SPARC

2021 UPDATE

SPARC Landscape Analysis and Roadmap for Action

September 2021

2021 UPDATE

SPARC Landscape Analysis and Roadmap for Action

SPARC

September 2021

Table of Contents

BACKGROUND

The increasing concentration of scholarly communications, courseware publishing, and data analytics into the hands of fewer commercial vendors continues to raise concerns, particularly in the absence of evidence that publishers have any interest in mind other than their short-term revenue and profit growth. The focus on protecting revenues even in the face of deep academic budget cuts, the relentless lobbying to protect "inclusive access" practices that limit student choice, and the reluctance to abandon practices that disadvantage researchers point to the conclusion that the academic community can protect its values only by increasing control of its own content and infrastructure.

The past year has seen more deals that led to more concentration, loss of diversity, and ultimately to the academic community's lessening control over its own destiny. However, there are also positive signs: a large merger failed, Invest in Open Infrastructure (IOI) was launched as a concerted effort to build a community-owned infrastructure, and some legislative progress was made. Much remains to be done, but the opportunity to tip the scales in favor of the interests of the knowledge community is significant and must be pursued.

This 2021 Update to the SPARC Landscape Analysis further explores these trends. Supplementing observations first published in the SPARC 2019 Roadmap for Action, this document suggests organizational changes in academic institutions to both (1) manage increasing strategic and ethical challenges and (2) deploy tools and analyze data to better understand the needs and protect the interests of individuals and communities. The recommendations underscore the need for the academic community to take control of its own content and infrastructure both to best serve its own interests and to protect and further its values of equity, inclusiveness, and academic freedom.

Report Authors:
Claudio Aspesi (Lead Author)
Nicole Allen
Raym Crow
Valorie Hollister
Heather Joseph
Joseph McArthur
Nick Shockey
Katie Steen
Report Editor: Marty Lentz

CONTINUING CHALLENGES

The *2020 Update: SPARC Landscape Analysis and Roadmap for Action* identified three key emerging trends that continue to significantly influence the academic publishing landscape:

1. Conflicts of Interest

The first of these trends was the strategic shift of some companies, including Elsevier (the leading scholarly publisher), continuing to build up their roles in research assessment. In the case of Elsevier, this is in addition to its traditional role in research dissemination. The academic community is just beginning to express concern about the conflict of interest inherent in being in these two businesses. The University of California Academic Senate passed a set of recommendations that addressed this issue in March 2019, when the university system prepared a comprehensive report on the use of research information management services (RIMS).[1] However, this scrutiny is the exception, rather than the rule. This issue is compounded by the fact that, while Elsevier claims to operate with the interests of researchers at heart, its actions collide with this narrative.

Public interactions with Elsevier's management during the first part of 2021 suggest that Elsevier itself continues to publicly downplay the conflicts of interest among its portfolio of activities. More broadly, little attention seems to be given to the conflicts that arise when Elsevier collects data from researchers and then sells research assessments to academic institutions, funding bodies, and governments.

For example, in September 2020, Brad Allen, chief architect at Elsevier, held a webinar organized by the Harvard Data Science Initiative. During the Q&A, which was open to the public, questions were asked about the ethics of artificial intelligence (AI) and about possible conflicts of interest that the use of AI could present.[2] Though the answers

1 https://senate.universityofcalifornia.edu/_files/reports/rm-jn-mb-rims.pdf

2 https://www.youtube.com/watch?v=cK3yKFhDyxs – the Q&A starts at 1:01:15

indicated that Elsevier is aware of the ethical issues affecting data science, the presenters offered no concrete steps the company has taken to address them, and this void has not stopped them from selling their products. When asked about conflict of interest when serving both researchers, funders, and governments, Mr. Allen allowed that his answer had not been on point and admitted he had not thought much about it.

Conflicts of interest are not limited to both publishing research and assessing it or to collecting individual researchers' data through productivity tools and selling those data to universities, funding bodies, and governments. Leslie Chan and George Chen have recently written extensively on the conflict of interest inherent in publishing research and contributing to university rankings.[3] Conversations with senior administrators of academic institutions often reveal the frustration engendered by university rankings, yet it is very difficult to find administrators who feel they are in a position to advocate for breaking the reliance on a system that negatively affects their own institutions.

Similarly, college bookstores are historically perceived as aligned with the interests of academic institutions and their communities. However, many campuses have outsourced their bookstore operations to Barnes & Noble and Follett, who are increasingly coordinating with publishers to promote the adoption of "inclusive access" programs that automatically bill students for digital course materials. These companies have an economic incentive to promote this model on campus, even if the prices and terms of service conflict with the interests of students and faculty.

The fact that Elsevier (and, potentially, other companies) would pursue interests that put them at odds with the interests of the academic community and tolerate internal conflicts of interest should not come as a surprise. The business of publishers is to make money; the "business" of academic institutions is to advance knowledge, not to enable publishers to achieve their commercial goals. Unfortunately, the responsibility for highlighting and resolving conflicts of interest falls squarely onto the academic community.

3 Chen, George, & Chan, Leslie. (2021). University Rankings and Governance by Metrics and Algorithms (Draft chapter available at https://zenodo.org/record/4730593#YI2FouspDOR. The final version will be available in the *Research Handbook on University Rankings: Theory, Methodology, Influence and Impact*, edited by Ellen Hazelkorn and Georgiana Mihut, forthcoming 2021, Edward Elgar Publishing Ltd.

2. Communal Distribution

The second trend the *2020 Update: SPARC Landscape Analysis & Roadmap for Action* highlighted was the intentions of leading publishers to launch communal research distribution services. The launch of Get Full Text Research (GetFTR) was ostensibly motivated by the desire of the publishers to facilitate researcher access to literature, and feedback from librarians suggests that GetFTR does address a real issue.

However, the launch of GetFTR also signaled a potential shift in the number of downloads of articles directly from the publishers' servers at the expense of legitimate alternative sources like repositories. After one year of service, GetFTR indicated it had signed up 10 publishers and 11 integrators (such as Semantic Scholar). GetFTR does not release any activity report, so there are no visible data on its impact. The architecture of GetFTR is designed to take place behind the scenes, with no ability for libraries or other users to opt out. The strategic issue highlighted in 2020 remains unaddressed: the risk of diverting users to publishers and away from legitimate repositories that choose to stay out of GetFTR, which would hollow out the value of those repositories.

GetFTR has the potential to undermine the effectiveness of services like Unpaywall and Google Scholar that radically threaten the current publishing ecosystem. Cooperating on some features, rather than competing, is in line with the interests of commercial entities as well as some not-for-profit players that equate their interests with those of commercial players. The functionality of GetFTR won't affect libraries' subscription decisions or faculty choices about where to publish. Sharing this kind of infrastructure is a rational course of action for players that benefit from the status quo to fend off those who wish to change it.

3. The "Bigger Deal"

A third concern was the emergence of requests to bundle publishing contracts, both transformative agreements (TAs) as well as traditional collections subscriptions, with the supply of data analytics services. This concern was spurred by the decision of two Dutch consortia to sign such a deal with Elsevier in May 2020. As of July 2021, no other deals

directly bundle these disparate products, though that does not guarantee that more will not be signed in the future.

It should be noted that bundling has largely favored publishers, whether it is bundling articles into journals (which improved the economics of printing, shipping, and selling); bundling journals into collections subscriptions (which put together important journals with less relevant ones, forcing libraries to pay for all of them); bundling reading and publishing activities in transformative agreements (which ensure high levels of spending and limit the opportunities for smaller publishers to compete); or bundling data analytics with subscriptions. In every case, some valuable offerings are packaged with lower-value ones, forcing customers to pay for everything, regardless of their actual need.

EMERGING CHALLENGES

In addition to the challenges highlighted in 2020, five emerging challenges merit consideration. Some of these are broad societal issues, while others are more strictly related to academic and knowledge activities.

1. Impact of the Pandemic on Inequities Within the Academic Community

The extent to which the academic community reflects the inequities of society at large is well documented, by both analysis and lived experience. An article published in March 2021, for example, shows how, before the onset of the COVID-19 pandemic, "after controlling for applicants' educational background, country of origin, training, previous research awards, publication record, and employer characteristics, Black applicants remained 10% less likely than White applicants to be awarded National Institutes of Health (NIH) research funding."[4] However, despite this disparate impact, the academic community has yet to respond meaningfully to address the personal and professional harm that continues.

The impact on underrepresented minority students in terms of mental health as well as expected graduation rates has also been documented.[5] Enrollment losses have disproportionately affected students from disadvantaged backgrounds. The latest available data from the National Student Clearinghouse Enrollment Center show the disproportionate impact of the pandemic on students attending different types of US academic institutions.[6] Virtually all the overall 2.5% enrollment decline centered on public 2-year

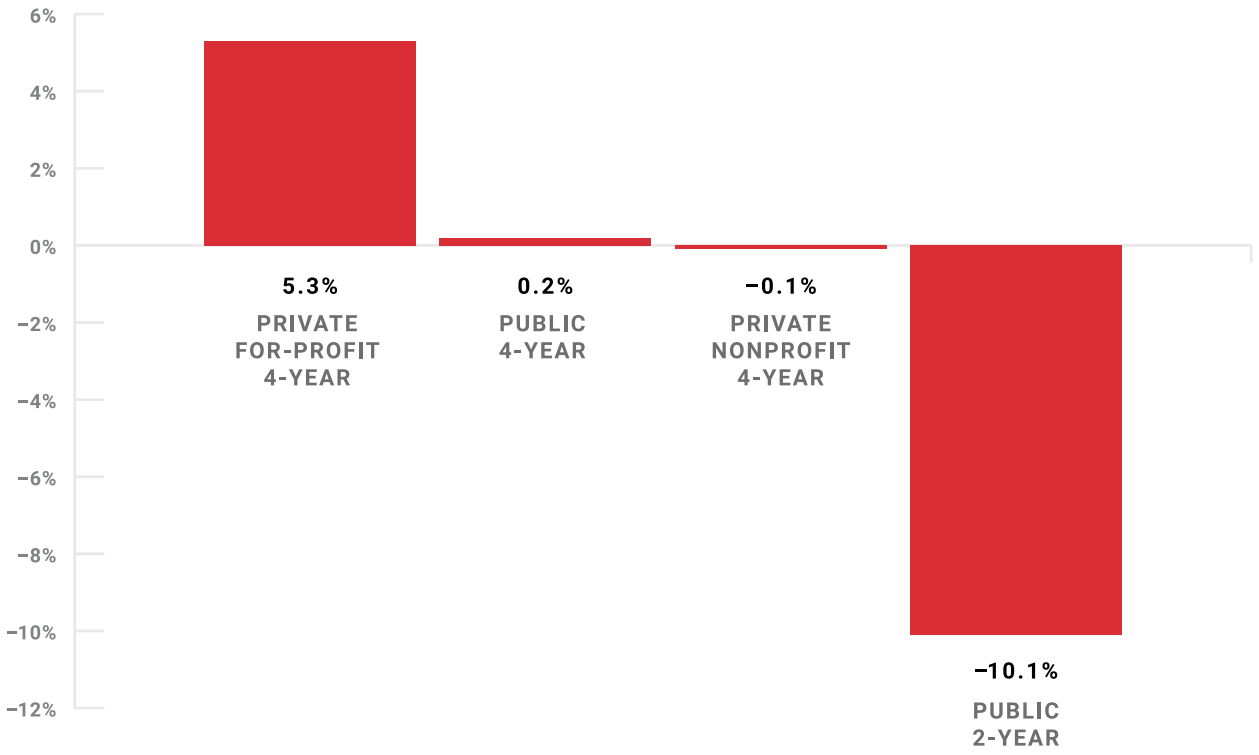
4 https://stm.sciencemag.org/content/13/584/eabe7189?_ga=2.40981414.505338189.1619101421-1688379423.1619101421

