

# Overview of STM Research Publishing and Licensing Practices

by Carlo Scollo Lavizzari

## Chapter 6



**LICENSING PRACTICES IN A GLOBAL DIGITAL MARKET**  
Published by the International Publishers Association  
**OCTOBER 2020**



# ABOUT THE AUTHOR

## **Carlo Scollo Lavizzari**

Carlo Scollo Lavizzari, a lawyer qualified in Switzerland, South Africa, England & Wales, specializes in intellectual property protection with more than 18 years of work experience with law firms in Africa, Europe, Switzerland and the United States of America. He is a strategist and problem solver in negotiations, conflict resolution and policy, participating actively in global advocacy for norm setting and in litigation, on all five continents.

Depending on client need, Carlo brings together proven specialists from a range of areas all across the world. Carlo is fluent in four languages: German, English, French and Italian.

### **COPYRIGHT**

© International Publishers Association, 2020

*All data, copy and images are subject to copyright and may not be reproduced, transmitted or made available without permission from IPA.*

*Designed by EDEP IMAGEN Y PRODUCCIÓN GRÁFICA S.L.*

### **ACKNOWLEDGEMENTS**

*Edited by José Borghino, Emma House and Anne Leer.*

*IPA is grateful for the valuable contributions from all those involved. A special thank you to the expert authors, IPA members, IFRR0, CMOs, the members of the advisory board and, to Kopinor for their expertise and financial support which has made this possible.*

### **MEMBERS OF THE ADVISORY BOARD**

**Kristenn Einarsson**, *Managing Director, Norwegian Publishers Association (NPA)*

**Caroline Morgan**, *Chief Executive, International Federation of Reproduction Rights Organisations (IFRR0)*

**Hugo Setzer**, *President, International Publishers Association (IPA)*

**Arne Vestbø**, *Secretary General, Norwegian Non-Fiction Writers and Translators (NFFO)*

# OVERVIEW OF STM RESEARCH PUBLISHING AND LICENSING PRACTICES

## CONTENTS

	EXECUTIVE SUMMARY .....	4
<b>1</b>	INTRODUCTION: WHAT IS STM RESEARCH PUBLISHING? .....	5
<b>2</b>	THE GLOBAL STM MARKET .....	6
<b>3</b>	BUSINESS MODELS AND LICENSING PRINCIPLES .....	8
	3.1. Journals .....	8
	3.2. Books .....	8
	3.3. Collective Licensing .....	8
	3.4. Access to scholarly and academic published material for higher educational and research institutions .....	8
<b>4</b>	PRINCIPLES AND PRACTICES IN LICENSING RESEARCH AND HIGHER EDUCATION INSTITUTIONS .....	9
<b>5</b>	ACCESS TO PUBLISHED MATERIAL OUTSIDE FORMAL LICENCE AGREEMENTS.....	10
	5.1. Case study examples from around the world .....	11
	5.1.1. Co-publishing Quality Content from South to North: .....	11
	5.1.2. Consortium and National Licensing: Examples Europe and South Africa and Brazil .....	11
	5.1.3. Collective licensing collaborations North and South: USA twinned with Ghana and Argentina.....	12
	5.2. Licensing for widest possible access: Individual initiatives .....	13
<b>6</b>	OPEN ACCESS AND TRENDS FOR BOOKS AND JOURNALS .....	14
	6.1. Offsetting deals .....	15
<b>7</b>	INITIATIVES FOR ACCESS TO RESEARCH INFORMATION AND PUBLISHING OPPORTUNITIES .....	15
<b>8</b>	STRATEGIC ISSUES BEYOND OPEN ACCESS AND GLOBAL ACCESS.....	16
	8.1. The changing nature of scholarly communications: Social Scientific Networks .....	16
	8.2. Access to research data and finding a reward model for making scientific data discoverable, curating data, verifying and assuring provenance, ethical collection of data .....	17
	8.3. Text and Data Mining .....	18
<b>9</b>	ARTIFICIAL INTELLIGENCE.....	19
<b>10</b>	SUMMARY AND CONCLUSION.....	20

# EXECUTIVE SUMMARY

The STM publishing industry (defined broadly to include academic, scholarly and professional publishing and related communication services) is a very substantial global business. It is estimated to be worth about US\$25.7 billion annually, with the books and journals segment by itself being worth US\$13-\$14 billion. The greatest share of global revenues comes from the USA, but China has overtaken the USA as the largest producer of research papers. It has been estimated that there are over 150 million different articles across all Web of Science databases and that there are between 10 and 16 million authors globally. According to CrossRef, there are more than 11,000 journal publishers responsible for around 60,000 individual journal titles.

Enabling access is the primary mission for STM publishers and licensing is the primary business model for journal publishing, given that the format has been predominantly digital for two decades. For book publishing, new business models are emerging space as digital formats become more popular.

Collective licensing is part of the business ecosystem of both journals and books. There are myriad types of licensing agreements and distribution models, including co-publishing, consortium and national licensing, document delivery and interlibrary loan.

A survey of licensing practice conducted in 2015 noted both that licensing allows cross-border access and that licences rarely refer to specific national exceptions to copyright. Other benefits of licensing include; long-term access for future generations to historic collections; authorised users permitted to download works for non-commercial use and structured/unstructured sharing; and wide availability of content for teaching and learning. Licensing builds co-operation between publishers and librarians; both are curators of scientific and more broadly academic content.

Publishers also provide a number of free uses to users, sometimes through joint statements

through the International STM Association and through initiatives such as Research4Life which provides free or low-cost access to journal articles to institutions in developing markets.

Open Access (OA) as a business model for journals started in earnest in 2002 with a broader expansion in 2010. There are now over 15,000 OA journals. Journals operate under the “Gold Open Access” model (charging Article Processing Charges – APCs – to authors or their institutions) and the “Hybrid” model, which combines subscriptions with APCs. Under Gold OA, journals are freely available to the end user; under the Hybrid model, publishers offer “offsetting agreements” which either reduce the APC costs or the subscription pricing. Some publishers are now also offering OA books (funded by Book Publishing Charges).

A significant development in the STM world has been the rise in Social Collaborative Networks (SCNs), which offer their platform services for new forms of academic communication and collaboration. Much of the content posted on these networks is without permission, which raises a number of concerns.

Other notable trends in this publishing sector are concerned with big data and increased computer power, fueling developments in Text and Data Mining coupled with the use of artificial intelligence to interpret the data. Discoverability and citability of data sets are also driving new initiatives. Underpinning these developments are the legal conditions which are being addressed through licensing and in some cases, legislation.

