Access to scholarly content:
gaps and barriers
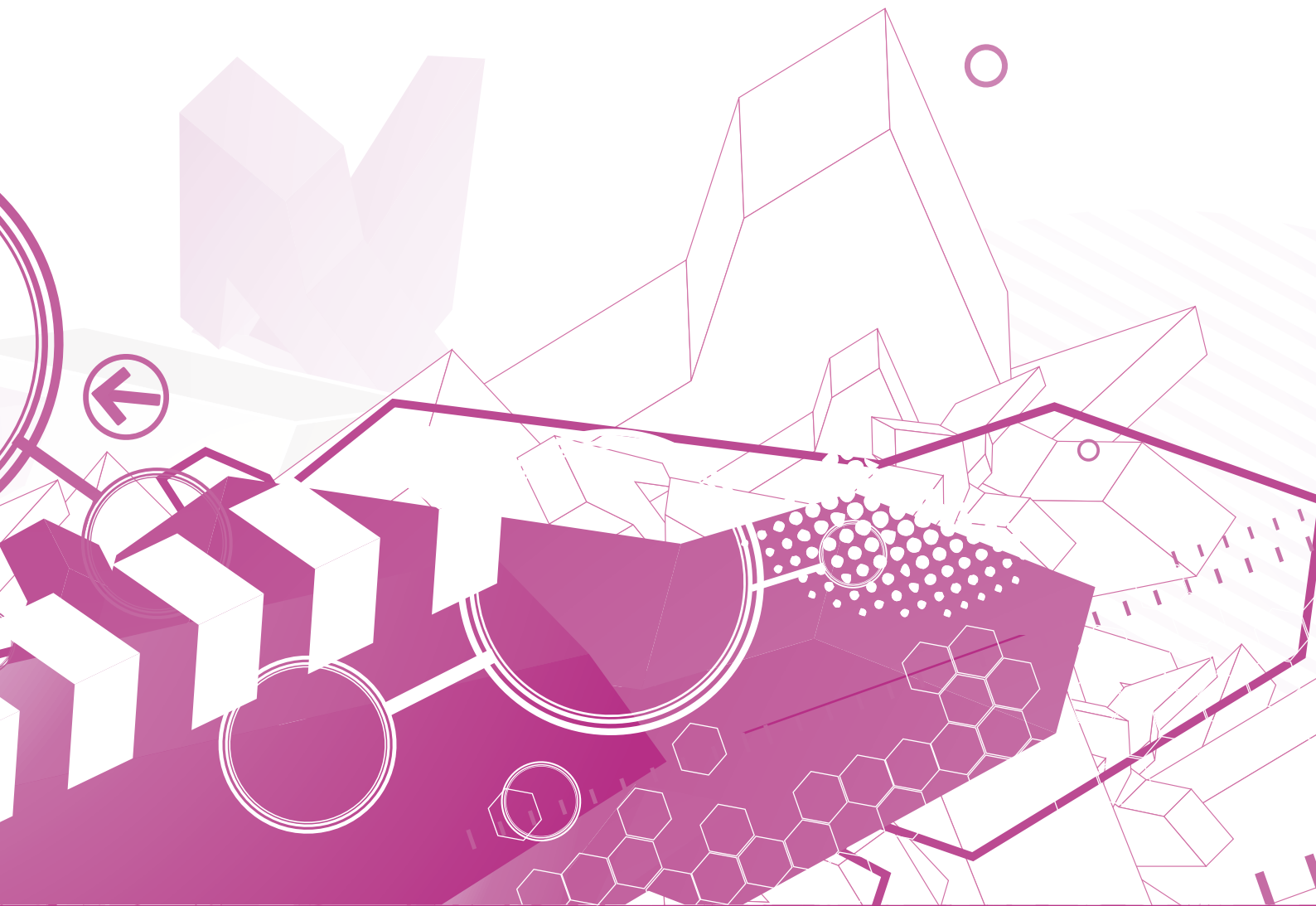
research
information
network

prc
publishing
research
consortium

JISC

Acknowledgements

Contents

Executive Summary

Aim and objectives of the study

Aim

The overall aim of this study is to investigate and quantify the extent to which members of different communities in the UK can gain ready access to formally-published scholarly literature, in particular journal articles and conference proceedings.

This study complements two other projects commissioned by the sponsoring group. Together these offer a substantial evidence base with which to better understand the dynamics of scholarly communication, now and in the medium term future.

Objectives

The objectives of this report are to:

- conceptualise and map the nature and extent of the gaps in and barriers to access to scholarly journal articles and conference proceedings as experienced by members of different communities in the UK, in academia and beyond;
- quantify, as appropriate and feasible, access gaps and barriers for different communities;
- provide an evidenced assessment of the significance of those gaps and barriers for different communities, in the light of their information requirements and levels of demand, their behaviours, and their working practices.

Methods

Much of the information presented here is based on an online survey of researchers and knowledge workers from UK universities and colleges, medical schools and health providers, industry and commerce, and research institutes. The sample frame comprised a balanced mix of UK journal authors (randomly selected from the Scopus database), plus lists derived from people who had registered to use Elsevier pay-per-view facilities, and personal subscribers to trade and professional magazines, again supplied by Elsevier. Invitations were sent to 20,000 individuals and 2,645 completions were received: a response rate of 13.2 per cent.

Questioning covered both digital and hard copy versions and was filtered so that respondents were not asked in detail about issues that were not important to them. It is important to remember this when reading the report, in order to avoid any misinterpretation. For example, 68.2% of higher education researchers complaining about insufficient journal titles (Figure 16) may seem very troubling, until we remember that the base for that question is only researchers who have difficulties in accessing journal articles, and not the entire researcher population. For this reason, we have tried to be very specific about the base for each question.

Other information in this report comes from a detailed analysis of the literature and secondary data analysis of the Labour Force Survey in an attempt to quantify the size of the UK professional knowledge worker sector.

Gaps and barriers to access

Gaps and barriers defined

The term `gaps' refers in this study to a situation where information is needed for a specific purpose but is not available from sources that someone is willing or able to use. Gaps exist when a researcher or knowledge worker knows they need something and cannot get ready access to it. An example might be a surgeon within the NHS who is unable to access up-to-date material on a particular surgical procedure. The term `barrier' refers here to the limitations which cause such gaps in access. In the situation with the surgeon described above, the barrier might be her library's lack of subscription to the journal in which the procedure is described.

Some of the barriers that researchers and knowledge workers face include:
- lack of awareness of the resources that are available to them;
- lack of access to appropriate hardware and software;
- broken links or other implementation issues;
- information not available digitally or in an inconvenient digital format;
- information being available only in an early version rather than the version of record;
- content in a format which is not suited to their needs;
- lack of membership of a library that has purchased a licence;

- a requirement to make a payment for access to the desired content at a level which the user considers disproportionate to the anticipated benefit;
- a burdensome purchasing process;
- conflict between the author's or publisher's rights and the desired use of the content;
- digital rights management or technical protection technologies that prevent the desired use of the content.

This study is important because it provides hard factual evidence on the size and significance of gaps and barriers to accessing scholarly information, for researchers within higher education and beyond.
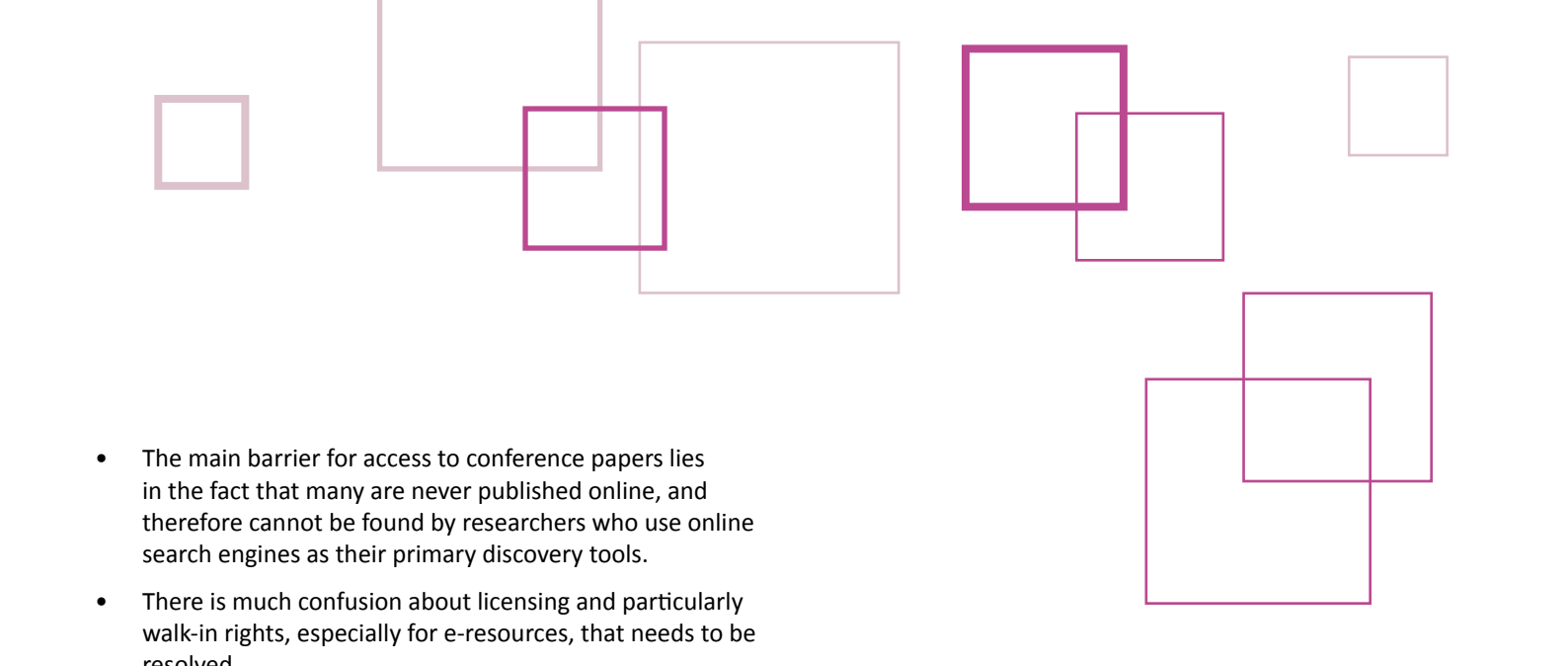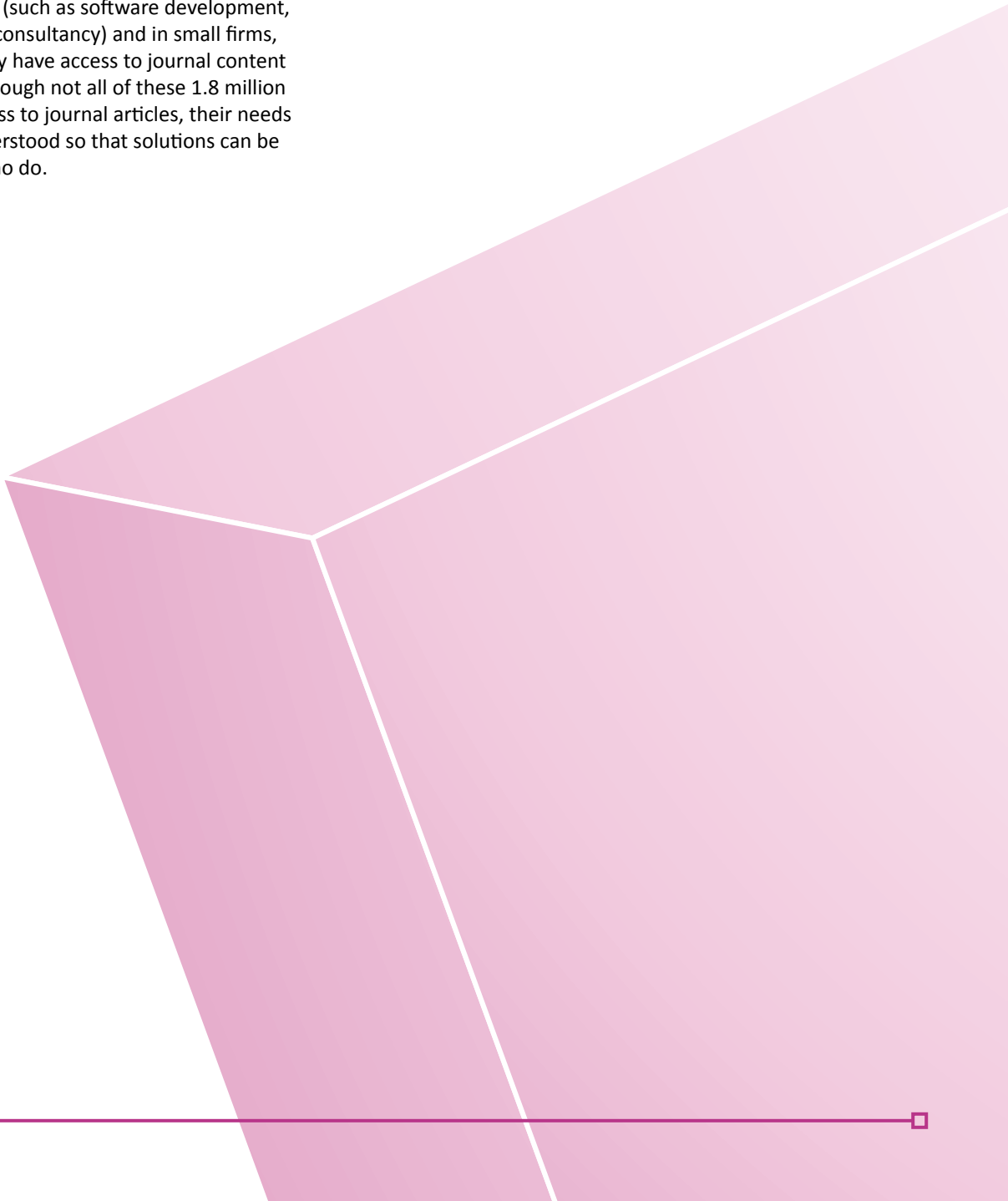
Key findings of this study

The key findings of this study are:

- Journal articles and conference papers are critical for advanced research and scholarship, and are rated as `important' by 90.4 per cent of survey respondents (for journal articles) and 58.2 per cent (for conference papers).

- Compared with other types of information resources, journal articles are relatively easy to access. This is especially true for survey respondents in universities and colleges, 93.1% of whom said that research papers were easy or fairly easy to access: the equivalent figure in industry and commerce was 79.1%. In a later question, put only to those researchers for whom journal articles are important, respondents in all sectors rated their access as somewhere between 'variable' and 'good'. Conference papers, on the other hand, were rated somewhere between 'variable' and 'poor'.

