

Where and how early career researchers find scholarly information

David Nicholas ^{1,2*} Cherifa Boukacem-Zeghmouri ³ Blanca Rodríguez-Bravo ⁴
Jie Xu ⁵ Anthony Watkinson,² A. Abrizah ⁶ Eti Herman ² and Marzena Świgoń⁷

¹Department of Russian History, Tomsk State University, Lenin Avenue 36, Tomsk 634050, Russian Federation

²CIBER Research Ltd, Newbury, Berkshire RG147RU, UK

³Université de Lyon, Université Lyon 1, 69100 Villeurbanne, France

⁴Biblioteconomía y Documentación, Universidad de León, 24071 León, Castilla y León, Spain

⁵School of Information Management, Wuhan University Ringgold Standard Institution, Wuhan, Hubei, China

⁶Faculty of Computer Science & Information Technology, University of Malaya, 50603 Kuala Lumpur, Malaysia

⁷Wydział Humanistyczny, Uniwersytet Warmińsko-Mazurski, 10-719 Olsztyn, Poland

ORCID:

D. Nicholas: 0000-0001-8046-2835

C. Boukacem-Zeghmouri: 0000-0002-0201-6159

B. Rodríguez-Bravo: 0000-0002-9476-7602

J. Xu: 0000-0002-9820-8066

A. Abrizah: 0000-0002-8224-5268

E. Herman: 0000-0001-8526-9081

*Corresponding author: David Nicholas

E-mail: Dave.Nicholas@ciber-research.eu

Abstract

This article presents findings from the first year of the Harbingers research project started in 2015. The project is a 3-year longitudinal study of early career researchers (ECRs) to ascertain their current and changing habits with regard to information searching, use, sharing, and publication. The study recruited 116 researchers from seven countries (UK, USA, China, France, Malaysia, Poland, and Spain) and performed in-depth interviews by telephone, Skype, or face-to-face to discover behaviours and opinions. This paper reports on findings regarding discovery and access to scholarly information. Findings confirm the universal popularity of Google/Google Scholar. Library platforms and web-scale discovery services are largely unmentioned and unnoticed by this user community, although many ECRs pass through them unknowingly on the way to authenticated use of their other preferred sources, such as Web of Science. ECRs are conscious of the benefits of open access in delivering free access to papers. Social media are widely used as a source of discovering scholarly information. ResearchGate is popular and on the rise in all countries surveyed. Smartphones have become a regularly used platform on which to perform quick and occasional searches for scholarly information but are only rarely used for reading full text.

INTRODUCTION

This paper reports the first-year findings of a 3-year, qualitative longitudinal study of 116 international early career researchers (ECRs). The Harbingers research project was initiated by the Publishing Research Consortium (PRC) to establish where and how 'digital natives' find scholarly information as part of the PRC's broader scholarly communications brief. More specifically, the Harbingers project set out to discover whether ECRs are taking advantage of the myriad discovery opportunities, such as the

ubiquitous smartphone, which are at their disposal in an increasingly open, borderless, and social scholarly information world.

Finding, searching, or discovering information, call it what you will (ECRs do not distinguish), needs to be understood in the context of the new world information order. We should be conscious that in a global, ever more open information world, where searching can be conducted seamlessly, anywhere and anytime, new (and old) researchers are going to be doing it differently. For a review of the recent (2010–2015) literature on researchers' information-seeking behaviours, see Spezi (2016); for a longitudinal view of the

Key points

- Google and Google Scholar are hugely popular with early career researchers (ECRs), irrespective of discipline, country, or language.
- Library platforms are used to obtain full-text documents, because often there is no other choice, but rarely mentioned by ECRs.
- The evergreen discovery tools, Web of Science and Scopus, are well known, popular, and trusted.
- Social media, online community platforms, and smartphones are used for finding scholarly information in all countries.
- Studies of the topic need to be cognizant that discovery is ubiquitous, seamless and, often, unconscious.

developments in this area, see the series of *Ithaka S+R US faculty surveys* and the *Ithaka S+R/Jisc/RLUK surveys of UK academics* (Housewright, Schonfeld, & Wulfson, 2013a, 2013b; Wolff, Rod, & Schonfeld, 2016a, 2016b). See also the NeoRef research group's national studies into the information-seeking behaviour of US academic scientists, reported in Hemminger, Lu, Vaughan, & Adams, 2007 and Niu *et al.*, 2010.

True, junior scientists, as arguably the most vulnerable populations in the scholarly community, usually cope with their challenging circumstances by following the norms of their chosen discipline (Harley, Acord, Earl-Novell, Lawrence, & King, 2010). Nevertheless, born digital as they are, they know no other reality than the ease of access to a large, diverse information universe, provided by digital resources, with catalogues and bibliographies sounding to them like foreign bodies and libraries a largely invisible presence. Thus, there is every reason to believe that they might be the harbingers of change in the scholarly enterprise, especially as among the many factors affecting the specific information-seeking behaviours of scientists, including demographic, psychological, role-related, and environmental factors, academic position was found to be the most important (Niu & Hemminger, 2012).

Also, importantly, searching is such a common event in the virtual space, for everything from holidays, clothes to journals, that it is not the conscious event it once was, to the point that we can now equate it with digital 'breathing'. This renders searching for information, at one and the same time, simpler and more complicated. Simpler because it is something (we think) we as scholars know all about; after all, we do it all the time and quite successfully too. Still, it is more complicated, too, because there is so much information around to be had so easily and in so many formats. What is important, not to say essential, and what is negligible? What is reliable and trustworthy and what is not? Plainly then, the same old questions, presumptions, and interpretations will simply not do, which is why questionnaire surveys so often provide confusing data on the

topic and leave us with more questions than answers. That is the reason why, when determining how best to study the subject, we chose long, deep, and repeat interviews and to ask questions in a broad scholarly context.