STM publishers are engaged in a wide range of access initiatives across the world ensuring access is enabled as broadly as possible for multiple uses. They rely on a copyright framework that provides for licensing and the development of future business models ■



# 1 INTRODUCTION: WHAT IS STM RESEARCH PUBLISHING?

---

'STM Research Publishing' is neither only about 'STM' (as in Scientific, Technical and Medical), nor solely about 'Research Publishing' (as in publishing a stand-alone book or journal article that can be placed on a shelf or displayed on a website). Instead, STM Research Publishing is understood to include any form of academic, scholarly and professional publishing and communication. This is well beyond the publication of individual titles of authors, and comprises a host of services around publications, including the underlying data, abstracts, referencing and cross-referencing with permanent, yet dynamic links, creating effectively a web of interconnected resources. Increasingly STM Research Publishing also offers decision tools, text and data mining applications, software, discovery and search algorithms and rankings, as well as videos explaining the research to interested audiences.

Moreover, 'STM' is not limited to the natural and life sciences, but also includes the Arts, Humanities and Social Sciences (sometimes referred to as AHSS)—a field that is of increasing importance with the digital developments of the Fourth Industrial Revolution, the so-called 'Industry 4.0' where new 'soft skills' are needed. Some commentators also include law, accounting and tax and indeed any professional publication in the broader concept of 'STM', while others draw the line at humanitarian law and transparent and ethical accounting and taxation rules.

Rather than a specific discipline of science, or a method of communication, what sets STM apart from its close cousin, the broader concept of tertiary educational publishing, is that STM is predominantly concerned with primary publications (and more recently, digital product solutions and services) as opposed to secondary publications. Primary publications are those that publish new scientific discoveries and scientific claims or present new evidence corroborating or falsifying state-of-the-art scientific hypotheses. In other words, STM publishing is not so much concerned with synthesising, summarising, re-publishing acquired state-of-the-art knowledge about a subject, but is concerned with new findings, which have never been published before.

# 2 THE GLOBAL STM MARKET

The value of the STM books and journals market is estimated at between US\$13 and US\$14 billion annually, while the broader STM information publishing market (including journals, books, technical information and standards, databases and tools, medical communications and some related areas) has been estimated to be worth about US\$25.7 billion annually, with a large proportion attributable to life sciences and medical publishing products and services.<sup>1</sup>

Geographically, about 41% of total global STM revenues are generated from the USA, 27% from Europe/Middle East, 26% from Asia/Pacific and 6% from the rest of the world. The industry employs directly an estimated 110,000 people globally, and in addition some 20-30,000 full-time employees are indirectly supported by the STM industry globally (not counting the supply chain).<sup>2</sup>

China has overtaken the US to become the largest producer of research papers globally, with a share of about 19%. The US accounts for 18% of global articles, while India has also seen rapid growth in recent years, and now produces 5% of global outputs, ahead of Germany, the UK and Japan, each on 4%.

## TOTAL GLOBAL STM REVENUES IN US\$ (BILLIONS)



As of mid-2018 a good estimate is that there were 33,100 scholarly peer-reviewed English-language journals and a further 9400 non-English-language journals, collectively publishing well over 3 million articles a year.<sup>3</sup> There are also many more journals that are not peer reviewed.

**CrossRef** is the STM industry body that applies Digital Object Identifiers (DOI) to publications and now also to data. Currently, the CrossRef database includes over 100 million Digital Object Identifiers, of which 73 million refer to journal articles from a total of almost 60,000 journals.

<sup>1</sup> It is important to realise that the function of journal articles and books form a core part of scientific research itself. Journals have been called the "minutes of science" and their function is not limited to disseminating peer-reviewed high quality information in certain fields, they also provide a mechanism for the registration of the author's precedence or priority claim to a scientific result; they also provide a fixed archival version for future reference. Books in the field of STM are either monographs, found frequently in AHSS fields, multi-author treatises or collections of standard reference works. The STM book market (worth between US\$3 to US\$4 billion annually, is evolving rapidly in a transition to digital publishing. As of 2016, about 1/3 of the book market consists of ebooks and this segment continues to grow rapidly.

<sup>2</sup> Rob Johnson, Anthony Wilkinson, Michael Mabe (Co-authors), *The STM Report: An overview of scientific and scholarly publishing* (2018), 50th anniversary edition (5th edition), Oxford/The Hague, pages 22 and 46.

<sup>3</sup> CrossRef Annual Report 2017-2018, page 12.













More broadly, Google Scholar was estimated to index between 100 and 160 million documents in 2014. **The Web of Science ‘Core Collection’ included about 70 million article records as of June 2018, out of a total of 150 million items across all Web of Science databases.**<sup>4</sup>

There are between 10 and 16 million authors world-wide. Most STM research publications have more than one author with the average being 4.5 authors per article. Most books are multi-author and perhaps even multi-volume works comprising multiple authors contributing to different chapters or sections. About half of articles have authors from at least two countries.

There are considerably more readers than authors of STM research publications and services. The degree of overlap between authors and readers will vary considerably between disciplines: in a narrow pure science field like theoretical physics there may be close to 100% overlap, but in a practitioner field such as nursing or medicine the readers will be many times more than the authors.

**Publishers: according to CrossRef, there are an estimated 11,000+ publishers responsible for some 60,000 journals.** The largest ten publishers (with number of journals) are: SpringerNature (>3000); Elsevier (2500); Taylor & Francis (2500); Wiley (1700); Sage (>1000); Wolters Kluwer (incl. MedKnow) (900); Oxford University Press (440); Hindawi (>400); Cambridge University Press (390); and Emerald (>300).

#### THE TEN LARGEST STM PUBLISHERS

 3000,0	 2500,0	 1700,0	 1000,0	
	 2500,0	 900,0	 440,0	 400,0
			 390,0	 300,0

The financial resources to publish may come from readers, authors, research institutions (such as university libraries), corporations with a significant R&D function, research funding bodies (government and non-government, philanthropic), or crowdsourcing. For ebooks specifically, there may be eLending models that can provide a source of funding. All funders have direct interest in STM research publishing and the underlying research output and its dissemination. Thus, the wider stakeholders in this publishing sector, alongside publishers, include libraries of universities, research funding bodies and the public at large.

<sup>4</sup> [https://clarivate.com/webofsciencegroup/solutions/web-of-science--core-collection/Crossref Annual Report 2017-2018](https://clarivate.com/webofsciencegroup/solutions/web-of-science--core-collection/Crossref%20Annual%20Report%202017-2018), page 2, available at: <https://www.crossref.org/pdfs/annual-report-2017-18.pdf>.