- Most researchers (71.5 per cent in the case of universities and colleges, 57.6 per cent in the case of industry and commerce) believe that access to journal articles has improved over the past five years. This increase in perceived access is probably due to innovations such as journal bundling or the `Big Deal', consortial purchasing by librarians, and the greater availability of information in digital form, including via open access.

- Despite these findings, there are specific areas where provision of access to journal articles is seen as less effective. The rest of this report focuses upon such areas, seeking to understand them as a precursor to providing some solutions to improve access where necessary.

- The UK industrial sectors reporting the poorest levels of journal access are the motor industry, utilities companies, metals and fabrication, construction, and rubber and plastics.

- The most common barrier to accessing journal articles in both academia and industry occurs when researchers must pay to access content. The majority of researchers for whom journal articles are important, in all sectors apart from industry and commerce , felt that they did not have access to enough titles through existing arrangements. It is possible that new discovery tools, especially gateway services like Google Scholar, PubMed, Scirus and the Web of Science, have exacerbated this problem by making it much easier to identify relevant literature, but not to subsequently access this literature.

Unlike previous finding aids such as library catalogues, these tools increase the amount of visible literature without promising access to that literature: consequently, researchers may see that the percentage of useful articles that they can access appears to have decreased.

- The findings suggest that information barriers can lead to significant non-productive activity and lost opportunities on the part of researchers and knowledge workers. Faced with a particularly hard-to-access journal article or conference paper, many researchers in both academia and industry simply give up and either look for another article with similar information, or do something else entirely.

- Researchers adopt a range of coping strategies to deal with articles they cannot easily access. For those in industry, the most common solutions (after giving up) are to look for an early version on the web, approach the author or order directly from the publisher. Researchers in academia are more likely to approach the author than to give up completely, and then use institutional solutions such as inter library loans and library-held hard copies, as well as online searches for early versions, to try and meet their needs.

- Most researchers feel that the current prices charged for individual journal articles are too high. Furthermore, a minority of researchers (26.3 per cent) have strong objections in principle to this mode of access.

- Nevertheless there are some indications of a potentially viable market for pay-per-view: 12.6 per cent of respondents say they might consider buying individual journal articles in the future, and this proportion rises to 43.8 per cent in the case of conference papers.

- Conference papers are less important than journal articles for many researchers, although this varies considerably between disciplines. In some, such as computer science, they are very important. Overall, 23.6 per cent of researchers rate conference papers as `extremely important' for their work. They are much more difficult to access than journal articles: 34.4 per cent of researchers and knowledge workers describe their current level of access to conference papers (in print or online) as `poor' or `very poor'.

- The main barrier for access to conference papers lies in the fact that many are never published online, and therefore cannot be found by researchers who use online search engines as their primary discovery tools.

- There is much confusion about licensing and particularly walk-in rights, especially for e-resources, that needs to be resolved.

- Based on an analysis of the Labour Force Survey, CIBER estimates that there are around 1.8 million professional knowledge workers in the UK, many working in R&D intensive occupations (such as software development, civil engineering and consultancy) and in small firms, who may not currently have access to journal content via subscriptions. Although not all of these 1.8 million necessarily need access to journal articles, their needs should be better understood so that solutions can be provided for those who do.

Context

Information environment

New technologies, business models and economic pressures have changed the information environment in which all researchers operate. Some of these changes have made it easier to access needed information, while others have added complexity to an already-challenging environment.

Growth of the scholarly literature

The annual volume of production of scholarly content has expanded consistently by around 3.5 per cent per annum for many years (see Figure 1), a growth rate consistent with the resources being put into public R&D but not with the growth in library budgets. More recently content overload has been exacerbated by channel overload. Whereas in the pre-digital era, the physical library was the first and possibly only port of call for research information, today's information environment is characterized by complexity: library systems, publisher and third party platforms plus a plethora of new social and other media.

Access not only to the published full text but also associated datasets, software, video links with text, etc., mean that the landscape is rich and varied, claiming greater attention of the limited time available to researchers.

Increasing specialisation

As the volume of knowledge grows, research effort tends to become more specialized. As a result, new disciplines and sub-discipline continually evolve and spawn new journals and conferences, further stoking the gaps between what is needed and what can be afforded. The same effect can be observed as a result of growing interest in inter-disciplinary research. Though there are several core journals in every discipline, these cannot meet all the specific information needs of a researcher. The increasingly specialized nature of scientific progress means that there is a need to keep up in other, possibly less familiar areas, as research becomes more global in outreach and multidisciplinary in approach.

Figure 1: Annual production of scientific articles: worldwide,1990-2010
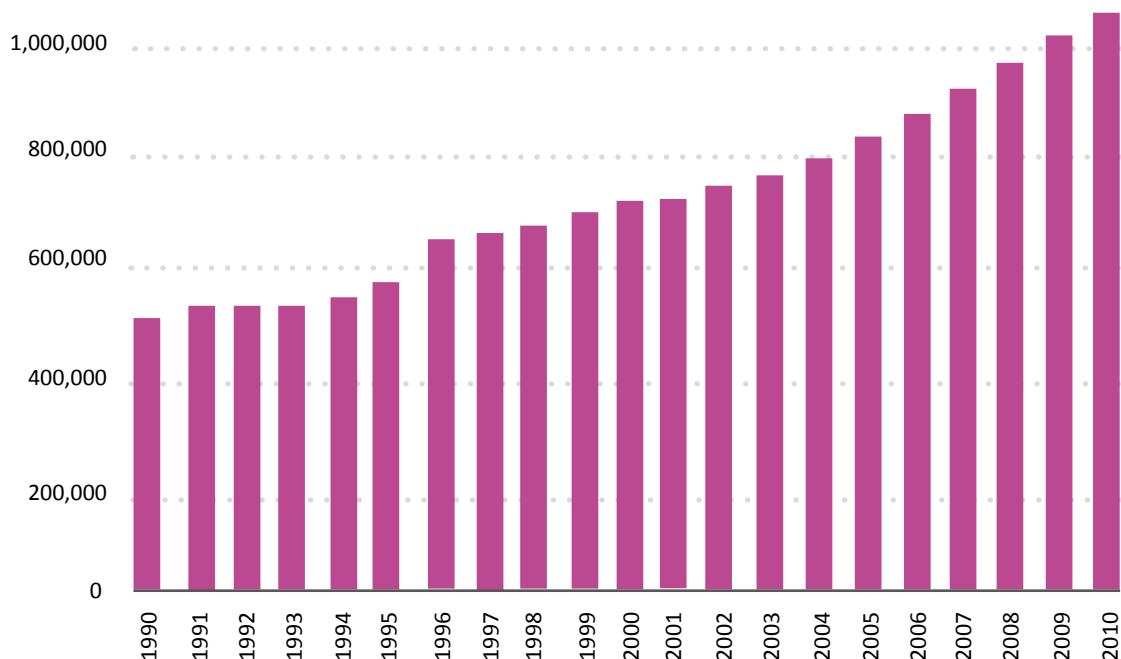
Source: CIBER analysis of Thomson Reuters Science Citation Index

Open access

Open access has been an important route to increasing availability of research articles that are free at the point of use. There are two main forms of open access. Gold open access relies upon journals which allow authors to make their articles freely available to readers immediately upon publication, usually by charging a fee - sometimes called an 'author pays' fee. Some journals are purely open access, while others have a so-called 'hybrid' model, where they operate on a subscription basis but allow researchers to pay a fee to make their individual article freely available.

Green open access consists of author self-archiving in institutional or subject based repositories. The repository will then make these copies freely available after an embargo period that is specified by the article's original publisher. This embargo period varies by discipline, but is usually 6-12 months. In some cases, researchers will also deposit an earlier version of their manuscript, often the 'submitted' version which has not been amended following peer review. These earlier versions are not usually subject to embargo periods.

Open access journals and repositories have grown significantly in number in recent years, and there has been a similar growth in the number of research funders who require publications arising from their grants to be openly available. There is relatively little reliable information on the number of OA articles available, but one estimate suggests that around 20.6% of ISI-indexed articles were available in open access form in 2008 via the green or gold routes.

Figure 2: Proportion of articles available via green or gold open access, or unavailable as open access, 2008
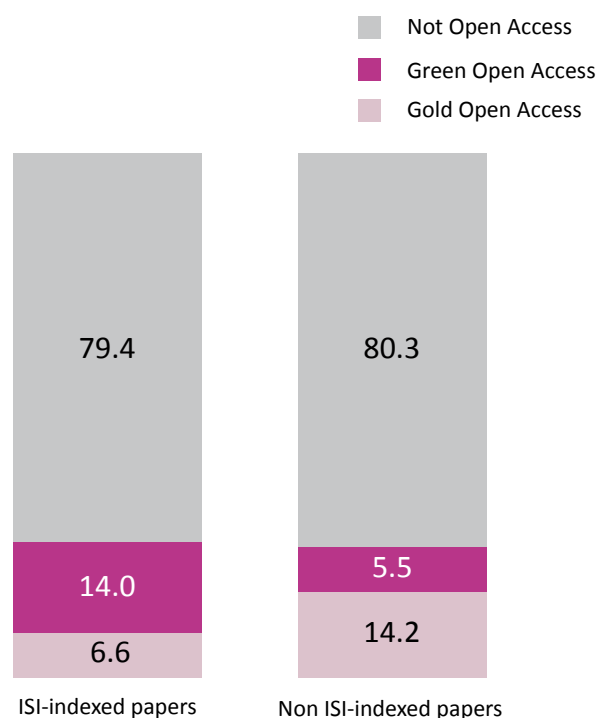
Source: Bjork and others, 2010

Researcher behaviour

In a world of rapidly growing literature, researchers have adapted by developing tactical online behaviours such as power browsing that maximize the efficiency and speed with which they scan and filter information. Such is the information flood that researchers are likely to need tools and algorithms to help them cope in the future: tools that treat the literature more as `data' to be mined.

The fragmented nature of research disciplines creates different search needs and habits among researchers: physicists differ markedly in their approach to scholarly communication to humanities researchers. Adopting procedures that shoe horn all scholars in the same approach could be counterproductive. Recent studies supported by RIN, JISC and Ithaka to investigate researchers in specific disciplines (life sciences, humanities, physical sciences) will be useful in helping to create more sensitively targeted and appropriate information support systems.
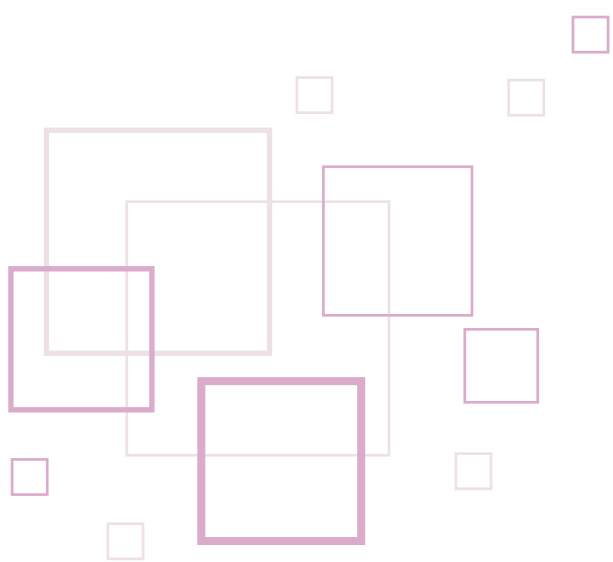
Knowledge workers

The term `knowledge worker' refers to those whose primary work involves developing and using knowledge in the workplace. Knowledge workers are found in professional, technical and managerial roles in all sectors of the economy, many working in organizations (or in a personal capacity) that do not have subscription access to scholarly e-content.   In the pre-Internet age, knowledge workers relied on membership of physical libraries to stay abreast of the latest developments reported in the scholarly literature. This has changed with the emergence of search engines that highlight the availability of relevant research material, and business models where content is licensed rather than owned. It is now easy for knowledge workers to locate the information that they need, but it is not always easy for them to access it, or to understand how they may reuse it.

Estimating the number of knowledge workers within the UK economy is problematic because of scoping issues and a lack of previous research. However, if we restrict our definition to numbers of professionals in various business sectors, then data from the Department of Business and Innovation Skills would lead us to conclude that there were 1.8 million such workers in 2009 in the UK. There are 130,000 IT staff involved in R&D activities alone, plus 78,000 civil engineers, 67,000 mechanical engineers, and so on. Of these an uncertain - but considerable - proportion are unaffiliated, without corporate library or information center support.

In addition there are large numbers of highly skilled individuals whose interests outside their employment make them potential consumers of scholarly information – so-called amateur scientists or hobbyists.

The British Library's BL Direct document request pattern gives some indication of latent demand among these 'unaffiliated' consumers. Though the majority of requests in 2010 – 10,500 – come from the academic sector, a further 3,300 came from professional sources, 2,200 from businesses and 3,100 from individuals. The 'long tail' of demand is beginning to make itself known, and will be further stimulated as a growing proportion of publications become `open'.

Detailed findings

Scholarly information needs in context

Figures 3 and 4 compare perceived importance of and perceived ease of access to a range of scholarly information resources, as rated by researchers in UK universities and colleges and knowledge workers in industry and commerce.

Respondents were asked to rate the importance to their research of each type of information using a 7-point scale, where 1=`not at all important` and 7=`extremely important'. We also record the percentage who rated each resource as being `fairly easy' or `easy' to access. These data are presented as column charts for universities and colleges (Figure 1) and for industry and commerce (Figure 2). There are problems, indicated by a serious discrepancy between importance and perceived ease of access, in the following areas:

Universities and colleges
- conference proceedings
- doctoral theses
- research datasets
- clinical guidelines
- legislative and regulatory information

Industry and commerce
- research articles
- reference works
- conference proceedings
- market research reports
- research datasets
- doctoral theses

These areas of provision are sub-optimal and deserving of further study , although mostly outside the focus of this report.

The analysis also shows the importance of original research and review papers in journals. Journal articles score from 5.67 on average in industry and commerce, to 6.32 in research institutes to 6.56 in medical schools and health providers to 6.86 in universities and colleges.

Compared with other scholarly and professional information resources, journal materials are relatively easy to access, with 93.1 per cent of university respondents rating ease of access as `easy' or `fairly easy' (81.6 per cent for research institutes, 74.9 per cent in the health sector and 70.1 per cent in the case of industry and commerce). Despite this finding, when asked which of the relevant resources on the graphics below they would most like to see access improved a large majority (38.5 per cent in the case of universities and colleges) identified original journal articles as their first choice, so it is clear that expectations around the availability and accessibility of journal articles are running very high. For many researchers, 'easy' access to most of the journal literature is not good enough.

Figures 5 and 6 display the same data in a different visual format, as scattergrams.