5 https://news.harvard.edu/gazette/story/2020/10/covid-carries-triple-risks-for-college-students-of-color/
https://dornsife.usc.edu/news/stories/3279/

6 https://nscresearchcenter.org/current-term-enrollment-estimates

institutions (**Exhibit 1**), while 4-year colleges registered largely flat enrollment, and for-profit colleges grew substantially (after several years of decline, sometimes in double digits). It should be noted that flat enrollment is not a particularly impressive performance, since courseware publishers historically assumed that a 1% rise in the unemployment rate would drive up college enrollment by 3%. The spike in US unemployment was brutally fast (unemployment rose from 4.4% in March 2020 to 14.7% in April 2020) and then declined quite rapidly. However, by August 2020, the unemployment rate still stood at 8.4%, and the additional 4 percentage points should have translated into a 12% rise in college enrollment. Clearly, the pandemic was like no other recent economic crisis in many ways, including in how it reverberated across the student population.

Exhibit 1: Changes in Enrollment, Fall 2019 to Fall 2020



Source: National Student Clearinghouse Enrollment Center

2. Demand for Faster Scientific Communication Channels

During the COVID-19 pandemic, the academic community did something out of the ordinary. Within four months of the first confirmed case, the research community had published an estimated 16,000 articles on the virus—almost 40% of which were preprints.[7] By the end of 2020, 100,000 articles had been published,[8] and according to the same analysis, preprints accounted for an estimated 17–30% of these articles.

As a yardstick, 1% to 2% of the articles listed in PubMed were initially made available as preprints.[9] While these data are an imperfect source of comparison, because one cannot know what percentage of COVID-19 preprints will become articles over time, the large number of preprints submitted in 2020 illustrates a sea change in the communication practices of the research community when faced with the urgency of responding to a major humanitarian health crisis.

This demand highlighted important—and diverse—pressure points in the current scientific communication ecosystem that merit close attention. Among these are the need to invest in technical and human infrastructure that can ensure rapid communication of scientific findings, along with the quality assurance, scientific integrity, and validation services the research community requires. Also crucial are both more robust research incentives for faster sharing of results outside of traditional journals and also rewards for contributions to validating, curating, and prioritizing research results on non-journal platforms.

3. Rising Privacy and Surveillance Concerns in Technology Used by Academic Institutions

Since the start of the pandemic, academic institutions have been confronted with new issues related to deployment of technology. In particular, three issues require attention

7 https://www.biorxiv.org/content/biorxiv/early/2020/05/23/2020.05.22.111294.full.pdf

8 https://www.nature.com/articles/d41586-020-03564-y

9 https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.3000151

from the academic community. Each of these raises different concerns:

- Insertion of tracking software in services sold to academic libraries
- Collection and sale of data by some commercial vendors with ties to the academic community to governments and law enforcement
- Risks and inequities of online exam proctoring tools

Tracking and Monitoring Software

In an October 2020 SNSI (Scholarly Networks Security Initiative) webinar, a group of scholarly journal publishers unveiled a plan to insert monitoring software on its platforms to protect copyright from cyber-attacks.[10] According to some reports, the SNSI initiative was not targeted to blocking or monitoring Sci-Hub,[11] but the agenda of the SNSI webinar explicitly included a presentation about Sci-Hub.

Critics of the initiative have pointed out the absence of evidence for some of the claims made by SNSI (from connections between Sci-Hub and Russian intelligence agencies to the use of Sci-Hub to steal passwords to access personal records or access to other databases).[12] Claims and counterclaims are difficult to adjudicate. On the other hand, the idea of academic libraries acquiescing to the deployment of software that monitors the behavior of their patrons and collects data with no conditions on what, why, and for what use, flies in the face of both long-held privacy expectations of library users and of academic freedom.

The SNSI presented an update at the STM (International Association of Science, Technical, and Medical Publishers) Conference in April 2021. The speakers reported seeing skepticism from the librarian community and indicated the need to reframe their

10 https://www.snsi.info/news-and-events/cybersecurity-landscape/

11 https://www.codastory.com/authoritarian-tech/spyware-in-libraries/

12 https://www.techdirt.com/articles/20201029/04034145611/to-prevent-free-frictionless-access-to-human-knowledge-publishers-want-librarians-to-be-afraid-very-afraid.shtml
https://netzpolitik.org/2020/news-from-elsevier-no-open-access-deal-but-spyware-against-shadow-libraries/

message from "protecting the publishers" to "protecting library patrons." Setting aside the communication strategy which appeared inadequate even to the speakers, the real issue seems to be the healthy skepticism of librarians. Publishers appeared surprised and out of touch, particularly considering the data resale activities of some of the companies in their ranks.

Collection and Sale of Data

RELX and Thomson Reuters operate businesses that collect vast amounts of data on more than 1 billion people, with a particular focus on the United States. While individual researchers have been highlighting this issue for years (Sarah Lamdan at CUNY, in particular[13]), concern about these surveillance businesses began to register with the academic community in 2020 and early 2021.The *Daily Bruin* (the UCLA student newspaper) called for a boycott of RELX and Thomson Reuters for selling data to the US Immigration and Customs Enforcement Agency (ICE).[14]

The intersection of data sales and academic institutions poses both ethical and practical challenges. Many academic institutions believe their fundamental values clash with some government policies. This issue is not confined to the US alone: in many countries, in past years academic institutions have clashed with governments or have been subdued and forced to acquiesce to government policies. Academic institutions should determine whether they want to be customers of and do business with companies that operate in activities that may be perfectly legal, but clash with their values.

There is precedent for this: in 2007, Reed Elsevier (as RELX was called then) divested its arms shows business after repeated calls to do so from activists as well as from *The Lancet*, one of the Elsevier journals.[15] At the time, Reed Elsevier pointed out that

13 https://socialchangenyu.com/review/when-westlaw-fuels-ice-surveillance-legal-ethics-in-the-era-of-big-data-policing/
http://www.inthelibrarywiththeleadpipe.org/2019/ice-surveillance/
https://www.jurist.org/commentary/2020/06/sarah-lamdan-data-policing/

14 https://dailybruin.com/2020/08/17/opinion-to-support-undocumented-students-uc-must-divest-from-companies-selling-data-to-ice

15 https://www.theguardian.com/business/2008/may/30/armstrade.weaponstechnology

all these shows were legal activities, underscoring that it is legitimate to ask companies to renounce, on ethical grounds, business with legally sanctioned or governmental programs. It is also important to underscore that a sale is an unsatisfactory remedy in most cases. First, the sale price of a business is a function of expected future earnings, so selling a business is a way to collect a significant part of future profits from the business. In addition, simply empowering someone else to continue undertaking an objectionable activity does not solve the issue. If activists want Thomson Reuters and RELX to stop selling data to ICE, they should ask for those businesses to be closed altogether, particularly because it may be difficult for other companies to replace the data made available by the two companies.

In addition to the ethical issues posed by the ICE business of Thomson Reuters and RELX, there is a practical one. SNSI openly lobbies libraries to install tracking software; according to sources, RELX already does so. ScienceDirect links directly to ThreatMetrix's (a RELX company) processing notice to describe how collected data are used: "As explained in our privacy policy, sciencedirect.com and Linkinghub.elsevier.com collect information through the use of cookies or similar technologies…Security cookies and related technologies, such as those provided by ThreatMetrix, are used to maintain online security and protect our website against fraud and abuse." There is no indication whether any of the information is made available to third parties.

The use of ThreatMetrix in ScienceDirect raises questions of whether the online activities of legal clinics in law schools through LexisNexis tools is monitored, and whether data are sold to government agencies, jeopardizing the rights of individuals and associations being assisted by legal clinics. In the past, RELX has denied doing so—but the only way to ensure that RELX does not combine information from LexisNexis with other databases is to abandon contracts with controversial agencies and governments.

Risks and Inequities of Online Exam Proctoring Tools

Exam proctoring raises significant issues of privacy and ethics. The sudden mass transition to online learning during the Spring 2020 term raised the issue of how to conduct exams and other assessments for a vast number of students who had limited or no

experience with distance learning. There was a rush to adopt online proctoring solutions that would allow an orderly conclusion of the semester for as many students as possible. These solutions were made available by both courseware publishers and independent companies, with little time to adopt sound rules to guard against violations of privacy and the risk of introducing biases that disfavor underrepresented and underprivileged communities.

Unsurprisingly, reports of issues quickly started to emerge,[16] particularly regarding software singling out some categories of people, like minorities, students with certain medical conditions, and parents with young children who cannot be left alone. In addition, proctoring software is deeply intrusive into students' personal computers and personal lives—intrusions that an increasing number of students are rightfully pushing back against.[17]

The issues posed by online proctoring fall into two broad categories: privacy and equity. The privacy issues are significant and easy to grasp: Private companies are collecting sensitive information on students' names, locations, and even physical appearance, exposing them to a number of risks. Since there have been several instances of use of surveillance technology for illegal purposes, adding a large group of students to the possible pool of victims appears ill-considered at best.

