Unit 04: Open Access Publishing

Open access publications and initiatives

Research community has been facing unprecedented changes, largely driven by technological developments accompanied by changes in research habits and scholarly communications. Research Institutes are grappling with fund scarcity on one hand and on the other hand the research output, citation metrics and visibility are becoming more and more important in the scientific community to get grants from funding agencies. Till recent times, one of the most important metric for the researchers was publication in journals. Researchers published their work in subscription based journals and also served on several editorial and peer review committees. But many a times they did not have access to their own work and the library could not afford to subscribe the same. This was one of the drawback of the traditional scholarly publishing model. One of the major arguments against the subscription based model of publishing has been that authors have given their work for free to publishers, worked on editorial boards and peer review committees without any financial gain, but still have to pay to get access to the same journals. Such type of business model is unique in itself and unsustainable. Taking opportunity from benefits that technology offered, the scholars, scientists and researchers have responded to this by working out alternative ways of sharing their research and the Open Access Movement came into existence.

Concept, Definition and Status

The concept of open access evolved during 1991 due to the realization of the need to facilitate scholarly communication. 'Open Access' to scholarly communication is viewed as a mechanism to address escalating journal prices, and as a means of circumventing growing limited access to the increasing volume of research literature. Other reasons for a move to 'Open Access' is the conviction that publicly funded research by rights should be more accessible to the taxpaying public; digital divide

between developing and developed world should diminish, that access to research by and in the developing world should be greatly improved; and that researchers at poorly funded institutional libraries will have increased access to the research literature. Open access publication is defined and described from a variety of perspectives:

A. Suber defined Open Access as "Open access (OA) literature is digital, online, free of charge, and free of most copyright and licensing restrictions. OA helps researchers directly, both as authors and readers. It helps the institutions that fund and supervise research, from universities and laboratories to foundations and governments. It widens the distribution of research literature and lowers costs at the same time, and does so without compromising peer review, preservation, indexing, or the other virtues of conventional publishing. Above all, because OA enhances research productivity and accelerates the pace of discovery, it helps everyone who benefits from research advances

B. Willinsky emphasized the access principle, viz., commitment to scholarly work carries with it a responsibility to circulate that work as widely as possible. In the digital age, that responsibility includes exploring new publishing technologies and economic models to improve access to scholarly work. Wide circulation adds value to published work; it is a significant aspect of its claim to be knowledge. The right to know and the right to be known are inextricably mixed.

C. OA serves the interests of many groups namely: authors, readers, libraries, universities journal publishers and funding agencies.

D. Proponents of 'Open Access' claim that it eases both the `serials crisis' and `permission crisis' thereby, facilitating the free exchange of information across borders as required by the scholarly community in the current times.

E. OA is compatible with copyright, peer review, revenue (even profit), print, preservation, prestige, quality, career-advancement, indexing, and other features and supportive services associated with conventional scholarly literature. The

legal basis of OA is the consent of the copyright holder or the public domain. OA focuses on literature that authors give to the world without expectation of payment.

The conception of the Open Access initiative traces back to several decades ago, but it managed to gain momentum only after the 1990s. The increased popularity and application of open access publication can be attributed to the advent of internet. Owing to the physical and economic barriers during the print age, Open Access publishing was almost impossible, even if an author was interested in getting wider audience for his article. With the rise in inflation, print costs have significantly increased which further raised the journal subscription prices.

However, the increased applicability of internet in scholarly publishing enabled authors to upload, download, print, and distribute digital data at low/no cost. This has offered an alternative to paper publishing and significantly benefitted the Open Access initiative. In accordance to the changing trends in the development of science and technology and the increasing volume of published knowledge, there is a growing demand for rapid exchange of scientific data. This demand can be met by transforming from pay-per-view printing to Open Access publishing.

The Budapest Open Access Initiative, Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities of 2003, Bethesda Statement on Open Access Publishing of 2003, and the Salvador Declaration on Open Access of 2005 are some of the milestones that further propagated the applicability of open access initiative in scientific publishing.

The Open Access initiative is also promoted by UNESCO as a means of contributing to the progress of global science. A clear mandate has been given by UNESCO, which states that UNESCO should 'maintain, increase and diffuse knowledge, by

assuring the conservation and protection of the world's inheritance of books, works of art and monuments of history and science' (Constitution, art, 1.2 c).

All these events helped in bringing about the public realization that Open Access can augment in the advancement of global science research by bringing together researchers, universities, libraries, institutions and scientific societies onto a common platform for the exchange of knowledge.

Budapest Open Access Initiative recommended two approaches to providing open access to the research literature: (a) open access journals (known as the “gold” road), and (b) institutional or individual selfarchiving in digital repositories. (Known as the “green” road). Berlin Declaration and Bethesda Statement on Open Access Publishing6, define Open Access as one that meets the following two conditions:

1) The author(s) and right holder(s) of such contributions grant to all users a free, irrevocable, worldwide, right of access to, and a license to copy, use, distribute, transmit and display the work publicly and to make and distribute derivative works, in any digital medium for any responsible purpose, subject to proper attribution of authorship (community standards, will continue to provide the mechanism for enforcement of proper attribution and responsible use of the published work, as they do now), as well as the right to make small numbers of printed copies for their personal use.
2) A complete version of the work and all supplemental materials, including a copy of the permission as stated above, in an appropriate standard electronic format is deposited (and thus published) in at least one online repository using suitable technical standards (such as the Open Archive Definition) that is supported and maintained by an academic institution, scholarly society, government agency, or other well-established organization that seeks to enable open access, unrestricted distribution, interoperability and long-term archiving.