Figure 3: Importance and ease of access: universities and colleges

Mean ratings (left axis), where 7=`extremely important' and % of users for whom access is `fairly easy' or `very easy' (right axis) (n=1,159)

Figure 4: Importance and ease of access: industry and commerce
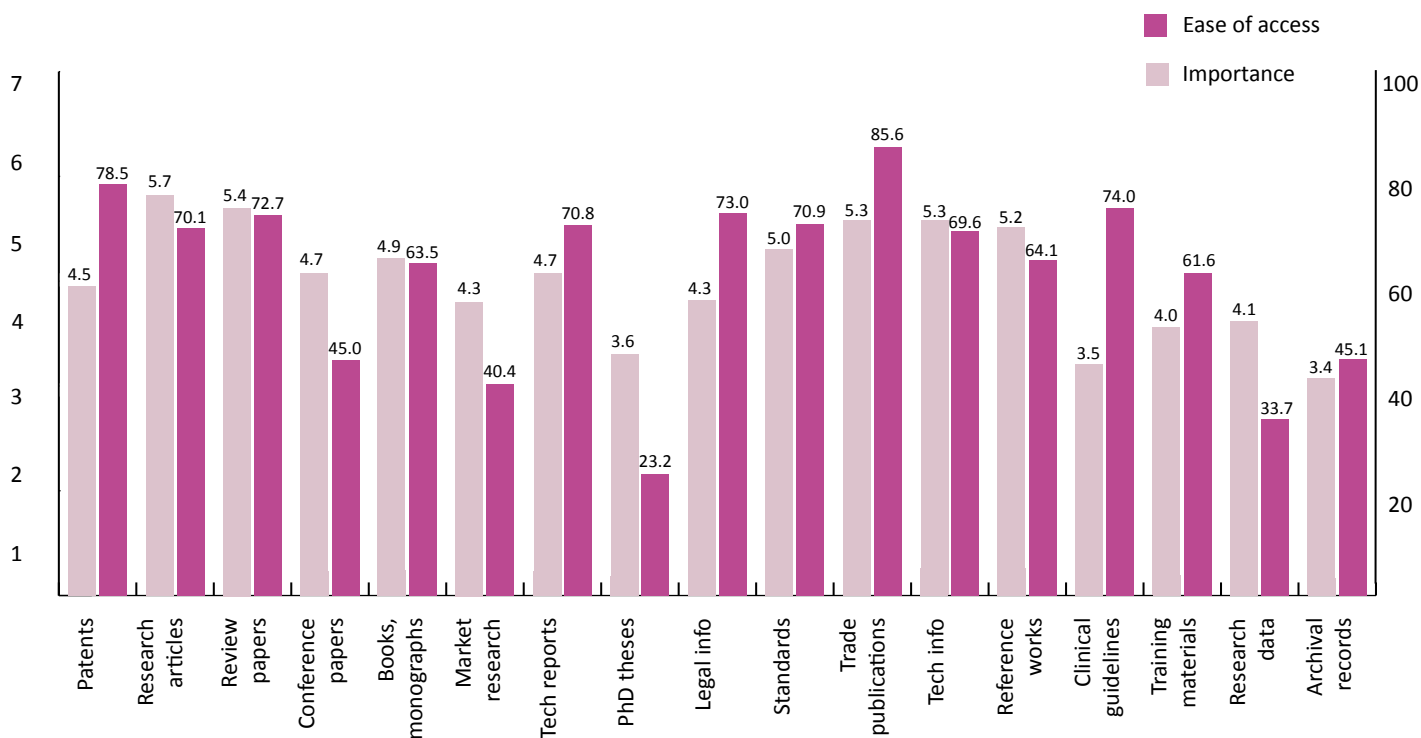
Mean ratings (left axis), where 7=`extremely important' and % of users for whom access is `fairly easy' or `very easy' (right axis) (n=699)

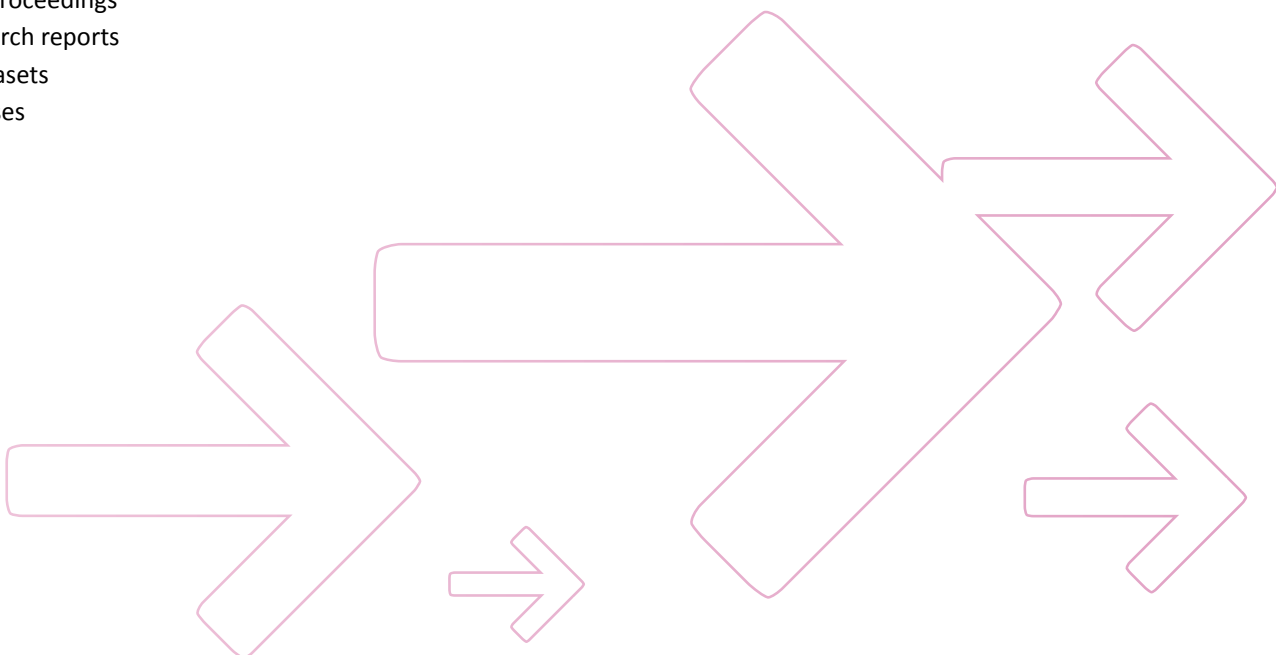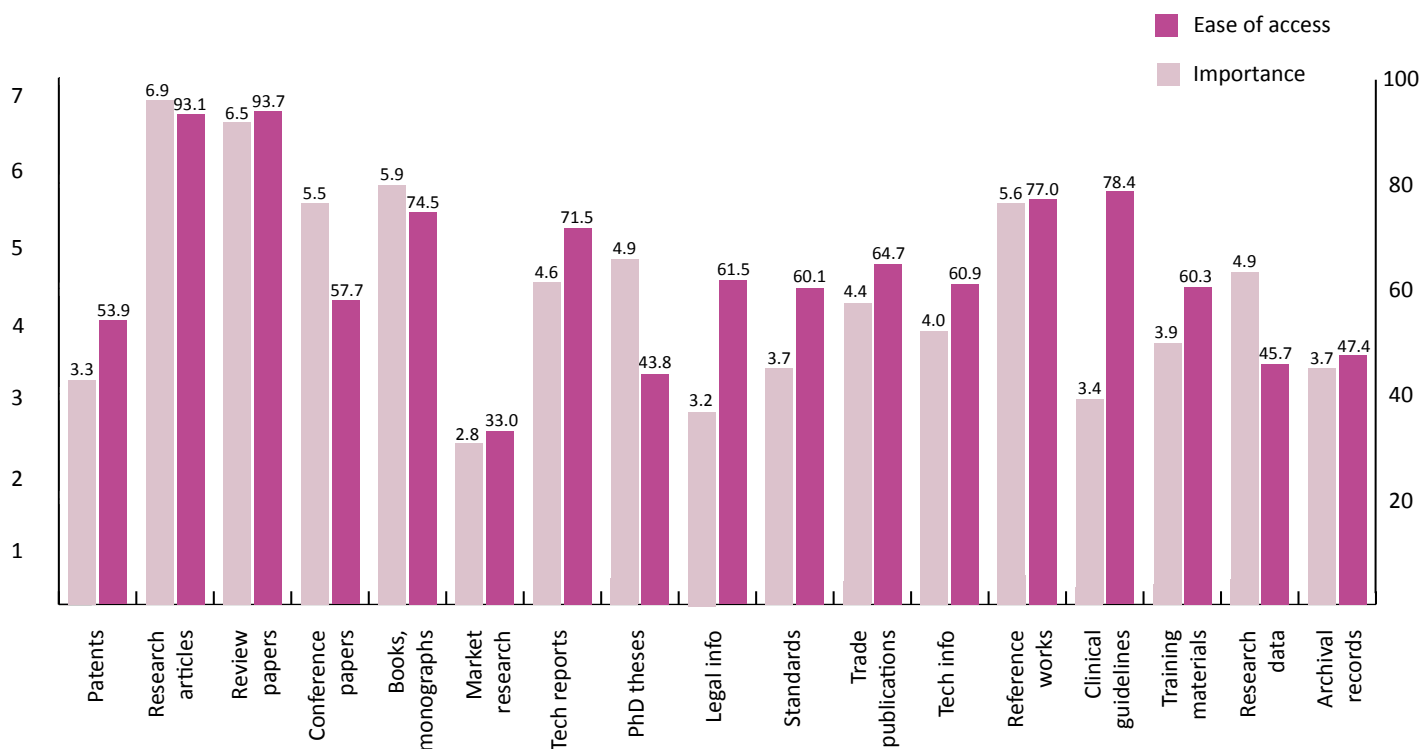Figure 5: Importance and ease of access: universities and colleges

Mean ratings (x axis), where 7=`extremely important' and % of users for whom access is `fairly easy' or `very easy' (y axis) (n=1,159)
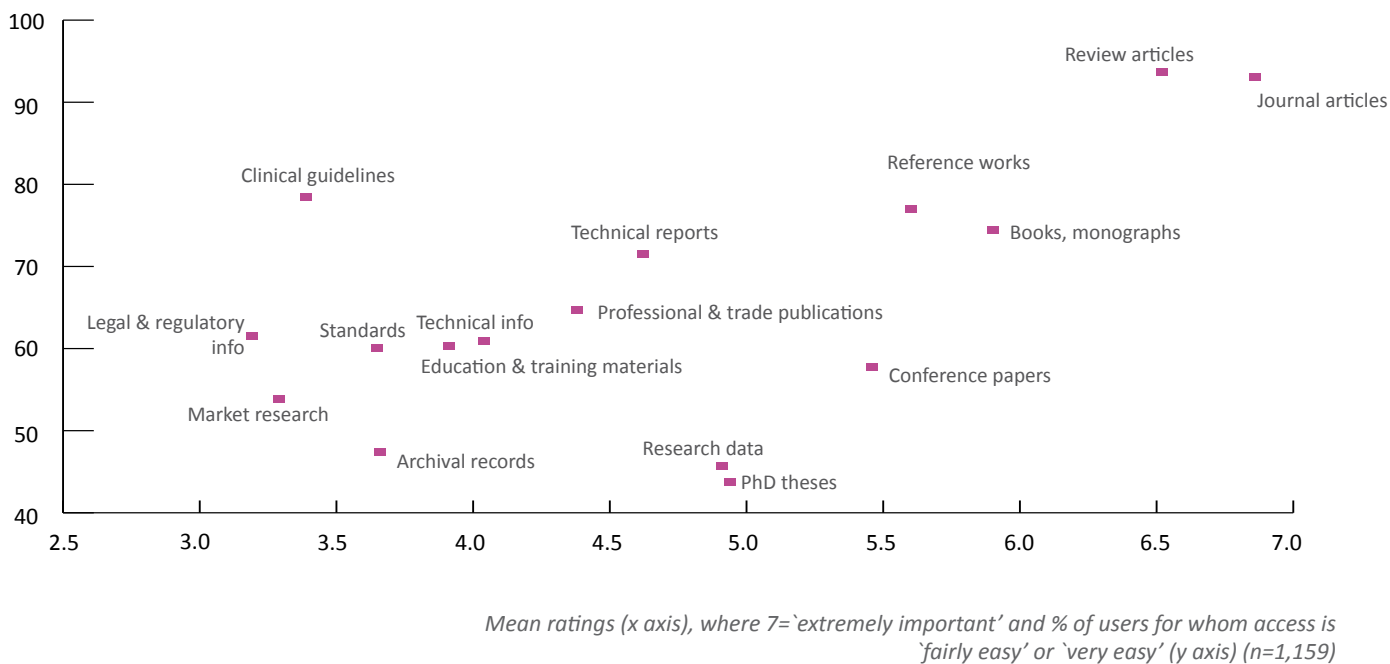
Figure 6: Importance and ease of access: industry and commerce

Mean ratings (x axis), where 7=`extremely important' and % of users for whom access is `fairly easy' or `very easy' (y axis) (n=699)
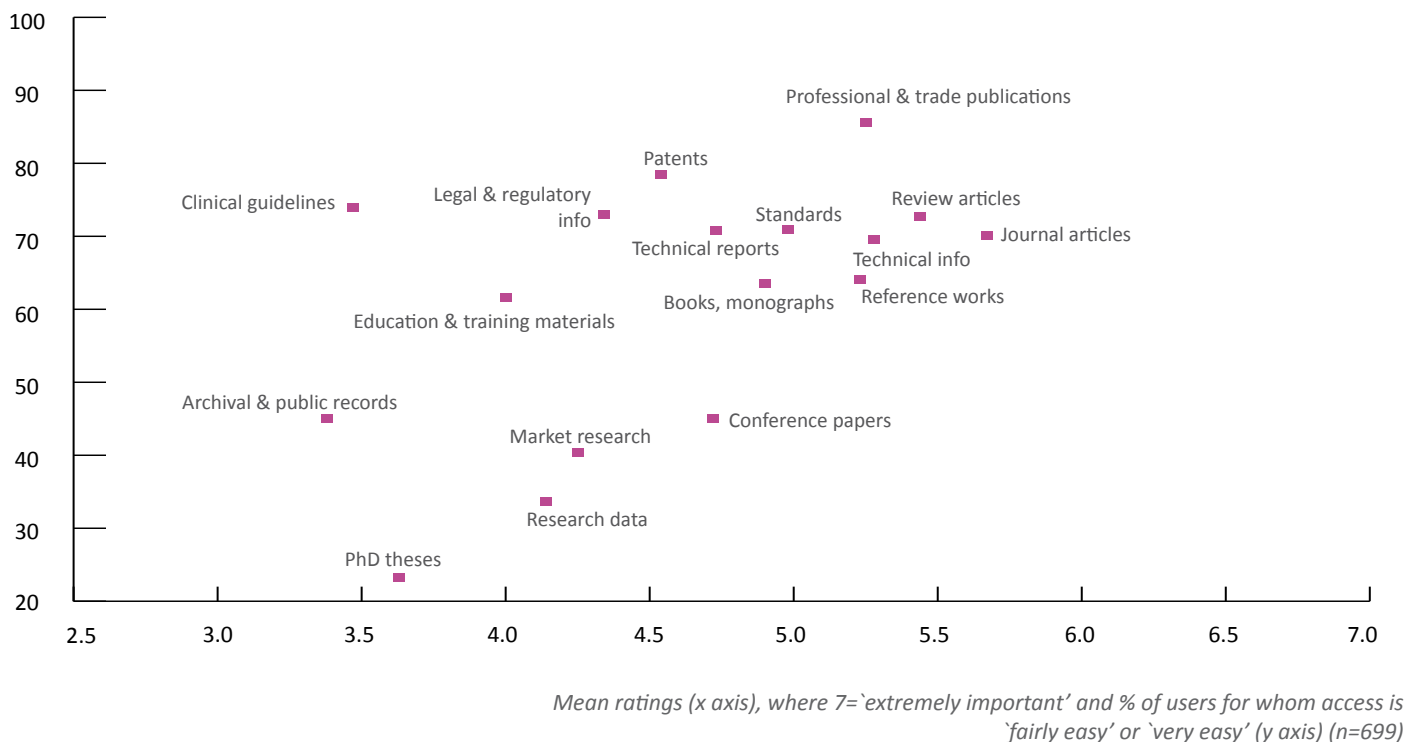
Current levels of access

There are around 25,000 peer reviewed journal titles in current production. No single library can possibly afford to acquire and process all of these titles, nor would it make sense for them to do so. One of the key characteristics of journal publication is that similar materials are brought together under one `cover'. The narrow scope and specialized nature of the journal means that for most researchers, a relatively small set of core titles can deliver a high proportion of what they need. But this approach cannot deliver everything to everyone, which means that gaps in provision are almost inevitable.