The issues related to inequity are even more complex. A major feature of online proctoring software is the use of algorithms to detect "suspicious" behavior that may indicate cheating or other academic integrity violations. This type of technology is problematic because it is bound to make mistakes, and those mistakes disproportionately harm students subject to the algorithm's biases. While some online proctoring tools

16 https://www.washingtonpost.com/technology/2020/04/01/online-proctoring-college-exams-coronavirus
https://hybridpedagogy.org/our-bodies-encoded-algorithmic-test-proctoring-in-higher-education/
https://onlinelearningconsortium.org/the-new-and-deeply-dissatisfied-users-of-online-proctoring/
https://www.insidehighered.com/news/2020/05/11/online-proctoring-surging-during-covid-19

17 https://www.newyorker.com/tech/annals-of-technology/is-online-test-monitoring-here-to-stay
https://sparcopen.org/news/2021/higher-education-reckons-with-concerns-over-online-proctoring-and-harm-to-students/

incorporate human review of flagged behavior, little is published about the protocols used, so there is no way for students or their advocates to know what behavior will trigger a human review and how the review is formulated. A vast literature in psychology demonstrates that how a question is formulated affects how people respond and that confirmation bias can play a role when determining guilt or innocence in forensic activities. It is possible that humans will review the evidence with a predetermined bias and find themselves more likely to decide against the student. Ultimately, the use of proctoring software poses serious ethical issues and, ideally, it should be phased out altogether. However, as long as it is deployed and in unique circumstances where institutions may believe its use is justified, the protocols used should be transparent and agreed upon with advocates of students and their families.

4. Continued Consolidation of the Publishing Industry

Consolidation of the publishing landscape continues to be a concern. In a rare win for activists in 2020, the merger of McGraw Hill and Cengage was halted, both in response to challenges by advocates and because of the subsequent conditions that regulators were likely to impose. Aside from this, however, a steady stream of consolidations continued apace, with academic publishing vendors announcing several significant deals.

In December 2020, Elsevier acquired Shadow Health, a developer of virtual simulation in healthcare and nursing education. In January 2021, Wiley disclosed it had reached an agreement to acquire Hindawi, once the largest Open Access (OA)-only publisher. In March 2021, Springer Nature Group announced the acquisition of Atlantis Press, an OA publisher. And finally, in May 2021, Clarivate announced it would acquire ProQuest.

One side effect of consolidation is the increased fragility of services made available to the academic community by third parties. In just a few months, the academic community learned it would lose two highly popular and prized services: PLOS's Article Level Metrics (ALM) platform and Microsoft's Academic Graph. In both cases there are alternatives, but as options decline, users should not assume that the owners will provide or maintain free access to their tools going forward. These two decisions show it is vital for the academic

community to take control of its own destiny, because dependency on the goodwill of others is not a viable strategy.

Leading higher education courseware publishers, hobbled by a declining market and in some cases high levels of debt, have been less active. The only notable acquisition was by Pearson in February 2021, when it announced it would buy Spotlight Education, a company that uses unique, proprietary technology to turn education data into personalized video reports.

5. The Expansion of "Inclusive Access" Courseware

The pandemic further accelerated the transition from print to digital courseware within US academic institutions. In 2020, Pearson derived 70% of its US higher education revenues from digital products. (It was about 60% in 2019.) For McGraw Hill, digital courseware accounted for 72% of FY 2021 US higher education revenues[18] (vs. 61% one year earlier), and for Cengage, for the nine months from April 1, 2020, to December 31, 2020, digital revenues accounted for 81% of the total (vs. 79% for the comparable period during the previous fiscal year).

"Inclusive Access" (IA)—digital course materials that are automatically charged to a student's tuition or fee bill—accounts for an estimated 20% of McGraw Hill's revenues. Cengage does not disclose the exact number, but IA and Cengage Unlimited combined account for about 15% of revenues. Pearson does not disclose the weight of IA. While the weight appears modest, the growth rate of these programs is much faster than for the overall business: At McGraw Hill, for example, IA net sales grew by 58% in FY2021, while US higher education billings grew by 5%. Publishers attach great importance to the IA model and have strenuously opposed consumer protection legislation in states, including California and Texas, which proposed a shift to "opt-in" student billing among other provisions.

For academic institutions, IA poses increasingly complex issues, because campus bookstores play a significant role in promoting this model to faculty, yet they also stand to

18 For McGraw Hill, Fiscal Year 2021 covers the period from April 1, 2020 to March 31, 2021

financially benefit from its adoption.[19] Not only does IA grant the campus bookstore an effective monopoly over student sales, but it also offers a potential bite at the apple of the vast amounts of student data generated through digital courseware (which SPARC has already discussed at length as an alarming trend for publishers). Particularly on campuses that have outsourced bookstore operations to companies such as Barnes & Noble and Follett, administrators considering IA need to grapple with these conflicts of interest and long-term implications.

Administrators may view bookstore sales of course materials as a means to an end, especially in cases where a contractual agreement shares course material-derived revenue with the institution. However, as course materials are increasingly digital, academic institutions need to shift their thinking away from how to sell textbooks toward how to sustainably manage the campus's teaching and learning content. Thus far, the expertise of academic libraries in the negotiation and procurement of digital materials has been underutilized in this transition, and it will need to become increasingly central as the adoption of IA models creates conflicts of interest for outsourced bookstore operations.

The increasing number of academic institutions that have adopted IA programs should also adopt policies that ensure student awareness and choice. The best option is to make programs "opt-in," so that students are charged only with their consent. Transparency measures also benefit students, such as the new law adopted by Texas in June 2021 requiring that institutions disclose information about IA fees to students up front in the course catalog.[20] Furthermore, in-depth scrutiny of contractual terms and conditions is crucial, as some of the legal agreements with IA vendors contain onerous clauses for students (e.g., charging students unless they actively "opt out" or linking discounts to threshold participation rates). SPARC has created a searchable database of more than 70 publicly available IA contracts as a resource for comparing these provisions.[21]

19 https://historynewsnetwork.org/article/180169

20 https://sparcopen.org/news/2021/texas-adopts-transparency-measure-for-automatic-textbook-billing/

21 https://sparcopen.org/our-work/automatic-textbook-billing/contract-library/

THE IMPACT OF 2020 ON THE LANDSCAPE

1. Scholarly Publishers

The commercial publishers covered in the *SPARC 2019 Landscape Analysis* fared reasonably well in the pandemic.[22] Scholarly journal publishers saw a small decline in revenue growth rate, but at a pace that barely affected their profitability. In negotiations with librarians, publishers showed little inclination to help alleviate the difficulties of their customers: most libraries were offered, at best, flat price increases for 2021. Any price cuts were generally linked to relinquishing significant rights (such as perpetual access rights to articles covered in the subscription), with the notable exception of Canadian Research Knowledge Network (CRKN), which demonstrated that it is possible to get a significant price cut from Elsevier without significant corresponding concessions.[23]

The decision to hold firm on pricing conflicts with the "partnership" that publishers claim to have with academic institutions. In fact, the impact of the pandemic on academic and library budgets has been significant. An Ithaka survey published in December 2020[24] indicated that 75% of respondents among US academic libraries had experienced cuts. These cuts clustered in three categories (1–4%, 5–9%, and 10–14%) representing in aggregate almost 60% of the libraries, with the final 15% experiencing even deeper cuts.

In part, revenue resilience is a function of the subscription model. In fact, shortly after the beginning of the pandemic, RELX (which is usually referred to as Elsevier) issued guidance on the expected revenue resilience of its science, technical, and medical (STM) business on the grounds that 75% of revenues are subscription based. Early in 2021,

22 Details about the financial performance are contained in Appendix I.

23 https://www.crkn-rcdr.ca/en/crkn-elsevier-license-renewal

24 https://sr.ithaka.org/publications/academic-library-strategy-and-budgeting-during-the-covid-19-pandemic/

Wiley indicated that subscription revenues are indeed affected by "modest pricing pressure," although this pressure is offset by rising OA revenues. Nonetheless, the publishers seem eager to continue privileging their short-term financial performance rather than providing relief when their customers are under stress.

2. Courseware Publishers

Textbook and courseware publishers fared less well, but still better than originally expected. At the beginning of the pandemic, Cengage revealed, for example, that it had prepared contingency plans based on various scenarios and that the most negative one assumed a 25% decline in revenues. Actual results in 2020 were much less dire than this extreme scenario: looking at the three leading US courseware publishers, their total revenue decline across all businesses ranged between –4 and –10%. These declines were particularly severe in other parts of their portfolios, like global assessment for Pearson (which declined by –14%), or K–12 for McGraw Hill (which declined by –11%). The US higher education businesses fared better, with US revenues declining by –12% at Pearson and rising by 5% and 4% respectively at MGH and Cengage.

3. Debt, Cash, and Equity

The *2020 Update: SPARC Landscape Analysis & Roadmap for Action* observed that cash conservation, access to liquidity and capital markets, and levels of debt of different vendors could become important drivers of what these companies would be able to do in the years to come. One year later, these concerns have fallen to the wayside—at least for now.

Cost of Debt

Because of their high levels of debt, the situation of MGH and Cengage is notably different from that of other companies. At the time of the June 2020 update, S&P rated both McGraw Hill and Cengage CCC (seven notches below investment grade). Unsurprisingly, the yields on the two companies' bonds spiked in the aftermath of the failed merger. In mid-February 2020, before the pandemic started to affect the financial markets, the

yields on McGraw Hill Education and Cengage bonds stood at about 9% and 10%, respectively. As of March 2020, when the pandemic had been factored in by financial markets, but the merger had not yet failed, yields stood at 15.3% and 18.5% for McGraw Hill and Cengage. By early May 2020, after the merger was abandoned, yields had risen to 29% and 37%, respectively, to then decrease again as financial markets recovered. As of July 9, 2021, yields on McGraw Hill Education and Cengage stood at about 5.75% and 9% respectively, broadly in line with the 6.5% average yield for CCC and lower-rated bonds (i.e., defaulted) in US dollars at the same date.