Open access publications: International initiatives

Internationally, a number of initiatives have been takenup for spreading the open access concept and the open publishing philosophy of scholarly communication. They can be divided into (i) the research initiatives, (ii) commercially driven initiatives, and (iii) collaborative projects. The researcher initiatives are author driven initiatives which include the e-print services such as the physics e-print archives arXiv first subject repository set up by Paul Gingsparg and the PubMedCentral [PMC], 'BioMed Central' (BMC) and the 'Public Library of Science' (PLoS) which provide open publishing facility for the biomedical researchers. 'Crossref', a publisher-linking service promoted by more than180 publishers the world over, is an example of commercially driven initiative. Collaborative endeavours include the ICAAP (International Consortium for the Advancement of Academic Publications), SPARC, High-Wire Press and many more such efforts. SPARC (Scholarly Publishing and Academic Resources Coalition) is an alliance of universities, research libraries, and organizations built as a constructive response to market dysfunctions in the scholarly communication system. Many commercial publishers offer open access to their publications after an embargo period or offer a choice to the researchers to make their article Open Access after paying necessary author processing charges (APC). Directory of Open access Journals (DOAJ) is a directory that indexes and maintains a list of high quality peer reviewed Open Access journals. SHERPA RoMEO is a searchable database of publisher's policies regarding the self- archiving of journal articles on the web and in Open Access repositories. SHERPA/ JULIET maintains a list of funding organizations' open access policies from around the world. SHERPA ROMEO/SHERPA Juliet help researchers clearly understand the publisher/funder OA policy. The Open Citation Project - Reference Linking and Citation Analysis for Open Archives has assembled a bibliography of studies on the effect of open access and downloads ('hits') on citation impact. Several funding agencies across the globe like NIH, Welcome trustetcare now mandating Open

access from the research resulting from grants given by them. In 2012, the European Commission encouraged all European Union (EU) member-states to put public-funded research results in the public domain in order to strengthen science and the knowledge-based economy. In the US, the National Institutes of Health (NIH) asks all its grantees to provide OA to the results of NIH-funded research within 12 months of publication. The Wellcome Trust requires OA to Wellcome-funded research within six months of publication, and the Research Councils UK also have a similar policy. Major research institutions in Australia, China, France, Germany, Greece, Hungary, Italy, Norway, Portugal, Switzerland, the UK, and the US have committed themselves to provide OA to their research output. 12 INDIAN SCENARIO Research in India is plagued due to two problems which relate to access and visibility. Both these problems can be solved by widespread adoption of Open access.

Open access publications: Indian initiatives

While the Open Access (OA) movement has been a topic of major debate and interest around the world, in India it is seen as an unprecedented opportunity to provide equality of access to essential research information and raising awareness of national research. Right since the initiation of the OA movement, the academic and the scientific community in India has been striving to promote unrestricted access to scholarly literature through Open Access. Several local, national, regional as well as international initiatives, have been taken up in different parts of the country, adopting open access software, configuring and commissioning of open archive harvester services, providing open course wares to the academic world, imparting training programmes on epublishing of journals as well as on institutional repositories etc. Some of the creditable activities such as the OA journals of the Indian Academy of Sciences (IAS), eprints@iisc, Librarian's Digital Library at DRTC, OpenMED and the IndMed services of NIC New Delhi, NISCAIR, IISC, efforts of MedKnow publications, the e-journal initiatives and archives at INSA, IIT Delhi,

Raman Research Institute, NIT Rourkela, Vidyanidhi. UGC has mandated the open access to theses and dissertations since 2009 and provided access through 'ShodhGanga'. These efforts deserve special mention. But the progress in the adoption of open access is slow.

SHERPA/RoMEO

Sherpa Romeo is an online resource that aggregates and presents publisher and journal open access policies from around the world. Every registered publisher or journal held in Romeo is carefully reviewed and analysed by our specialist team who provide summaries of self-archiving permissions and conditions of rights given to authors on a journal-by-journal basis where possible.

The policy information provided through this service primarily aims to serve the academic research community. Since the service launched over 15 years ago, publisher policies and the open access sector have changed a lot. Open access policy can be complex and varies according to geographical location, the institution, and the various routes to open access – all of which affects how and where you can publish your research.

How to use it

The database uses a colour-coding scheme to classify publishers according to their self-archiving policy. This shows authors whether the journal allows pre-print or post-print archiving in their copyright transfer agreements.