SCOPE AND DEFINITIONS

There are different and conflicting definitions of early career researchers circulating (Poli, 2016), and they vary from country to country. We opted for the following, which obtained broad agreement from all our national partners:

Researchers who are generally not older than 35, who either have received their doctorate and are currently in a research position or have been in research positions, but are currently doing a doctorate. In neither case are they researchers in established or tenured positions. In the case of academics, they are non-faculty research employees of the university.

The main focus of the study is on ECRs in the sciences and social sciences, which is where the main funder's (PRC) priorities lie and, also, where the vast majority of ECRs come from (see, e.g. Higher Education Funding Council for England [HEFCE], 2015). The study aimed also to obtain a wide geographical reach as we wished to support research on issues facing the Sciences, Techniques and Medicine (STM) industry globally. Balancing the need for representativeness (with regards to size, importance, level of development, and language) with PRC interests and the availability of interviewees on the ground, ECRs from the UK, USA, China, Malaysia, Poland, Spain, and France were recruited.

METHODOLOGY

A longitudinal, 3-year investigation is being undertaken, asking the same ECRs the same questions each year in order to map attitudes and behaviour and identify any changes to them. Structured interviews are used to establish a personal link with ECRs in order to obtain their full cooperation over the 3-year period. Interviews were conducted face to face and/or remotely (Skype or telephone). A detailed interview schedule was compiled and sent to interviewees ahead of the interview. The structure and scope of the interview and the nature of the questions to be used were informed by two focus group meetings held prior to the start of interviewing, one with publishers and the other with ECRs recruited through the aid of the aforesaid publishers. The interview schedule contained around 60 questions, and the whole interview took between 60 and 90 min. However, for the purposes of this article, we are only interested in three main questions (although some of the others provided important contextual data):

- How do you find the scholarly information you need? Google, library catalogues, online networks, and so on?
- Do you search for and read scholarly papers on your smartphone?
- Do you use social media in your scholarly activities to find out information and (if so) from what media?

The full list of questions can be found in the online report, *Early career researchers: The Harbinger of change* at http://ciber-research.eu/download/20160916-Harbingers-research_instruments.pdf.

The project was funded to follow around 100 ECRs, but anticipating wastage as the project proceeded, 116 ECRs were recruited from the case study countries (Table 1). In reaching this number, interviewers for the case countries were given a recruitment quota of 20–29 for the UK and USA (the larger number a reflection on the importance of these communities to publishers) and 10–15 for the other countries. Within this number, the general guidance was to build the sample along the following lines: (1) two-thirds science and one-third social sciences (to reflect the larger numbers of ECRs

TABLE 1 Numbers and nationalities of ECRs.

| Country | No. |
|----------|------------|
| Poland | 10 (8.6%) |
| Spain | 18 (15.5%) |
| France | 14 (12.1%) |
| Malaysia | 12 (10.3%) |
| China | 13 (11.2%) |
| UK | 21 (18.1%) |
| USA | 28 (24.1%) |
| Total | 116 (100%) |

TABLE 2 Subject representation of ECRs.

| Subject | Total | Percent (%) |
|----------------------------|-------|-------------|
| Biology and agriculture | 25 | 21.6 |
| Medicine and health | 15 | 12.9 |
| Engineering and technology | 12 | 10.3 |
| Chemistry | 10 | 8.6 |
| Computer science | 10 | 8.6 |
| Physics | 8 | 6.9 |
| Psychology | 6 | 5.2 |
| Other social sciences | 22 | 19.0 |
| Sciences | 88 | 75.8 |
| Other sciences | 8 | 6.9 |
| Social sciences | 28 | 24.2 |

TABLE 3 Gender, age, and status of ECRs.

| | Females | | Males | |
|------------------|----------|----|----------|----|
| | N | % | N | % |
| Gender | 49 | 42 | 67 | 58 |
| Age | Twenties | | Thirties | |
| | 36 | 31 | 80 | 69 |
| Doctoral/Postdoc | Doctoral | | Postdoc | |
| | 28 | 24 | 88 | 76 |

in science), (2) a representative balance of men and women, and (3) a range of ages within the 20s and 30s age groups. The 116 ECRs come from 81 institutions; there are more men in the sample (mainly because there are just more of them, especially in the sciences), and it is generally skewed towards the sciences. See Tables 2 and 3 for more details of the sample.

Recruitment was undertaken in a number of ways because of convenience and national preferences as to what was the best way to ensure maximum cooperation and compliance. The basic methods were to enlist publisher and learned society help in getting in touch with their authors residing in the countries covered (UK, USA, Spain) and to use university and researcher networks (UK, Poland, France, Malaysia, China). In some cases, these methods were supplemented by personal contacts, workshop attendances, and by the ECRs themselves (the invitations going viral).

Interviews are conducted by national interviewers in their own languages, except in the case of Malaysia (where English was used due to the proficiency of the ECRs). Such personal interviewing procedures were used to build a relationship between the ECRs and the investigator as we need to keep in contact with the ECRs for 3 years. The proceedings of the interviews were taken down in note form as it was felt we were going to get a better response this way. A transcript of the interview was returned to the interviewee for validation and further data collecting purposes, which was necessary to plug the inevitable gaps in the interview record. The record was then translated into English for all non-English speaking countries and then manually coded using a heuristic approach and a standardized thematic framework.

RESULTS AND DISCUSSION

The principal question asked of ECRs was *How do you find the scholarly information you need?* To this, there was added the prompt: *Google, library catalogues, online networks, and so on?* The motivation for asking this question was to establish whether ECRs use university libraries and their web-scale/index-based discovery systems, and/or the popular search engines, and/or the even newer, social media-based online community networks.