In part, this is due to the strong reaction by monetary and—to some extent—fiscal authorities. Central banks have responded to the pandemic by substantially expanding their monetary policy. Large purchases of both government and corporate bonds have injected substantial amounts of liquidity into the economies of both the US and many European countries, contributing to lowering (or, in the case of Eurozone countries, maintaining) low interest rates. In addition, many countries have seen or are expected to provide substantial support to their economies through aggressive fiscal policies. The initial shock was significant: the VIX index, which measures the expected volatility of US financial markets,[25] spiked in March 2020, when it reached levels last seen in October 2008, at the height of the financial crisis triggered by subprime loans. Since then, the index has steadily declined (with some corrections associated with political uncertainty), as both monetary and fiscal policy have been adjusted to fight the economic impact of the recession.

Cash

All the companies highlighted in the *2020 Update: SPARC Landscape Analysis & Roadmap for Action* launched cash conservation and cost-cutting programs. The financial actions ranged from suspending stock buybacks to securing additional lines of credit

25 The VIX index measures expected market volatility (i.e., the expectation that prices can change up or down dramatically in a short period) and is calculated from the prices of options on the S&P 500 for the following 30 days. The VIX index, which was created by the Chicago Board of Trade, started trading in March 2004, but its performance has been back calculated and is commonly available since 1990. In general, spikes in the index are associated with recessions, although the predictive power of the VIX has been questioned.

or refinancing bonds and credit facilities to extend their maturity into a further future, as well as taking advantage of lower interest rates. Both Pearson and RELX suspended stock buybacks in 2020, and virtually all companies launched significant cost-cutting programs aimed at either permanently lowering their cost base, postponing some expenditures until later in 2020 and into 2021, or both.

Every company secured additional lines of credit, and the ones that had debt nearing maturity renegotiated their facilities both to take advantage of low interest rates and to extend maturities. In general, courseware publishers took stronger action to reduce leverage and conserve or add cash to their balance sheets than did companies like RELX and Wiley that are less dependent on actual student enrollment (**Exhibit 2**). It is important to underscore that, in addition to cash at hand, all these companies have access to lines of credit which they can use to draw down additional cash as needed. For example, Pearson had on December 31, 2020, almost £1.1 billion (US$1.5 billion) in cash and cash equivalents at hand, but it reported total liquidity (which would typically include these credit facilities) of £1.9 billion (US$2.6 billion).

Exhibit 2: Changes in Net Debt/EBITDA and Cash at Hand for Selected Publishers

	NET DEBT/EBITDA		CASH AVAILABLE	
	12/31/2019	12/31/2020	12/31/2019	12/31/2020
Pearson	1.3x	0.8x	£437m	£1097m
McGraw Hill	5.2x	4.3x	$392m	$488m
Cengage	6.9x	5.6x	$287m	$445m
RELX	2.5x	3.3x	£138m	£88m
*Wiley**	1.8x	2.2x	$117m	$91m

Source: Company reports and presentations

**Data as of 1/31/21*

New Equity

In June 2020, Springer Nature Group (SNG) had just failed to list itself on the Frankfurt Stock Exchange through an Initial Public Offering (IPO) in the spring of the year, because the financial markets went into a deep dive as the pandemic came to dominate the headlines. The company attempted an IPO again in the early Fall of 2020 and it failed again, as the markets corrected downwards again, this time as the second wave of the pandemic hit the Northern Hemisphere. At the time of publication of this report, however, markets have recovered and have reached new highs, and yet the IPO has not been resurrected—which suggests more fundamental concerns in the financial markets about the long-term value and sustainability of the business.

A little-noticed article published in the *Financial Times* in December 2020 indicated that BC Partners was entertaining the option of selling its stake in SNG to a new fund controlled by BC Partners itself,[26] and more details emerged in March 2021 on Bloomberg.[27] The sale was finally announced officially on June 10, 2021. There are sound reasons for what may appear a bizarre action. Private Equity's (PE) strategy is to resell stakes in the several businesses that are held in a fund—either through an IPO or an outright sale. If selling one or more businesses becomes impossible, most fund rules allow the PE company to distribute the shares in any company that has not sold to the investors. Investors, however, have no desire to hold a small number of shares in a company that fails to find a buyer, and a distribution of shares is viewed as a negative mark on the performance of a PE investor. Hence, BC Partners chose to raise additional funds in a new Single Asset Acquisition Fund from other investors led by asset manager Neuberger Berman (as well as from itself) and acquire the SNG from itself. In this way, investors in the BC Partners fund that previously held the SNG stake can receive the proceeds of the sale, while BC Partners (alongside the new investors) can hope it will be able to finally sell the stake at a later stage.

26 https://www.ft.com/content/537ee5cc-2a74-4397-bdfb-4d846e6b8200

27 https://www.bloomberg.com/news/articles/2021-03-31/bc-partners-draws-neuberger-to-springer-nature-after-shelved-ipo

Historically, the academic community—SPARC included—has viewed with some satisfaction the repeated failures of SNG to go public. However, this satisfaction should be tempered by additional considerations. The first is that the funding deriving from the sale of shares to institutional investors would allow SNG to reduce its debt. In turn, if SNG had a lower debt burden, the company could expand into data analytics in a way it has not been able to do. (In fact, Digital Science remains a separate business owned directly by Holtzbrinck, the largest shareholder in SNG). The conflict of interest between the publishing activities and the data analytics business of Elsevier remains a significant concern, and though launching SNG into data analytics would pose the same issues, it would also allow mounting a direct competitive effort that could limit the competitive position of Elsevier.

In June 2021, Apollo Global Management (the PE company that acquired the education business from McGraw Hill in 2012) sold the company to Platinum Equity (another private equity company) for $4.5 billion. The valuation was possibly lower than Apollo had hoped for: an article published by Bloomberg in March 2021 indicated that Apollo was considering a sale of the company at a valuation (including the value of debt) of $5- to $6 billion.[28]

Finding a buyer for McGraw Hill proved more difficult than Apollo may have originally expected. Repeated attempts at an IPO failed, the proposed merger with Cengage failed when the US Department of Justice and United Kingdom regulators demanded onerous remedies, and a strategic buyer from outside the courseware industry did not emerge. Press releases were vague about the terms of the deal, and in particular they did not indicate whether the $4.5 billion valuation included the debt of the company ($1.765 billion on March 31, 2021). Depending on whether debt was included, the valuation was either a prudent 10.2x or an expensive 14.2x **EBITDA**.

EBITDA *refers to earnings before interest, taxes, depreciation, and amortization.*

28 https://www.bloomberg.com/news/articles/2021-03-25/apollo-is-said-to-weigh-6-billion-sale-of-mcgraw-hill-education

MGH's **net debt/EBITDA** has been lowered from 7.1x to 4.0x in two years, but that still leaves little room to add sufficient additional debt to boost equity returns and—after almost 10 years under the ownership of Apollo—additional cost savings are likely to be modest. Platinum Equity, therefore, faces some challenges in adding value to its investment. Since the debt will be refinanced, Platinum does have the option of raising again the debt of MGH to levels seen in the past (although that could be a risky move in light of fears that interest rates will rise to tame a possible return to inflation). Alternatively, MGH could embark on a series of technology acquisitions aimed at adding revenues and improve competitiveness in the core higher education market.

Net debt/EBITDA ratio *indicates a company's ability to pay off its debt. The lower the ratio, the higher the ability of a firm to pay off its debt. Many analysts consider ratios lower than 3 acceptable and higher than 4 a possible indicator of future distress.*

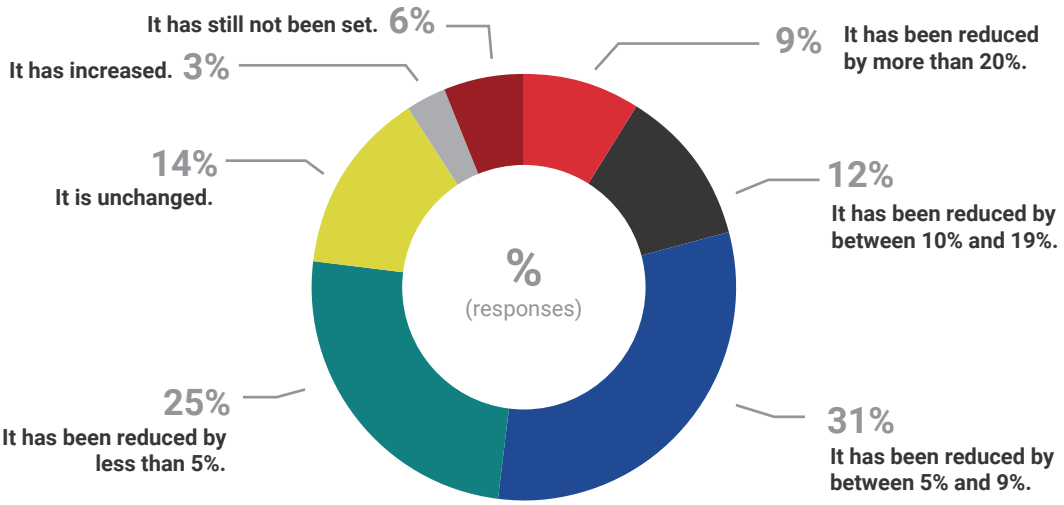
ACADEMIC LIBRARIES AND THEIR HOME INSTITUTIONS

1. Responses to Budget Cuts

Academic institutions are responding with a wide array of strategies, ranging from signing transformative agreements that are largely favorable to entrenched publishers to unbundling their big deals with the goal of drastically reducing spending with legacy publishers and reinvesting the funds in other initiatives. SPARC conducted a survey of its membership on this in early 2021. About half of the member libraries responded, with most indicating that they were facing budget cuts driven by COVID-19. Slightly more than half (56%) of respondents have to cope with cuts below 10%, a little less than a quarter (21%) with cuts above 10%, and the remaining 23% either had no visibility yet or had unchanged budgets (**Exhibit 3**).

