



STM 报告

科技与学术期刊出版综述

庆祝期刊出版 350 周年



Mark Ware

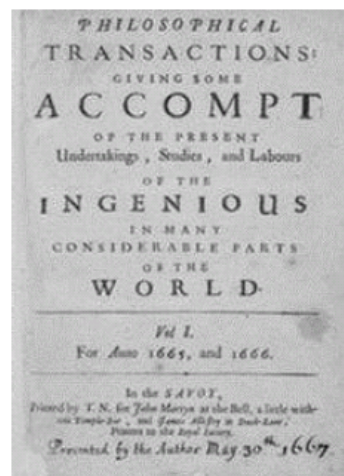
Mark Ware 顾问公司

Michael Mabe

国际科学、技术与医学出版者协会

第 4 版

2015 年 3 月



© 2015 STM: 国际科学, 技术与医学出版者协会

出版

国际科学、技术与医学出版者协会

Prins Willem Alexanderhof 5, The Hague, 2595BE, The Netherlands

STM 简介

STM 是学术和专业出版者的全球领先的贸易协会。STM 拥有分布于 21 个国家的 120 余个成员, 每年出版的期刊论文占全球论文总量的近 66%, 并出版数万种专著和参考书目。STM 的成员包括学术协会、大学出版社、私人公司、新生出版机构和成熟的出版机构。

STM 的目标和任务

- 协助出版机构及其作者在科学、技术和医学领域进行科研成果的传播活动;
- 协助在电子网络时代关注改进科学、技术和医学信息的传播、存储和获取的地区和国际性机构及传媒业;
- 联合国际出版商协会 (IPA) 及相关的地区出版者协会和政府机构或专业团体开展上述工作。

STM 参与信息识别与电子版权管理系统的研发。STM 成员可通过信息快报、STM 网站、电子邮件等途径及时快捷地了解行业信息。STM 组织研讨会、培训课程和行业会议。

Mark Ware 咨询公司 为 STM 及企业间电子商务 (B2B) 提供出版咨询服务。详细信息请参见 www.markwareconsulting.com。

目 录

内容提要	1
1. 学术交流	8
1.1 研究周期.....	8
1.2 学术交流的类型.....	9
1.3 学术交流体系的变化	9
2. 期刊	11
2.1 什么是期刊?	11
2.2 期刊的出版周期.....	11
2.3 销售渠道和模式	15
2.4 期刊经济与市场规模	19
2.5 期刊和论文的数量及趋势	23
2.6 全球科学产出的趋势	32
2.7 作者和读者	40
2.8 出版商	42
2.9 同行评议.....	42
2.10 阅读方式.....	50
2.11 学科差异	53
2.12 引用和影响因子	56
2.13 期刊出版的成本	66
2.14 作者的行为、认知和态度	70
2.15 出版道德.....	74
2.16 版权和许可	75

2.17 长期保存.....	82
2.18 转换准则.....	84
2.19 研究人员对期刊的获取.....	84
3. 开放获取.....	90
3.1 开放获取驱动力.....	91
3.2 开放获取的商业模式.....	92
3.3 OA 期刊的类型.....	100
3.4 延迟 OA.....	103
3.5 通过自存档实现 OA（绿色 OA）.....	103
3.6 其他的 OA 衍生模式.....	118
3.7 SCOAP3.....	119
3.8 学术书籍的开放获取.....	120
3.9 公共获取.....	121
3.10 全系统和经济视角研究.....	121
3.11 开放获取的其他进展.....	123
3.12 转型和可持续发展问题.....	126
3.13 自存储对期刊的影响.....	131
3.14 开放获取对论文使用的影响.....	134
4. 学术交流新进展.....	137
4.1 “科学 2.0”或“开放科学”.....	138
4.2 FORCE 11 和“科学转型”.....	139
4.3 出版平台和应用程序接口（API）.....	139
4.4 社交媒体.....	140
4.5 移动接入和应用程序.....	142
4.6 研究数据.....	144

4.7 语义网和语义富集.....	149
4.8 新的文章格式和特征.....	152
4.9 文本和数据挖掘.....	153
4.10 可重复性.....	154
4.11 大数据和分析.....	155
4.12 身份识别和消歧.....	156
4.13 研究管理和分析.....	157
4.14 FUNDREF.....	158
4.15 图书馆出版.....	159
4.16 开放注解.....	159
4.17 学术团体.....	160
4.18 作者服务及工具.....	161
4.19 合作创作和分享工具.....	161
4.20 开放笔记本科学.....	162
5. 结 论.....	163
6. 信息资源.....	170
6.1 出版商组织.....	170
6.2 全球统计与趋势.....	170
6.3 开放获取.....	170
6.4 出版产业研究和分析.....	171
7. 参考文献.....	172

