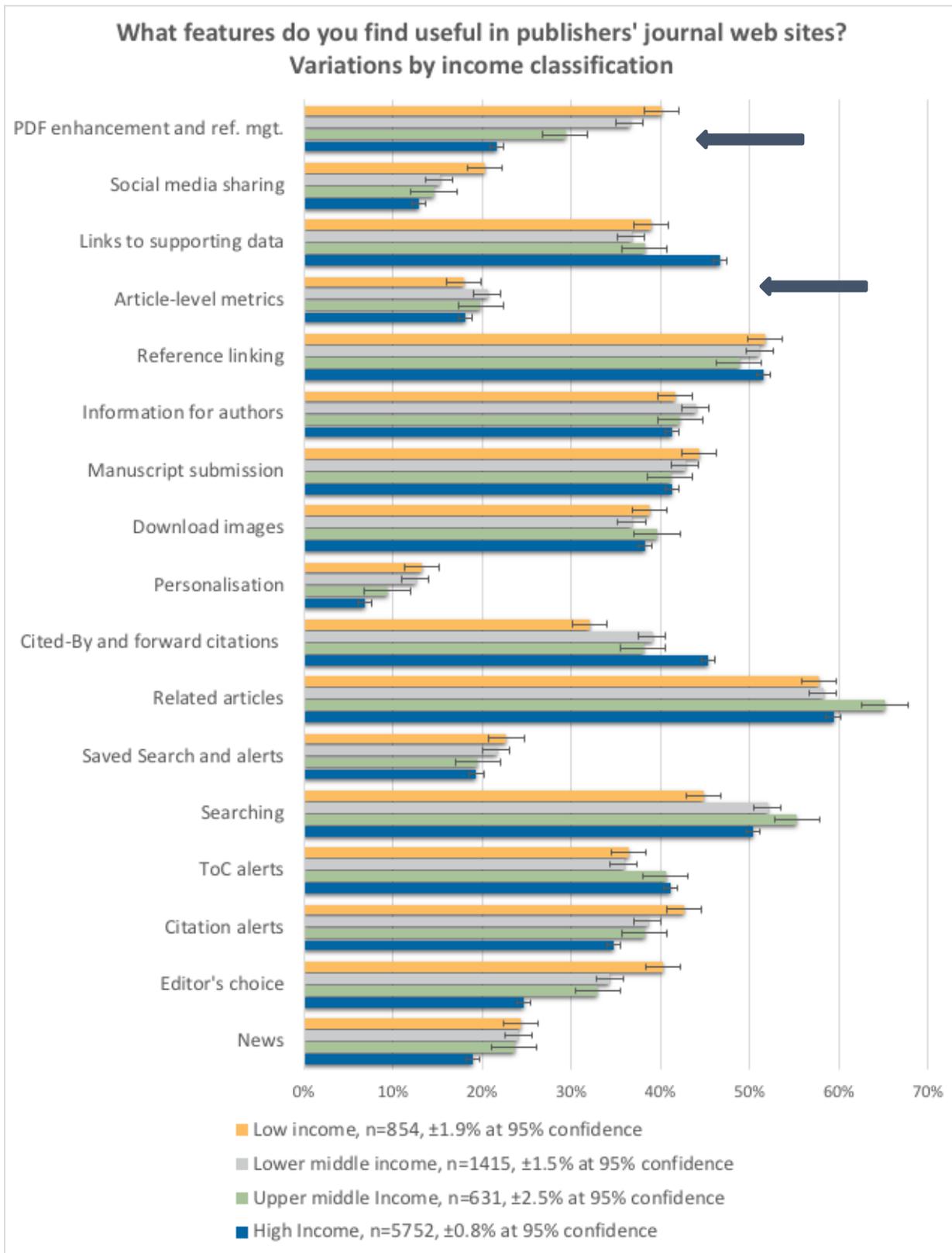


Figure 44 - Features, income, 2018

Figure 45 shows the usefulness of features across broad subject areas. We can see there are many significant differences. People in HSS value Related Articles more than people in STE and Medical subjects – around 70% of respondents said they found Related Articles useful. They also still value ToC alerts slightly more than people in other subject areas, perhaps because continuous publishing is not as prevalent in Humanities as the sciences.



People working in STE subjects find author information, manuscript submission, links to supporting data and cited by features useful.

Related articles and reference linking are the most useful features for people in medical subjects.

Related articles and reference linking is important across all subject areas with around half of all respondents indicating they found it useful.

Social media sharing and article level metrics were stated as being useful by less than 20% of respondents.