# 3 BUSINESS MODELS AND LICENSING PRINCIPLES

---

## 3.1. Journals

---

STM publishers are at the forefront of electronic and online publishing and there are various ways in which publishers provide content to their readers. Licensing is the legal vehicle, be it for purchase by individual readers, libraries, research or educational institutions, corporations or by other entities. Also, where publishing is otherwise funded, publishers license its use on Open Access terms which is free to the intended readers and the wider public.

Licensing partners on the user side are sometimes entire consortia, or access is provided to an entire region or country through a national or regional licence. Typically, journals are not licensed individually, but as part of a bundle of journals. STM publishers also offer agreements that provide for article rental (for a day, week, or month), for some form of document delivery or inter-library loans (either directly or through licensing agents or collective management organisations).

## 3.2. Books

---

The rapidly growing ebook market is seeing new business models emerging with access made possible through a subscription agreement for an individual title or for an ebook collection. Other access models include the rental of full ebooks and/or on a chapter by chapter basis. Some publishers license ebooks for online access only, while others offer to ship a paper copy on request for a small fee.

A significant difference between books and journals is that academics are far more likely to purchase the books themselves to remain well ahead of supply via libraries, whereas articles are mainly obtained from library e-collections. However, for publishers the library market remains crucial for journals, and ebooks account for a growing proportion of book budgets as libraries move from print to electronic collections.

## 3.3. Collective Licensing

---

Individual licensing, to an individual library, university or a consortium of institutions, remains the primary way of providing legal access, free of charge at the point of use for the institution's researchers. STM publishers also widely rely on the collective licensing network and systems offered by IFRRO (International Federation for Reproduction Rights) members, the collective management organisations. Moreover, some IFRRO members such as CLA (UK) and CCC (USA) offer bespoke services in the fields of document delivery of individual documents or course packs or chapter licensing services to students. Content is also made available in limited proportions and to less developed parts of the world in a convenient way. As such collective licensing is part of the lifeblood of STM and research publishing alongside individual licensing.

## 3.4. Access to scholarly and academic published material for higher educational and research institutions

---

**Digital is the main medium for science publishing, communication and sharing.** Print remains important in social science and humanities and for preservation, but in all fields the primary tools for access are the electronic versions.



Enabling access to scholarly and academic published material is at the core of the STM publishers' mission and licensing is the primary business model. However, this cannot be done without the copyright framework and the freedom of contract.

Competition between publishers, as well as individual negotiations between publishers and their customers will invariably result in there being different contract terms between different contracting parties. Many licence terms are responses to demands for more rights in the licensing of electronic content. In this regard, subscription licences are no different to most other cross-border supply contracts.

## 4 PRINCIPLES AND PRACTICES IN LICENSING RESEARCH AND HIGHER EDUCATION INSTITUTIONS

In 2015, the author of this chapter conducted a qualitative survey<sup>5</sup> of licensing agreements and the findings are summarised below.

**The chief role of online licensing is to grant access to a defined user group, not necessarily in a defined territory.** For academic institutions, most licences are not inherently cross-border only, for example, for travelling professors or faculty who may enjoy remote access through a Virtual Private Network (VPN). For corporate and multinational licences, access and user rights are almost invariably cross-border and use takes place among employees and contractors affiliated either remotely or as part of the corporate intranet.

**Licensing allows cross-border uses implicitly** where those are needed on institutional networks via the institutional secure internal network or VPN, for the main uses, such as searching, browsing, printing, downloading of individual articles for authorised users. For greater clarity, some uses are explicitly dealt with and it is made clear in what way the uses are permitted, for example inter-library loans.

**Licences rarely refer to specific national exceptions.** However, this does not mean that the licences do not deal with key customer needs for academic and scientific content. In this sense licences need to be understood as 'exceptions plus', rather than as mechanisms to erode statutorily identified, socially beneficial uses or liability shields and safe harbours, namely exceptions.

Most of the licences reviewed contained lists of permitted uses that clarify for the user what practices are allowed under the agreement, regardless of possible interpretations or expectations regarding education or research copyright exceptions. In some cases, publishers and their licensees spell out how they are to apply exceptions to the specific case of their agreement, for example by setting numerical limits of copies allowed to be made or supplied and by agreeing on reporting and record keeping obligations (number of copies), or the agreements refer to applicable local guidelines that industry or custom, may have locally created (for example **CONTU** guidelines in the USA). In this sense, licensing agreements closely define the parties' understanding of permitted uses under general categories of exceptions.

The survey found no evidence in licensing agreements surveyed that those would declare national exceptions inapplicable by express agreement. Claims that STM publishers attempt to do this are unfounded.

The supply of copies of articles covered by a licence to other libraries or to end users other than persons affiliated with the licensed institution, licensing terms can deal with:

<sup>5</sup> Accessibility of STM published material to libraries and their patrons: Report on STM Qualitative Survey on Publishers' Supplying by way of Licensing Journals, Books, Databases and other Published Material to Higher Education and Research Institutions, 5 June 2015.

(1) **Inter-library Loan (ILL)**; based on customary practice or local law. The vast majority of all the licences surveyed expressly permit ILL in some form, while some licensing agreements are more specific about ILL than others. The distinguishing feature of ILL is that it is undertaken for non-commercial purposes and in a fairly unsystematic or occasional manner (reacting to ad-hoc requests) and that the overall quantities supplied do not interfere with the normal exploitation of publishers, namely subscription purchase procedures and individual article supply licensing to other libraries and consortia for their bespoke access needs;

(2) **Document Supply**; (also known as ‘document delivery’) is often described as a systematic delivery of copies of journal articles that goes well beyond non-profit occasional inter-library loan and includes the supply of documents to end-users on a commercial as well as a systematic non-commercial basis. The right of a licensee library to carry out Document Supply is usually negotiated separately in licensing agreements and is often permissible based on special transactional or ‘token’ agreements.

**Licensing enables long-term access for future generations to historic collections and back-files as archival copies.** Licensing conditions sometimes offer very long-term or even perpetual access guarantees. Licensing agreements also allow back-up copies, some more explicitly than others, and provide for preservation either by the library or by joint publisher-library initiatives such as **Portico**<sup>6</sup> and **CLOCKSS**<sup>7</sup> and others in case of catastrophic events or bankruptcies. Thus, STM publishers understand very well that they are working not just in the present, building and curating the ‘record of science,’ but also for posterity, for the general public, and for future generations of scientists.

**Licensing enables ‘authorised users’**, as defined (that is, students, researchers, employees, contractors, and unaffiliated ‘walk-in users’ of libraries) to download (for their own non-commercial needs) individual works (articles) or extracts (chapters of books) permanently. Some licences also explicitly allow for unstructured sharing of content with researchers and research groups, provided the use is non-commercial and unsystematic and it does not create substitutes for additional subscriptions or licensed content. Some agreements allow even structured sharing on or via social networks.