Exhibit 3: SPARC Survey

1. Since the outbreak of COVID in March 2020, what percentage change have you seen in your overall library budget?



Source: SPARC survey

Of the libraries facing cuts above 10%, most (87%) have already decided to seek discounts from publishers, as have 68% of those facing cuts of less than 10% (**Exhibit 4**). The sum of those that have decided to seek discounts and those reporting they will likely ask totals more than 96%. Unbundling at least one big deal is decided or likely for 87% of libraries with cuts above 10% and for 68% of libraries with cuts below 10%. Exercising a financial hardship clause is decided or likely for 33% of libraries with cuts above 10% and for 28% of libraries with cuts below 10%. According to the SPARC survey, cutting staff

Exhibit 4: SPARC Survey

14. As a result of COVID-related budget pressure, how likely are you to:

		Very unlikely	Somewhat unlikely	Same	Somewhat likely	Very likely	We have already chosen to pursue this strategy	Responses
Seek discounts from publishers	COUNT	2	0	2	10	24	89	127
	ROW %	1.5%	0.0%	1.6%	7.9%	18.9%	70.1%	
Unbundle a big deal	COUNT	9	8	19	28	29	35	128
	ROW %	7.0%	6.3%	14.8%	21.9%	22.7%	27.3%	
Make significant cuts to a large journal package	COUNT	5	14	15	34	25	35	128
	ROW %	3.9%	10.9%	11.7%	26.6%	19.5%	27.3%	
Exercise a financial hardship clause	COUNT	41	30	18	20	8	10	127
	ROW %	32.3%	23.6%	14.2%	15.7%	6.3%	7.9%	
Leverage a financial hardship clause	COUNT	33	30	16	19	14	16	128
	ROW %	25.8%	23.4%	12.5%	14.8%	10.9%	12.5%	
Cut staff positions	COUNT	38	27	14	18	7	23	127
	ROW %	29.9%	21.3%	11.0%	14.2%	5.5%	18.1%	
Pursue new contract arrangements (publish & read agreements) with publishers	COUNT	10	14	32	28	25	19	128
	ROW %	7.8%	10.9%	25.0%	21.9%	19.5%	14.8%	

Source: SPARC survey

was seen as likely or decided by a sizable minority of libraries (about 41%, regardless of the depth of the cuts, and even 28% of libraries with no cuts).

The impact of cost reduction plans has started to be visible. The *SPARC Big Deal Cancellation Tracker* shows that several academic and research libraries have started to act. In the first half of 2021, 12 institutions unbundled from big deal packages (with 11 of those cutting Elsevier titles).

At the other end of the spectrum, as shown in Exhibit 4, many libraries mention read and publish (R&P) agreements (agreements structured to convert library subscription payments to payments for OA articles by affiliated authors) as a possible response to budget cuts—62% of libraries with large cuts and 55% of those with smaller cuts have either decided or are likely to pursue at least one transformative agreement (even if it is unclear whether librarians really see transformative agreements as a source of savings or a support to OA policies and faculty preference for OA publishing).

2. Role of Transformative Agreements

Though SPARC does not closely monitor these agreements, a large database of deals is available through the ESAC (Efficiency and Standards for Article Charges) Transformative Agreement Registry.[29] There is no standard definition for what constitutes a transformative agreement, nor do all academic institutions accept them. However, as common themes from these agreements emerge, it is useful to highlight some of their characteristics and implications. For libraries, institutions, and consortia that pursue these agreements, the intent is typically to convert subscription spending into publishing spending, with the goal of minimizing double-dipping and introducing price competition, as costs per article become clearer. However, these agreements have several consequences, often unintended and undesirable, for the participating institutions.

- Since all these agreements are based on article processing charges (APCs), they value that model at the expense of others. Participating in these agreements favors

29 https://esac-initiative.org/about/transformative-agreements/agreement-registry

well-funded institutions, and STEM disciplines are favored over SSH disciplines and less well-funded institutions.

- Aggregate library spending may be expected to rise, because most publishers will seek to maintain their current revenues. The corporate sector contributes an estimated 15% of total subscription revenues for some of the leading STM publishers, but this contribution would probably decline to near zero in a complete shift to open access, and publishers would look at academic funding as the most obvious source of alternative revenues. Some marquee institutions may receive attractive deals, but several others are effectively being asked to compensate for these discounts with higher spending elsewhere.

- Financial resources that could be reinvested in supporting community-owned academic communications infrastructure continue to be spent to support the incumbent vendors, thus stifling competition, innovation, and change.

- The transformative agreement model implicitly redistributes aggregate costs from "reading-heavy" to "publishing-heavy" institutions. This shift creates tiers of winners and losers, and it is likely to concentrate losses in a relatively small number of institutions that will be affected much more severely than expected. It also signals that open "read" access has no economic value, which undermines alternative open models that seek support from "read" institutions.

- Many transformative agreements are based on average APCs, bundled APC prices that conflate "must-publish-in" journals with journals that are not as prestigious and relevant. These bundles negate (or curtail severely) the possibility that competition will lead to declining prices and margins for lower-impact-factor journals. Smaller publishers that do not have such marquee journals can be particularly disadvantaged.

- To the extent that the pricing structure of both traditional subscriptions and transformative agreements remains opaque and cannot be directly compared with those of others, they continue to lack transparency.

- Publisher concentration is likely to increase. Larger publishers have more resources, both to support the lengthy and complex negotiations required (leaving smaller publishers unable to complete as many deals) and also to provide supplementary

services (like APC administration and accounting). Moreover, publishers entering R&P agreements gain a meaningful competitive advantage in attracting authors, further increasing concentration.[30]

- Inequities in research access would be replaced by inequities in publishing opportunities, as APCs are already too high for many researchers, not only in low- and middle-income countries but also in less-funded institutions in high-income countries. Waivers programs are discriminatory, both because they formalize different categories of authors and also because their actual operations are inconsistent, something which is recognized even within STM (International Association of Science, Technical, and Medical Publishers).[31]

- Publishers like Elsevier that also operate a data analytics business may be handed vast amounts of additional data on grants (and on grants spending patterns), further contributing to consolidation of the data analytics sector into a small number of competitors with excessive market power.

One final word: R&P agreements will likely prove transitional, rather than transformative. European and global funders that participate in Coalition S will disqualify hybrid journals by 2024. However, the trend in the US is toward achieving shorter embargo periods for OA and ultimately for eliminating them altogether. In this context, R&P agreements may prove less relevant to the economic model of scholarly communications than is perceived today.

30 The Impact of the German 'DEAL' on Competition in the Academic Publishing Market. https://www.dice.hhu.de/fileadmin/redaktion/Fakultaeten/Wirtschaftswissenschaftliche_Fakultaet/DICE/Discussion_Paper/360_Haucap_Moshgbar_Schmal.pdf

31 https://scholarlykitchen.sspnet.org/2021/04/19/guest-post-apc-waiver-policies-a-job-half-done/?informz=1

HOW TO RESPOND

The *SPARC 2019 Roadmap for Action* accompanying the *SPARC 2019 Landscape Analysis* outlined a three-pronged approach to addressing issues stemming from the increasing pursuit by commercial vendors of data and data analytics in the academic community. Of particular concern is that the academic community is fragmented at many levels, hindering its capacity to respond to challenges posed by the deployment of approaches and technologies that the academic community is unprepared to manage.

At the base of SPARC's recommendations were three critical ideas that supported the suggested way forward:

1. Base actions and plans on principles.

It is vital to identify a structured set of principles that represent a foundation and a compass for action. SPARC has identified principles used across many organizations when dealing with artificial intelligence,[32] complemented it, and given examples of how these principles can be translated into actual contractual clauses (**Exhibits 5 and 6**). A robust debate around these principles will make the list even more useful.

Exhibit 5: Principals of Data Analytics Usage*

- Clear areas of application
- Equity
- Transparency
- Strong privacy protection
- Accountability
- Human control
- Customization
- Governance
- Avoidance of conflicts of interest

*Based on *Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI Source* published under the auspices of the Berkman Klein Center at Harvard University

32 In particular, please see https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3518482

Exhibit 6: Possible Contractual Terms and Conditions/Code of Ethics Requirements

PRINCIPLES OF DATA ANALYTICS USAGE	POSSIBLE CONTRACTUAL TERMS/DATA POLICIES
Clear areas of application	Upfront exclusion of specific uses
Equity	Right to demand correction of biases
Transparency	Right to demand third party inspection/audit/ evaluation
Strong privacy protection	Explict control over use of data; right to erasure/ correction
Accountability	Ability to appeal; human responsibility for use and outputs
Governance	Effective input of all stakeholders
Human control	Opt-out of automated decisions; human review of recommendations

2. Increase coordination, alignment and—where possible—cooperation within the knowledge community.

The academic community is fragmented into in a variety of institutions that struggle to operate as a coherent group. Academic institutions operate in different countries and regions, balance research and teaching differently, and often compete for funding, for faculty, and for students. In addition to this fragmentation is a long tradition of independence and self-management of schools and departments within large universities, as well as a distinctive separation between academic institutions and learned societies.

A significant change to this culture of decentralized decision making is unlikely. However, important opportunities exist for both closer cooperation and community realignment. It is certainly difficult to ask schools and departments that are historically independent to abandon that position, but these entities will likely support coordination if it helps them. Better procurement processes, help and support in the management of data sets and flows, and updated policies that focus on the best way to support authorized uses of

data (as opposed to merely protecting from unauthorized access, as is often the case today) could prove popular even in highly decentralized institutions. Similarly, in recent years learned societies have operated at arm's length from (and sometime in conflict with) academic institutions.

3. Support structured processes involving as many stakeholders as possible to define responses.

Because of fragmentation, each institution will likely define its own approach and develop actions and programs to fit its specific culture and priorities. Many of these decisions will require the in-depth understanding of specific legal, ethical, business, computer science, and economics issues and solutions. It is just as important to bring in the voices of all stakeholders, including those in communities that may be least protected from the impact of the actions and programs.

Specific Actions

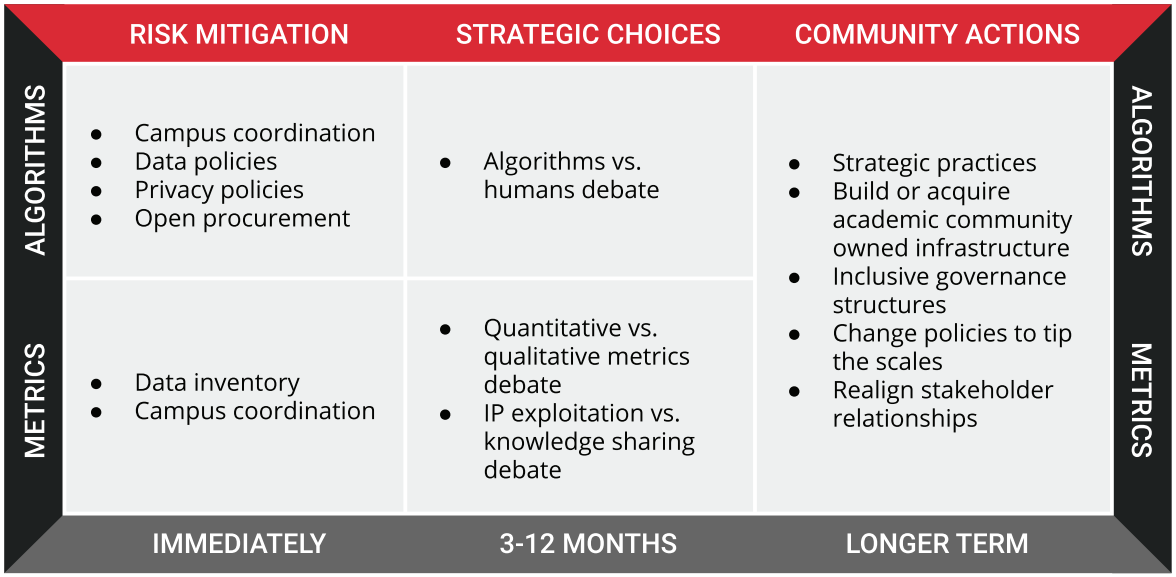
The *SPARC 2019 Roadmap for Action* that accompanied the original *SPARC 2019 Landscape Analysis* also proposed several specific actions to manage the impact of data analytics and AI on academic institutions and their communities. The proposed actions (**Exhibit 7**) fall into three categories: Risk Mitigation, Strategic Choices, and Community Actions.