*The features people find useful on a publisher website varies significantly across demographics. Related articles and reference linking are generally the most useful features. Social media sharing and article level metrics are not particularly popular in any demographic studied*

# Wiley Online Library

Accelerating research discovery  
to shape a better future

WILEY

Wiley Online Library hosts an integrated set of features that will help researchers, professionals, and other readers discover and engage with the content that is most relevant to them.

- 1,600+ peer reviewed journals (over 4 million articles)
- 21,000+ online books
- 200+ multi-volume references and handbooks
- 18 Current Protocols (Laboratory Manuals featuring over 18,000 protocols)
- 13 databases (chemistry & evidence-based medicine)

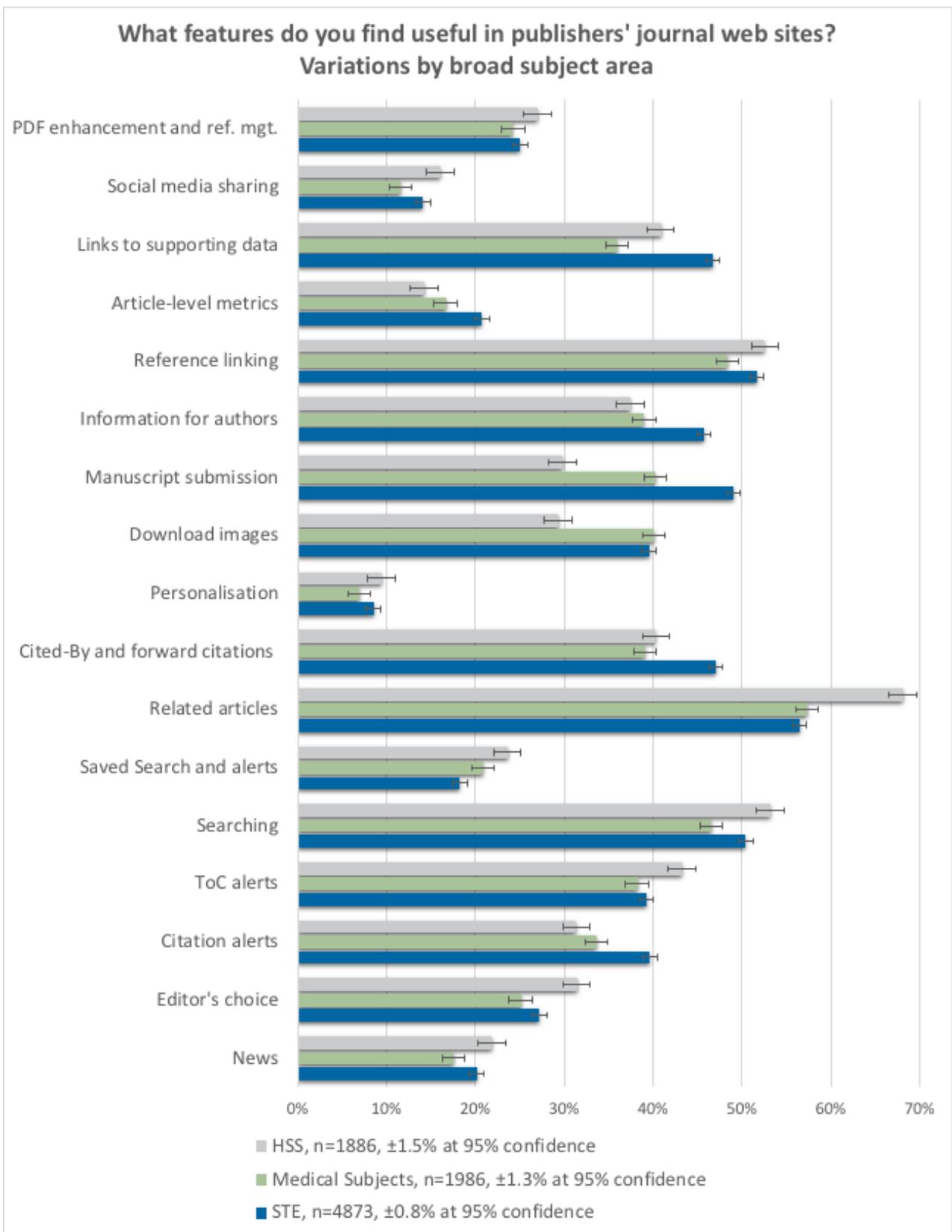


Figure 45 - Features, subject, 2018

## 8. METHODOLOGY

This research carries on from, and expands upon, previous research undertaken in 2005, 2008, 2012 and 2015 (also by Simon Inger and Tracy Gardner) and attempts to follow the trends in behaviour over that period. 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 two key questions on reader behaviour (broadly speaking, 'search' and 'browse') 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 2018. Two questions have been retired since the research in 2015.