RoMEO Colour	**Archiving policy**
Green	Can archive pre-print *and* post-print or publisher's version/PDF
Blue	Can archive post-print (ie final draft post-refereeing) or publisher's version/PDF
Yellow	Can archive pre-print (ie pre-refereeing)
White	Archiving not formally supported

Software tool to identify publications developed by SPPU

Predatory journals find most of their prey in developing countries, and in particular, among emerging economies where research output is rapidly growing. Savitribai Phule

Pune University (SPPU) of India plays a key part in India's anti-predatory journals efforts.

The University Grants Commission, the statutory organization established by the Government of India for the coordination, determination, and maintenance of standards of teaching, examination, and research in university education. To deal with the problem of predatory journals, UGC created a white list of quality journals as a proactive step. However, due to some flaws during the list's creation and implementation, it was polluted with poor-quality journals and faced severe criticism. A large number of poor-quality journals were included in the UGC approved list, which opened the floodgates for desperate authors. Nevertheless, the UGC India was the only one of the ten most common funders who provided guidance about journal selection on its website .

An effort to prepare 'Guidelines for Research Publications' was undertaken at SPPU. In 2017, a Center for Publication Ethics was established and created a group of like-minded academicians and developed a robust protocol to analyze the UGC list, and found that over 88% of journals recommended by universities for inclusion in the UGC list were of poor quality. As of now the UGC approved list of Journals stands canceled and is replaced with UGC-CARE Reference List of Quality Journals (UGC-CARE List).

While Indian science has shown a marked growth in high-quality scientific publications, it has also been reported that the percentage of research articles published in predatory journals is high in India. Despite good intentions, the regulatory provisions of UGC seemed to have triggered a sudden spurt in predatory journals, giving way to 'publish or perish' culture. An ineffective monitoring of research quality coupled with desperation to publish for regulatory compliance have led to massive growth of predatory publications in India. To understand this rapid penetration of predators in the Indian academic community, it is necessary to understand the vastness, diversity and

complexity of the Indian higher education system. India still follows an affiliating system of British origin, where, by and large, over 40,000 colleges deal with undergraduate instruction, while over 900 universities are entrusted with post-graduate education and research. As per All India Survey on Higher Education (AISHE) 2018-2019 data, about 1.285 million teachers are in the system to take care of 36.6 million students, of which 4.14 million are in Masters programs and 161,000 in doctoral programs. In 2018, about 34,400 students were awarded PhD degrees. The Government of India awards nearly 10,000 research fellowships every year. According to Scopus data, 147,537 articles were published by Indian authors in 2018. The majority of them are from over one hundred institutes of national importance and a large number of national laboratories managed by different research councils. A typical Indian university is meant for degree education, whereas national institutes and laboratories are mandated for research. Such bifurcation seems to be a major reason for poor research culture in most Indian universities. The situation is now rapidly changing for good through systematic efforts by the government to promote inter-institutional collaborations and transdisciplinary research, and by creating more awareness regarding academic integrity and publication ethics through initiatives like UGC-CARE.

The Indian academic community, including teaching faculty, scientists, and research scholars need to ensure that the journals and/or conferences they choose follow standard ethical policies. They also need dependable and credible guidance. Rather than having UGC undertake this task in isolation, it is decided to form a consortium involving reputed research councils and national academies of science, engineering, medicine, social sciences, and humanities. Strong support to pursue this idea was received from the Ministry of Human Resource Development, Government of India, Chairman of UGC and the key academic leadership. There was a consensus that unethical practices leading to ‘pay and publish trash’ culture in India needs to be thwarted immediately. It was agreed that any attempt to compromise academic integrity

should be challenged, questioned, shamed, and de-recognized at all levels. To bring these ideas to reality, in early 2019, UGC decided to set up a 'Consortium for Academic Research and Ethics' (CARE) to promote academic integrity and publication ethics, and to improve the quality of research in Indian universities. The UGC appointed an empowered committee to steer the activities of CARE. The CARE invited over 30 statutory councils and government bodies across disciplines as members of the consortium to identify, continuously monitor, and maintain a reference list of quality journals across disciplines and identified reputable universities from four regions to facilitate submission and preliminary screening of journals for consideration of inclusion in the UGC-CARE list.

Proposals for any journal to be considered for inclusion in UGC-CARE list need to be submitted by editors, publishers or individuals following a process as illustrated in the flowchart. Faculty members and Internal Quality Assurance Cell (IQAC) from academic institutions need to prima facie ensure that the journal they are submitting is of good quality and follows ethical practices. All such proposals and journals submitted by CARE members are analyzed as per a validated protocol by the UGC Cell for Journal Analysis established at Center for Publication Ethics, SPPU.