While the question seems on the surface to be very straightforward, it is, in fact, far from that. First, because it transpires that many ECRs interpret the phrase 'scholarly information' as publications (aka journal articles), which is hardly surprising as journal

articles have repeatedly been found to be the most frequently used source of information used by academics to assist with their work (Tenopir, King, Christian, & Volentine, 2015). However, as the inclusion of online networks in the prompt indicates, the research team had a wider interpretation of the term in mind, which includes scholarly contacts, ideas, and data. We had anticipated this to a certain extent by including another discovery question later on in the survey, asked in the context of social media use: *Do you use social media in your scholarly activities to find out information and (if so) from what media?*

Second, because there are so many players involved in providing access to journals, it is not easy for ECRs to appreciate whose services they are using (e.g. publishers', aggregators', libraries', online communities', software houses'). What confuses and compounds the situation is the increase in mediating services and the continued, but diminishing/invisible, role of the library. Also, muddying the water is the fact that so much information is available in an open access (OA) form in institutional repositories (IR), online communities, and personal websites, all conveniently shepherded together by Google robots to produce the one-stop access users want. This state of affairs is recognized in the recent Ithaka S+R Surveys of US and UK faculty, according to both of which there are noteworthy increases in researchers' opting for alternative sources to their library's collection, including materials that are freely available online (Wolff *et al.*, 2016a, 2016b). No wonder then that ECRs find themselves lost in a maze of searching opportunities. To conclude that they do not have the 'foggiest' about these opportunities might be to exaggerate a little, but it would not be an exaggeration to say that they do not care less who/what enables their access to full-text papers as long as they have it. They are blinkered in the chase for the full-text paper.

Third, the finding process may be seen as multi-staged, first, obtaining a reference and, second, getting access to the full-text. For instance, sometimes ECRs might obtain a reference from the paper they are reading (or, less commonly, via Twitter) and then go directly to Google, Google Scholar (GS), PubMed, or the library portal to fetch the article. If this does not work, then they might go to ResearchGate (RG) to see if the sought-after paper is there. Interestingly, this is not too popular an option among researchers generally, at least not yet: according to the aforementioned 2015 Ithaka Surveys, where only a very small share of their respondents indicated that they would request a copy of some material needed for research and teaching using social media (Wolff *et al.*, 2016a, 2016b). It may, however, be indicative that US scholars from younger age cohorts (who are the subject of our study) were found to be more inclined to ask a friend at another institution for a resource as compared to those from older cohorts (Wolff *et al.*, 2016a).

Sources used

Surprisingly, perhaps, for an international study that covers such diverse countries as Poland and China, there is much consensus as to where to go for information, demonstrating the global nature of scholarly activity. Table 4 shows this as well as the diversities. There are two components to the ratings of resources given in the

table: (1) numbers of ECRs mentioning they used the resource and (2) how heavily they use the sources. Categorizing and identifying sources in a world of hubs, interfaces, mediators, and aggregators is not easy for ECRs who are not knowledgeable about ownership or origin. So, as a result, we have adopted a 'belt and braces' approach to classification, and the data should be regarded as approximations. Generally speaking, the table shows the universal importance of Google/GS and those hardy perennials, the Web of Science (WoS) and Scopus, and the supporting role of the library. PubMed and ScienceDirect are also generally popular, and RG is clearly a rising star. US and UK ECRs are very similar, so they have been bracketed together for the purposes of the table.

Search engines

Consistent with prior evidence as to researchers' use of search engines (Borrego & Anglada, 2016; Housewright *et al.*, 2013a, 2013b, Jubb, Look, & Sparks, 2007; Niu *et al.*, 2010; Wolff *et al.*, 2016a, 2016b), inclusive of young researchers (Haglund & Olsson, 2008), Google and GS are found to be universally popular with all the 116 ECRs, irrespective of country, language, and discipline. In most countries, they are the one ECRs go to first. GS is especially rated highly in the USA, where two-thirds of ECRs say it is their top source.

To examine the reason for the success of Google and GS, it is helpful to consider what the ECRs report in the different countries. Within France, ECRs highly rated GS for its 'meta-searching' opportunities. It is thought to be a search engine that harvests the web's rich resources and searches everywhere, including repositories (thematic and institutional) and academic social media. In a period when French libraries unsubscribed to many resources, it is also thought invaluable because it is:

- A source that gives the number of citations, whatever the type of document, and thus makes it easier to make quality judgements about the quality of scholarly information.
- A tool that provides the PDF wherever it is and allows the ECR to access the article wherever they are located.

Illustrating interesting connectivity in the digital environment, where all things are related and linked, a good number (5 of 14) of French ECRs discovered RG as a result of GS searching because GS ranks RG and similar community platforms first before the publisher platforms. Having realized that GS is providing greater visibility for scholarly social media, a number of French ECRs decided to create an account on RG.

In China, scholarly search engines, including GS, are seen to be useful at the start of a search and for searching for general scholarly information. However, Chinese ECRs are alert to their weakness too, particularly their inability to always lead to the full-text document. Interestingly, GS is blocked through China, so theoretically, it is unavailable for Chinese ECRs, which has an impact on their searching behaviour. As a result of the prohibition, many ECRs resort to what are generally acknowledged as second-rate national services, such as Baidu Scholar Search, which is good for Chinese language. As a result, Google/GS, while

TABLE 4 Main categories of scholarly resources used by ECRs to find information.