One of the key limitations of the older surveys was their reach. However, the 2012 to 2018 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 five 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, and France, but their clientele is 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.

10976 people embarked on the survey, and 10752 of them completed the demographics stage. At the end of the demographics, 9116 people elected to take the journals survey, 7520 elected to take the books survey, and 3405 took the video survey.

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, books and academic videos 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

---

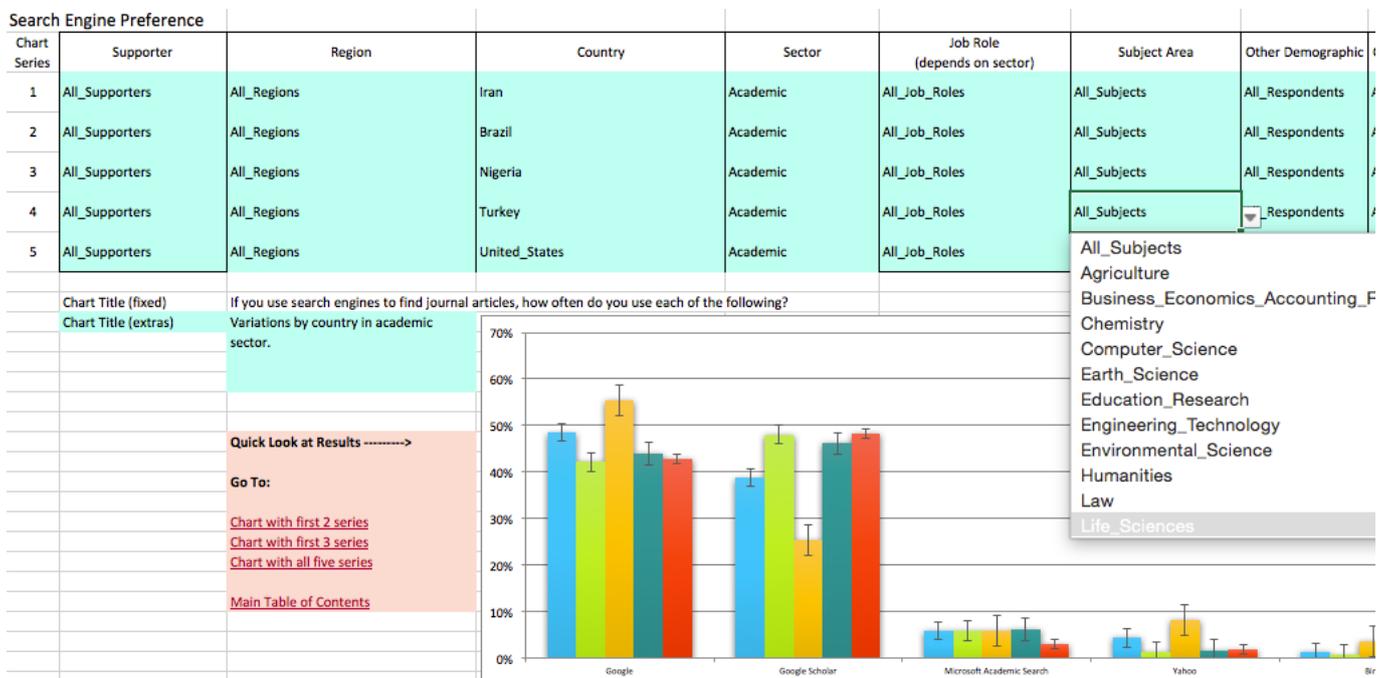
Analysis was performed using Excel. Many of the results were reduced to binary arrays to make further comparisons easy to calculate. A tool was developed so that a number of comparisons could be made for each survey question easily. Figure 46 shows how the tool works, creating 5 series of data, each 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 46 – 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, 2015 and 2018 data we have created a random sampling of the respondents to the 2012-2018 data down to the 2008 levels. This random sampling included a probability calculation which created broadly the same demographic breakdown in the 2012-2018 “down-sampled” sets as in the 2008 set. For example, in 2015, approximately 1 in 30 of the responses of those in Humanities were used, and 40% of them would be in North America. In 2018, that same sample was achieved by taking 1 in 16 of the responses in Humanities, of which about 45% are 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 each of 2012, 2015, and 2018 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 earlier surveys are now considered redundant and have been dropped.