We asked researchers and knowledge workers to describe their perceptions of their current level of access (in print or online) to journal articles and conference papers, using a five-point scale where 1=`very poor' and 5=`excellent'. Figure 7 gives an analysis of responses by broad sector. This shows the average response to that question, with 95 per cent confidence intervals around the mean. The data suggest that researchers in universities and colleges currently perceive themselves to have significantly better levels of access (and greater variation) than their colleagues in any of the other three sectors, each of which falls short of a `good' rating in terms of their aggregate response. Confirming the findings from the earlier figures, satisfaction with current levels of access to conference papers lags very significantly behind journal articles for all groups.

Figure 7: Current level of access to journal articles and conference papers by broad sector (only those for whom the resource is important)

	conference papers	journal articles
Universities and colleges	2.87	3.64
Medical schools and health providers	2.63	3.45
Industry and commerce	2.63	3.09
Research institutes	2.84	3.45

Very poor = 1 · Poor = 2 · Varies = 3 · Good = 4 · Excellent = 5
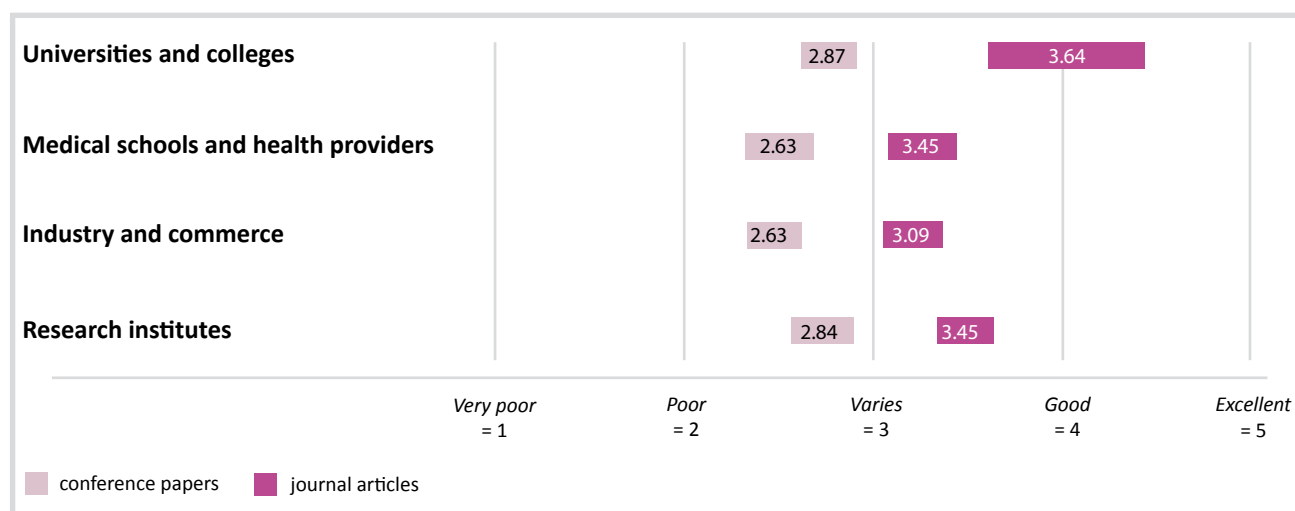
Mean ratings (n=2,473)

Current levels of access by subject

Journal articles

It is widely recognized that disciplinary differences play an important role in researchers' information behaviours. Therefore, we looked at the perceived levels of access to journal articles by discipline, both in higher education and in industry.

For researchers in universities and colleges, there is relatively little difference between subject disciplines in relation to current levels of access (Figure 8), with perceptions falling consistently between `varies' and `fairly good'.

The picture is much more variable in industry and commerce (Figure 9) where there are some large differences between disciplines. For example, environmental science scores much lower than neuroscience or pharmacology and toxicology. As well as clearer differences between subjects, there is considerably more variation within disciplines, as indicated by the generally larger confidence intervals around the means. In other words, provision is more patchy.

Figure 8: Current level of access to journal articles by subject: universities and colleges (only those for whom journal articles are important)

Subject	Mean rating
Arts and humanities	3.46
Biological sciences	3.66
Chemistry and chemical engineering	3.67
Earth sciences	3.88
Environmental sciences	3.71
Health sciences	3.61
Life sciences	3.61
Materials science and engineering	3.62
Maths and computer sciences	3.58
Neuroscience	3.77
Pharmacology and toxicology	3.83
Physics	3.78
Social sciences	3.65

Very poor = 1   Poor = 2   Varies = 3   Fairly good = 4   Very good = 5
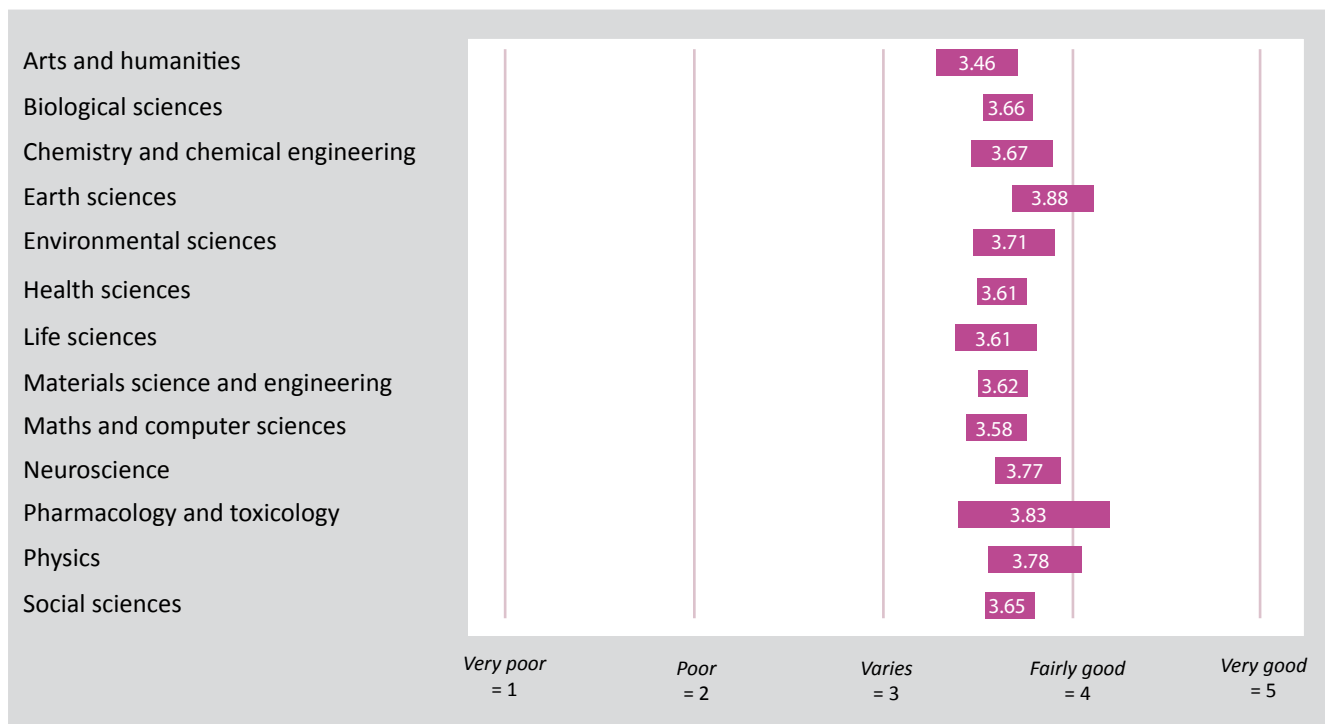
Mean ratings (n=1,171)

Figure 9: Current level of access to journal articles by subject: industry and commerce (only those for whom journal articles are important)

Subject	Mean rating
Arts and humanities	3.17
Biological sciences	3.24
Chemistry and chemical engineering	3.14
Earth sciences	3.24
Environmental sciences	2.69
Health sciences	3.30
Life sciences	3.19
Materials science and engineering	3.00
Maths and computer sciences	2.96
Neuroscience	3.75
Pharmacology and toxicology	3.47
Physics	2.92
Social sciences	2.78

Very poor = 1   Poor = 2   Varies = 3   Fairly good = 4   Very good = 5
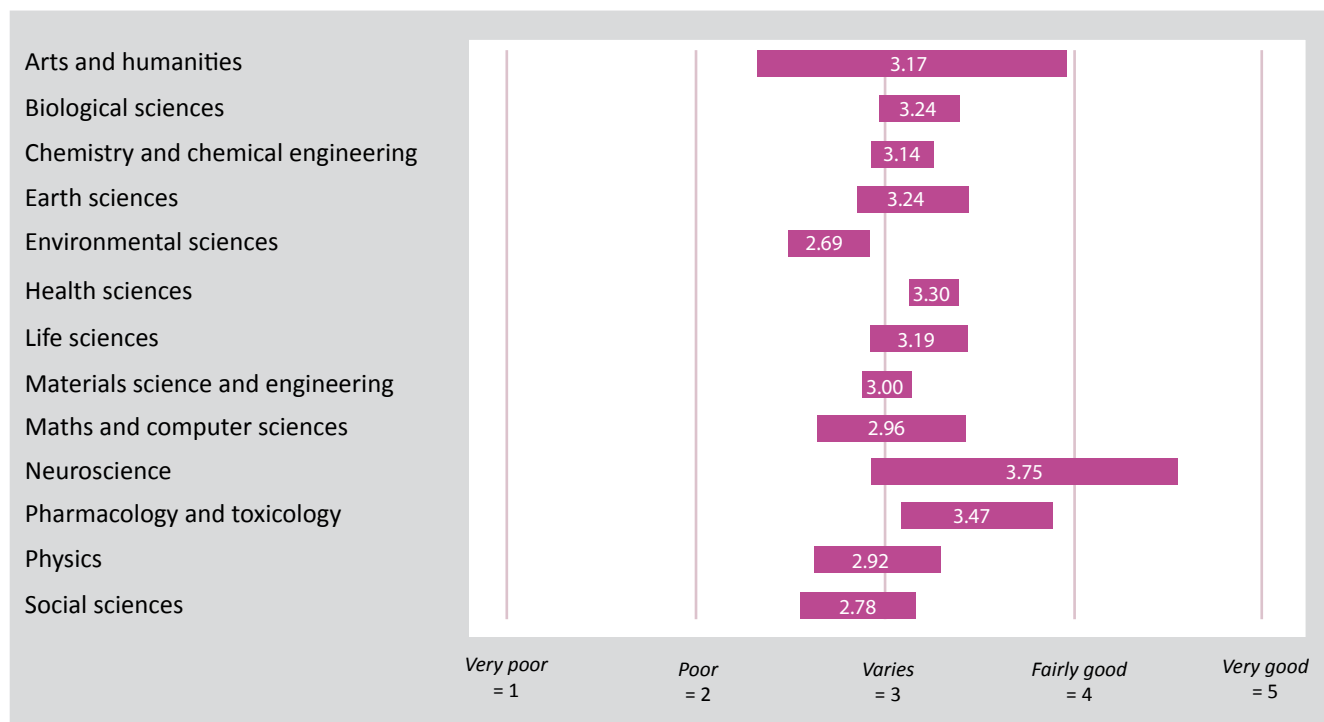
Mean ratings (n=762)

Conference papers

Compared with journal articles, conference papers score lower in terms of perceived current levels of access for researchers in both universities and colleges (Figure 10) and in industry and commerce (Figure 11).

All of the means reported here are considerably smaller than for journal articles and all fall around the interface between `variable' and `poor'. However, there are fewer obvious differences between disciplines, either in higher education or in industry.