In June of 2019, the first edition of UGC-CARE Reference List of Quality Journals was released accompanied by a Public Notice on Academic Integrity. In addition to the Reference List of Quality Journals, the CARE website provides useful resources including relevant publications, audio visual materials, videos, weblinks etc. The CARE website also provides FAQs, and information on feedback and grievance redressal mechanisms. The UGC has warned that the Indian academic community must avoid publication in predatory/dubious journals and participation in predatory conferences. They must not be associated as Editors/Advisors or in any other capacity with journals, publishers, or conferences involved in fraudulent, dubious, and deceptive practices. The Vice Chancellors, selection committees, research supervisors/guides, and other experts

involved in academic evaluation and assessment should ensure that their decisions are primarily based on quality of research work and not merely on number of publications. Henceforth, any publications in predatory journals or presentations in predatory conferences shall not be considered for academic credit for selection, confirmation, promotion, performance appraisal, award of scholarship or academic degrees, or credits in any form. The CARE website and UGC-CARE List have created more awareness and helped the cause of promoting academic integrity and ethical publishing. We hope that these steps by the UGC should discourage authors from choosing predatory publishing as an easy way to earn academic benefits and should curtail the article flow to predators. The UGC-CARE project has been widely praised among academic fraternity in India and abroad. Reputed journals like *Nature, The Journal of Alternative and Complementary Medicine,* and *Current Science* have recognized this effort.

What are the main challenges the UGC-CARE Project is facing so far? Are there any plans to address these challenges?

As a preventative step, it is important to create more awareness to ensure that the academic fraternity (especially young researchers) stay away from predatory journals. This is very challenging and needs continuous efforts, especially with the vast expanse of the Indian higher education system. We organize seminars and conferences to discuss the importance of academic integrity. The UGC has also proposed a course on academic integrity as a mandatory requirement of pre-PhD course work. This will greatly help young scholars to know the possible adverse consequences of publishing in predatory journals and compromised academic integrity on their careers. Keeping the UGC-CARE list up to date is one of the main challenges.

We have observed that a few journals have changed their behavior after being listed by UGC-CARE. They start behaving in predatory manner and try to lure authors by

aggressively advertising their listing in UGC-CARE. As a corrective step, we have created a feedback and grievance registering mechanism where anyone can report unethical practices of journals or publishers. We re-evaluate and scrutinize such journals. If they are found to be to be involved in any unethical practice, the journal is removed from UGC-CARE list. We update the UGC-CARE list every quarter so that new good journals can be added and poor-quality journals, if any, are removed. To decide the quality of journals in the humanities, arts and Indian languages remains a major challenge. We are trying to modify our existing protocol to suit the special requirements of such disciplines. Capturing citations of journals in Indian languages is also a problem. A team from the UGC Cell for Journal Analysis and the Center for Publication Ethics at SPPU are continuously working to address these challenges

Tools for Finding a Journal for Publication

Selecting the right journal to publish scholarly research papers is as vital as the research work for authors. It is because only the right journal serves the ultimate purpose of it reaching the target audience. Only then the hard work of research will get its due credit and used by other researchers. With technology development, the selection of journals is made easy with different tools.

Like a search in a database for references, the journal search tools require the author to provide keyword or abstract. It will enable it to identify potential journal titles. Publishers and organizations create many of these tools. The journal metrics of these tools enable the authors to identify the highly ranked reputed journals to submit their research papers. If there are no journal metrics in the journal finding tool, use Scopus or Journal Citation Reports databases to identify the highly ranked journals.

How to use a journal finding tool to identify the merits of journals?

Enter the research paper's keywords, title, and abstract in the search engine of the journal finding tool. Based on the manuscript, it will suggest three curated journals appropriate to it.

Before submitting the research paper to one of the suggestions of the journal finder, it is pertinent to check

- Review the Overview and Aims & Scope of the journal to check its match with the research papers
- Read the ”Instructions to authors’ to recheck whether the research paper is written as per the instructions to avoid rejection
- Read articles in the journal to find out whether it matches the scope of the research paper and also for the editors’ preferences
- Check the ”Author Compliance Tool’ of the journal finder to meet institution or funder requirements

Few best journal finding tools include

Elsevier Journal Finder:

Elsevier Journal Finder using smart search technology and field of research specific vocabularies to match the input with the Elsevier journals to publish them. For authors, getting their research paper published can be a challenge and it is even more challenging when their paper is rejected by a journal because it is out of scope. It can often add months to the publication process slowing career progress. We know that nearly one third of the visitors to Elsevier’s Authors’ Home are trying to choose a journal for their paper. For Editors, dealing with out of scope papers can substantially add to their workload. In a bid to help authors The Elsevier launched *Journal Finder* tool, accessible from www.elsevier.com/authors which

The tool is designed to:

- Help less experienced authors select suitable journals for their papers
- Enable authors working across multidisciplinary fields to identify possible journals
- Highlight journals that offer open access options, and provide information on publication speeds and impact factors

How does the tool work?

Authors enter their paper title, abstract or keywords and the tool creates a list of Elsevier journals that match the topic of their article. They can then order the results based on their priorities, such as highest impact factor or shortest editorial time. The selection contains links to each journal's homepage and Elsevier Editorial Submission (EES) page.

JANE- Journal/Author Name Estimator (PubMed):

The Journal/Author Name Estimator (JANE) is a free online bibliographic journal selection tool. Journal selection tools, also known as journal matching or journal comparison tools, are popular resources that help authors determine the most appropriate in scope journal to publish their manuscripts. JANE is one of the earliest journal selection tools, debuting in 2007. The resource is web-based and allows users to input keywords, abstract text, or author names and view related articles based on user-supplied terms. At the time of this writing, no formal mobile app or browser extension has been developed to utilize the resource. There is an application programming interface (API) freely available in beta version that is available to users who want to integrate JANE into their own applications.