**Licensing enables teaching**; content can be used by teachers and learners in the academic setting, and the electronic content can be used to produce electronic or print course packs.

**Licensing secures electronic platforms and content and builds co-operation between publishers and librarians who are both custodians of scientific content.** Rather than creating liability for libraries, many licensing contracts contain clauses whereby the licensee is held inculcable by the publisher in respect of claims of infringement for authorised uses. For unauthorised uses, libraries are not held liable for any prejudice or damage in the licensing agreements surveyed but rather are encouraged to co-operate with the publisher to find and prevent recurring forms of unauthorised access and infringements.

## 5 ACCESS TO PUBLISHED MATERIAL OUTSIDE FORMAL LICENCE AGREEMENTS

Apart from their participation in access partnerships such as Research4Life, covered elsewhere in this chapter, STM publishers, particularly those that make their published content available by

<sup>6</sup> <https://www.portico.org/>

<sup>7</sup> <https://clockss.org/>

way of subscription licensing agreements, have signed up to a number of statements organised through the International STM Association allowing certain uses free to the user:

- a. **STM Text and Data Mining Declaration**<sup>8</sup>, allowing access to subscribed content for text and data mining, free of a copyright fee if it is for a non-commercial purpose.
- b. **Safe Harbour Provisions for the Use of Orphan Works for Scientific, Technical and Medical Literature**<sup>9</sup>.
- c. **Safe Harbour Provisions for the Digitization and Making Available of “Out of Commerce Works”** forming part of the Scientific, Technical and Medical Literature<sup>10</sup>.
- d. **Statement on Document Delivery to Qualifying Institutions under the Research4Life Program in United Nations-Designated Least Developed Countries**<sup>11</sup>, allowing document delivery to qualifying institutions in Least Developed Countries free of permission and without a copyright fee.

## 5.1. Case study examples from around the world

### 5.1.1. Co-publishing Quality Content from South to North:

**Taylor & Francis**, incorporating the **Routledge** imprint, publishes a suite of top academic research journals with **NISC** of Grahamstown, South Africa.

Benefiting from state-of-the-art international and local publishing services, these journals are available to purchase as online and print, or in online-only editions, with differential pricing available for Africa. Nine of the ten journals are ranked in the Thomson-Reuters Journal Citation Reports and all are fully accredited by the South African Department of Higher Education and Training.

Many titles deal with the distinctive African environmental heritage, including marine science, forestry, and ornithology. **Ostrich**, founded in 1930, competes for the place of oldest South African journal. Other NISC journals are dedicated to music and language in African contexts and key health challenges such as HIV/AIDS and mental well-being.

The NISC-Taylor & Francis collaboration optimises the best of the local and global.<sup>12</sup>

### 5.1.2. Consortium and National Licensing: Examples Europe and South Africa and Brazil

**Europe:** The ‘**EU Big Deal Licensing Report: A mapping of Major Scientific Publishing Contracts in Europe**’ presents data from 28 negotiating consortia in 2016 and 2017. The survey focused on the functions and working process of consortia, as well as on the conditions of contracts for big deals concerning scientific periodicals, databases and ebooks. The results of the survey show that consortia broadly represent the interests of relevant stakeholders from the university and library sectors and are largely driven by researchers’ needs. The updated report conducted in 2017-2018 gathers data from 31 consortia covering an unprecedented 167 contracts.<sup>13</sup>

8 [http://www.stmassoc.org/2013\\_11\\_11\\_Text\\_and\\_Data\\_Mining\\_Declaration.pdf](http://www.stmassoc.org/2013_11_11_Text_and_Data_Mining_Declaration.pdf)

9 [http://www.stmassoc.org/2014\\_03\\_04\\_Safe\\_Harbor\\_Provisions\\_for\\_the\\_Use\\_of\\_Orphan\\_Works.pdf](http://www.stmassoc.org/2014_03_04_Safe_Harbor_Provisions_for_the_Use_of_Orphan_Works.pdf);

10 [https://www.stm-assoc.org/2014\\_03\\_04\\_Safe\\_Harbor\\_Provisions\\_for\\_the\\_Use\\_of\\_Out\\_of\\_Commerce\\_Works.pdf](https://www.stm-assoc.org/2014_03_04_Safe_Harbor_Provisions_for_the_Use_of_Out_of_Commerce_Works.pdf)

11 [https://www.stm-assoc.org/2013\\_10\\_07\\_STM\\_Statement\\_on\\_Document\\_Delivery\\_to\\_least\\_developed\\_countries.pdf](https://www.stm-assoc.org/2013_10_07_STM_Statement_on_Document_Delivery_to_least_developed_countries.pdf)

12 <http://niscjournals.com/>

13 <https://eua.eu/resources/publications/829:2019-big-deals-survey-report.html>.

**South Africa: SANLIC** the South African National Library and Information Consortium negotiates on behalf of Institutions of Higher Learning in South Africa, and on request also for other institutions. Its 2019 model licences<sup>14</sup> and its list of current members<sup>15</sup> are both available for download.

**Brazil:** The **CAPEL Portal de Periódicos** is the Brazilian national electronic library consortium. It is a virtual library that aggregates high-quality content, provided through publishers and international scientific associations. The program is maintained by CAPES, a public foundation attached to the Ministry of Education, whose mission includes the consolidation of the post-graduate system in Brazil. Its role also includes access to, and communication of scientific production. More than 3 million faculty members, researchers, graduate and undergraduate students and technical personnel from more than 400 institutions, have free access to the full text of more than 32,000 leading journals and the premier multidisciplinary and subject databases covering all areas of academic activity.

### 5.1.3. Collective licensing collaborations North and South: USA twinned with Ghana and Argentina

**USA: Copyright Clearance Center** has, over many years, supported several RROs in their developmental work. The support has been organised under CCC's ongoing International Advancement Program and has been deployed for the benefit of selected RROs in Africa, the Caribbean, Latin America and Asia. Each program is bespoke, designed to meet the specific needs of each RRO, and to complement other developmental support offered by organisations such as IFRRO and NORCODE.