Figure 10: Current level of access to conference papers by subject: universities and colleges (only those for whom conference papers are important)

Subject	Mean
Arts and humanities	2.96
Biological sciences	2.93
Chemistry and chemical engineering	2.64
Earth sciences	3.07
Environmental sciences	2.65
Health sciences	2.80
Life sciences	2.71
Materials science and engineering	2.67
Maths and computer sciences	3.09
Neuroscience	2.80
Pharmacology and toxicology	3.00
Physics	3.00
Social sciences	2.95

Very poor = 1 Poor = 2 Varies = 3 Fairly good = 4 Very good = 5
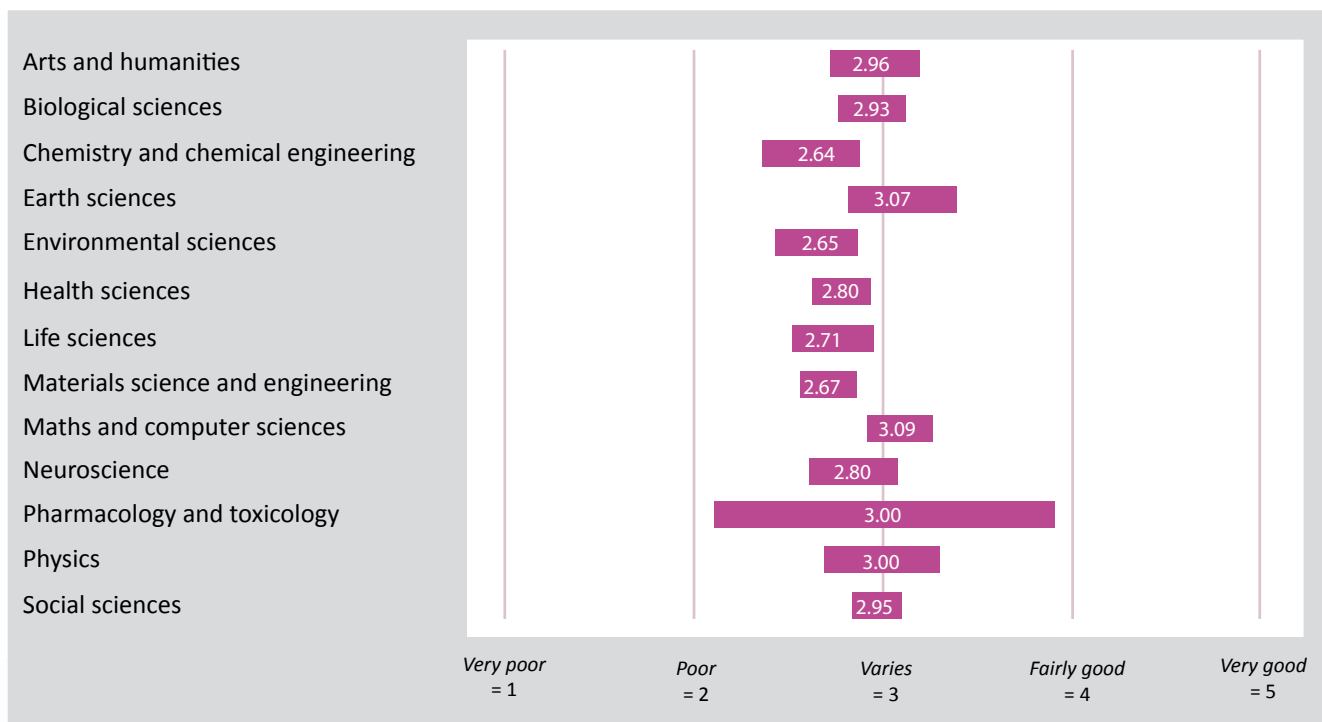
Mean ratings (n=1,209)

Figure 11: Current level of access to conference papers by subject: industry and commerce (only those for whom conference papers are important)

Subject	Mean rating
Arts and humanities	2.67
Biological sciences	2.72
Chemistry and chemical engineering	2.54
Earth sciences	2.39
Environmental sciences	2.49
Health sciences	2.62
Life sciences	2.44
Materials science and engineering	2.64
Maths and computer sciences	2.78
Neuroscience	2.75
Pharmacology and toxicology	2.79
Physics	2.58
Social sciences	2.73

| Very poor = 1 | Poor = 2 | Varies = 3 | Fairly good = 4 | Very good = 5 |
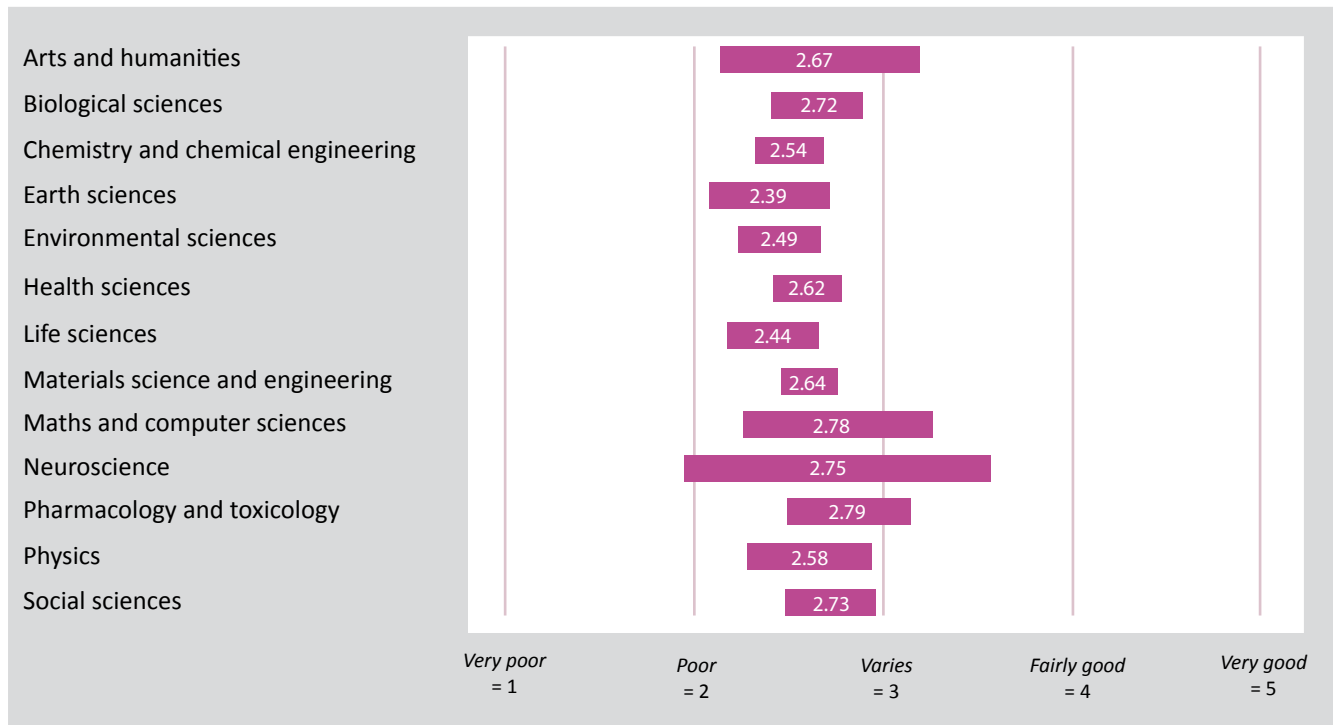
Mean ratings (n=769)

Current levels of access by researchers and knowledge workers in industry and commerce

Journal articles

Focusing specifically on knowledge workers in industry and commerce, the sectors with the highest levels of current access to journal articles (Figure 12) are pharmaceuticals, agriculture, computing and telecommunications. Those with the least satisfactory provision are the motor industry, utility companies, metals and fabrication, construction, and rubber and plastics. These fall statistically significantly (at the five per cent level) below the other sectors in terms of expressed satisfaction.

Once again, the general picture is that provision is far from optimal, hovering in all cases somewhere between 'poor' and 'variable'.

Statistically, there is no difference in the importance attached to journal articles by researchers in small and medium-sized enterprises - those that employ fewer than 250 staff - and larger organisations. There is however a significant difference in perceived levels of current access, with 68.9 per cent of knowledge workers in SMEs saying that this is 'fairly' or 'very easy' compared with 77.6 per cent of those in large companies.

Figure 12: Current level of access to journal articles by industrial classification (only those for whom journal articles are important)

Sector	Rating
Pharmaceutical	3.40
Agriculture	3.28
Computing	3.27
Telecommunications	3.27
Medical and precision instruments	3.24
Public administration	3.17
Chemicals	3.13
Food and drink	3.13
Oil and gas	3.12
Media publishing and printing	3.11
Aerospace	3.10
Machinery and equipment	3.06
Transport	3.05
Rubber and plastic	2.96
Construction	2.96
Finance	2.88
Metals and fabrication	2.80
Electricity and gas supply	2.77
Water supply	2.65
Motor vehicles	2.57

Very poor — Poor — Varies — Good — Excellent

average across the whole survey

Mean ratings, industry and commerce respondents only (n=1,020)

Conference papers

When we turn to conference papers in Figure 13, utilities companies (electricity, gas and water supply) emerge as being unusually poorly served by comparison with the other sectors. Any differences between sectors here are less clear cut than Figure 12 might trick the eye to believe, since the standard deviations on this question are wide.

This means that averages alone should not be taken too seriously: the real message here is that even within sectors, access to conference papers is variable and uneven.

There is no difference in the relative importance attached to conference papers between those working in SMEs or larger companies.

Figure 13: Current level of access to conference papers by industrial classification (only those for whom conference papers are important)

Sector	Mean rating
Telecommunications	3.18
Media publishing and printing	2.92
Computing	2.89
Motor vehicles	2.86
Food and drink	2.79
Finance	2.74
Aerospace	2.74
Public administration	2.74
Transport	2.71
Construction	2.70
Machinery and equipment	2.67
Medical and precision instruments	2.65
Agriculture	2.60
Metals and fabrication	2.60
Chemicals	2.59
Pharmaceutical	2.54
Rubber and plastic	2.54
Oil and gas	2.52
Electricity and gas supply	2.49
Water supply	2.41

Very poor — Poor — Varies — Good — Excellent

average across the whole survey

Mean ratings, industry and commerce respondents only (n=1,025)

Access now compared with five years ago

We asked survey respondents `How does your current level of access to journal articles and conference papers compare with five years ago?' The following figures and tables represent perceived levels of access, which may reflect a variety of factors including (but not limited to) more available content, better search and discovery systems, changing individual needs and changing methods of scholarly communications.

Journal articles

The direction of travel in Figure 14 is clear: the perceptions are that access to journal articles is easier now than it was in 2005. A large majority (71.8 per cent) of university and college researchers feel that access is `a little' or a `lot better' now than before.

Of the subject communities, researchers in the neurosciences report the most progress over the past five years (mean=4.26), physicists the least (mean=3.68).

There are no significant differences by industrial sector.

Conference papers

In contrast, most respondents did not perceive an improvement in access to conference papers. This applies across all sectors, as shown in Figure 15. Table 1 highlights the differences between perceived improvements to access for conference papers and journal articles, by comparing the difference between the percentage of researchers who felt access had got easier and those who felt it had got more difficult. In the table, the + sign before responses suggests that more researchers thought access had improved than thought it had declined.

Subject or industrial sector do not make significant differences to perceived changes in access to conference papers.

Table 1: Sectoral differences in perceptions of change in access

	Net difference between 'a little' or 'a lot easier' and 'a little' or 'a lot more difficult'	
Broad sector	*Journal articles*	*Conference papers*
Universities and colleges	+ 64.1	+ 31.1
Medical and health	+ 47.6	+ 27.9
Industry and commerce	+ 43.9	+ 23.7
Research Institutes	+ 46.2	+ 19.8

Percentage point differences
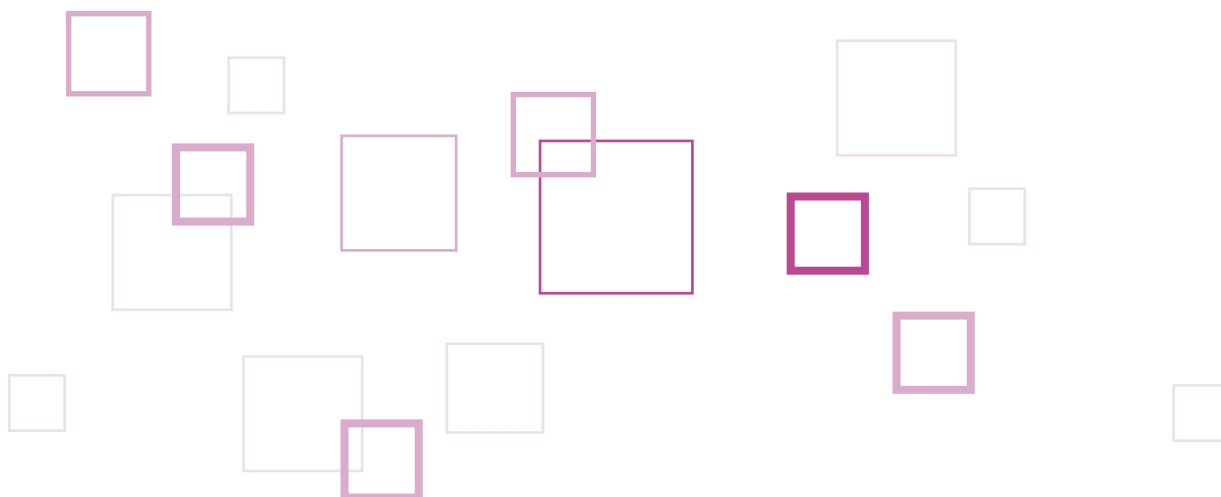
Figure 14: Perceived levels of access to journal articles now compared with five years ago (only those for whom journal articles are important)

- A lot more difficult
- A little more difficult
- About the same
- A little easier
- A lot easier

Percentages within broad sector (n=2,242)
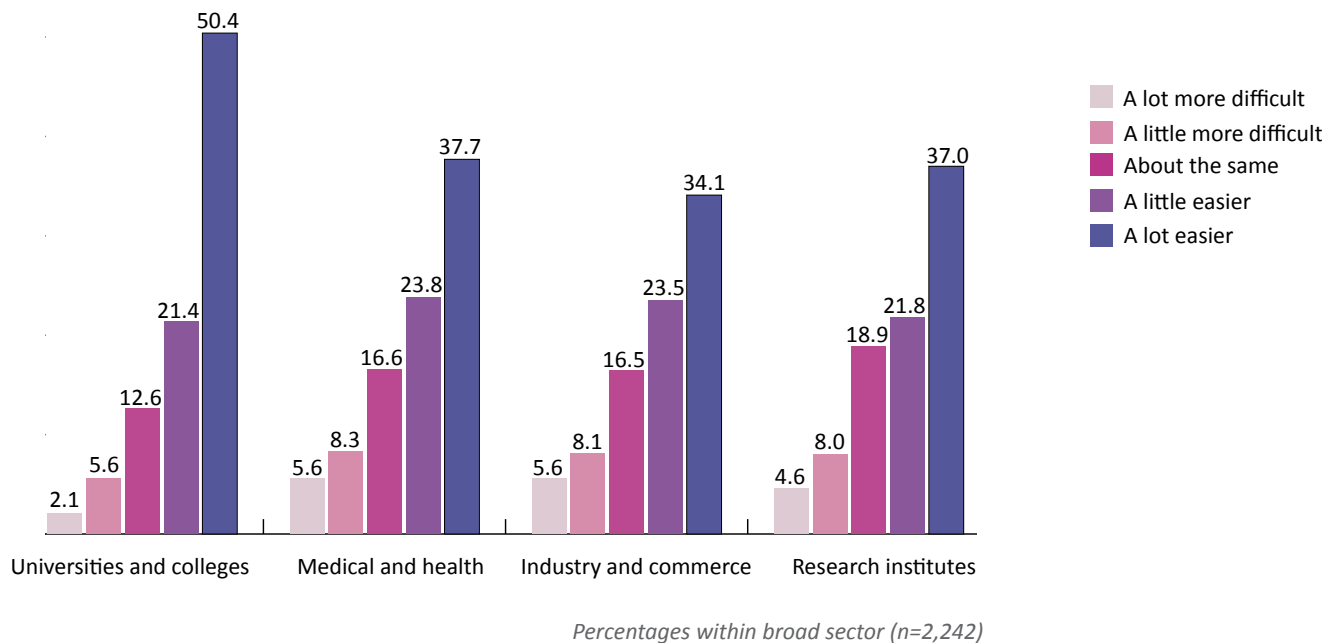
Figure 15: Perceived levels of access to conference papers now compared with five years ago (only those for whom conference papers are important)