| Resource | France | China | Poland | Spain | UK/USA | Malaysia |
|--|--------|-------|--------|-------|--------|----------|
| <i>Search engines</i> | | | | | | |
| Google | ++++ | ++ | ++++ | ++++ | +++ | ++++ |
| Google Scholar/citations | ++++ | ++ | ++++ | ++++ | ++++ | ++++ |
| Sci-Hub | ++ | - | - | - | - | + |
| <i>General databases</i> | | | | | | |
| Scopus | ++ | ++ | +++ | +++ | - | ++++ |
| WoS | ++ | +++ | ++++ | ++ | + | ++++ |
| EI | - | +++ | - | - | - | + |
| CNKI | - | ++++ | - | - | - | - |
| Wanfangdata | - | ++ | - | - | - | - |
| <i>Specialized databases^a</i> | | | | | | |
| PubMed NCBI | ++ | + | - | +++ | +++ | ++ |
| SciFinder | + | - | - | - | + | + |
| Publishers' platforms | | | ++ | | | |
| ScienceDirect | ++ | +/- | - | - | ++ | ++ |
| SpringerLink | ++ | + | - | - | - | + |
| IEEE | + | ++ | - | - | - | + |
| Libraries and their platforms ^b | ++ | ++ | ++ | + | + | ++ |
| <i>Social media</i> | | | | | | |
| ResearchGate | +++ | ++ | ++ | +++ | ++ | +++ |
| Academia | - | - | + | + | + | + |
| YouTube | + | - | + | - | - | - |
| Wikipedia | - | - | - | - | - | + |
| Disciplinary SM (e.g. Math Stack Exchange) | + | - | - | - | - | - |
| WeChat | - | ++++ | - | - | - | - |
| QQ | - | ++ | - | - | - | - |
| Weibo | - | ++ | - | - | - | - |
| LinkedIn | - | + | - | ++ | + | + |
| Twitter | - | + | - | ++ | ++ | + |
| Facebook | - | - | - | + | +(USA) | - |
| <i>Repositories</i> | | | | | | |
| arXiv | ++ | + | - | + | - | ++ |
| HAL | ++ | - | - | - | - | - |
| E-LIS | - | - | - | + | - | - |

++++, Used often first by many ECRs; +++, used often by many ECRs; ++, used moderately by middling numbers of ECRs; +, used occasionally by a few ECRs; -, not mentioned. CNKI, China National Knowledge Infrastructure; EI, Engineering Village; ECRs, early career researchers; E-LIS, E-prints in Library & Information Science; HAL, Hyper Articles en Ligne; IEEE, Institute of Electrical and Electronics Engineers; SM, social media.

^aThe use of specialized databases is affected by the discipline of the recipients - for example, Polish ECRs are unlikely to use PubMed much because there are no medical scientists in the sample.

^bBecause ECRs are not always conscious that they are using library services and their discovery platforms, we have rolled them together and used data obtained elsewhere to interpret what they said.

important, are not so important as they are in other countries. Even so, 11 ECRs mention them as popular sources. They, like the French ECRs, rate their internationality highly, but they also raise the fact that their currency is a real asset too. Thus, one ECR explained that they usually use GS rather than the digital library resources that are not as current. There are several reasons why Google and GS are used by Chinese ECRs, although they are theoretically unavailable in the country. The main reason is that many ECRs use proxy servers to get around the prohibition, and also, many ECRs used GS in China before it was banned in 2014 and so had built an allegiance. Thirdly, ECRs who have an overseas study background are used to GS and also Facebook (also banned). On returning to China, they find it difficult to break the GS habit and so resort to proxies for access. Finding this validated the use of personal interviews as it is unlikely that ECRs would have been willing to declare this 'illegal' use on a questionnaire.

Elsewhere, all Malaysian ECRs mention Google/GS as a major source. Indeed, for them, it has overtaken traditional systems (i.e. library sources and services and publisher platforms) as the most used source of scholarly information. For Malaysian ECRs too, Google is thought to be a great starting point. This, again, is because of its enormous reach and the fact that it allows for a borderless search, which offers much in the way of access to OA sources. They mention another advantage nobody else did, albeit an obvious one: it is constantly available without the necessity of complex logging-in procedures. In Spain, GS, known as Google Académico, is the most popular source of scholarly information. Indeed, Spanish researchers mention that it is good not just for finding documents but also for finding other researchers.

General databases

The literature consensus is that general databases are the preferred starting point for a bibliographic search (Borrego & Anglada, 2016; Housewright *et al.*, 2013a, 2013b; Niu *et al.*, 2010; Wolff *et al.*, 2016a, 2016b). This is not much different for ECRs: the relatively old stalwarts, the WoS and Scopus, are very popular and still hold their own, with WoS shading it in terms of popularity (but not in Spain, see below). Interestingly, neither database rates highly in the UK or USA, although WoS has the edge with four mentions. Malaysian ECRs, on the other hand, are particularly reliant on citation index databases, rating WoS a close second to Google/GS, and it is worth reviewing why this is the case. The first reason is all the ECRs come from research-intensive universities, which mandate publishing in WoS and Scopus-indexed journals and can afford the high costs of subscribing to them. The second reason is that ECRs have been exposed to WoS and Scopus early in their research careers. Third, it has to do with the perceived quality of the journals covered, as one ECR said, 'I don't want to miss out on potentially relevant and important sources in WoS'. They mainly use WoS and Scopus for a whole range of information needs – recent publications, hot topics, Journal Impact Factors, citation data and other metrics, research trends, and potential collaborators. Fourth, and this may explain low use in the UK/USA, they are preoccupied with impact factors when it comes

to publishing, and so, it is natural that they should continue to use these brands when it comes to searching for information. UK/US ECRs are interested in impact factors but not as preoccupied with them, and this may explain their (relatively low) use of sources. Malaysian ECRs also find it very helpful that WoS and Scopus have linked their bibliographic information to OA articles.

In Poland, the WoS vies with GS for popularity, with nearly everyone using it. In France, however, as in the UK/USA, WoS is not as popular, only being named twice. In the cases where WoS is used, it is also used as a starting point for research and is similarly appreciated for the GS functionality 'Look Up Full Text' offered that allows the user to go directly to the full text identified by GS. Having said that, even those ECRs with access to WoS prefer using GS, considering it to have wider coverage. As mentioned earlier, Scopus is more popular in Spain than WoS, coming second only in popularity to GS, which is a little surprising given that the licence to use WoS is paid for all public institutions by the Science Ministry.