## 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. Some of the A&Is now effectively allow for a browse function, through faceted search, narrowing results perhaps to a title and date range, somewhat similar to a journals homepage, and we have seen a rise in importance of these resources in browse presumably as a consequence.,

### 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. Many libraries have reduced budgets for A&Is and have cut their purchase from peripheral or under-funded subjects, driving discovery to library resource discovery tools instead.

### 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. With millions of links, or in some cases millions of documents stored within them, they have become an important alternative to other discovery methods, with the added advantage of linking the content to individuals with profiles in the databases. 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. There may well be a degree of brand familiarity, and hence trust, that helps drive their importance.

## 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. However, as content is increasingly published ahead of the issue publication date ('online first') table of contents alerts by definition provide a rather late alert for much of the content, whereas subject or search based alerts will be more current.

## 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 retried quickly and is a frequently cited reason for their popularity. Google changed its policy on the indexing of pay-walled content several years ago, 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 over the years a clear division has emerged in the use cases for each resource, and the academic search engines provide more consistent results from reader to reader, whereas engines like Google learn an individual's interests and tailor the results to them. This may be good for some, and bad for others.

## 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.

## DELIVERY SITES

---

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%	70.2%
Between 30% and 70%	39.2%
Less than 30%	9.8%
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

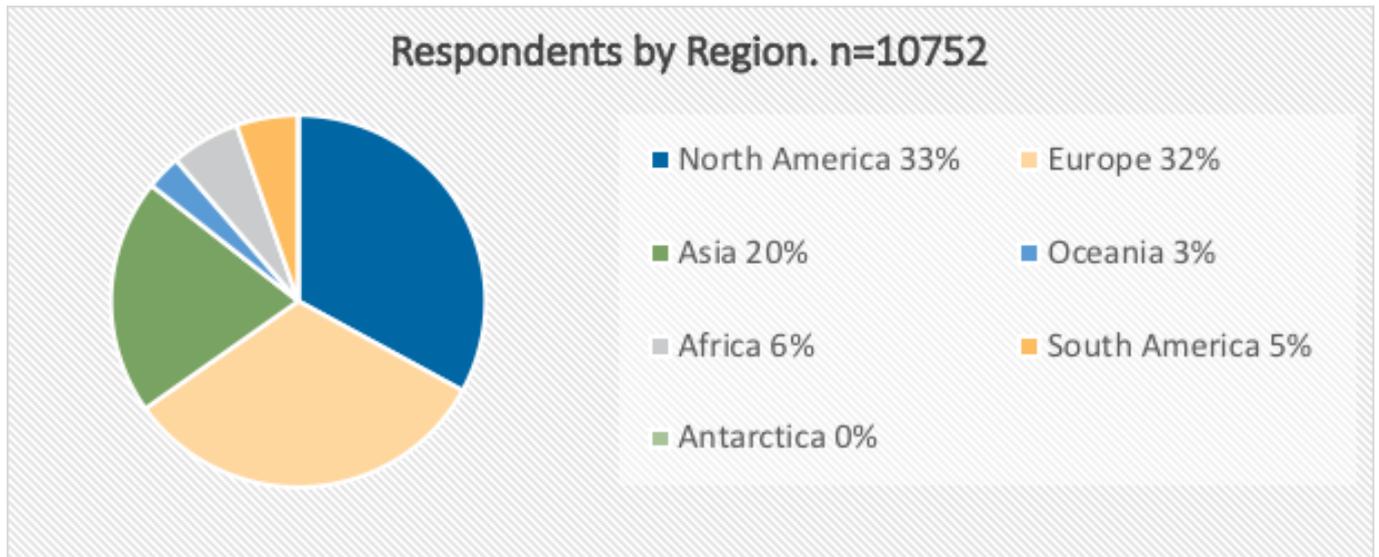


Figure 47 – Respondents by region, 2018

As shown in Figure 47, 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. 10976 people told us their country (and hence region).