- A lot more difficult
- A little more difficult
- About the same
- A little easier
- A lot easier

Percentages within broad sector (n=2,019)
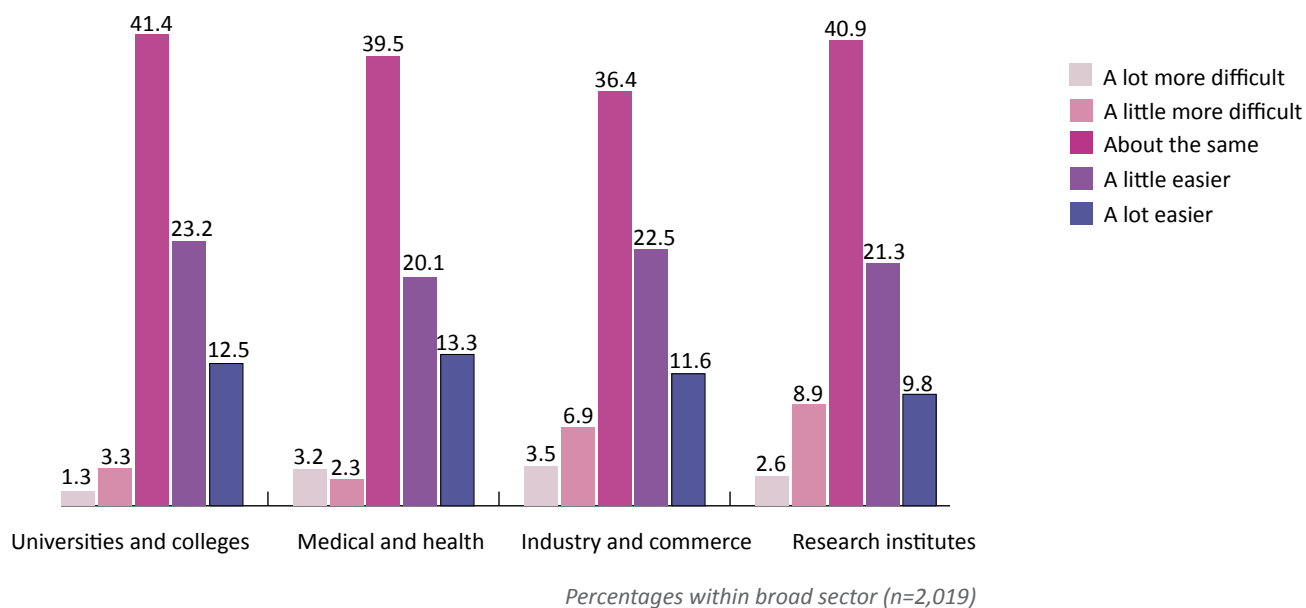
Limited access to journal articles and conference papers

We earlier defined a gap as a mismatch between a known information need and a researcher's ability to meet that need. Discovery tools have made it much easier for researchers to find relevant information – in other words, to identify their own information needs. But this can lead to real frustration if researchers cannot then access the full text of desired articles. We asked respondents who said that they would like to see access to journal articles or conference papers improved to say why, using a free comments box. Their comments were coded using a grounded theory approach, allowing the main themes to emerge from their text rather than being imposed using pre-coded categories.

Journal articles

The main complaint of researchers in universities and colleges (68.2 per cent) is that they do not have access to a sufficiently wide range of titles. Researchers in medical schools and health providers (52.8 per cent) and research institutes (57.4 per cent) also felt this to be a major concern, but only 17.7 per cent of researchers in industry and commerce mentioned it as a problem.

This is particularly interesting given the general decline in special libraries and corporate information centres, which provided important research support to businesses.

The results, as shown in Figure 16, highlight a secondary concern: sometimes older material is not available online.

Conference papers

As Figure 17 shows, the issues for conference papers are different. Most problems here seem to arise because conference papers are not routinely published, made available online or organized in a way which permits easy discovery.

Figure 16: Gaps in article provision (only those for whom journal articles are important and who want to see access improved)
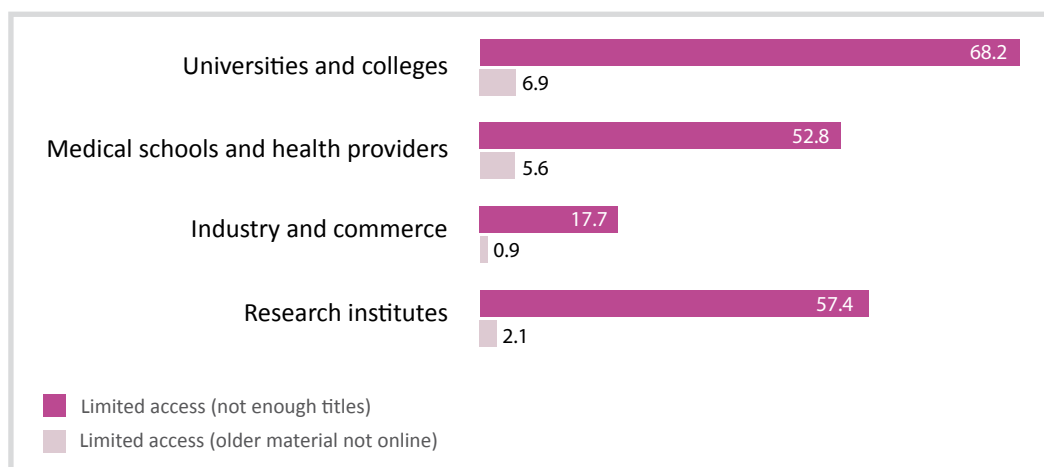
Percentages of respondents mentioning limited access to journals (n=521)

Figure 17: Gaps in conference paper provision (only those for whom conference papers are important and want to see access improved)

Universities and colleges
- 44.6
- 24.3
- 14.9

Medical schools and health providers
- 42.9
- 28.6

Industry and commerce
- 18.9
- 51.4
- 10.8

Research institutes
- 20.0
- 10.0
- 10.0

- Full text not available online
- Difficult to identify relevant papers
- Never published or published too slowly
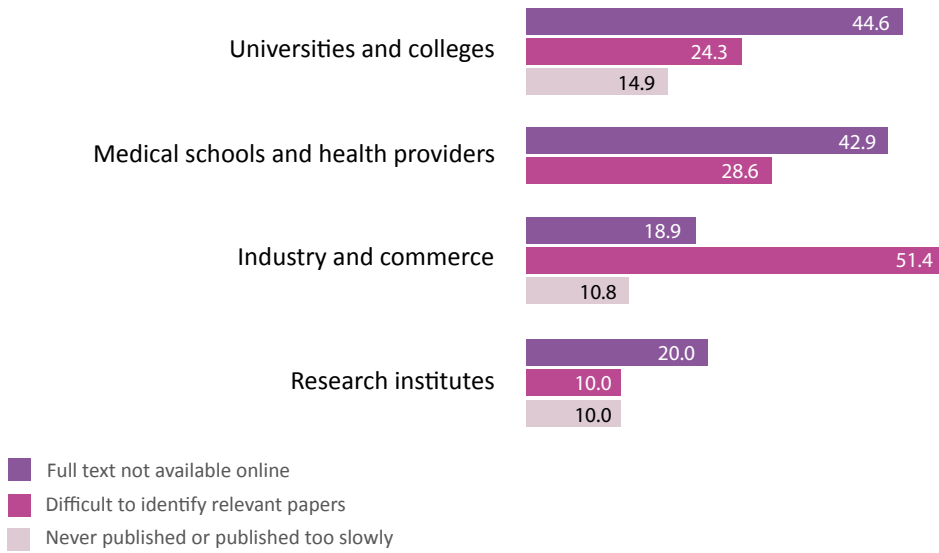
Percentages of respondents mentioning limited access to conference papers (n=128)

Assessment of gaps and their significance

Figure 18 shows, for universities and colleges and for industry and commerce, how often researchers fail to find the information that they need.

Our survey design took respondents through a series of filters. We first established how important journal articles and conference papers were to respondents' research, in order to exclude those for whom they are not actually relevant. Relevance was defined as a rating greater than 4 on a 7-point scale where 1=`not at all important' and 7=`extremely important'. For journal articles, 98.6% of academics and 78.8% of industry and commerce workers fell into this category. For conference papers, these percentages were 77.1% and 60.0% respectively. The figure presents this group as a base population of 100 researchers.

We then asked them to rate the perceived difficulty or ease of their current access. The percentage of those who rate it `fairly' or `very difficult' forms the first row in the tables opposite, represented as a percentage of the original 100 researchers for whom access is important.

Next we asked the respondents whether they could recall a recent experience (a `critical incident') where access to the full text of a paper had been problematic. This forms the

second level of the graphic and it is expressed using the same baseline (i.e. all researchers for whom articles or conference papers are important).

Finally, we asked whether they eventually managed to secure the full text. Those who could not are shown as the final level in the graphic.

The findings highlight considerable differences between journal articles and conference papers, with the latter being very much more problematic. In the case of journal articles they also highlight major sectoral differences, with gaps being more frequently encountered in industry and commerce, compared with academia.

It is impossible to estimate the impact of these failures using the data within this study. The consequences might be trivial, perhaps merely a cause of mild frustration, or they might be very serious. It is possible to imagine, for instance that an experiment might be repeated pointlessly or a funding or commercial opportunity missed due to a researcher or knowledge worker missing something really critical in the literature.

Figure 18: Gaps and their significance

For every **100** researchers or knowledge workers who rate access to **journal articles** as important

In universities and colleges	In industry and commerce
5.4 describe their current level of access as 'fairly' or 'very difficult'	**24.6** describe their current level of access as 'fairly' or 'very difficult'
5.2 recalled a recent access problem	**20.1** recalled a recent access problem
2.3 recalled a recent access problem that was not eventually resolved	**17.1** recalled a recent access problem that was not eventually resolved

For every **100** researchers or knowledge workers who rate access to **conference papers** as important

In universities and colleges	In industry and commerce
32.1 describe their current level of access as 'fairly' or 'very difficult'	**32.0** describe their current level of access as 'fairly' or 'very difficult'
15.3 recalled a recent access problem	**17.0** recalled a recent access problem
11.7 recalled a recent access problem that was not eventually resolved	**12.3** recalled a recent access problem that was not eventually resolved

Survey responses rescaled per 100 researchers who say that journal articles or conference papers are important

Barriers to accessing journal articles and conference papers

In order to establish the nature and extent of barriers to access to scholarly content, we used a `critical incident' approach: asking respondents to focus on a recent event when they had had problems accessing a specific journal article. Figure 19 shows why they said they needed that particular article. Research was the main intended use across all sectors, but researchers in medical and health sectors also use articles for training and personal study to a much greater extent than those in other sectors. This suggests that the impacts of gaps in access affect not just research capacity but also education and personal interests and development.

Journal articles

The main barrier to access is an unwillingness to pay for an article at the prices currently being quoted (Figure 20). This is followed by the lack of a hard copy in the library or the unavailability of the article in a digital format.

Differences in responses between the four sectors are insignificant, except that university and college researchers are much more likely (at the five per cent level) than the others to report problems accessing material from home or not being able to find a physical copy in their library.

Knowledge workers in industry are much more likely to report technical problems paying for an article once they find it, which probably reflects the greater incidence of pay-per-view activity in this sector.

Conference papers

Pay walls offer a less substantial barrier to accessing conference papers, but this is probably because, as the two other main issues suggest, many papers are either not available online, or cannot easily be discovered there.

Again, the responses to this question (Figure 21) are similar across the four sectors, although knowledge workers in industry were much more likely than others to say that the conference paper was available online but in a format (e.g. a flat PDF) that did not allow for further manipulation of the content.

Figure 19: Reasons for needing critical incident journal article (only those who experienced a recent problem)

Percentages within sector (n=315)

Figure 20: Critical incident barriers to access: journal articles (only those who experienced a recent problem)

Barrier	Universities and colleges	Industry and commerce
Found the article online but had to pay to access the content I needed	79.4	77.4
Could not find the article in the library (library did not have a physical copy)	25.4	5.8
The article is not available in an online version	25.4	19.4
Tried to access the article from home but discovered I could only access it from work	19.0	1.9
Searched online but could not find the article	17.5	11.0
Could not access the article online because I could not remember or did not know the password	12.7	3.9
Found the article online, had to pay for it but had technical difficulties paying	9.5	21.3
Found the article online but it was not in a useful format (e.g. a flat pdf)	6.3	8.4
Could not remember the exact name of the author and/or title of the article	4.8	3.2
Could not access the article because of a broken link	4.8	8.4

Universities and colleges
Industry and commerce

Percentages within sector (n=315)

Figure 21: Critical incident barriers to access: conference papers (only those who experienced a recent problem)

Barrier	Universities and colleges	Industry and commerce
The paper is not available in an online version	43.8	33.3
Searched online but could not find the paper	27.3	23.9
Found the paper online but had to pay to access the content I needed	26.6	34.7
Could not find the paper in the library (library did not have a physical copy)	15.4	4.8
Could not access the paper because of a broken link	5.0	7.8
Tried to access the paper from home but discovered I could only access it from work	4.2	2.3
Found the paper online, had to pay for it but had technical difficulties paying	3.5	4.6
Could not access the paper online because I could not remember or did not know the password	3.0	3.4
Found the paper online but it was not in a useful format (e.g. a flat pdf)	2.5	4.8
Could not remember the exact name of the author and/or title of the paper	2.1	4.1

Universities and colleges
Industry and commerce

Percentages within sector (n=821)

Barriers to re-use

Many researchers access scholarly resources in order to re-use content in some way; for teaching, to extract tables, figures or other materials for use in their own publications, to share with colleagues, or to undertake text or data mining. As Figure 22 shows, this latter activity appears to be particularly prevalent in industry and research institutes. But this may be because respondents took `text or data mining' to mean anything on a spectrum from reviewing the literature to cut and paste.

Figure 23 is the result of a content analysis of barriers to re-use article content. Only a minority of respondents in universities and colleges (31.4 per cent) and even fewer in research institutes (17.5 per cent) said that they generally had few or no problems re-using article content. However, in the medical and health sector (58.1 per cent) and in industry and commerce (56.5 per cent) those experiencing such problems were in a majority. As suggested above, it may be that definitions of 'reuse' differ between the various sectors.