In China, given the language differences, things are inevitably a little bit different with CNKI (China National Knowledge Infrastructure), Wanfang (an integrated knowledge service platform providing full-text scholarly journal articles, standards and patents, policies and regulations), EI (Engineering Village), and WoS being the most popular sources. CNKI is the biggest and most popular full-text database for Chinese-language scholarly journals, but because its charges are thought to be rising too quickly, this, according to our Chinese interviewer, has resulted in around 10 top universities boycotting it during 2016. To use WoS and EBSCO is not a result of boycotting CNKI. Because top universities have budget to buy expensive international databases such as WoS and EBSCO, they don't rely too much on CNKI, so they led the boycotting.

Specialist databases

PubMed deserves a mention here as it is popular both as a source of full-text publications and as a search engine in biology and chemistry, especially in France. It is often used in tandem with Google and GS. However, even for French ECR biologists who believe PubMed is peerless, Google and GS are the most common starting points for research. The only country that did not make use of PubMed is Poland (Table 4), but this is easily explained by the absence of bio-medics in their sample. In the UK and USA, where there is a good proportion of bio-medics, around a quarter of ECRs rated PubMed as their first or equal first port of call.

Publisher platforms

ECRs use scholarly journals, but they are rarely conscious of the publishers of these journals and even more rarely visit their websites. If publisher platforms are mentioned at all, it is for publishing purposes (e.g. author guidelines). The main exception to this being ScienceDirect, which is, probably, a bigger brand name than that of the publisher (Elsevier). It obtained six name checks in Malaysia, three in Spain, two in the UK/USA, and just one in

Poland and France. This takes us to 13 name checks for scholarly publishers from 116 researchers, not many really. However, in addition, two Polish ECRs did say they used 'remote access to publishers' databases' and 'websites of publishing houses', and two Spanish ECRs mentioned 'journal providers' in general. Publisher alerts are mentioned in passing but not ranked highly, and table of content (TOC) alerts are 'usually binned'. It is of interest to note the lack of interest in TOC alerts, previously one of the mainstays of discoverability. The lack of awareness or use of publisher platforms for discoverability may not come as a surprise to publishers. According to one major journal publisher (I. Bannerman, personal communication, October 23, 2016), they do not expect scholars to regularly come to their sites to look for things, but they do expect them to come regularly to their sites to get things. Similarly, many marketing efforts are shifting from promoting journals to promoting articles.

Libraries and their platforms

For most ECRs, university libraries are largely seen as study spaces for undergraduates and not places to go to for discovering research information. This echoes prior evidence, which shows that physically visiting the library is rarely a path taken by researchers (Borrego & Anglada, 2016; Niu *et al.*, 2010; Schonfeld, 2014; Wolff *et al.*, 2016a, 2016b). Indeed, the Ithaka S+R Surveys of US and UK faculty report that for the second cycle, now about one-seventh of their respondents believe that institutions should redirect money spent on library buildings and staff to other needs. Of course, as a distribution channel for freely accessible full texts from publishers, libraries (aka their websites) cannot be wholly ignored as the US and UK Ithaka S+R Surveys testify: the buyer role of the library, its paying for resources needed, academic journals and electronic databases, has remained the most highly important role for faculty members overall. These findings are consistent with those of the survey among Catalan faculty, according to which, 86% of the respondents rated the buyer function of the library as highly important (Borrego & Anglada, 2016). Of course, ECRs are not party to, or familiar with, the buying function of the library, and this explains a good deal of what they told us about libraries and their discovery platforms.

Thus, questions about the role of the library in discovery are complicated and confused by the fact that, although ECRs use its services, they are often unaware the library is involved in anything but in a minor capacity. Library platforms largely go unnoticed and unmentioned, with ECRs passing through them blindly – more focussed on the branded services to which the library gives them access. It follows, of course, that they are unable to distinguish between library catalogues, library portals, and hosted databases. They view libraries through a traditional prism. Chinese researchers in particular are confused, with most of them having to use the library to get their journal articles on the one hand and, on the other hand, over half saying they did not use the library for discovery purposes. French ECRs are another case in point. Many, again, are unaware that the access they get to

publishers' resources is, in fact, a consequence of library subscriptions. Others, begrudgingly, mention the library portal, which obliges users to go through it as a means of accessing WoS. Indeed, if French ECRs can help it, they do not use the library at all – physically and digitally. When asked why not, ECRs explain that discovery is too messy, too difficult, and that they spend too much time on getting something they could get more easily elsewhere (French university websites are said to be 'labyrinthian'). In the case of Chinese ECRs too, around one fourth (three ECRs) expressed negative views about the library services, complaining about the lack of currency of services (compared to GS) and the poor interface.

Social media

There is evidence to indicate that social media/online communities have come of age as places to find scholarly information (Spezi, 2016). Thus, 11 services/platforms are mentioned by our 116 ECRs, and one, RG, is particularly popular in all seven countries. For really significant use of the social media, we really need to look to China. In China, three quarters of all the source mentions are to a veritable legion of social media and online community sources. WeChat, QQ, Weibo, LinkedIn, Twitter, and RG are the names mentioned the most (as already noted, Facebook is blocked and not supposed to be used but may still be). Chinese ECRs tend to follow senior experts and colleagues on social media, and once something is posted, they are informed, and then, they use it. They use social media as a selective dissemination of information (SDI) service in this regard, pulling information rather than pushing for it.