Though this diagram is not intended to be prescriptive, the algorithms describe some examples of potential steps to promote open infrastructure, and the metrics describe how to measure the success of those efforts.

Metrics (what to measure) should be clearly differentiated from algorithms (how to measure). Metrics should be determined only by the academic community, while algorithms can come from a variety of sources, although their use must be subject to the principles we outlined earlier.

Exhibit 7: The SPARC Framework for Action

Open Data Infrastructure: A Roadmap



In addition to the actions proposed in 2019 and 2020, there are three additional courses of action to consider this year: **organizational change**, **cooperation**, and **community-controlled infrastructure**.

- **Introduce necessary organizational change.** Institutions have sometimes treated data and infrastructure as an afterthought, but one lesson of 2020 is that understanding what data and information are collected, by whom, for what purposes, and with what protocols, is a necessity. It is also critical to decide when and how the information should be used, what principles and codes of ethics should apply, and what control the academic community should exercise over the process. This points to the growing need for institutions to establish key roles: chief data officer and university (or college) ethicist.

  The chief data officer (CDO) role is becoming a familiar one. For example, Title II of the Open Government Data Act[33] (passed into law on December 31, 2018 and

33 This became Public Law No: 115-435. https://www.congress.gov/bill/115th-congress/house-bill/4174

signed on January 14, 2019) requires every federal agency to appoint a chief data officer and defines a list of strategic tasks. Educause (the largest community of technology, academic, and campus leaders advancing education through IT) has a higher education chief data officer working group. In many institutions, this role may need to be upgraded: the position's objective should include managing the strategic uses of data and the principles that should be adhered to when using data tools in addition to the operational ones, while collaborating with IT departments on IT issues. Many CDOs today are instead focused on more technical activities around data collection, preservation, and security.

The proper role of the CDO is not managing data (which should continue to be the responsibility of individual offices and departments), but rather developing strategies, policies, procedures, and guidelines, as well as transferring best practices. Most current institutional data policies are almost exclusively oriented towards security, focusing primarily on limiting the risks of unauthorized access to different types of data, and they do not address the strategic uses of data and the non-negotiable principles which should be required for authorized data access. It is also critical that CDOs work with their peers to identify—and demand that vendors correct—the systemic biases that characterize virtually all algorithms.

The role of university/college ethicist is a newer concept. Several issues posed by the adoption of data analytics have no single, straightforward answer that suits every campus. Is it acceptable for an academic institution to screen applications with software or to monitor the online behavior of applicants? Is it fair to use software to detect possible online cheating in spite of reports that it may disproportionately single out minorities, women, and some people with disabilities? Is it acceptable to use software that predicts possible violent behavior of students and staff and take action for acts that have not been yet committed? Should research data in relevant disciplines be aggressively harvested for commercial purposes to make up for falling income from other sources, or should data be made open for society at large to profit? Institutions will legitimately come to diverging conclusions.

The role of a university/college ethicist is to lead and facilitate the institution's response to these and the many additional issues that data analytics and AI will

pose. Ideally, answering these questions would be based on collecting inputs from disciplines such as ethics, law, economics, and computer science and then consulting representatives of all the categories of people affected, such as faculty, staff, students, and administration. However, this process must be properly organized and managed, and having an individual explicitly tasked with leading this process is crucial. In addition, a university/college ethicist would be available to consult with anyone seeking ethical advice when studying an academic initiative. This role should be distinct from those of both an ombudsperson (who is generally charged with collecting and investigating allegations of misadministration or violation of rights and codes and who may also be alerted to violations of the law) and that of a compliance officer (who is primarily tasked with reducing legal risks for an academic institution).

The role of chief ethicist is a novel concept, but it is not a completely unknown one. In the corporate world, by 2019, several companies had identified individuals to steer corporate decisions in accordance with values.[34] Some pioneering examples are also found in the academic community (e.g., Penn State and UC San Diego[35]). However, a review of several current academic searches for these positions indicates that many academic institutions view this role as overlapping with the management of compliance, narrowing the role to support of legal and audit activities and reducing its visibility.

In a prescient article published in 2003, John B. Bennet argues that ethics officers in a diminished role could do more harm than good[36] by conveying to the faculty and staff that they do not have to exercise personal ethical judgment (since someone else is now in charge) and by failing to influence institutions if they do not have the ear of governing boards. Bennet's article underscores the need for these appointments to be made at senior level, be focused on ethics rather than compliance and auditing, and have regular and unfiltered communications with governance boards.

34 https://www.forbes.com/sites/insights-intelai/2019/03/27/rise-of-the-chief-ethics-officer/

35 Please see https://news.psu.edu/story/631684/2020/09/14/administration/penn-state-names-corporate-leader-attorney-university-ethics and https://chancellor.ucsd.edu/about-the-office/judy-bruner

36 https://scholars.fhsu.edu/cgi/viewcontent.cgi?article=1008&context=alj

- **Pursue cooperation.** Individual academic institutions are generally not able to negotiate from a position of strength with publishers that have access to much more information, such as the prices paid by comparable institutions, or how many different customers the publisher may have in a country or institution for their different products.

  Investing in community-controlled infrastructure is the most obvious next step, but not the only one possible. Academic institutions could work with learned societies, for example, to lure them away from their dependence on subscription revenues from publishers. Many societies must be wondering whether Elsevier's decision to sign a memorandum of understanding with the University of California that has no "reading" revenues and includes society journals in their OA scheme should alarm them. Their leaders, regardless of existing short-term agreements with publishers, must be wondering whether—over time—they can count on Open Access revenues to replace what they earned through their share of subscription revenues. This event presents a unique opportunity to launch programs aimed at aligning the interests and capabilities of academic institutions and societies (and perhaps some publishers).

  Other opportunities for collaboration include, for example, pursuing advocacy on specific themes of common interest (such as surveillance and the sale of data to third parties), supporting litigation and antitrust actions, funding and developing open educational resources, and lobbying for student protections against inclusive access and for digital circulation rights.

- **Invest in community-controlled infrastructure.** Corporations move fast—often much faster than academic institutions. Since the November *SPARC 2019 Roadmap for Action*, the pandemic has understandably set back plans for community investment in infrastructure. However, commercial players have continued to advance their plans for leveraging data analytics and further entrenching themselves in critical academic processes. Senior leaders of academic institutions still have an opportunity to mobilize the financial resources and talent necessary to develop community-owned infrastructures that both support open and equitable dissemination and preservation of research communications and the attached metadata, and that also allow analyzing those metadata to help senior decision makers manage their institutions by their own priorities.

Considering the benefit to the community, the resources required to fund such a project may be a wise investment. Building a fully functioning research dissemination and data analytics company may require an investment of less than $40–50 million, but this money must be raised, and that leads to questions of whether this is best accomplished by partnerships between the academic community and the private sector, between the academic community and NGOs, or between the academic community and governments. In turn, this requires understanding if there is an opportunity to build and operate a sustainable community-owned infrastructure, how it should be funded, and whether the intellectual and knowledge output of academic institutions should generate financial resources to fund this infrastructure. The launch of Invest in Open Infrastructure (IOI) provides appropriate coordination for the academic community to develop a full community-controlled infrastructure. Alternatively, leaders from research institutions around the world should commit to building this infrastructure, with the support of funding bodies, if necessary. This leadership group would commit to designing the infrastructure to further the interests of the global academic community, and not just those of wealthy countries or institutions.

The choice between open and closed data and knowledge has implications along a spectrum of issues extending beyond funding academic knowledge infrastructure. For example, open data raises national security and economic competitiveness issues, as well as questions about academic freedom, academic priorities, and even the fundamental goals of academic institutions. Launching a structured process to analyze these implications appears a critical step that leaders of academic institutions need to take sooner rather than later.

4. Learn from the pandemic.

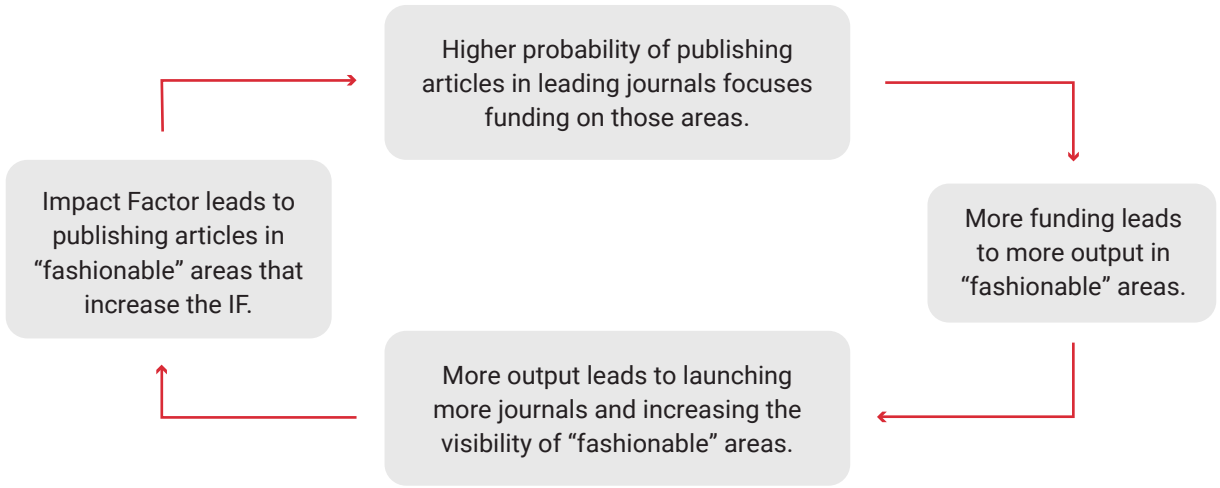
In addition to the initiatives SPARC identified as part of the *SPARC 2019 Roadmap for Action*, the subsequent 18 months demonstrated both the value of science and knowledge and the necessity of fostering open science practices. These two broad societal themes must be pursued, both because the pandemic has demonstrated the need to ensure equitable access to progress in health care practices and because other looming

challenges such as climate change and the loss of biodiversity require major advancements in research.