**Ghana:** The Ghanaian RRO, **CopyGhana**, identified two priorities as it sought to launch its first educational licensing program. First, it sought to include in its annual copyright licence the large repertoire of rights from American publishers and authors represented by CCC. US content is used intensively by teachers, students, and librarians in Ghana's universities and polytechnics, so the inclusion of the US repertoire was an essential ingredient for the success of the CopyGhana licence. CCC deployed a type of bilateral agreement specially prepared for such circumstances. Under the terms of this agreement, CopyGhana was entitled to include CCC's vast repertoire of US works in its annual academic licence and to retain for a fixed period the royalties it generated that would otherwise have been payable to CCC's rightsholders. The agreement stipulated a list of milestones that CopyGhana agreed to meet in order to renew the agreement beyond the initial term. Second, CopyGhana's management asked for a bespoke training course to support its efforts to launch and sell its annual licence. Two senior CopyGhana staff travelled to CCC's offices in the United States for a period of intensive training, organised, delivered and funded entirely by CCC. A close, continuing relationship between CCC and CopyGhana was maintained over several years. This has contributed to CopyGhana's great success in launching and selling its annual copyright licences to Ghana's universities and polytechnics and led to the first payment of royalties to US rightsholders in 2018.

**Argentina:** CCC's support for the RRO in Argentina, **CADRA**, was tailored to the needs of an established RRO seeking to diversify into new markets and to modernise their communications. CCC helped fund a market study that sought to quantify the prospects for business licences in Argentina and provided staff to train CADRA's sales personnel in Buenos Aires. CCC designed and produced promotional brochures for several different markets and supported the production of several podcasts and a short promotional video for CADRA.

<sup>14</sup> <https://sanlic.org.za/for-publishers/>

<sup>15</sup> <https://sanlic.org.za/about/#currentmembers>.

## 5.2. Licensing for widest possible access: Individual initiatives

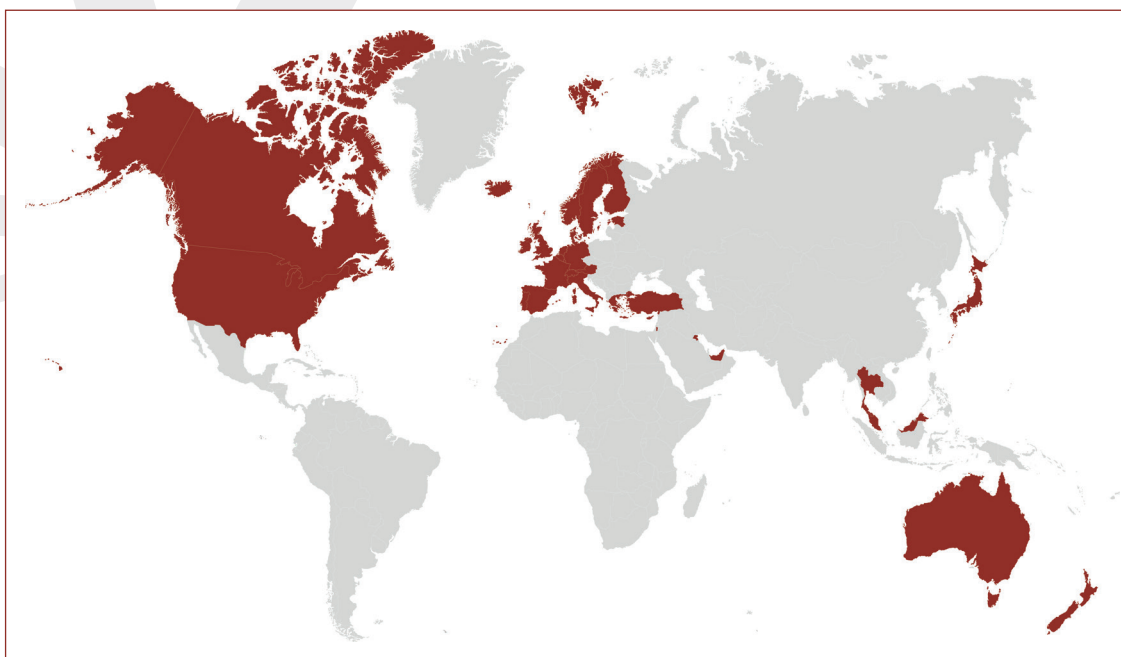
Two examples illustrate publishers' commitment to providing digital access:

### SpringerNature's ebooks and MyCopy

**SpringerNature** offers one of the most complete collections of peer-reviewed eBook subject collections spanning Science, Technology and Medicine (STM) and Humanities & Social Sciences (HSS) available today. Institutions, organisations, and businesses worldwide trust SpringerNature ebooks to help excel research and learning for their library users. For every level of learning, education, and research, SpringerNature's ebook collections offer a rich mix of book types such as textbooks, proceedings, monographs, reference works, handbooks, and more.

**MyCopy** allows library patrons to order their own personal print copy of a SpringerNature ebook for US\$24.99. The MyCopy service continues to expand and is currently available in the following countries:

#### *MYCOPY SERVICE PROVISION AROUND THE WORLD*



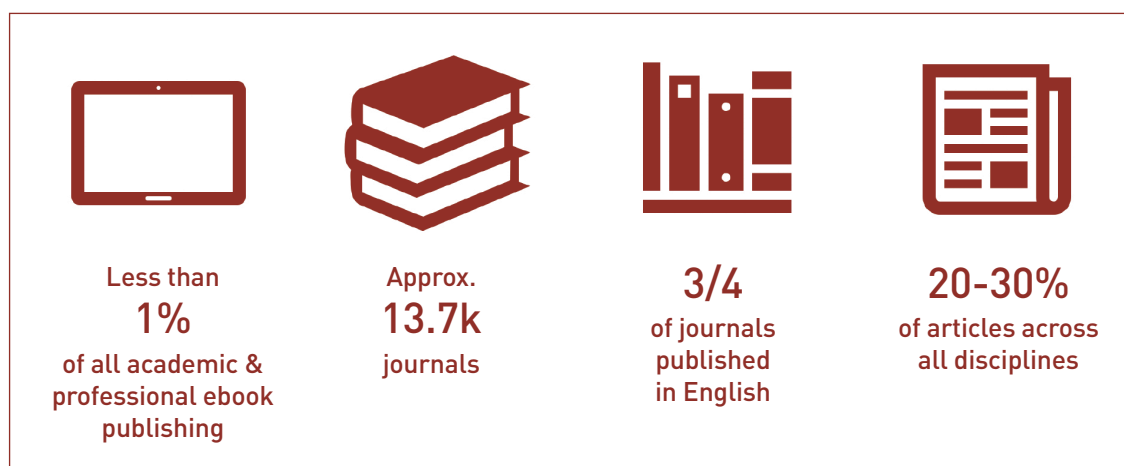
### Subito: Licensing and Document Delivery Internationally

**Subito** is a German library consortium of 30 libraries in Germany, Austria and Switzerland. Through a comprehensive collection of licensing agreements with both individual publishers and also the German CMO, **VG Wort**, as well as the CCC subsidiary **RightsDirect**, Subito is able to legally supply individual articles to libraries for use by patrons and faculty members of receiving libraries (Subito Library Service) and also to individuals (Subito Document Delivery).