内容提要

学术交流与 STM 出版

1. STM 出版发生于范围更为广泛、包括正式和非正式形式的学术交流系统。学术交流在科学研究周期的不同阶段起着不同的作用，并且（如出版）正经历着技术驱动的变革。学术交流方式有一对一、一对多、多对多、以及口头与书面等，这种分类为分析技术对学术交流的潜在影响提供了有用的框架（P12）。
2. 期刊是学术交流过程中的核心部分，并且是科学研究整体的一个组成部分。期刊不仅传播信息，而且为作者的优先权提供登记机制、通过同行评议维持学术质量、提供一个固定的文档版本供将来的参考等。期刊还为研究人员浏览不断增加的出版材料提供了一个重要的途径（P16）。

STM 市场

3. 英语 STM 期刊出版在 2013 年度的收入约为 100 亿美元（2008 年为 80 亿美元，年均增幅约为 4.5%），2013 年度更广义的 STM 信息出版市场价值大约为 252 亿美元。全球约 55% 的 STM 市场收入（包括非期刊 STM 产品）来自美国，28% 来自欧洲和中东，14% 来自亚太地区，4% 来自其他地区（P23）。
4. STM 出版行业的全球雇员约为 11 万人，其中约 40% 在欧盟地区。此外，除生产供应链方面的就业外，约有 2 万至 3 万全职人员间接从事全球 STM 产业（P24）。
5. 虽然本报告主要侧重于 STM 期刊，但 STM 图书市场（每年的价值约为 50 亿美元）正在快速地向数字出版转型发展。2012 年度电子书约占 17% 的市场份额，其增幅远高于 STM 图书和 STM 市场总体的增幅（P24）。
6. 估计全球有 5000~10000 个期刊出版机构，其中约有 5000 个出版机构的期刊被 Scopus 数据库收录。主要的英语语种的商业和专业学协会中总共有大约 650 家期刊出版机构，所出版的期刊约为 11550 种，约占期刊总量的 50%。其中约 480 个出版机构（73%）和约 2300 种期刊（20%）为非营利性的（P45）。
7. 2014 年末全球出版的同行评议英文版期刊约为 28100 种（另有 6450 种非英文版期刊），每年总计出版的文章约为 250 万篇。两个多世纪以来，论文数量和期刊数量均呈稳步增长的态势，年均增幅分别约为 3.0% 和 3.5%，

并且近年来呈加速增长的迹象。该增幅与科研人员的数量以每年约 3%上升是一致的，目前科研人员的数量在 700~900 万之间，基于不同的界定而变化，但其中只有 20%多次以作者的身份出现（P27）。

8. 美国以约 23%的份额继续主导全球研究论文的产出，但中国和东亚的论文产出增幅最为迅速。中国的论文产出数量已超过 15 年以两位数增长，以占全球总量的 17%而位居第二，紧随其后的是英国（7%）、德国（6%）、日本（6%）和法国（4%）。然而，各国论文被引用总数的排序以美国占 36%在全球遥遥领先，中国以占 6%而位列第 11 位（P38）。

研究行为与动机

9. 虽然期刊的出版方式发生了转变，但研究人员发表论文的主要动机没有大的改变，仍然是为了获取基金资助和个人的职称晋升（P69）。
10. 阅读模式正在发生改变。研究人员阅读的论文更多，每年平均为 270 篇，（医学和自然科学研究人员的阅读量相对较多，人文和社会科学研究人员的阅读量相对较少），但阅读每篇论文所花费的时间更少，由 20 世纪 90 年代中期的每篇 45~50 分钟下降到目前的 30 分钟多一点。科研人员越来越多地通过搜索而不是浏览来获取和选定文章。目前很少有证据表明社交媒体推荐是文献获取的主要来源（与消费者新闻网站不同），不过新的科学社交网络可能会改变这种状况。研究人员通常只花很少的时间在出版商的网站“徜徉”，只是采集一些将来可能需要参考的信息（P52）。
11. 尽管同行评议有不足，但学术界仍然将其视为学术交流的基础并致力于从事这方面的工作。通常情况下，研究人员每年评审 8 篇稿件，每篇稿件花费约 5 个小时。然而，研究产出的增长，包括来自新兴经济体的快速增长，可能会导致文章来源和审稿人之间暂时的不平衡，从而使得同行评议承受着一定的压力（P45）。
12. 同行评议正经历着很多创新，并且演进式的评议方式比激进式的变革更能获得支持。例如，一些开放的同行评议形式（如在论文发表前或发表后公开审稿人的姓名，与论文一起发表审稿意见，等等）正变得越来越普遍。级联评议（Cascade review，即审稿意见随同稿件一起在期刊之间转投）甚至独立于期刊的评议（journal-independent，“便携式”）正在形成一席之地。然而，同行评议实践中最显著的变化是被巨型期刊 *PLoS ONE* 及其效仿者采用的“合理但没有重要意义”的审稿准则已经被广泛采用。取代传统同行评议的发表后评议得到的支持很少，但发表后评议作为传统同行评议的补充已引起一些人的兴趣，例如，公共医学共享（PubMed Commons）作为