- **Foster equitable open science practices.** Open science cannot be equitable if research is inequitably focused on the most privileged members of society. The weight accorded to leading journals because of their impact factors (IF) has given these journals the incentive to operate a *covert science policy*: publishers and editors have incentives to maintain or raise their IF, and this leads them to prioritize publishing articles that are likely to be widely cited. This means they will prefer to publish articles in areas that are "fashionable" and of wide interest, and this focus of the leading publishers in turn affects funding and the priorities of funding bodies (**Exhibit 8**). Unfashionable disciplines and approaches (like those affecting rare diseases or people in disadvantaged communities) are structurally disadvantaged by these dynamics.

  SPARC has been aware of these issues for a long time because of its global work, but vast groups within the academic community have not yet focused on them. The academic community must acknowledge that it cannot be held hostage to impact factors, not just because of their limitations in assessing researchers

Exhibit 8: The Vicious Circle Triggered by the Impact Factor



when making key tenure and promotion decisions, but also because they foster fundamental inequities.

When the Declaration on Research Assessment (DORA) was released in 2012, the distortions caused by the IF were well understood,[37] yet little changed over the next eight years. In 2019 the National Academies of Sciences launched a roundtable to define new incentives for open science that align tenure and promotion decisions to virtuous behavior. It is vital that the academic community support this process and translate the roundtable's recommendations into policies and actions.

- **Raise societal investments in knowledge as a critical priority.** The academic and research community has achieved major accomplishments since the onset of the pandemic in early 2020, leading to the fastest identification of a new virus and development of vaccines, treatments, and protocols ever seen. At the same time, most academic institutions were able to provide continuity in teaching, learning, and research services. More broadly, knowledge has helped mitigate the risks of a pandemic that some expected but few were properly equipped to tackle. The academic community should build on this success to demand much more support by the rest of society.

  Politicians and regulators increasingly recognize that open knowledge dissemination accelerates progress. As part of its infrastructure plan, the Biden administration initially proposed an ambitious increase ($250 billion over several years) in the government's research and innovation investment. Though this seems like a large amount, it is less than 10% of the total additional infrastructure spending the administration initially proposed, and *it remains to be seen what will happen as the plan goes through Congress*.

37 https://science.sciencemag.org/content/340/6134/787.full

In this report, we have pointed out how the past year has seen more deals that led to even more concentration, to loss of diversity, and ultimately further eroded the academic community's control over its destiny. We have also highlighted some positive signs: a large merger failed, Invest in Open Infrastructure was launched as a concerted effort to build a community-owned infrastructure, and some legislative progress has been made.

Much remains to be done, but we see many emerging signals that the academic community understands that regaining control over its content, its data, and its infrastructure is vital to achieving its objectives and staying true to its values. We look forward to continuing our efforts to support the knowledge community as it regains control over these critical elements.

APPENDIX

The performance of key companies in the past 12 months

RELX

In 2020 the company continued to show modest growth in its less-cyclical businesses, with 2% underlying revenue growth (i.e., excluding the impact of acquisitions and divestitures closed in the previous 12 months and changes in foreign exchange rate) for the combination of scientific, technology, and medical (STM) business—usually referred to as Elsevier—which grew 1%. The Risk and Business Analytics segment grew 3%, and Legal grew 1%. The Exhibitions business was deeply affected by the pandemic and lost 71% of its revenues, leading to a combined –9% revenue decline for RELX overall. The company's three more stable businesses grew their combined adjusted operating profit by 4%, while Exhibitions swung to a £164 million (approximately US$211 million) loss from a £331 million (approx. US$425 million) profit in 2019, leading to an overall –18% decline in the total adjusted operating profit for the company. The STM business grew headline revenues by 2%, but only by 1% when the impact of exchange rate and acquisitions/divestitures is stripped out. (This impact is almost entirely attributable to foreign exchange variations.) **Headline adjusted operating profit** was up 4%, but—once again—when stripped of the impact of foreign exchange variations and acquisitions/divestitures, profits grew by only 1%.

Headline adjusted profit *uses a company's income from operations, trading, and investments, excluding exceptional items like write-offs or acquisitions. It is therefore a way to assess how well a company is doing during "business as usual."*

In the presentation of annual results, RELX's management indicated that STM revenues from digital products grew by 3% in 2020, while print declined at a pace double that of recent years. The company also took credit for opening its content in response to the COVID-19 pandemic. According to its presentation of 2020 results, RELX "mobilized content and data analytics expertise in support of its global response to Covid-19 pandemic" [by

providing] "50,000+ articles and 200 million downloads."[38] Predictably, it failed to discuss its initial reluctance to open up content—reluctance which was shared by other publishers. For 2021, management issued guidance to another year of modest underlying revenue growth. There was no discussion of the resilience of the subscription model during downturns, in contrast to the messages management had communicated in previous updates during 2020. The business remains very profitable, with an adjusted operating margin of 37.2%.

While the overall financial position of RELX deteriorated visibly, it remains strong. Free cash flow declined by –28%, or £481 million (approx. US$617 million). Over 85% of the decline was driven by the £415 million (approx. US$533 million) decline in adjusted operating profit. Broadly speaking, all other sources and uses of cash were unchanged relative to 2019. However, in 2019 the company used £600 million (approx. US$766 million) to buy back shares. The suspension of the share buyback after the first quarter led to a much smaller cash outflow (£150 million, equal to about US$193 million). The company, however, did increase spending on acquisitions (£878 million/US$1,127 million, vs. £416 million/US$531 million in 2019). This increased outflow, coupled with the lower cash flow from operations and the fact that currency translations did not contribute to cash inflows in line with 2019, means that year-end net debt rose by 11.4% from £6,191 million (US$8172.1 million) to £6,898 million (US$9415.8 million). **Net debt/EBITDA** rose from 2.5 at the end of 2019 to 3.3 at the end of 2020. While not excessive, this level of debt is higher than the 2.5 to 3.0 target net debt/EBITDA ratio the company has historically pursued. Management seemed to regard this unusually high level of debt as a temporary event driven by the decline of the Exhibitions business; this statement seems to rule out an issuance of new shares to repay some debt and return the company to its target

EBITDA *refers to earnings before interest, taxes, depreciation, and amortization.*

Net debt/EBITDA ratio *indicates a company's ability to pay off its debt. The lower the ratio, the higher the ability of a firm to pay off its debt. Many analysts consider ratios lower than 3 acceptable and higher than 4 a possible indicator of future distress.*

38 https://www.relx.com/~/media/Files/R/RELX-Group/documents/results/results-presentations/2020-results-presentation.pdf

debt range. As usual, management provides very little guidance to the financial markets for 2022; it only expects to have another year of revenue and profits growth (and it further qualifies this guidance by excluding the Exhibitions segment).

Wiley

Relative to other companies, Wiley had a very good year. Overall revenues grew by 4%, with growth recorded in two out of three reporting segments. In Research, organic (i.e., excluding the impact of foreign exchange variations and acquisitions/divestitures) revenues grew by 3%; in Education Services, organic growth was 7%; finally, in Academic & Professional Learning, organic revenues declined by –3%. Good revenue growth led to significant improvements in profitability: EBITDA grew by 16%, cash from operations grew by 25%, and free cash flow grew by 48%. These significant improvements in profitability reflect both the aggressive cost saving programs that Wiley, in line with all publishers, executed early in the pandemic as well as the high operating leverage of mostly digital businesses: in the transition from print to digital, publishers drastically reduce their variable costs, and additional revenues flow almost directly into EBITDA, earnings, and cash flow.

The research business of Wiley did very well during fiscal year 2021. Publishing (which accounts for 96% of research revenues) grew by 5%, and platforms by 7%. Underlying activity was also strong. Wiley reports that it published 15% more articles than in fiscal year 2020, and that usage (i.e., downloads) grew by 25%. In the past, Wiley reported revenue growth for subscriptions and open access, but it has discontinued this practice. All that is known is that OA revenues grew in aggregate by 38%, but the number is not meaningful, since Wiley closed its acquisition of Hindawi on December 31, 2020, and therefore the OA numbers include one quarter of Hindawi revenues. The business continues to be profitable: the business has operated for years at about a 35% **EBITDA margin** and a 29% contribution to profit (largely comparable to RELX's 37% operating margin).

EBITDA margin *measures a company's operating profit as a percentage of its revenue without considering interest, taxes, debt, and amortization. It is a widely used way to compare companies based on what they earn, since it strips out the impact of a company's capital structure and tax policies and focuses on operating performance. EBITDA margins are best used to compare companies in the same industry.*

Education and professional services revenues declined by –3%, but with diverging performances for Education Publishing, which grew by 2%, and Professional Learning, which declined by –8%. This segment also remains quite profitable (albeit less so than the research segment) with an EBITDA margin of 25%.

The financial position of Wiley remains strong, with a net debt/EBITDA ratio of 1.7x (including the impact of the Hindawi acquisition).

Wiley provides detailed guidance, unlike some of its peers. The company offered a numeric range for revenues, EBITDA, EPS (Earnings Per Share), and cash flow. Taking the actual 2021 results and comparing them to the midpoint of the guidance, management guides these growth rates: 7.4% revenue growth, 1.4% EBITDA growth, flat earnings per share and a –18.3% decline in cash flow. The expected muted growth of earnings is attributable to the return to spending for in-person business and for some investments, and higher capital expenditures also explains the large decline in expected cash flow generation.

Pearson

The 2020 results were poor, but unsurprising. The company reported a 10% revenue decline and an adjusted operating profit of £313 million ($US402 million).