# 6 OPEN ACCESS AND TRENDS FOR BOOKS AND JOURNALS

OA started in earnest as a business model in 2002. From about 2010, a second phase of OA has seen a broader expansion of Open Access with most traditional scholarly publishers adopting (to a greater or lesser extent) Open Access publishing practices, and new Open Access journals becoming well-established within their disciplines.

## OPEN ACCESS AND TRENDS FOR BOOKS AND JOURNALS



As of September 2020, the number of Open Access journals was over 15,000 (of which three quarters were published in English) according to the Directory of Open Access Journals (DOAJ)<sup>16</sup>. 70% of journals in the DOAJ do not charge author-side fees and their sustainability may therefore not be assured. Among the current accepted Open Access sustainable funding models is Gold Open Access as well as the hybrid model. Both essentially rely on Article Publication Charges (APCs). In 2018, Open Access articles as a proportion of all articles across all disciplines ranged between 20% and 30%. When Green OA, which still essentially relies on the subscription model is excluded, that proportion fell to about 15-18%, but it is anticipated that journal articles available through Open Access will continue to grow as a share of all articles for the foreseeable future.<sup>17</sup>

By contrast, despite strong year-on-year growth, the OA book market is still less than 1% of all scholarly and professional e-book publishing. According to some estimates there were only around 10,000 e-book titles available in 2016, with humanities and social sciences (HSS) accounting for almost three quarters. Among the more successful OA book publishers are **The University of California Press** and the **University College London Press (UCL Press)**. Among commercial publishers, **Ubiquity Press** has an important role in supporting university and society-based publishing and driving growth in the sector. Ubiquity also works outside English-speaking countries where growth in OA publishing among not-for-profits is slower.

For commercial book publishers the dominant OA business model is charging Book Publishing Charges (BPC). A typical BPC is currently around £10,000. **Palgrave**, which is the humanities and social sciences imprint of SpringerNature charges BPC ranging from £11,000 to £17,000. **Taylor & Francis/Routledge** also has OA book options. These charges and the current lack of

<sup>16</sup> <https://doaj.org/>

<sup>17</sup> <https://doaj.org/oainfo>



funding at this level for book authors is hampering the spread of open access in the academic book market. An alternative solution, pooling library resources, was promoted by **Knowledge Unlatched**. French platform *OpenEdition* operates on a 'freemium' model, offering a catalogue of over 6000 ebooks in the humanities and social sciences. Books are made freely available online, but libraries can choose to pay for premium services such as downloadable file formats.<sup>18</sup>

## 6.1. Offsetting deals

Over the years publishers have globally discounted the subscription rates of journals which also benefit from APC revenues, however, uneven take-up of these 'hybrid' journals to date has resulted in some countries and institutions experiencing increased costs. Meanwhile the corresponding savings on global subscriptions are widely distributed and may be obscured by price changes arising from inflationary pressures, increasing article volumes and a range of other factors. Thus, in response, several publishers have entered into local offset agreements designed to reduce the overall cost faced by research organisations or consortia. Under an offset agreement, OA publication costs are offset by lower subscription costs. There are different approaches to achieving this.<sup>19</sup> Some offset agreements reduce the cost of APCs and some reduce the amount an institution pays for a subscription in proportion to the amount it pays for APCs. Some publishers offer credits against future APCs when subscriptions are taken out; others offer credits against future subscription payments when APCs are paid; a third approach bundles subscriptions with future APCs for modest additional payments.

# 7 INITIATIVES FOR ACCESS TO RESEARCH INFORMATION AND PUBLISHING OPPORTUNITIES

**Research4Life**<sup>20</sup> is the central element in STM's Responsibility programme. The original drive for setting up Research4Life came from a need to empower authors of less developed countries to write manuscripts that meet the standards of global scientific practices. Without the ability to cite and research prior work, it is hard to find the correct shoulders of giants to stand on. In other words, without offering access to state-of-the-art research, it is almost impossible to craft a manuscript that would be accepted in a peer-reviewed journal. Research4Life has gone a long way to addressing this imbalance. **It is a partnership of United Nations agencies, leading universities and publishers for the following programs HINARI, AGORA, OARE, ARDI and GOALI (since 2019) that make journal articles available for free or at very low cost to institutions in developing countries.** The partnership's original goal was to help attain six of the UN's eight Millennium Development Goals by 2015, reducing the scientific knowledge gap between industrialised countries and the developing world. The programmes collectively provide some 8,500 institutions in 118 developing countries with free or low-cost access to some 20,000 journals, 69,000 books and 120 other information resources from 180 publishers.

While the progress towards Open Access in well-funded, developed economies is gathering pace, there are concerns that a business model based upon payment for publication rather than for consumption will reinforce existing inequities in the global research community. Thus, the original mission of Research4Life empowering authors and researchers of the less-developed markets to

<sup>18</sup> <https://resource-cms.springernature.com/springer-cms/rest/v1/content/16216770/data/v1>, page10.

<sup>19</sup> See for instance SpringerNature : <https://scholarlycommunications.jiscinvolve.org/wp/2016/10/24/offsetting-models-update-on-the-springer-compact-deal/> ; or Wiley : <https://openaccess.mpg.d46e/2336450/deal-contract-with-wiley-signed>.

<sup>20</sup> <http://www.research4life.org>



contribute to world science could inadvertently become compromised. Moreover, poorly funded disciplines may struggle to include the costs of dissemination in their budgets, while researchers in the less-developed markets risk being excluded completely from adding their work to the international corpus of scientific output. A rapid shift to Open Access without considering these risks could, unintentionally, replace a barrier to access with a barrier to publishing opportunity.

Offering Article Publication Charge (APC) waivers is not a long-term sustainable solution because they simply spread the costs of publication across fewer contributors, add to the unpredictability of the model and may actually dis-incentivise publishers from accepting articles from authors who are unable to pay. Indeed, any model which places the cost burden on the individual author is unlikely to be accepted by the market in any part of the world, while poorly funded institutions in the less-developed markets would find it virtually impossible to create a special fund to cover the payment of APCs for their staff. More work is needed in this area as Research and Innovation Publishing evolves into the 21st century and Research4Life offers a discussion platform to find appropriate solutions.