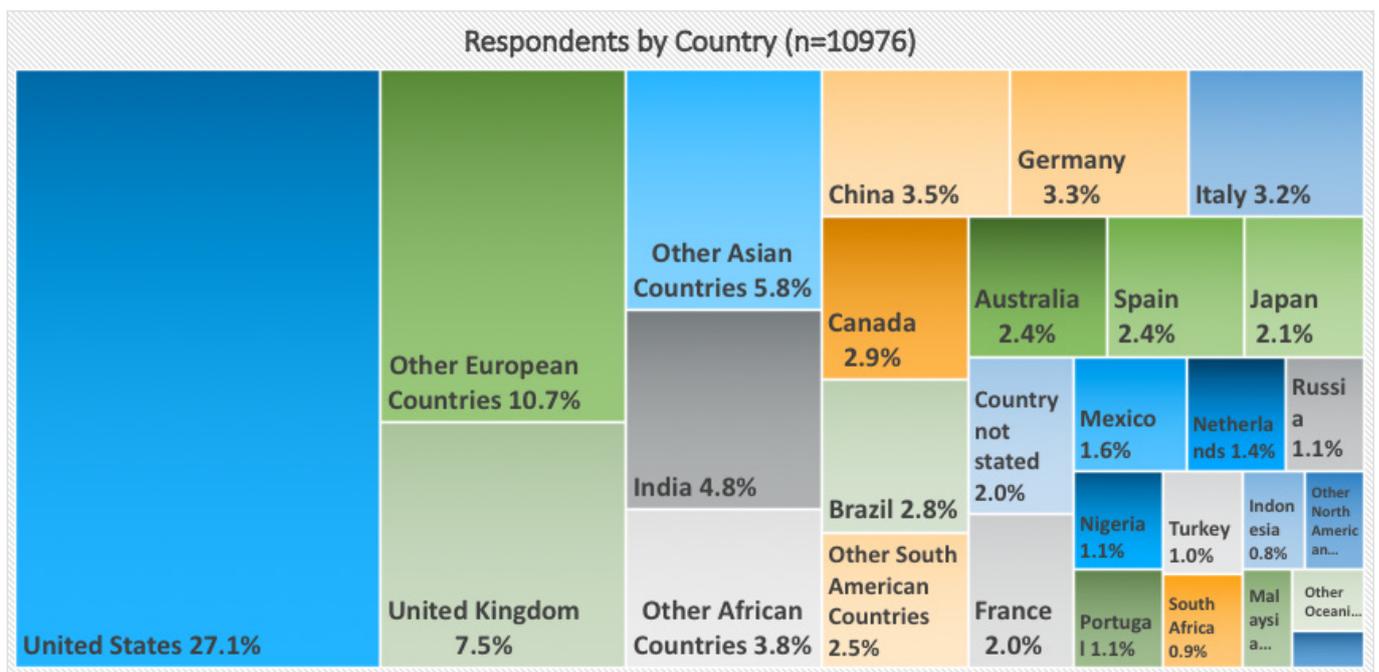


Figure 48 – Respondents by country of origin, 2018

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.