WeChat, established in 2011, is a particularly interesting and novel service and as such deserves some explanation. It is a little different in being a cross-platform, instant messaging service. WeChat supports users to register as an official account, which enables them to send feeds to subscribers, interact with subscribers, and provide them with services. Generally, it provides scholarly information via the so called 'Official Accounts'. Official Accounts could be operated by an individual or institution and provide both post-publication full-text papers and popular, general interest science information. There are about 1,000 Official Accounts that can be deemed to be scholarly. Interviewees mentioned, for instance, these accounts as being especially useful: MedSci_cn, which provides clinical medicine information, run by a group of medical scientists; HuanqiuKexue, operated by the *Scientific American* Magazine, which provides general information about science; and CAS-iop, which provides information about physics and is run by the Chinese Academy of Science. WeChat turned out to be an especially popular reading source for ECRs, with 8 of 13 interviewees saying they use it.

Interestingly, and unusually, there is also an eBay-like website named Taobao (www.taobao.com) in China where people sell downloaded full-text papers on the web, but no ECR admitted to using it. This business model involves obtaining full-text download permission to get a full paper and then to sell it to those who have no access to the full text.

UK ECRs also widely use social media to find information; 18 told us so as compared to 3 who did not. RG is used most (nine mentions) and is followed by Twitter (seven) and LinkedIn and Academia (both two). RG is used to find out what people are doing, and Twitter is used as a cut down version of e-mail by which information is given and received. In the case of the US, there are almost similar proportions of adherents, with 19 ECRs using social media and nine not doing so. Again, RG is the most popular service (13). One ECR mentioned that even if you do not start your search on RG, Google will take you there. Facebook (six) is an equal second and is used to keep US ECRs up to date. However, none of the UK researchers use it. Twitter is also mentioned six times. This is a lower proportion than for the UK, and for some reason, three ECRs made a particular point that they do not use Twitter for scholarly information seeking. LinkedIn (three) and its networks also have followers.

Social media are reasonably well used in Poland where six ECRs use social media for finding information, and RG is, yet again, especially popular. In Spain, RG is the fourth most popular source, and LinkedIn and Twitter are also rated highly by some ECRs. Facebook and YouTube also get mentions. French ECRs are not quite as convinced of the value of the social media, but it is the same story regarding RG, which is a popular platform. Usage is pragmatic, with the function being to access papers (obtain PDFs) and make connections with their colleagues and peers. As French ECRs see that more and more people are going to RG, they go there more frequently or are tempted to join. Disciplinary social media, such as Math Stack Exchange, are much appreciated by French ECRs and provide a more focused platform to share information, documents, and expertise.

In Malaysia, ECRs do not use social media much to find scholarly information but occasionally use RG and Academia.edu in this context. However, two ECRs find Wikipedia especially useful. One mentioned, 'I refer to sources such as blogs and Wikipedia to get ideas about issues, people's opinions, definitions, related to my research topic'; and another offered, 'actually, Wikipedia is a perfect tool for latest information because the contents are authored by a large community of contributors. Maybe not accurate, but of use'.

OA sources

It has been calculated that, at least in some fields and some countries, half of the journal literature is already openly available in some shape and form, whether that is through OA journals, IR, online communities, or authors' websites (Archambault *et al.*, 2013), all nicely shepherded together in a GS search. As one Malaysian ECR reminded us, 'There is no better channel for finding OA publications than GS'. ECRs, working in this increasingly open scholarly information environment in which there are no paywalls and subscription agreements, unsurprisingly tend to think that OA is a very good thing. This, of course, partly explains the huge popularity of Google and GS with them. Spanish ECRs (eight) are particularly vocal about the benefits of OA. For example, one participant mentioned, 'OA permits me to find many

documents needed for my research from more researchers'; another said, 'OA gives access to contents to researchers without access to institutional subscriptions to databases or journal providers'.

Malaysian ECRs also feel that OA had a big impact on finding/discovering/accessing scholarly sources. Thus, one pointed out, 'Many people will ignore the journals with restricted access, and this will place the pressure on the authors to deposit their work on RG so others can gain access to it'. Another said that even when they access WoS, the relevant articles would be those that are OA. Others offered, 'It is really helpful; the OA button in WoS! I can now directly access and read full-text right from there', and 'OA journals are less established, but thanks to WoS, I can be sure of the goods once there, I don't have to go far to chase them down'.

ArXiv.org, mentioned by physicists and mathematicians, is the main OA resource for French ECRs. They use it intentionally as a starting point of their research. Interestingly, the French national repository, HAL, is not mentioned as a starting source of scholarly information by French ECRs. This is largely explained by the fact that ECRs have noted that GS indexes repositories – particularly HAL – and other academic social media more comprehensively than anyone else and, consequently, use it as their central source for finding all things OA. As in other countries, French ECRs do appreciate an information environment where 'research funded by public money is accessible without paying a penny'. They also appreciate the fact that OA is more embedded in discovery systems (Google, GS and WoS), and that provides more visibility to a wider range of documents. An ECR who works on interdisciplinary research relies intensively and deliberately on OA resources (HAL and international repositories, OA journals) to search and find articles from different disciplines his library cannot reasonably afford.

Nobody, anywhere, mentioned IR, but they might have arrived at an IR unknowingly via a Google search, just as the French researchers unknowingly found the resources of HAL.

Smartphones

Smartphones have clearly established themselves as platforms to locate scholarly information, especially via Google. However, as Nicholas and Clark (2013) showed through their analysis of the usage logs of a major cultural multimedia website, Europeana.eu, mobile searches are information 'lite': typically shorter, less interactive, and with less content viewed per visit. Indeed, mobile devices are providing the ultimate 'information take-away' with all the evidence showing that we use them for information bites and snacking – more bouncing, more new visits, shorter visits, and simple and less productive searching are a feature of their use. This is very much the case where our ECRs are concerned.

Generally, at least half of all ECRs use smartphones to locate scholarly information, but this does vary by country. In Malaysia, where all 12 ECRs use them for this purpose, they use them because they are always furnished with an internet connection wherever they go and whenever they want. However, not many

read articles on their smartphones. As one ECR explained, 'My smartphone provides easier, faster, more reliable access to the web when I am travelling. But I still need my laptop or tablet to download research papers'.