## Other schemes:

1. **HighWire Press** offers free access for developing countries to a list of journals, based simply on software that recognises from where the user is accessing the site. Some publishers offer similar schemes independently, for example the Royal Society of Chemistry and, the National Academies Press;
2. **INASP<sup>21</sup>'s SRKS scheme** ended in 2018 and its replacement programme, SERKS (strong and equitable research and knowledge systems), will embrace their **AuthorAID** and local **Journals Online** (JOLS) activity, but will significantly downsize the licensing and access component which was the cornerstone of the previous **PERII** and **SRKS** five year programmes;
3. **EIFL (Electronic Information for Libraries)** partners with libraries and library consortia to build capacity, advocate for access to knowledge, encourage knowledge sharing and initiate pilot schemes for innovative library services.

## 8 STRATEGIC ISSUES BEYOND OPEN ACCESS AND GLOBAL ACCESS

### 8.1. The changing nature of scholarly communications: Social Scientific Networks

One of the most significant developments in the STM landscape is the rapid rise of **Social Scientific** or **Collaborative Networks** (SCNs), most notably Mendeley, SSRN, ResearchGate, Scholix, Figshare and Academia.edu, which offer their platform services for new forms of science communication, collaboration and for integrating science, data, clinical evidence and experiences into an exciting new mix of scientific discourse and engagement.

Typically, SCNs are platforms aimed at connecting researchers with common interests. Users create profiles and are encouraged to list their publications and other scholarly activities, upload

<sup>21</sup> <https://www.inasp.info/about/history>

copies of manuscripts they have authored, and build connections with scholars they work or co-author with. However, much of this content is posted without permission and contrary to the STM Principles (see below).

**Academia.edu** is probably still used more by researchers in the social sciences and humanities but there is pressure to pay for a premium service that is off-putting to many researchers. Some academics have thus created *ScholarlyHub*<sup>22</sup>, a non-profit OA repository that gives access to academic papers, research projects and researchers. The platform aims to become a member run and owned SCN that aggregates research, teaching and other professional resources. Another network, *Colwiz* (collective wisdom)<sup>23</sup>, launched in 2011 and provided interactive digital collaboration and free reference management services for researchers in academia, industry and government globally. Colwiz also developed the ACS Chemwork platform for the American Chemical Society. In 2013, Taylor & Francis incorporated Colwiz's interactive PDF reader into their journals' platform and in 2017 its parent, Informa, acquired the whole company. In 2016, the company also developed the *wizdom.ai* research intelligence product. At the time of writing, Colwiz functionality was being merged into *wizdom.ai* to develop an intelligent research assistant under the *wizdom.ai* brand.

Bibliography management software, such as *Endnote* (Clarivate Analytics),<sup>24</sup> *Flow* (Proquest)<sup>25</sup>, *Pages* (SpringerNature), *Zotero*<sup>26</sup> etc., also allows users to share their research libraries with other users but typically the sharing is inherently one-to-one or one-to-few, or restrictions on the numbers of users with whom content may be shared are explicitly enforced.

The popularity of SCNs is perhaps an indication of the way in which authors prefer to share their articles. However, uncertainty over the copyright status of academic papers hosted on social networking sites raises concerns over the persistence of such content. To counter this, the STM Association has developed **the 'STM Voluntary Principles for Article Sharing'**<sup>27</sup>. The principles were developed between 2014 and 2015, and provide a very useful baseline for any SCN or platform offering SCN services. The website <https://www.howcanishareit.com> provides more information and a tool to check the status of published works and links to the sharing policies of individual publishers.

## 8.2. Access to research data and finding a reward model for making scientific data discoverable, curating data, verifying and assuring provenance, ethical collection of data

One of the over-arching trends in research publishing is linked to the global trends of big data and increased computing power, which also underlies the developments in text and data mining and artificial intelligence discussed below.

In order to understand the role of data, STM's Eefke Smit, Director for Innovations and Standards has coined the term '*Data Pyramid*' to structure the discussion. At the top of the pyramid sits the 'data' that forms part of research publications; the layer below includes processed data and

22 <https://www.scholarlyhub.org/>

23 <https://www.wizdom.ai/>

24 <https://endnote.com/>

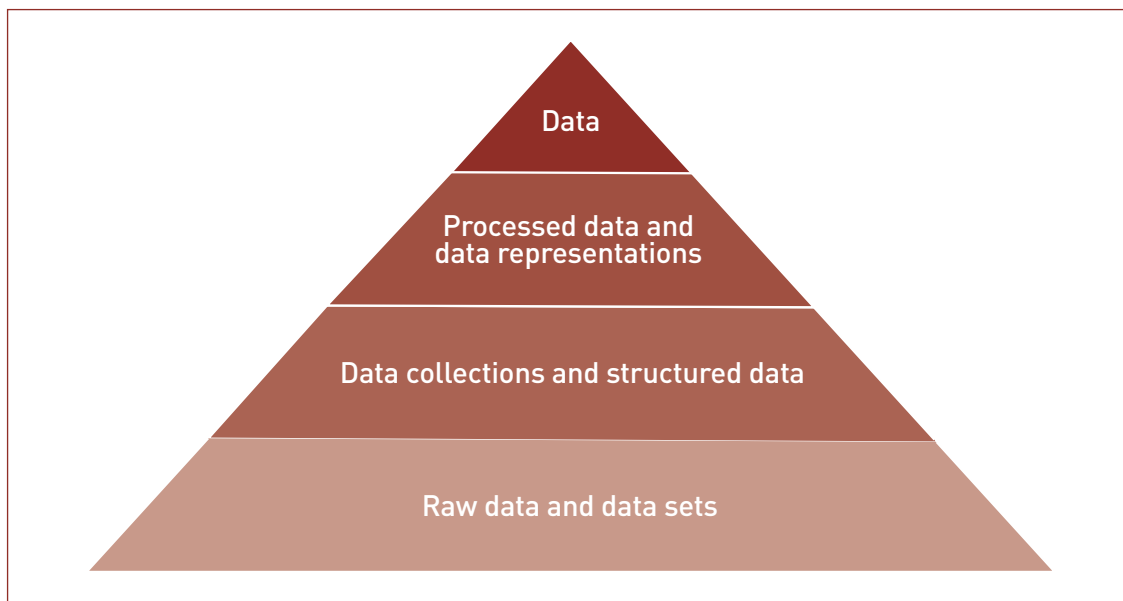
25 <https://flow.proquest.com/>

26 <https://www.zotero.org/>

27 [https://www.stm-assoc.org/2015\\_06\\_08\\_Voluntary\\_principles\\_for\\_article\\_sharing\\_on\\_scholarly\\_collaboration\\_networks.pdf](https://www.stm-assoc.org/2015_06_08_Voluntary_principles_for_article_sharing_on_scholarly_collaboration_networks.pdf)

data representations (which form the basis for data in publications); that pyramid layer is in turn supported by data collections and structured data; which in turn relies and is sourced from both 'raw data' and data sets.