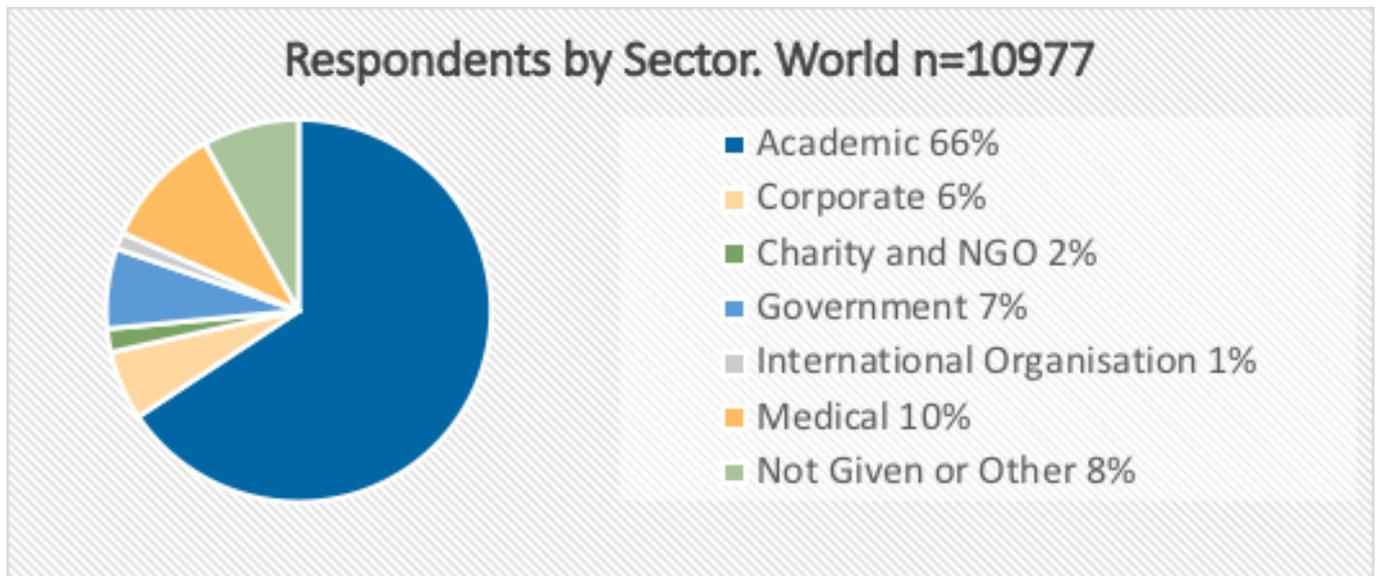


Figure 49 – Respondents by sector, 2018

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. 10977 people told us which sector they worked in. Although the academic sector dominates, even the 6% of respondents from the Corporate sector represents over 650 people.