JANE interfaces directly with the PubMed operating from the PubMed/ MEDLINE data set, meaning both MEDLINE-indexed journals as well as articles deposited into PubMed Central can be retrieved when searching the resource. JANE's indexing criteria include journals from PubMed/MEDLINE that contain abstracts published within the past ten years. JANE does not search categories that are not viewed as original research. For example, editorials, newspaper articles, comments, conferences, directories, retractions, errata, and so on are omitted. During a search, JANE uses the Lucene open source search engine to search for the most similar fifty articles based on user input and assigns similarity and confidence scores, which determine the search result order.

Besides journal comparison functions, other uses of the JANE resource include convenient identification of related articles that authors can read and/or cite in their

manuscripts, as well as aggregation of authors who could potentially serve on journal review boards. For example, publishers who need to fill editorial or peer-reviewer positions can search JANE, using the author search to identify relevant subject specialists.

Springer Journal Suggester:

It uses semantic technology to identify the right journal for the research paper from 2,600 Springer publications and BioMed Central journals. Enter your abstract, description of your research, or a sample text and the Springer Journal Selector provides a list of relevant journals. You can refine the results based on requirements for Impact Factor or publishing model, including an option to match to journals that are fully open access or have open access options. You will find them listed by subject area and then alphabetically. Detailed instructions for authors, information about the aims and scope and the types of papers that are published in a specific journal can be found on that journal's homepage.

IEEE Publication Recommender:

It searches 170 + periodicals and 1500 + conferences to compare critical points such as impact factor and submission to publication time for the best match for the scholarly research paper.

EndNote Manuscript Matcher:

EndNote Manuscript Matcher requires registration through Web of Science or EndNote register login using the ECU email to give the right journal match for the keywords, title and abstract. Along with the appropriate journal, it also gives the journal impact factor and also related articles.

Unit 05: Publication Misconduct

Rightly, the public expects scientists, researchers, clinicians and journal editors to be honest and trustworthy. Failure to live up to these ideals can result in science being corrupted, patients harmed and financial sponsors deceived. While the majority of research is conducted properly and reported honestly, a depressing series of scandals

shows that there is a dishonest minority. In the worst cases, data have been invented or manipulated to reach fraudulent conclusions. But there are also lesser or more subtle degrees of scientific and publication misconduct.

Misconduct by editors, publishers and peer-reviewers

Authors are not the only ones who may be guilty of misconduct. Editors, publishers and peer reviewers also have responsibilities: for example, peer reviewers have a duty of confidentiality pre-publication; they have a duty not to allow professional or personal jealousy or rivalry to influence or determine the advice they offer editors; and they have a duty not to cause undue delay to the processing of a submitted paper Editors have a prime duty to their readers to maintain the integrity of the scientific record. This must take precedence over their other duties, for example, making sure their journal is readable and profitable (or, at least not a financial burden for the society, academic institution, governmental body or publisher to whom they are responsible). Therefore, they should follow good practice guidelines, such as those published by the International Committee of Medical Journal Editors (ICMJE) or the Council of Science Editors.

Important functions include correcting significant inaccuracies or misleading reports by publishing corrections; ensuring that proper ethical standards have been followed in the conduct of research or clinical practice forming part of submitted or published papers and paying strict regard to patient confidentiality. Editors can access advise from the Committee on Publication Ethics (COPE) by way of flowcharts devised from the organisation's experience over 8 years of handling allegations of misconduct. If a satisfactory explanation cannot be supplied by authors, then editors should normally report any reasonable concerns about research misconduct to their institution(s) or those who funded their study so that they can investigate and publish a notice of concern where the initial case looks strong, followed by retraction when there is a finding of fraud or a major error which, if left to stand, would significantly distort the scientific record. Editors and their publishers must make sure that their journal is open and transparent in its instructions to authors (advice to contributors), especially with regard to describing the peer review process as well as its definitions for authorship and requirements for declaration of competing interests. They should have a well defined appeals procedure and an independently supervised complaints process. Publishers, themselves cannot escape responsibility, if only because they may be required to investigate and adjudicate on complaints against editors or editorial boards. Some publishers have accepted that responsibility. For example Wiley Blackwell provides a set of ethical guidelines which it expects its journal editors to follow. Additionally, publishers should not attempt to interfere with editorial freedom unless there are

exceptional circumstances whereby an editorial board or other responsible body produces cogent evidence that an editor has misused that freedom.

Types of misconduct

Submission of fraudulent data

The extent of fraudulent research data is not known, although many experienced editors believe that undiscovered fraud is much more common than is supposed. It is rarely easy to detect. An editor or associate editor processing a paper may be suspicious that the results are «too good to be true» but without specific expertise in the topic, he or she cannot be certain. Statistical analysis of a paper will sometimes demonstrate that data must have been manipulated. Likewise, reviewers sometimes express concerns about the honesty of a paper.