More than half the UK ECRs in this study use smartphones occasionally for scholarly purposes, but very few use them most of the time for most purposes. Moreover, when probed, a small minority read on a smartphones – they use them to find papers through alerts. If they are not in the lab, the smartphone is their contact with the Internet. Half the US ECRs we spoke to use smartphones but, again, only occasionally.

In Spain, too, nearly half of the ECRs in this study said that they do use smartphones occasionally, mainly for searching purposes when they are travelling, commuting, or out of the place where they live. They are a little less popular in Poland where 3 of 10 ECRs said they use smartphones to read or search for articles, with one saying, 'Yes...Recently, I have used it more often. I use Mendeley application or I jump to the Websites, i.e. Google Scholar, Facebook etc'.

Asking people about something they do almost unconsciously/automatically, like searching, can lead to some strange answers, and this was very much the case when we asked Chinese ECRs about finding scholarly information on their smartphone. Thus, three say they do, and six say they do not, but those who said no also admitted to using WeChat, a smartphone-based app. Furthermore, the four who said they do not use smartphones and WeChat admitted to using online research networks and social media, something ECRs tend to do using their smartphones. So, there is a good case for saying a majority of Chinese ECRs probably use smartphones for scholarly purposes.

In France, take-up is lower with four (out of 14) ECRs say they use smartphones, although two of them use them intensively. They use smartphones mostly to receive current awareness information from PubMed, GS, and arXiv rather than for reading articles.

Searching style

Searching styles tend to boil down to tracking down the full text as fast as possible. Thus, French ECRs use Hubs (GS, Google and WoS) and try to find, as one interviewee said, their 'PDF' as quickly as possible. If the PDF is not found via the first links to publisher platforms, OA repositories, or social media, they go to Sci-Hub. If that is not successful, some send an e-mail to the author, and others give up. All French ECRs complain about information overload, and this information-seeking path is intended to deal with it. Chinese and Malaysian ECRs are also driven by the same need to obtain the full text as quickly as possible from the information mountain but are, perhaps, more tenacious. So, if they cannot get access to the full text, they will change tools, and if there is no full text at all, they will check the references and try to find the full text of similar papers. From all this, it follows that ECRs look for articles rather than journals as our aforementioned publisher rightly assumed.

For Chinese researchers, there are essentially two broad paths to finding scholarly information depending on the type of institution to which the researcher belongs. ECRs in research-intensive universities and government-funded research institutes have well-resourced libraries, which can afford expensive discovery systems, which provide full text to researchers. So, ECRs from these universities and institutes, some bereft of Google/GS, tend to search for information through the library website and then, maybe, move to a publisher's portal. It is only if they cannot obtain the information that they want, they would turn to GS (if they had access) and other open sources or social media. The ECRs from teaching-intensive (less well-resourced) universities typically go first to GS if they can. If they cannot find the full text they want there, they will search for the specific items on the university library's website and, maybe, also go to the publishers' database for content.

Besides at an institution level, disciplinary differences are also an important variable when searching for information. So, Chinese STM researchers frequently use GS and the main bibliographic databases, while some social scientists prefer to use Chinese language-based search engines and Chinese language academic databases, such as CNKI and Wanfangdata. Also, physicists and information scientists tend to favour publishers' database, and biomedical scientists and computer scientists tend to favour open sources. ECRs also say that that senior/older colleagues are more likely to opt for digital libraries and publishers' databases, while younger ones prefer to use GS.

Finally, Spanish ECRs place much reliance on e-mail alerts to keep the abreast of new things. Publishers of journals and RG provide them with these alerts. So, there is a serendipity factor in finding interesting material. GS provides suggestions as well. It is when ECRs need to gather information for a project or a paper that they perform their searches.

CONCLUSIONS

This is the first year's findings of a 3-year long project, and we shall be following our researchers for another 2 years, asking them the same 'under the bonnet' questions about discovery in order to determine whether ECRs are going to be the harbingers of change. Nevertheless, we believe we have more than enough evidence already to suggest some interesting things are taking place, so there is plenty of justification for continuing to study ECRs. The popularity of Google/GS can come as no surprise, but the fact that it is universally so (even when banned in China), and in most cases, ECRs' most popular source, has to be. Google is the new reference point for searching for scholarly information, and the other players are increasingly dependent on its algorithm to put them in a good position on its hits page. Visibility is everything in today's scholarly environment, especially with regards to reputation, and that is what is powering Google. What is also helping Google is the increase in the amount of scholarly information appearing freely and in OA form; it is closer than ever to the one-stop information shop. Close to Google/GS in popularity

are those two hardy perennials, WoS (not a bad performance for a citation index) and Scopus (boosted by all Elsevier's full-text papers).

The fact that physical libraries are hardly used at all, and the library has become invisible to most ECRs, looks like bad news for university libraries. It gets worse because their library platforms and their much-hoped for web-scale discovery services are typically rated a poor third or fourth to Google in the popularity ranks. While ECRs often, but not always, have to go to the library website to access the 'free' full-text journal and citation services, very few mention the library website as a means of discovery. Libraries, if thought of at all, are seen as facilitators for access and not for discoverability.

ECRs are sympathetic to OA publishing and its intentions, particularly appreciating its value when they search GS for full-text papers. The combination of GS and OA publishing leads to a general detachment of ECRs from publisher (and library) platforms. Social media/online communities are, potentially, the new intermediary and now widely used as a source of scholarly information in most countries. RG is popular everywhere and looks to be a major player, and Twitter, LinkedIn, and Facebook have their places. Smartphones have become a platform to search for scholarly information in a number of countries, most notably Malaysia and UK.

China shows us that discovery is far from a straightforward process because, as we have seen, political and economic factors intervene in what one might have considered to be a very a neutral 'library' space. Google is banned, as is Facebook, but ECRs have ways around it. Indeed, it has almost generated an underground economy. Also, there are some very interesting, innovative discovery models and services in the Chinese marketplace, for instance, WeChat.