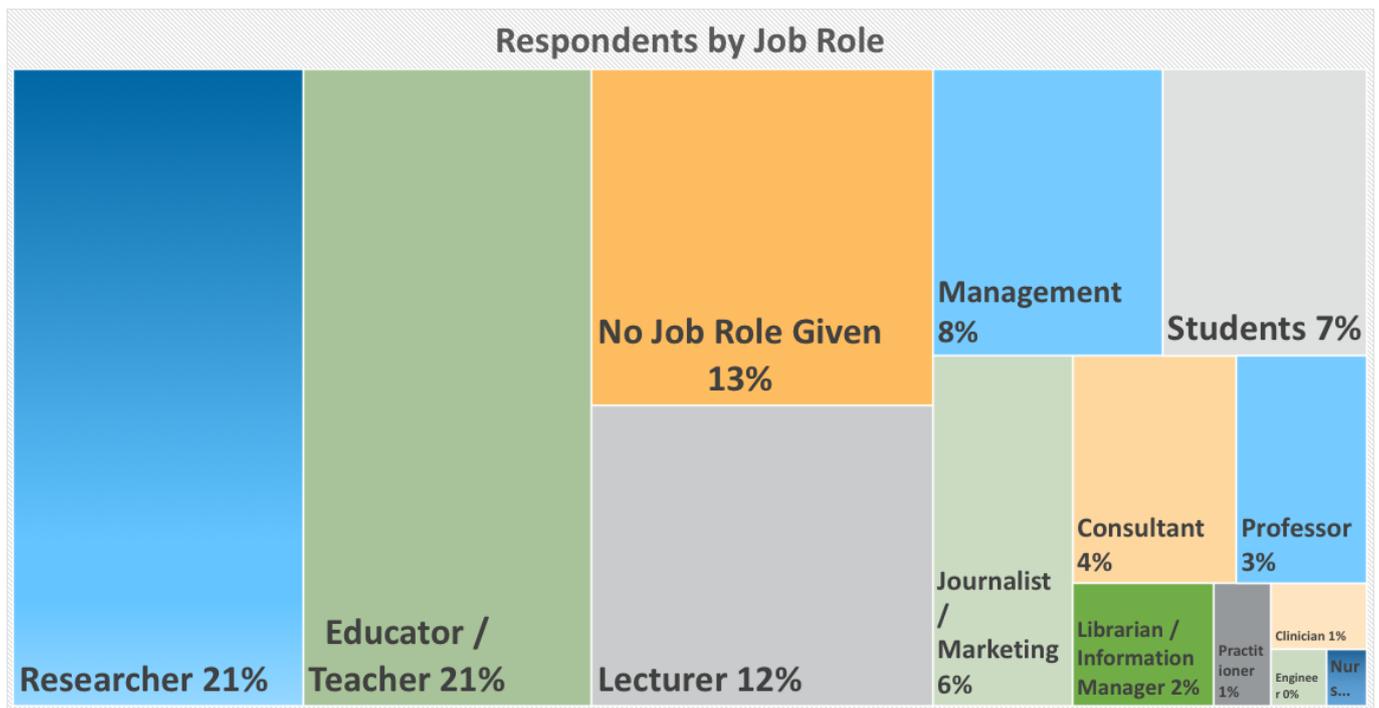


Figure 50 – Respondents by job role, 2018

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



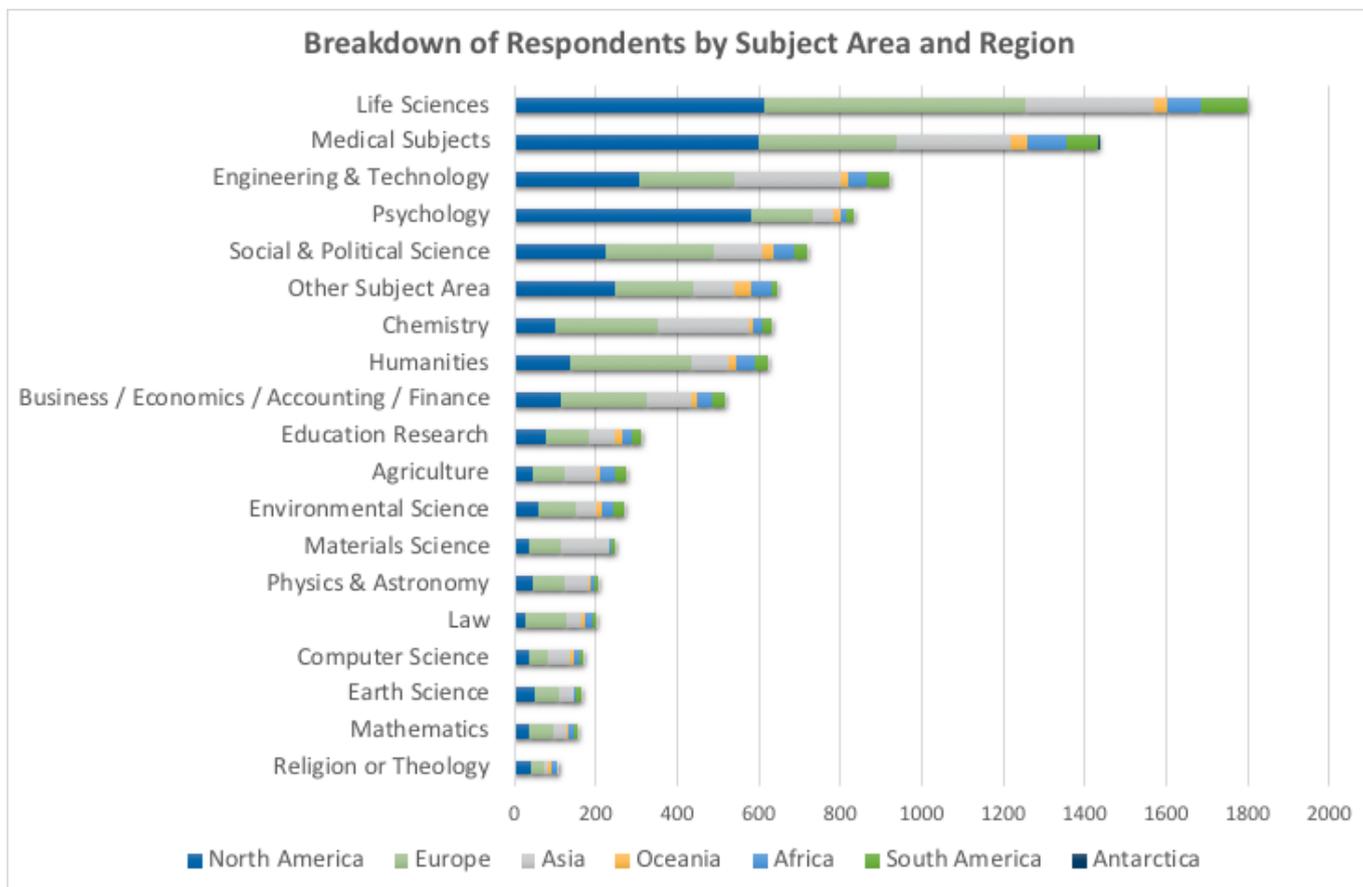


Figure 51 – Respondents by subject area and region, 2018

When embarking on this project, our aim was to get around 600 responses in each subject area, so that detailed demographic analysis would be possible within subjects with minimal error-bars. This was achieved for eight of the subject areas. Computer Science, Earth Science, Religion and Mathematics 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.