Incomplete or improperly processed data

The reliability of the scientific record can be disturbed by conduct far short of fraud. For example, it is commonplace that inconvenient data are sometimes excluded from a study or that the most advantageous statistical analysis is performed, especially if the results can be used, for example, to increase prescribing rates or enhance the chance of further research funding.

Use of plagiarism software

Many people think of plagiarism as copying another's work or borrowing someone else's original ideas. But terms like "copying" and "borrowing" can disguise the seriousness of the offense:

According to the Merriam-Webster online dictionary, to "plagiarize" means:

- to steal and pass off (the ideas or words of another) as one's own
- to use (another's production) without crediting the source
- to commit literary theft
- to present as new and original an idea or product derived from an existing source

In other words, plagiarism is an act of fraud. It involves both stealing someone else's work and lying about it afterward

All of the following are considered plagiarism:

- turning in someone else's work as your own
- copying words or ideas from someone else without giving credit

- failing to put a quotation in quotation marks
- giving incorrect information about the source of a quotation
- changing words but copying the sentence structure of a source without giving credit
- copying so many words or ideas from a source that it makes up the majority of your work, whether you give credit or not (see our section on "fair use" rules)

Most cases of plagiarism can be avoided, however, by citing sources. Simply acknowledging that certain material has been borrowed and providing your audience with the information necessary to find that source is usually enough to prevent plagiarism.

Plagiarism detection tools

In the past two decade, several plagiarism detection tools have been developed. The major ones are Turnitin and Urkundu.

Urkund

The tool was created by the Swedish company in 1999 and can be used only by educational institutions, not by individuals, e.g.: teachers, professors, or students. The checker is accessible through the Internet, email or integration with LMS's of choice by universities or colleges. It checks across online sources and digital databases.

How it works:

If the email version is chosen, the Urkund support manager assigns special email addresses to educators, which then can be used to invite students and receive checked student works.

All the assignments students send to their teachers' emails are automatically scanned against the Internet sources and digital databases. As a result, the educator gets already scanned writings (in other words, plagiarism reports) with text duplications highlighted on the left and original text from external sources on the right. It's up to the educator to choose whether to share the report with his or her students or not.

Pros:

- Students have free access to the checker
- The tool recognizes characters replaced by cheaters
- It's possible to compare checked and original texts side by side
- Check results are accurate

- It supports a great number of file formats

Cons:

- Users aren't allowed to choose types of checks other than the Internet & databases scan
- Simultaneous document checks aren't supported
- Text matches on original sources aren't highlighted
- Students can't see check results unless their educators decide to email them

Turnitin

Turnitin is used by a lot of universities and colleges across the globe. It was created only for corporate use and has a rather extended functionality similar to that of LMS (having a calendar for assignments, separate discussion tabs, the "Join" tab for inviting students to join a particular class, and more).

The tool requires both instructors and students to register before getting access to their accounts. The detector scans writings through the Internet, university or teacher's databases and publisher's databases. Before assigning tasks to students, instructors can choose what type of check to use. The checker is available online or via integration with LMSes.

How it works:

As it was mentioned, each user needs to register first. When access is received, the process of assigning tasks usually goes like this: the instructor creates a class and adds an assignment, then invites the student to join the class. After the student is enrolled, he/she completes the assignment and submits it to the instructor. At this stage, the educator evaluates the checked writing and shares a report with his/her student.

Regardless of this simple description of educator-student interaction, the checker's functionality is rather complex, that's why instructors go through a series of training before they are ready to start using the system.

Pros:

- The tool can check several files at once
- Scanning documents in the background mode is possible
- Multiple file formats are supported
- It's possible to share reports via email
- It can be used as a kind of LMS

Cons:

- Free trial isn't available
- The tool can't be customized
- No pricing mentioned
- The system considers repeated submissions as plagiarized
- Citations detection is not always accurate
- Customer support works slowly

Data bases

Whether you are writing a thesis, dissertation, or research paper it is a key task to survey prior literature and research findings. Especially, you will be looking for trusted resources, most likely peer reviewed research articles. Academic research databases make it easy to locate the literature you are looking for.

Citation databases

A citation database is a form of bibliographic index which provides a record of citations between publications, enabling a user to see which publications have cited which other publications. Such a database will show which authors have cited a publication and how many times an author has been cited.

Citation databases have been developed as a means of evaluating publications, allowing a user to establish citation counts and to check, for example, which publications and authors are the most cited.

Citation analysis and bibliometric indicators have been made possible by such databases. However, citation count in itself should not be taken as a guarantee of quality and there can be many reasons for a particular citation (e.g. negative citations, self-citation).

Citation databases tend to focus on journal articles but may cover other material such as books, conference papers, dissertations or reports. No citation database covers all publications. Note also that some disciplines (e.g. the sciences) are more heavily covered than others (e.g. the arts). Citation databases do not tend to provide a user with full-text access to the publications which have been indexed.

1. Scopus

Scopus is one of the two big commercial, bibliographic databases that cover scholarly literature from almost any discipline. Beside searching for research articles, Scopus also provides academic journal rankings, author profiles, and an h-index calculator.