This analysis also shows the large number of researchers and knowledge workers who complain that both articles and abstracts are often unfit for purpose. Abstracts may be so vague that they discourage researchers from going to the next step and accessing the full text. This is a particular problem when users are faced with a pay wall and are uncertain whether the investment will be worthwhile. There were also many complaints about the limitations of full text articles, centering around two issues: poor writing and editing and tight word limits. Some articles are so poorly written that it is impossible to replicate an experiment or, in the worst cases, fully to understand what is being reported. Similarly, the typical 5,000-7,000 word limit can hinder communication: simply not enough detail is available to evaluate what has been done. This unexpected finding shows that access to full-text content is not necessarily the end of the story: poorly written, overly terse, or incompletely documented work can present further barriers to usefulness.

Figure 22: Intended re-use of critical incident journal article (only those who experienced a recent problem)

Percentages within broad sector (n=313)

Undertake text or data mining on this and other documents: 41.7, 29.4, 61.7, 80.0

Extract tables, figures or other materials for use in a publication of your own: 38.9, 23.5, 19.1, 20.0

Share a copy with colleagues: 27.8, 35.3, 25.5, 30.0

Share a copy with students: 19.4, 23.5, 2.1, 0

- Universities and colleges
- Medical schools and health providers
- Industry and commerce
- Research institutes
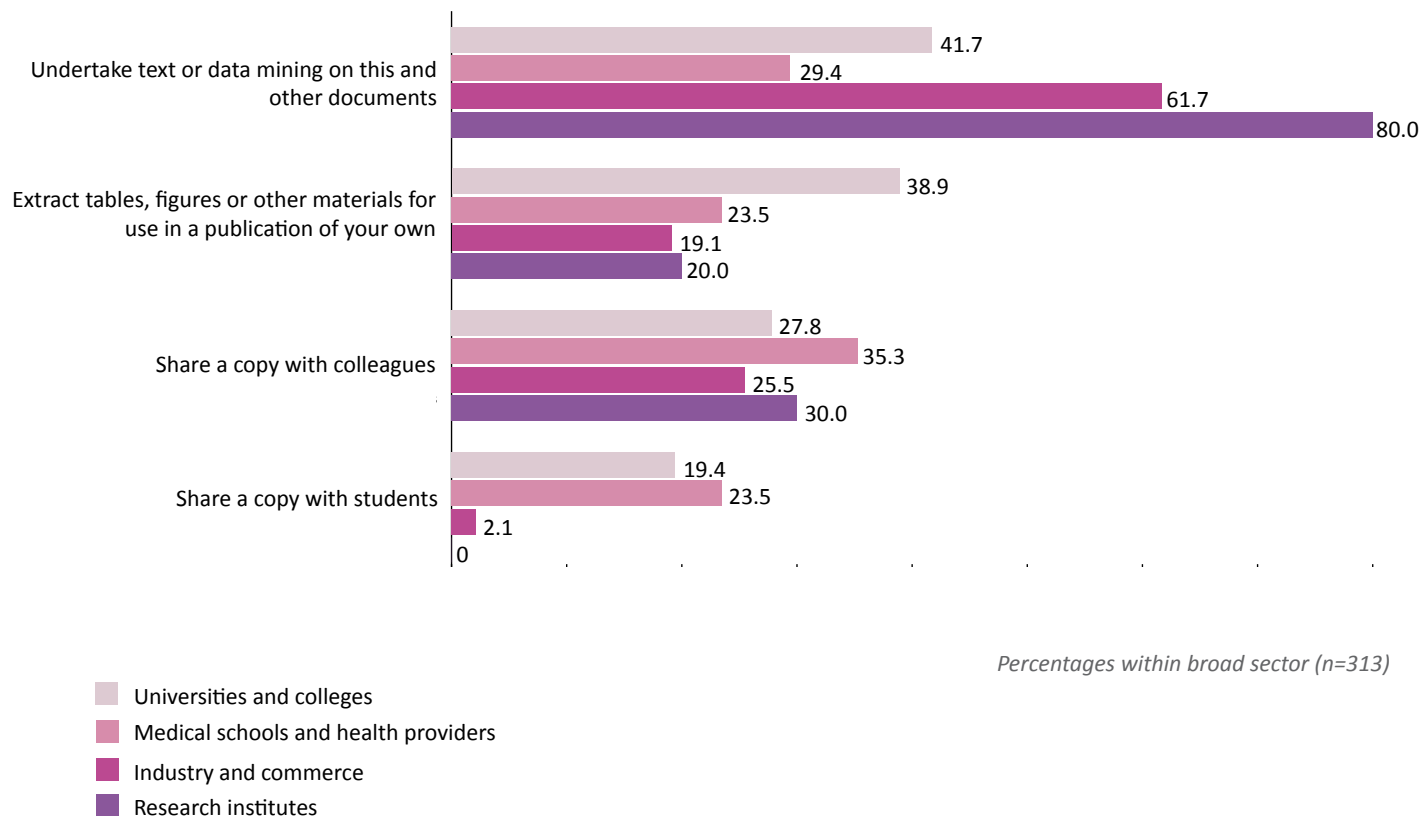
Figure 23: Barriers to re-use of journal articles (only those for whom journal articles are important)

Barrier	Universities and colleges	Medical schools and health providers	Industry and commerce	Research institutes
No	31.4	58.1	23.8	56.5
Poor quality writing and editing makes it difficult to replicate work	17.5		7.0	
Copyright issues (e.g. fair dealing)	13.1	19.4	29.4	15.2
Difficulty or confusion over getting permissions	10.2	6.5	6.3	
Difficulty to access underlying data or supplementary materials	9.5	3.2	5.6	10.9
Problems with e-formats (e.g. read-only pdfs)	8.8	3.2	8.7	4.3
Unable to post a copy on a web site or unsure	3.6	4.8	8.7	6.2
Foreign language issues	2.2	3.2	4.0	
Proprietary or restricted access information	2.2		4.0	
Hard copy availability only	1.5	1.6	0.8	4.3
Information not in a suitable form for text mining			0.8	

Content analysis of free text responses, percentages within sector (n=371)

Paying for scholarly content

As we have seen, pay walls may constitute a disincentive to article use if the prices are too high. Figure 24 is a classification tree that relates the answers to two questions: satisfaction with current levels of access to the journal literature, and whether the respondent had exercised the option to pay per view over the last twelve months. The branches split at points where clusters of respondents occur that are statistically significantly different from one another.

Where access is judged to be `excellent', the prevalence of pay-per-view activity is very low (7.9 per cent of all respondents). This rises to 22.2 per cent in the case of `good' access and 36.8 per cent in the case of `very poor', `poor' or `variable' levels. The message is both clear and obvious: pay-per-view models are used as a coping strategy in the event of a gap or access barrier, not as a general strategy for accessing the literature except perhaps in the case of certain pockets within industry.

Nevertheless, the analysis in Figure 25 suggests that pay-per-view business models could still play an important part in addressing gaps and barriers. As well as quantifying the extent to which researchers are already using these services, it reveals considerable latent demand (those that have not yet `but might in the future' use these services).
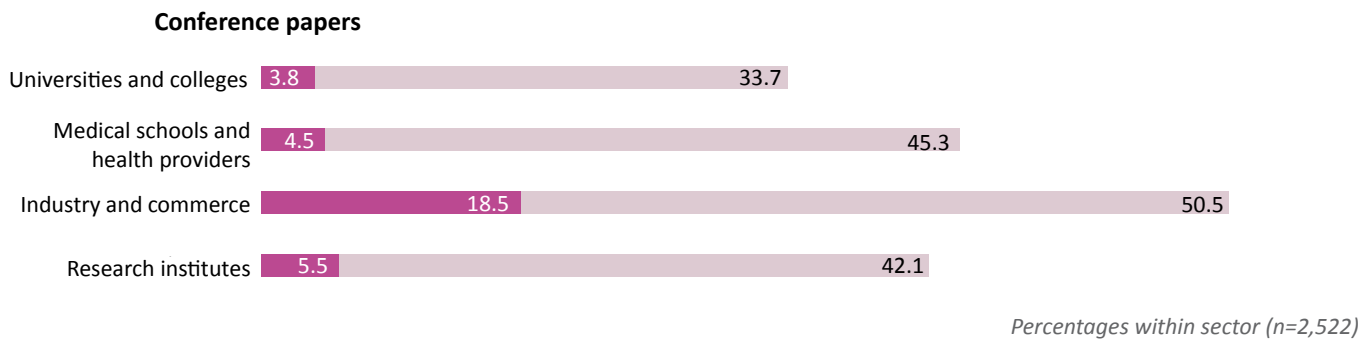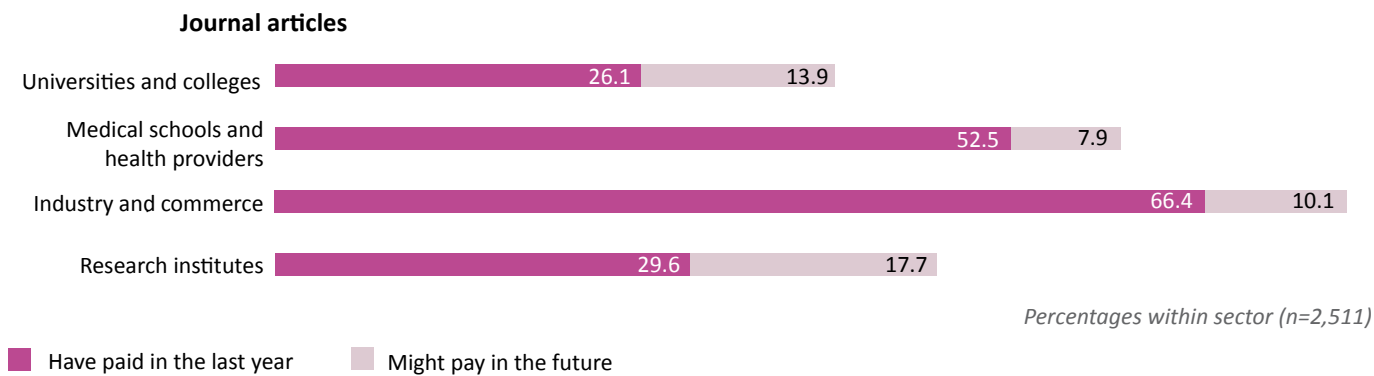
In terms of subject or discipline, researchers in the relatively well-funded areas of the health and life sciences and pharmacology and toxicology are the most likely to be current pay-per-view customers. The industry sectors with the highest concentrations of pay-per-view use are: food and drink; medical and precision instruments; and media, publishing and printing.

Figure 24: Purchasing of article level content and level of access to the literature (only those for whom journal articles are important)
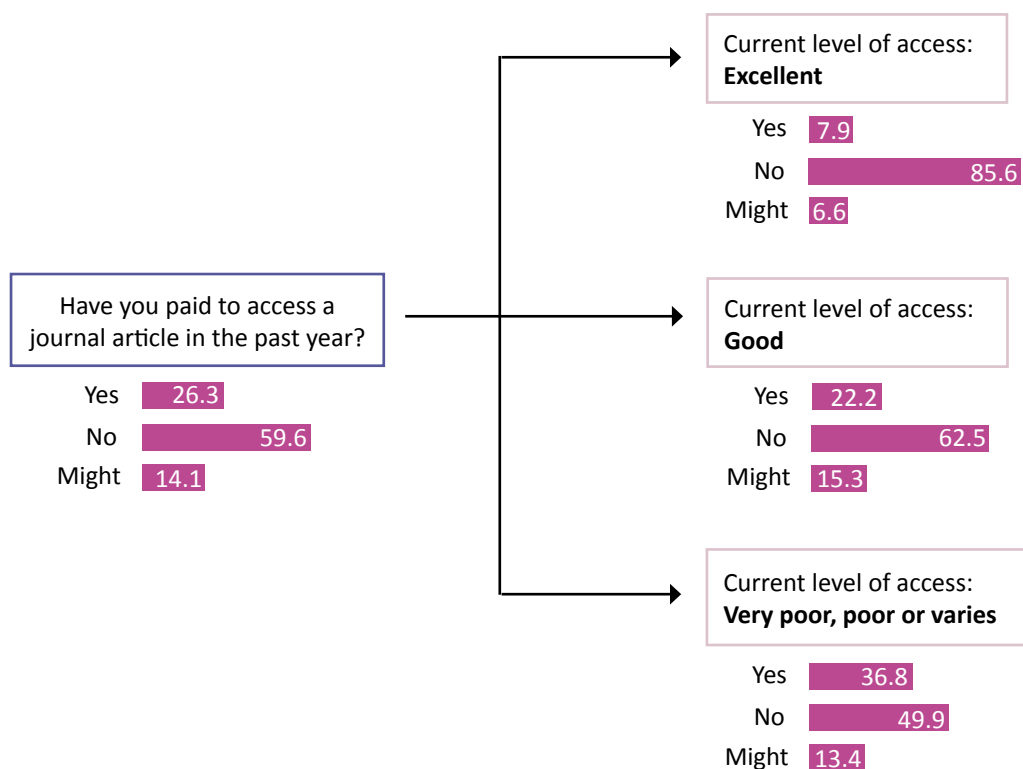
Percentages of respondents in all sectors (n=2,421)

Figure 25: Paying for access: journal articles and conference papers (only those for whom journal articles and conference papers are important)

Journal articles

	Have paid in the last year	Might pay in the future
Universities and colleges	26.1	13.9
Medical schools and health providers	52.5	7.9
Industry and commerce	66.4	10.1
Research institutes	29.6	17.7

Percentages within sector (n=2,511)

■ Have paid in the last year　　　■ Might pay in the future

Conference papers

	Have paid in the last year	Might pay in the future
Universities and colleges	3.8	33.7
Medical schools and health providers	4.5	45.3
Industry and commerce	18.5	50.5
Research institutes	5.5	42.1

Percentages within sector (n=2,522)