Pearson's higher education courseware business declined by –12% in the US and by –13% in North America overall. It is worth underlining that this performance was visibly worse than that of Cengage (+1% in the March–December 2020 period) and McGraw Hill (+2.5% in the March–December 2020 period). Even if the periods are not directly comparable, Pearson appears to have lost significant share. In part, this share loss is driven by the lower percentage of revenues deriving from digital products (70%, vs. 81% at McGraw Hill and 83% at Cengage) versus print textbooks. Because print textbooks sales are generally higher priced, the loss of print sales affects revenues disproportionately. Looking ahead, Pearson expects US higher education revenues to rise in 2021 and 2022 on the back of increases in enrollment, the erosion of the "secondary" book market determined by the declining availability of used textbooks and the hope to stabilize or regain

some of the lost market share. Whether these forecasts will prove correct remains to be seen, of course.

The company continues to have a strong financial position. Net debt/EBITDA declined from 1.3x at the end of 2019 to 0.8x at the end of 2020 in spite of the much lower EBITDA (–45% compared to 2019) and operating cash flow (–25%). In large part, the debt reduction is due to the disposal of the remaining 25% stake in Penguin Random House (PRH), which contributed £530 million (US$681 million). In addition, management decided in March 2020 to stop returning money to shareholders (over and above the ordinary dividend, which was left stable) through a share buyback. At the time of the decision to halt the buyback, the company had £183 million left to spend. Adding up the proceeds from the PRH disposal and the share buyback halting, the total accounts for 129% of the net debt reduction. In other words, the PRH disposal and the halting of the share buyback went into debt reduction and into covering the lower cash generation from operations. Going forward, the company plans to focus on further reducing operating costs, including through a symbolic downsizing of its central London headquarters.

Pearson offered generic guidance for 2021, something that was not a given in light of current uncertainties. The company expects to achieve revenue growth in 2021 (with a smaller decline in the US higher education courseware business relative to 2020). There was no guidance on profitability for 2021.

Pearson's CEO Andy Bird made a reference to sustainability in his presentation, focusing on the company's carbon footprint and the goal to become carbon neutral by 2030. This is not a particularly ambitious goal, since the company expects to phase out print products altogether in the next five years, and no reference was made to social equity issues beyond the generic statement that "education should be affordable for everyone" at the beginning of the presentation.

The new Pearson strategy comes as no surprise. From the very beginning, Andy Bird was clearly bound to turn the company into a much more consumer-oriented company. Bird, who has an extensive background in consumer media, announced earlier in 2021 several senior appointments of people coming from the same mold. This is not an entirely

new idea. Four years ago, previous CEO John Fallon had hired a chief strategy officer who came from the market research industry in the hope of focusing the company on consumers. This is a welcome change from the old Pearson culture of viewing boards of education and faculty as their customers, although it remains to be seen what this will mean in substance.

Bird has also reorganized the company along new reporting lines, both for the purpose of managing the businesses and to provide financial information to the investment community. The company has reorganized several times since 2006. Surprisingly, John Fallon announced the last reorganization just at the beginning of 2020, when he had already announced he would be leaving. Investors have seen geographic divisions, products divisions, matrix organizations, and even a mix of geographic and product divisions. Bird is returning Pearson to a pure product divisions structure, with the goal of moving as many costs as possible into the divisions rather than centrally (except where it is financially unjustified).

Bird also announced in March 2021 that the higher education courseware business will be merged into one global division called Higher Education, but that all the higher education businesses outside the US (including Canada) are effectively for sale (although a formal decision to do so has not been made, and these activities will still be reported for the time being as part of the higher education business). Since the US accounts for almost 90% of all higher education courseware revenues, the divestiture of the non-US businesses would allow management to focus further into the only market where Pearson has real scale.

Over the years, investors have been treated to many widely diverging messages about the sources of Pearson's competitive advantage: it was, at different times, a company rooted in sound educational science and research, or at the forefront of the digital classroom revolution, or a global digital platform on which local content could be easily overlaid, or the operator of virtual classrooms (both in K–12 and college). Now it is the turn of the consumer-oriented company that looks at Spotify and Netflix as the examples of what consumers want. Investors are right to be confused.

It is definitely true that failing to understand that students were decision makers alongside faculty, and that they could rebel against paying high prices for textbooks, was a fatal weakness. However, failing to understand—at least until recently—that the quality of content still matters to faculty, and that students could pressure their faculty to adopt less expensive materials like open education resources (OER) could leave Pearson in a position where it does not know how to compete. Pearson recently chose to launch Pearson+ and compete directly with Cengage Unlimited in offering deeper discounts, in spite of having no obvious incentive to do so as the market leader. This decision may well be driven by the continued share losses, but it raises questions as to whether Pearson can return to sustainable revenue growth in US higher education courseware, as deeper discounts may negate over time any progress obtained by eliminating the secondary textbook market.

McGraw Hill

Total company revenues declined by –4% for the fiscal year which ended on March 31, 2021, in large part because of the sharp decline in print revenues (–21%), which were almost totally offset by a 10% increase in digital revenues. It should be noted that, differently from Cengage, which earns 12% of its revenues from high schools, K–12 accounts for 35% of MGH's 2021 revenues (and 38% in 2020). The adoption calendar for states, as well as the mix between print and digital adoptions, drive a significant volatility both in terms of the total weight of K–12 revenues and the mix between print and digital. Excluding K–12, digital revenues for MGH grew by 17% in FY 2021. In spite of the overall revenue decline, the aggressive cost savings program executed to manage the impact of the COVID-19 pandemic allowed MGH to post a significant rise in EBITDA, which grew by 18%. Aggressive cash conservation actions increased the cash flow of the company by 50%.

The US higher education business performed well in the past year. Total billings grew by 5% as a result of a further sharp decline in print revenues (25%), which was more than offset by a 15% increase in digital sales; the EBITDA of the higher education business grew by 28%. As in the case of Cengage, MGH reported market share gains (120 basis points per share vs. 100 bps for Cengage). Though numbers are not available for all

publishers, Pearson reported a 12% revenue decline for its US higher education business. The numbers are not perfectly comparable, because Cengage and MGH both close their fiscal year in March, and the first quarter of the calendar year is syphoning away revenues from Q4 of the previous calendar year (this shift is the consequence of print sales being recorded at the time when books are shipped, while digital sales are recorded when licenses are activated). Nonetheless, the decline of Pearson's revenues in calendar year 2020 is likely to represent a major source of market share gains for both MGH and Cengage. Very much as in previous years, MGH attributes a major role to Inclusive Access in the growth of its digital business and in its market share gains.

The *2020 Update: SPARC Landscape Analysis & Roadmap for Action* spotlighted the high levels of debt incurred by MGH, particularly in the aftermath of the failed merger with Cengage. The company has worked very hard to reduce debt, and its net debt/EBITDA ratio stands at 4.0x, a marked improvement compared to 5.4x at the end of 2019 and 7.0x at the end of 2018. While 4.0x is still a relatively high number, it is now starting to approach a comfortable level even in case financial markets were to tighten (there have been periodic scares in the spring of 2021 about possible, unanticipated interest rates increases because of the resurgence of inflation). It is unclear how much the sharp improvement in cash generation achieved in 2020 can be sustained, because cost cutting and lower capital expenditures can easily be reversed, but even maintaining the current revenue/spending ratio should allow MGH to reduce its net debt/EBITDA ratio to close to 3.0x in a year. In June 2021, Apollo Global Management sold McGraw Hill to Platinum Equities, and the deal closed on August 2, 2021. It will then be the new owners who will refinance the debt of the company. At that point, they will decide whether to continue to pay down debt or increase leverage again to pay for some of the $4.5 billion they agreed to pay for the equity of the company.

Cengage

Company cash revenues declined by –6% for the full fiscal year (which runs from April 1 through March 31). This performance covers the entire course of the COVID-19 pandemic until now, as the impact in Q1 of calendar year 2020 was minimal. The –6% decline is a much better result that management may have feared one year ago: at the

time, management indicated it had developed a number of scenarios for revenue decline, and the most pessimistic scenario was based on a –25% revenue decline. The decline was primarily driven by international higher education, which declined by –19%, by the secondary school business (–16%) and by English language teaching (–31%). Earnings before preprint expenditures were flat at $315 million, a particularly strong performance in a business that is largely based on fixed cost businesses (like most media companies, and particularly so with digital media businesses). The flat earnings were largely achieved through aggressive cost cutting, and it remains to be seen how sustainable these cuts will prove over time.

The US higher education business cash revenues grew by 2%, and net sales grew by 4% for the full year. This growth, coupled with cost cutting and the steady migration from print to digital courseware, led to a 14% increase in EBITDA. Cengage reports this commonly used metric by subtracting prepublication costs, since management views prepublication as akin to investment rather than as an ongoing expense. In its call with investors, management underlined the continued focus on Cengage Unlimited (CU) and on inclusive access: Cengage does not break down the individual components of what it defines as "institutional sales" (a category that includes both CU and inclusive access), but it communicated that the entire category grew by 40% over the preceding 12 months. Finally, management estimates it grew its US higher education market share to 26.1%, a 100 basis points increase in the preceding 12 months. Since McGraw Hill has also guided to continued market share gains in the same time frame, it is reasonable to assume that share has been taken away from Pearson, which had a poor start in the first part of 2020, and possibly from smaller publishers.

In the case of Cengage, high debt has been a cause of concern, particularly early in 2020 when yields on Cengage's debt spiked under the twin effects of uncertainty triggered by the pandemic and by the failed merger with McGraw Hill. At the end of March 2021, Cengage had reduced its net debt/EBITDA ratio to 5.6x from 6.1x one year ago. This ratio is still quite high, and further reduction will likely have to come from significant revenue growth. The company has done a great deal to reduce costs, and further cuts will be more difficult to achieve. In addition, cash flow growth in 2021 was significant: unlevered free cash flow (i.e., cash generated irrespective of whether it is used to pay interest) grew

by almost 24%, but the bulk of this growth was obtained by reducing working capital. In turn, some of this is driven by the transition to digital (as physical inventory shrinks and individual digital sales can be monetized sooner). Cengage already stands at 83% digital sales in higher education, and opportunities for further gains will likely decline going forward. Management does believe that further revenue growth is coming, particularly in US higher education, thanks to a combination of enrollment rebound, stimulus spending raising participation rates, the phaseout of the used books market through digital contracts and, possibly, further market share gains.

September 2021