- Coverage: approx. 71 million items
- References: 1.4 billion
- Discipline: Multidisciplinary
- Access options: Limited free preview, full access by institutional subscription only
- Provider: Elsevier

2. Web of Science

Web of Science also known as Web of Knowledge is the second big bibliographic database. Usually, academic institutions provide either access to Web of Science or Scopus on their campus network for free.

- Coverage: approx. 100 million items
- References: 1.4 billion
- Discipline: Multidisciplinary
- Access options: institutional subscription only
- Provider: Clarivate (formerly Thomson Reuters)

Research Metrics

Research metrics are the fundamental tools used across the publishing industry to measure performance, both at journal- and author-level.

For a long time, the only tool for assessing journal performance was the Impact Factor - more on that in a moment. Now there are a range of different research metrics available. This “basket of metrics” is growing every day, from the traditional Impact Factor to Altmetrics, *h*-index, and beyond.

Citation-based metrics

What is the Impact Factor?

The Impact Factor is probably the most well-known metric for assessing journal performance. Designed to help librarians with collection management in the 1960s, it has since become a common proxy for journal quality.

The Impact Factor is a simple research metric: it's the average number of citations received by articles in a journal within a two-year window.

The Web of Science Journal Citation Reports (JCR) publishes the official results annually, based on this calculation:

Number of citations received in one year to content published in *Journal X* during the two previous years, divided by the total number of articles and reviews published in *Journal X* within the previous two years.

Number of citations received in 2017 to content published in *Journal X* during 2015 and 2016, divided by the total number of articles and reviews published in *Journal X* in 2015 and 2016.

How can I get an Impact Factor for my journal?

Only journals selected to feature in the Science Citation Index Expanded (SCIE) and Social Sciences Citation Index (SSCI) receive an official Impact Factor.

To be eligible for coverage in these Web of Science indices, journals must meet a wide range of criteria. You can find out more about the journal selection process on the Clarivate website.

For many journals, the first step to receiving an Impact Factor is to feature in the Emerging Sources Citation Index (ESCI). For more information on the ESCI, read our introduction here.

What are the disadvantages of the Impact Factor?

- **The Impact Factor is an arithmetic mean and doesn't adjust for the distribution of citations**.

This means that one highly-cited article can have a major positive effect on the Impact Factor, skewing the result for the two years. Most journals have a highly-skewed citation distribution, with a handful of highly-cited articles and many low- or zero-cited articles.

- **The JCR doesn't distinguish between citations made to articles, reviews, or editorials.**

So that the Impact Factor doesn't penalize journals that publish rarely-cited content like book reviews, editorials, or news items, these content types are **not** counted in the denominator of the calculation (the total number of publications within the two-year period). However, citations to this kind of content **are** still counted.

This creates two main problems. Firstly, the classification of content is not subjective, so content such as extended abstracts or author commentaries fall into an unpredictable gray area. Secondly, if such articles are cited, they increase the Impact Factor without any offset in the denominator of the equation.

- **The Impact Factor only considers the number of citations, not the nature or quality.**

An article may be highly cited for many reasons, both positive and negative. A high Impact Factor only shows that the research in a given journal is being cited. It doesn't indicate the context or the quality of the publication citing the research.

- **You can't compare Impact Factors like-for-like across different subject areas.**

Different subject areas have different citation patterns, which reflects in their Impact Factors. Research in subject areas with typically higher Impact Factors (cell biology or general medicine, for example) is not better or worse than research in subject areas with typically lower Impact Factors (such as mathematics or history).

The difference in Impact Factor is simply a reflection of differing citation patterns, database coverage, and dominance of journals between the disciplines. Some subjects generally have longer reference lists and publish more articles, so there's a larger pool of citations.

- **Impact Factors can show significant variation year-on-year, especially in smaller journals.**

Because Impact Factors are average values, they vary year-on-year due to random fluctuations. This change is related to the journal size (the number of articles published per year): the smaller the journal, the larger the expected fluctuation.

What is CiteScore?

CiteScore is the ratio of citations to research published. It's currently available for journals and book series which are indexed in Scopus. CiteScore considers all content published in a journal, not just articles and reviews.

CiteScore was produced by Scopus in December 2016 and you can easily replicate it via the Scopus database. In addition to CiteScore, Scopus also publish additional rankings, such as the CiteScore percentile based on subject categories, and a monthly CiteScore tracker.

The CiteScore calculation is:

Number of all citations recorded in Scopus in one year to content published in *Journal* X in the last three years, divided by the total number of items published in *Journal X* in the previous three years

Journals that publish a large amount of front matter (such as editorials or peer commentaries) will perform worse by CiteScore than by Impact Factor because this front matter is rarely cited.

SNIP - Source Normalized Impact per Paper

SNIP is a journal-level metric which attempts to correct subject-specific characteristics, simplifying cross-discipline comparisons between journals. It measures citations received against citations expected for the subject field, using Scopus data. SNIP is published twice a year and looks at a three-year period.