一个建立在 PubMed 良好信誉上的发表后评议平台而引人关注。类似地，“替代计量学”（altmetrics）作为对同行评议的一种可能有效的补充，以及在其他方面的影响力而引起人们的兴趣。一项发表后评价的新技术是开放注释，即使用一个新的网页标准允许可被引用的评论显示在任何网站上（P47）。

13. 科研与出版伦理继续受到关注，这一方面体现在出版伦理委员会（COPE）等类似组织机构的重要性不断增加，另一方面体现在检测诸如抄袭剽窃等学术不端行为的技术解决方法的发展。过去 10 年来期刊撤稿的数量大幅增长，但一致的看法是这更可能归因于人们相关意识的提高，而不是不端行为的增加（P73）。

技术

14. 目前几乎所有的 STM 期刊都可以在线获取，并且，出版机构通常都将早期的纸版期刊数字化，以便可以回溯到第一卷。只订购电子版期刊的用户比例大幅上升，这可部分地归因于采用折扣的期刊打包订购模式的驱动。目前绝大多数专业期刊的使用都是基于电子版期刊，有些一般性期刊同时也提供纸版，包括学协会会员期刊和某些领域（如人文领域和某些执业领域）的期刊。已发展成熟的研究领域（如非执业领域）的纸版期刊数量在未来的几年有可能会加速下降（P30）。
15. 社交网络和其他社交媒体对学术交流的影响要赶上它们对更广泛的消费者网络的影响，尚需时日。社交媒体使用扩增的主要障碍在于缺乏清楚的令人信服的证据说明采用这种交流方式的好处大于真实的成本（如时间）。此外，质量和信任也是有关的因素：研究人员对于不经过同行评议且缺乏公认归因方式的学术传播方式的使用仍然十分谨慎。虽然面临这些挑战，社交媒体确实似乎逐渐变得更加重要，如：新近出现的科学社交网络（Academia、Mendeley、ResearchGate）会员的快速增长，在公众中的发展趋势，以及社交特性与出版平台和其他软件的融合等（P72；134）。
16. 同样，对移动设备（智能手机和平板电脑）快速而广泛的应用尚未显著改变大多数研究人员与大多数期刊内容的互动方式：2014 年通过移动设备获取占大多数 STM 平台的流量不足 10%（尽管在临床医学等某些领域该比率明显更高），但这种情况正在改变。医生和其他医疗专业人员采纳此类服务用于专业目的是最快的，通常是获取概要的辅助服务、参考资料或学习材料，而不是发表原始研究论文的期刊。然而，对于大多数研究人员而言，“真正的工作”仍然是完成于笔记本电脑或个人电脑（P24；30；139）。