### DATA PYRAMID



A key element that the STM Association has collaborated on with Datacite<sup>28</sup> (the main organisation driving the discoverability and citability of data) is to make the different layers of the pyramid more interconnected. DataCite in turn is working with Crossref.

There are various competing initiatives that try to define or imagine the 'article of the future' as a collection or node rather than a fixed document. In this way, it is possible to envisage the future article really to be a 'knowledge stack'.

## 8.3. Text and Data Mining

TDM (text and data mining) has the potential to transform the way scientists use data. TDM draws on natural language processing and information extraction to identify patterns and find new knowledge from collections of textual content. Semantic enrichment and tagging of content are likely to enhance TDM capabilities. At present TDM is most common in life sciences research, especially within pharmaceutical companies, but relatively little used elsewhere.

Until recently, the main challenges for more widespread adoption of TDM were legal uncertainties as to what was permitted; the lack of an efficient licensing regime; technical issues such as standard content formats including basic common ontologies; the need for content aggregation to permit mining cross-publisher corpuses; the costs and technical skills requirements for mining; limited incentives for researchers to use the technique; and a lack of understanding on the part of publishers.

These challenges have been addressed through various initiatives:

- The STM Association and **PDR**, the leading pharmaceutical sector research infrastructure

<sup>28</sup> <https://datacite.org/>

group, issued an updated joint sample licence<sup>29</sup> in 2012 that includes a TDM clause. TDM is one of the developments that has been prevalent in the corporate R&D world for some time and as it became a more used practice, the academic sector approached STM and governments to address licensing issues;

- STM publishers issued a statement in November 2013 committing its signatories to implementing the STM sample licence clause, or otherwise to permit non-commercial TDM of subscribed-to content at no additional cost; to develop the ability to mine content; and to develop platforms to allow integration of holdings across institutions for TDM purposes. The statement has been subsequently updated, with most recent version dating from 2017<sup>30</sup>.
- CrossRef's text and data mining tools (originally **Prospect**)<sup>31</sup>: this offers a metadata API and services that can provide automated linking for TDM tools to the publisher full text, plus a mechanism for storing licence information in the metadata, and optionally, a rate-limiting mechanism to prevent TDM tools overwhelming publisher websites;
- Copyright Clearance Center (CCC) offers a service targeted at life science companies. **RightFind XML for Mining**<sup>32</sup> provides access to approximately 10 million articles in XML content from more than 60 STM publishers with normalised metadata, and consistent licensing terms for mining the content for internal research. The system reduces the necessity for one-off licensing negotiations, along with the associated administration costs, while providing additional royalties to rightsholders when their content is used for text mining;
- STM publishers continue to develop the field and also to accompany the implementation of the EU's Digital Single Market (DSM) Directive that provides an exception for reproductions necessary as part of TDM with certain safeguards and for 'Research Organisations' as defined, if they have lawful access to the content in question (for example, a subscription). The DSM Directive also makes it clear that commercial uses generally and non-commercial uses falling outside scientific research or of non-subscribed content remains subject to licensing<sup>33</sup>.

## 9 ARTIFICIAL INTELLIGENCE

Artificial Intelligence (AI) emerged as one of the hot topics of 2019, as predicted by the Future Lab of STM's STEC back in 2017/2018. AI here is simply understood as any human-machine co-production in which the machine performs tasks it has learnt through any type of data analysis. Whilst itself a fast-evolving subject, all practical AI to date rests on the availability of 'clean data' which is fit for use in reinforced, supervised or unsupervised machine learning, either immediately or after pre-processing. As further elaborated above, 'data' can be both completely unstructured (like plain text and video) and highly structured (like metadata, tables and well-formatted records), any combination thereof or anything between that falls into one of the layers of the 'data pyramid'.

**For at least the next three years, it must be anticipated that individual STM publishers will rightly seek to realise the tremendous opportunities but also be faced with great risks**

29 <https://p-d-r.com/wp-content/uploads/2019/08/2012-Model-Licence.pdf>

30 [https://www.stm-assoc.org/2017\\_05\\_10\\_Text\\_and\\_Data\\_Mining\\_Declaration.pdf](https://www.stm-assoc.org/2017_05_10_Text_and_Data_Mining_Declaration.pdf)

31 <https://www.crossref.org/education/retrieve-metadata/rest-api/text-and-data-mining-for-researchers/>

32 <http://www.copyright.com/business/xmlformining/>

33 <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32019L0790&from=EN#d1e953-92-1> (See Articles 3 and 4 of the DSM as adopted)

**associated with AI.** Publishers' valuable services and skills in generating, curating and validating data need to be seen by all relevant stakeholders for what they are: an indispensable piece of the puzzle to humanise and make AI useful and ethical.

As a sector, STM will focus immediately on the creation of an enabling framework for responsible human-machine co-production, including establishing standards and regulations as well as processes for testing and validation of data, taking into account:

- The fast-evolving nature of AI;
- The need for data markets for training, testing and validation data;
- The risks and efficiencies associated with intensive and wide data usage;
- The need to balance negative incentives for data harvesting, generation and marketing with positive incentives;
- The need to take public security, safety and the legitimate interests of third parties into account;
- The need to enable diverse and parallel usability of data as input, output, by-product and training data;
- The need to consider network effects and benefits and risks of data pooling.

## 10 SUMMARY AND CONCLUSION

---

Licensing of electronic content is the lifeblood of STM publishing. Licensing as a vehicle is incredibly flexible and able to accommodate all kinds of business models, from outright purchase to rental, whether itemised or as collections, databases; tailor-made for all shapes and sizes of users, be it the scholarly community in a large or small country, in the developed or developing world, or a country or region in transition; or commercial enterprises, whether SMEs or large multi-national corporations. In addition, licensing allows the provision of access to copyright-protected content in a commercial setting and for segmentation to non-commercial, reduced-rate or nil-rate segments. Licensing also permits the development of new licensing models such as Open Access (author-pays) or hybrid 'offsetting deals'.

Finally, as this chapter illustrates, STM publishers are engaged in a wide range of access related initiatives across the world. By leading and/or actively participating in access projects STM's members demonstrate their commitment to delivering the highest level of sustainable access to high-quality content to the widest range of stakeholders ■



**IPA**  
international  
publishers  
association

**INTERNATIONAL PUBLISHERS ASSOCIATION**  
Promoting and protecting publishing worldwide

International Publishers Association  
Avenue de France 23, 1202 Geneva, Switzerland  
[www.internationalpublishers.org](http://www.internationalpublishers.org) ■ Tel: +41 22 704 18 20  
[secretariat@internationalpublishers.org](mailto:secretariat@internationalpublishers.org)  
🐦 @intpublishers