The SNIP calculation is:

Journal citation count per paper, divided by citation potential in the field.

SNIP normalizes its sources to allow for cross-disciplinary comparison. In practice, this means that a citation from a publication with a long reference list has a lower value.

SNIP only considers citations to specific content types (articles, reviews, and conference papers), and does not count citations from publications that Scopus classifies as "non-citing sources". These include trade journals, and many Arts & Humanities titles.

IPP - Impact Per Publication: Also known as RIP (raw impact per publication), the IPP is used to calculate SNIP. IPP is a number of current-year citations to papers from the previous 3 years, divided by the total number of papers in those 3 previous years.

SJR - Scimago Journal Rank

The SJR aims to capture the effect of subject field, quality, and reputation of a journal on citations. It calculates the prestige of a journal by considering the value of the sources that cite it, rather than counting all citations equally.

Each citation received by a journal is assigned a weight based on the SJR of the citing journal. So, a citation from a journal with a high SJR value is worth more than a citation from a journal with a low SJR value.

The SJR calculation is:

Average number of (weighted) citations in a given year to Journal X, divided by the number of articles published in Journal X in the previous three years.

What is the *h*-index?

The *h*-index is an author-level research metric, first introduced by Hirsch in 2005. The *h*-index attempts to measure the productivity of a researcher and the citation impact of their publications.

The basic *h*-index calculation is:

Number of articles published which have received the same number of citations.

For example, if you've published at least 10 papers that have each been cited 10 times or more, you will have a *h*-index of 10.

What are the advantages of the *h*-index?

- **Results aren't skewed**

The main advantage of the *h*-index is that it isn't skewed upwards by a small number of highly-cited papers. It also isn't skewed downwards by a long tail of poorly-cited work.

The *h*-index rewards researchers whose work is consistently well cited. That said, a handful of well-placed citations can have a major effect.

What are the disadvantages of the *h*-index?

- **Results can be inconsistent**

Although the basic calculation of the *h*-index is clearly defined, it can still be calculated using different databases or time-frames, giving different results. Normally, the larger the database, the higher the *h*-index calculated from it. Therefore, a *h*-

index taken from Google Scholar will nearly always be higher than one from Web of Science, Scopus, or PubMed. (It's worth noting here that as Google Scholar is an uncurated dataset, it may contain duplicate records of the same article.)

- **Results can be skewed by self-citations**

Although some self-citation is legitimate, authors can cite their own work to improve their *h*-index.

- **Results aren't comparable across disciplines**

The *h*-index varies widely by subject, so a mediocre *h*-index in the life sciences will still be higher than a very good *h*-index in the social sciences. We can't benchmark *h*-indices because they are rarely calculated consistently for large populations of researchers using the same method.

- **Results can't be compared between researchers**

The *h*-index of a researcher with a long publication history including review articles cannot be fairly compared with a post-doctoral researcher in the same field, nor with a senior researcher from another field. Researchers who have published several review articles will normally have much higher citation counts than other researchers.

What are altmetrics?

Alternative metrics (or "altmetrics") help you to measure the impact of a journal by looking at the social activity around it. They use quantitative and qualitative data alongside traditional citation- and usage-based metrics to provide an insight into the attention, influence and impact of academic research.

The most common method of reporting on altmetrics is the Altmetric Attention Score. This tool tracks a wide range of online sources to capture the conversations happening around academic research.

How is the Altmetric Attention Score calculated?

Altmetric monitors each online mention of a piece of research and weights the mentions based on volume, sources, and authors. A mention in an international newspaper contributes to a higher score than a tweet about the research, for example.

What are the advantages of the Altmetric Attention Score?

- **Receive instant, trackable feedback**

Altmetric starts tracking online mentions of academic research from the moment it's published. That means there's no need to wait for citations to come in to get feedback on a piece of research.

- **Get a holistic view of attention, impact and influence**

The data Altmetric gathers provides a more all-encompassing, nuanced view of the attention, impact, and influence of a piece of research than traditional citation-based metrics. Digging deeper into the Altmetric Attention Score can reveal not only the nature and volume of online mentions, but also who's talking about the research, where in the world these conversations are happening, and which online platforms they're using.

"**G-index** is introduced as an **improvement of the h-index** of Hirsch to measure the global citation performance of a set of articles. If this set is ranked in decreasing order of the number of citations that they received, the **g-index is the (unique) largest number such that the top g articles received (together) at least g^2 citations**" (Egghe, L., "Theory and practice of the G-index". Scientometrics, vol. 69, no. 1, (2006), pp. 131-152).

Keep in mind:

- The g-index gives **more weight to highly-cited articles** (whereas h-index is insensitive to it)
- Someone's g-index will always be **equal to or greater than the h-index**
- You can consult your g-index through **Publish or Perish** (free downloadable software for academic citations analysis)

The **i10-index** was created by Google Scholar (Google Scholar Blog, 2011):

i10-Index = the number of publications with at least 10 citations

Keep in mind:

- It is **only accessible from Google Scholar Citations** (author profile) (you need a Google account to manage your profile)
- The i10-index is easy to calculate