This paper is the first report from this project and has provided us with valuable insights into the ways in which ECRs search and discover scholarly information.

Limitations

This study is based on a relatively small sample of ECRs and so might not be representative of the ECR population at large. It is also based on personal interpretation of the questions and willingness to be reported honestly and objectively, all of which may bias some of the answers that we received. However, we feel that these limitations were compensated for by using personal interview techniques and asking in-depth questions and the collection and analysis that has enabled us to disentangle a topic that requires a broader and more nuanced investigation than is traditionally undertaken.

ACKNOWLEDGEMENTS

The research reported here has been funded by the Publishing Research Consortium, the National Social Science Foundation of China, No.15CTQ026, Ministry of Higher Education Malaysia (UM.C/HIR/MOHE/FCSIT/11), and the Tomsk State University Competitiveness Improvement Program.

REFERENCES

- Archambault, E., Amyot, D., Deschamps, P., Nicol, A., Rebout, L., & Roberge, G. (2013). *Proportion of open access peer-reviewed papers at the European and world levels—2004–2011* (Science-Matrix report). Montréal, Canada: Science Matrix Inc. Retrieved from http://recolecta.fecyt.es/sites/default/files/contenido/documentos/First%20report_OA%20Peer-Review%20Papers%202004-2011.pdf
- Borrego, Á., & Anglada, L. (2016). Faculty information behaviour in the electronic environment: Attitudes towards searching, publishing and libraries. *New Library World*, 117(3/4), 173–185. doi: <http://dx.doi.org/10.1108/NLW-11-2015-008>
- Haglund, L., & Olsson, P. (2008). The impact on university libraries of changes in information behavior among academic researchers: a multiple case study. *The Journal of Academic Librarianship*, 34(1), 52–59. doi:<http://dx.doi.org/10.1016/j.acalib.2007.11.010>
- Harley, D., Acord, S. K., Earl-Novell, S., Lawrence, S., & King, C. J. (2010). *Assessing the future landscape of scholarly communication: An exploration of faculty values and needs in seven disciplines*. Berkeley, CA: Center for Studies in Higher Education, UC Berkeley. Retrieved from <https://escholarship.org/uc/item/15x7385g>
- Hemminger, B. M., Lu, D., Vaughan, K., & Adams, S. J. (2007). Information seeking behavior of academic scientists. *Journal of the American Society for Information Science and Technology*, 58(14), 2205–2225. doi:10.1002/asi.20686
- Higher Education Funding Council for England. (2015). *Staff employed at HEFCE-funded HEIs: Trends and profiles*. Retrieved from <http://www.hefce.ac.uk/analysis/staff/ecr/>
- Housewright, R., Schonfeld, R. C., & Wulfson, K. (2013a). *Ithaka S+R US faculty survey 2012*. New York, NY: Ithaka S+R. Retrieved from www.sr.ithaka.org/research-publications/us-faculty-survey-2012
- Housewright, R., Schonfeld, R. C., & Wulfson, K. (2013b). *UK survey of academics 2012: Ithaka S+R|Jisc|RLUK*. New York, NY: Ithaka S+R. Retrieved from http://www.sr.ithaka.org/wp-content/uploads/2015/08/UK_Survey_of_Academics_2012.pdf
- Jubb, M., Look, H., & Sparks, S. (2007). Researchers' use and perceptions of discovery services. *Learned Publishing*, 20(2), 147–153. doi:10.1087/174148507X185126
- Nicholas, D., & Clark, D. (2013). The second digital transition: To the mobile space—An analysis of Europeana. *Learned Publishing*, 26(4), 240–252. doi:<https://doi.org/10.1087/20130402>
- Niu, X., & Hemminger, B. M. (2012). A study of factors that affect the information-seeking behavior of academic scientists. *Journal of the American Society for Information Science and Technology*, 63(2), 336–353. doi:10.1002/asi.21669
- Niu, X., Hemminger, B. M., Lown, C., Adams, S., Brown, C., Level, A., ... Cataldo, T. (2010). National study of information seeking behavior of academic researchers in the United States. *Journal of the American Society for Information Science and Technology*, 61(5), 869–890. doi:10.1002/asi.21307
- Poli, S. (2016, 10 May). The annual conference of the German Association for HE Research. [Web log post]. SRHE - The Society for Research into Higher Education Blog. Retrieved from <https://srheblog.com/2016/05/10/the-annual-conference-of-the-german-association-for-he-research/>
- Schonfeld, R. C. (2014). *Does discovery still happen in the library?* New York, NY: Ithaka S+R. Retrieved from http://www.sr.ithaka.org/wp-content/mig/files/SR_Briefing_Discovery_20140924_0.pdf

- Spezi, V. (2016). Is information-seeking behavior of doctoral students changing? A review of the literature (2010–2015). *New Review of Academic Librarianship*, 22(1), 78–106. doi:<http://dx.doi.org/10.1080/13614533.2015.1127831>
- Tenopir, C., King, D. W., Christian, L., & Volentine, R. (2015). Scholarly article seeking, reading, and use: A continuing evolution from print to electronic in the sciences and social sciences. *Learned Publishing*, 28(2), 93–105. doi:<https://doi.org/10.1087/20150203>
- Wolff, C., Rod, A. B., & Schonfeld, R. C. (2016a). *Ithaka S+R US faculty survey 2015*. New York, NY: Ithaka S+R. Retrieved from <http://sr.ithaka.org/?p=277685>
- Wolff, C., Rod, A. B., & Schonfeld, R. C. (2016b). *UK survey of academics 2015: Ithaka S+R|Jisc|RLUK*. Retrieved from <http://www.sr.ithaka.org/publications/uk-survey-of-academics-2015/>