Respondents who indicated that journal articles or conference proceedings are of no interest are excluded
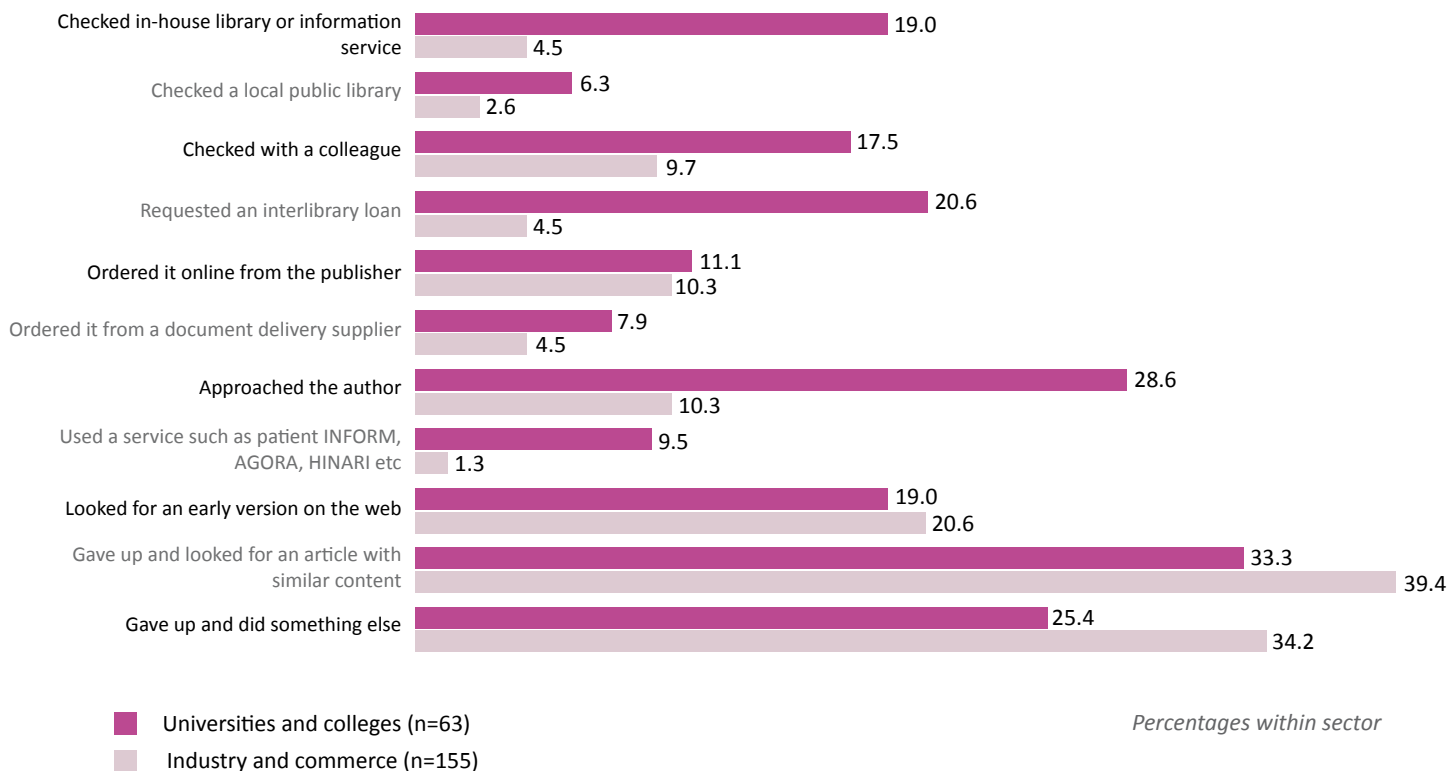
Coping strategies: journal articles

The critical incident component of the survey also explored what people did to try to acquire the full text of the last article they had difficulty accessing. Researchers in universities and colleges employ a wide range of coping strategies (Figure 26) which include checking the library for a hard copy, approaching the author, requesting an interlibrary loan, and looking for an early version in a repository or on the web. The modal responses, however, were to give up and look instead for another article with similar content, or to give up and doing something else entirely.

Researchers and knowledge workers in industry, perhaps even more pressed for time than university academics, employ much more limited coping mechanisms and are very much more likely to seek alternative content or do something else. Early versions appear to offer an important safety valve.

These findings suggest that considerable amounts of non-productive activity follows in the wake of failure at the terminal or library shelf.

Figure 26: Journal article coping mechanisms (only those who experienced a recent problem)

Checked in-house library or information service — 19.0 / 4.5
Checked a local public library — 6.3 / 2.6
Checked with a colleague — 17.5 / 9.7
Requested an interlibrary loan — 20.6 / 4.5
Ordered it online from the publisher — 11.1 / 10.3
Ordered it from a document delivery supplier — 7.9 / 4.5
Approached the author — 28.6 / 10.3
Used a service such as patient INFORM, AGORA, HINARI etc — 9.5 / 1.3
Looked for an early version on the web — 19.0 / 20.6
Gave up and looked for an article with similar content — 33.3 / 39.4
Gave up and did something else — 25.4 / 34.2

Universities and colleges (n=63)
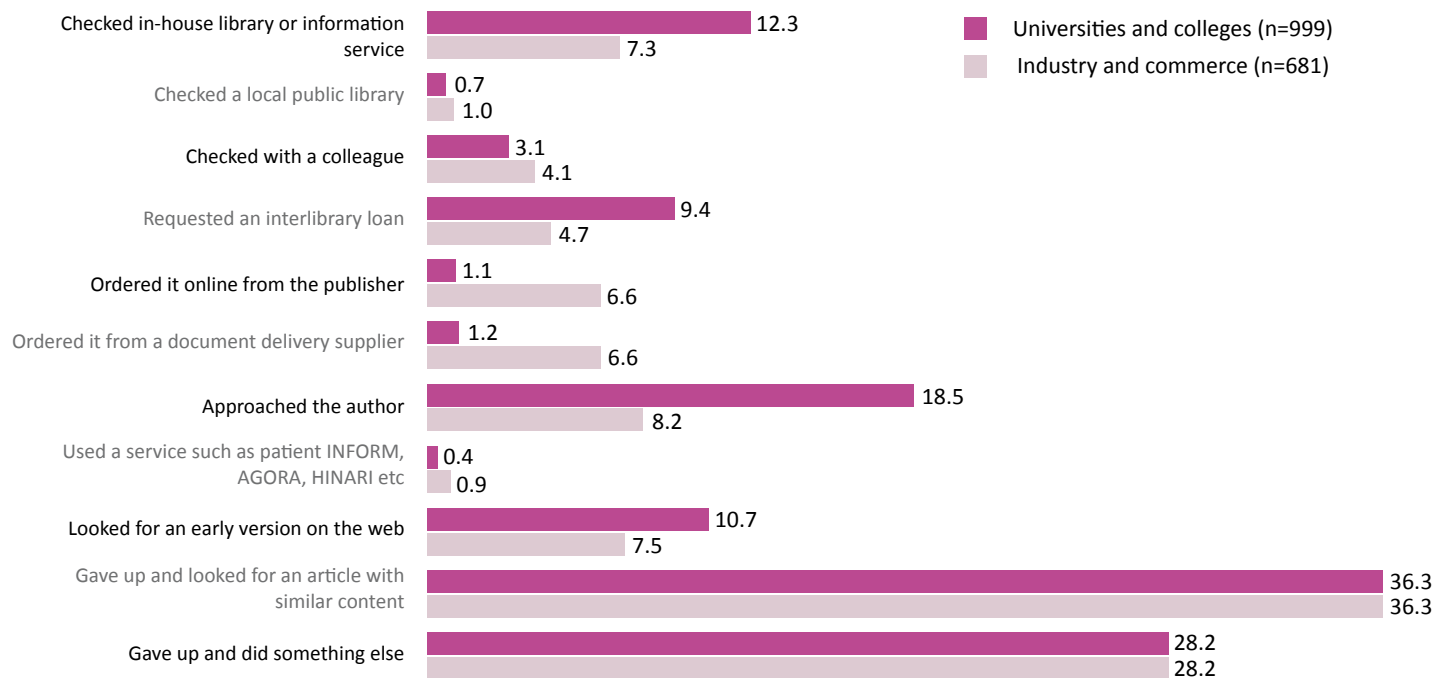Industry and commerce (n=155)

Percentages within sector

Coping strategies: conference papers

When it comes to securing the full text of conference papers, researchers use a wider set of coping strategies. This reflects a theme running throughout this report, that the conference literature is problematic in terms of both gaps and barriers to access.

The modal response for university academics is again to contact the author direct, followed (in decreasing order of frequency) by requesting an interlibrary loan, looking for an earlier version or checking the library.

The two least popular options: ordering the paper online from the publisher or from a document delivery supplier turn out to be the two most popular routes for knowledge workers in industry.

Figure 27: Conference paper coping mechanisms (only those who experienced a recent problem)

	Universities and colleges (n=999)	Industry and commerce (n=681)
Checked in-house library or information service	12.3	7.3
Checked a local public library	0.7	1.0
Checked with a colleague	3.1	4.1
Requested an interlibrary loan	9.4	4.7
Ordered it online from the publisher	1.1	6.6
Ordered it from a document delivery supplier	1.2	6.6
Approached the author	18.5	8.2
Used a service such as patient INFORM, AGORA, HINARI etc	0.4	0.9
Looked for an early version on the web	10.7	7.5
Gave up and looked for an article with similar content	36.3	36.3
Gave up and did something else	28.2	28.2

Percentages within sector

3

Conclusions and recommendations

Perceptions of access

Journal articles and conference papers are rated as 'important' by 90.4% of survey respondents (for journal articles) and 58.2% (for conference papers).

Journal articles are considered relatively easy to access compared to other information resources, especially in universities and colleges. There are patches of particularly poor access in certain UK industrial sectors, including the motor industry, utilities companies, metals and fabrication, construction, and rubber and plastics.

Researchers feel that they do not have enough journal titles available to them within their institution. This is a particular problem when they discover something that they think looks useful but which is held behind a paywall. It is possible that new discovery tools have exacerbated this problem by making it easier to identify relevant literature, but not to subsequently access this content.

Conference papers are considered less important overall, but they remain important in some disciplines such as computer science. They are much more difficult to access than journal articles.

Barriers to access

The main barrier to access for both industry and academic researchers is a paywall. The other barriers that were important reflect the circumstances of each group. For instance, lack of a hard copy in the library was much more important for academic researchers than for those in industry, while technical problems during a payment process were more important for industry researchers than academics. To an extent, this reflects the coping strategies adopted by researchers in each sector when faced with an inaccessible article (see below). Academic researchers were much more likely than those in industry to experience problems when trying to access an article at home (as opposed to their workplace), presumably reflecting different working patterns.

The main barrier to access for conference papers appeared to be that they are either unavailable online, or cannot be found online using researchers' preferred search tools.
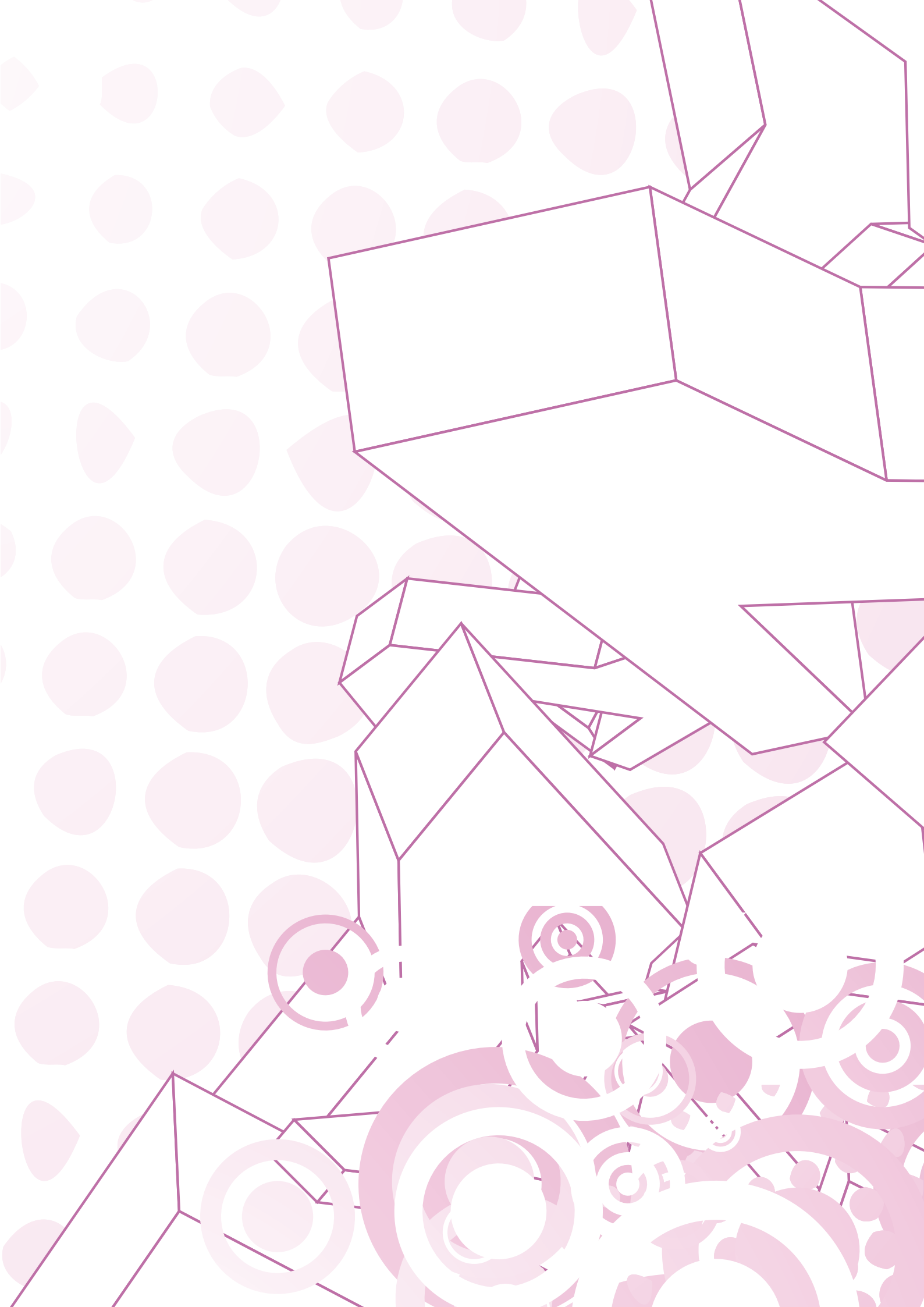
Coping strategies

Coping strategies differ in industry and academia. The most common response for both groups, when faced with an inaccessible article, is to give up and look for another article with similar content. For those in industry, the next most common solution is to give up entirely, followed by looking for an earlier version on the web, approaching the author directly or ordering from the publisher. Those in academia would approach the author, then give up completely, and then use institutional solutions such as inter-library loans and library-held hard copies, as well as online searches for earlier versions, to meet their needs.

For conference papers, the two overwhelming responses to lack of access are to look for another article with similar content, or to give up and do something else. In both academia and industry researchers would then approach the author, look for an earlier version on the web or check in-house library or information services: there was less diversity in the preferred coping strategies for this type of information.

Recommendations

This report has presented new data about the extent and nature of gaps in, and barriers to, access to research information. However, it does not add new information to the already substantial literature on how to resolve these gaps and barriers, and any recommendations for action based purely upon this data would necessarily be speculative and, thus, unhelpful.

We therefore recommend that further research should present the findings of this study to a range of stakeholders, including (but not limited to) publishers, librarians, aggregators, repository managers, funders, authors, researchers and other bodies with an interest in scholarly communications. Their expertise can be used to derive a series of recommendations for action that might address the gaps and barriers identified within this study, with a focus upon 'low-hanging fruit'. These recommendations should then be tested among a wider group of stakeholders to establish the relative ease and importance of each action, in order to determine the priorities.

Design by Catherine Gray