## Respondents by World Bank Country Income Classifications

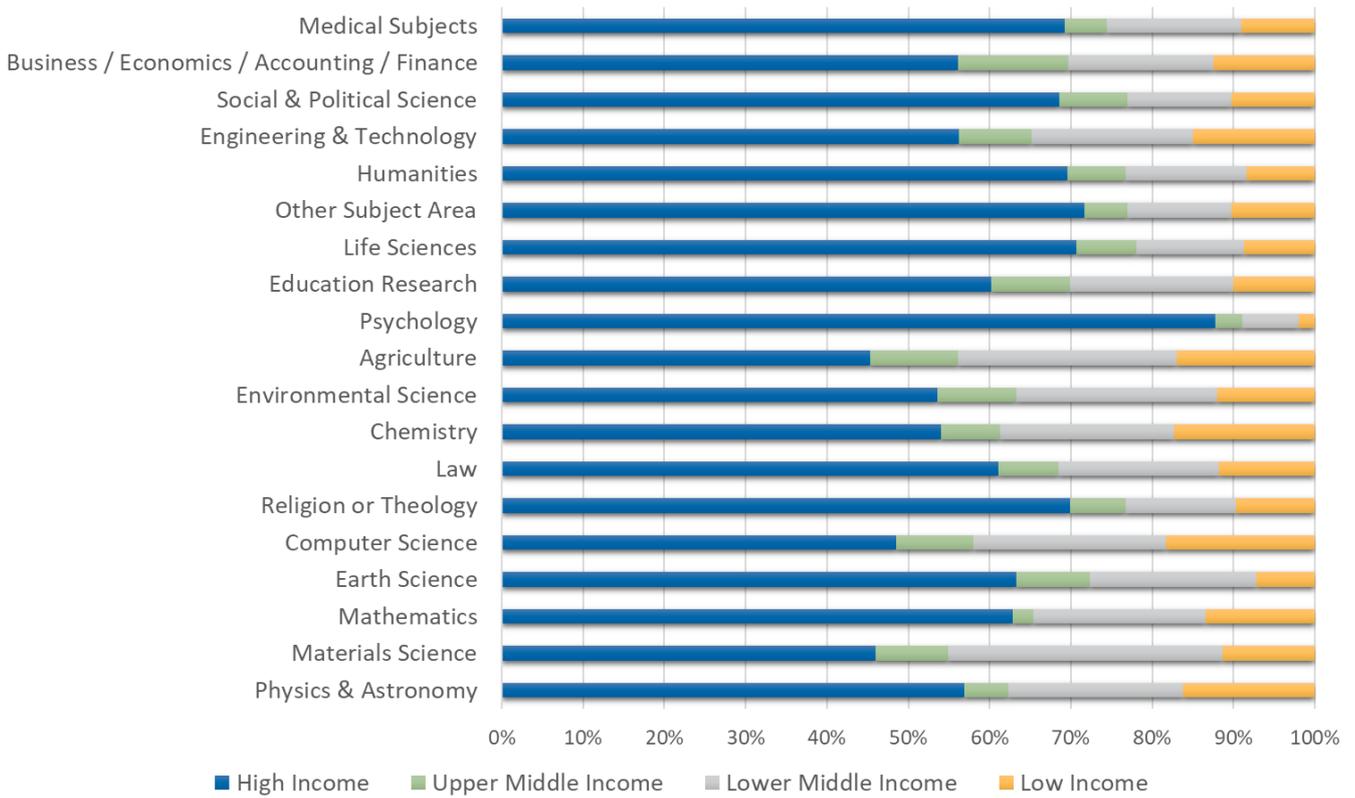


Figure 52 – Respondents by World Bank income categories and subject (2018)

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.



## 10. CONCLUSIONS

There is no one dominant player in content discovery. Whilst the large and influential A&Is and search engines continue to lead, other methods of discovery continue to be important. A reminder to publishers that all bases need to be covered when thinking about their content distribution strategy – it certainly isn't simply all about Google.

Perhaps somewhat surprisingly, A&Is are holding their ground as a discovery resource in the academic and medical sectors although academic search engines are the most important resource for academic researchers. Library discovery appears to have peaked and is only really important for people in HSS.

Social Media has not grown in importance as people perhaps expected it to, although certainly HSS subjects have caught up with STEM since 2012. The important thing to note is that social media is less important across many demographics than other discovery resources.

Video content is generally found via Google or YouTube but medics in particular are finding video content within the journals they read.

In terms of delivery we estimate around 60% of the time people are reading articles from a “free” resources. We think PMC is responsible for much of this in the medical sector. It stands to reason that if Google Scholar (and PubMed in medicine and life sciences) are they key discovery channels, and they make the “free” incarnations easy to navigate to from their sites, then people will be finding and accessing the free incarnations of content. SciHub is responsible for around 5% of downloads in the wealthier nations, and must pose a special threat since the content incarnations are normally the final published version.

There continues to be very significant regional and sectoral differences in how readers discover content.

# 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.