17. 数据密集型研究的暴发促使出版机构创建新的解决方案来链接出版物和数据（反之亦然），以便于数据挖掘，以及将数据作为一个潜在的出版单元来管理。随着研究数据联盟（Research Data Alliance，2013年推出）的领衔和协调，这方面的变化依然迅速：大多数研究资助机构已经引入或加强政策要求数据存储或共享；数据知识库在数量和类型方面均发展成熟（包括针对“孤儿”数据的知识库）；以及推出使研究数据可被引用、可见和可获取的 DataCite。同时，出版机构的对策包括：与许多社群导向项目的密切合作、创建期刊的数据存储和共享政策及引进数据引用政策、链接或整合数据、创建一些开创性的数据期刊和服务、诸如汤森路透（Thomson Reuters）的数据引用索引的数据发现服务的发展等等（P138）。
18. 文本和数据挖掘已开始生命科学产业的应用中显现一席之地，有改变科研人员使用文献方式的潜力。其重要性预计会因为以下因素而增强：更大数字语料库的驱动、不断增强的计算机功能和易于使用的软件、更广泛获取的内容。大量的倡议（如 CrossRef 的 TDM 工具、PLSclear、Copyright Clearance Center 和 Infotrieve 为 TDM 集成论文内容的服务）目前已经开始基于以下方面出现：授权许可的框架（如 STM 的许可条件）和程序（权利许可，如 PLSclear 和 CrossRef 的 TDM 工具）、TDM 的内容获取和集成（如 CrossRef 的 TDM 工具，Copyright Clearance Center 和 Infotrieve 的集成服务）、标准化内容格式（如 CCC 的导航服务），但这个领域仍处于起步期（P80；146）。
19. 对于资助者和机构而言，越来越重要的研究评估和计量已经体现于诸如基于元数据分析（使用、引用等）的研究分析这样的信息服务的增长，以及诸如 CRIS 工具（当前研究信息系统）这样新的软件服务的增长（P150）。
20. 语义技术已在 STM 期刊中成为主流，至少对于大型出版机构和平台供应商而言是如此。内容的语义富集（通常是使用软件工具自动提取元数据并识别和链接实体）已被广泛用于增强搜索和发现、提升用户体验、加强新的产品和服务、改善内部生产力。全面的语义网络仍然比较遥远，但出版机构开始使用链接的数据，使用语义网络标准使内容更容易被发现和再利用（P143）。
21. 虽然出版商一直提供诸如同行评议和文字编辑这样的服务，有关作者、研究全球化、新技术应用等竞争的增加正在推动作者服务的扩展并更加专注于提高作者的用户体验。其中一个可能出现的新领域是在线协作写作工具：许多初创公司已经开发出相关的服务并且有一些大型出版商声称正在探索这一领域（P153）。

商业模式与出版成本

22. 供给方和需求方都采取聚集方式已成为常态，即期刊以打包形式销售给图书馆联盟（见下文的“开放获取”）。电子书也开始出现类似的销售模式（P19）。
23. 虽然“大额交易”（Big Deal）和类似的打包折扣销售正在不断拓宽研究人员获取期刊的种类，同时也在不断降低单个产品的平均价格，但捆绑销售模式使图书馆承受着需要更大的灵活性和掌控力、需要更合理的定价模式及实际上更低价格的压力。尽管如此，捆绑销售的益处似乎足使其保持一段时间的重要地位，但可能在适用范围（例如捆绑或抵销开放获取费用）和新的定价模式方面不断演化（P21；69）。
24. 研究人员对于学术内容的获取正处于历史高点。内容捆绑及相关的联盟协议模式继续提供前所未有的获取水平，每年的全文下载估计为 25 亿篇，并且篇均下载成本处于历史低点（对于很多大用户而言远低于 1 美元）。各类调查表明科研人员将其获取期刊的状况划归为“好”或“很好”，并且声称他们的获取已有所提高。然而，这些研究人员仍然将期刊论文列为改进获取的第一选择。看来以往优异的获取水平可能不再能满足当前的需要，并且更大量的可发现内容（例如通过搜索引擎）有时也使人们发现并不是所有可发现的内容都能即时获取，从而可能导致挫败感（P83）。
25. “Research4Life 计划”为发展中国家的研究人员提供了免费或非常低成本的获取途径。这个“计划”还在继续扩大，正致力于增加可获取内容的数量和范围以及注册的机构数和用户数（P86）。
26. 最常被提及的获取壁垒是成本障碍和价格，但调查中提到的其他获取壁垒还包括：对可获得资源的认识不足；繁琐的购买程序；数字出版的附加税（VAT）；格式和 IT 问题；图书馆会员的不足；作者或出版者的权益与预期的内容使用之间的冲突（P84）。
27. 通过识别并解决某些特定的获取壁垒或障碍以扩大获取持续受到关注。虽然开放获取已经受到最多的关注，其他探索的思路还包括增加国家许可获取基金以拓展至国家层面的可获取性覆盖范围；通过公共图书馆的预约获取（2014 年英国试点的全国性方案）；诸如中央和地方政府部门、志愿部门、企业等许可获取的发展（P84）。
28. 每篇文章的平均出版成本基于一系列的因素有很大的不同，这些因素包括退稿率（会导致同行评议的成本）、内容的范围和类型、编辑服务的水平

等。据消费者教育与保护协会（CEPA）估计，2010 年基于订阅的纸版和电子版期刊出版每篇论文的平均成本大约是 3095 英镑（不包括非现金支付的同行评议费用）。有关开放获取对成本节约效果的潜力已有很多讨论，而纯开放获取期刊出版商基于他们的财务报告推测了平均的论文出版成本，为这方面的案例研究提供了证据，结果表明开放获取论文出版成本在 290 美元（Hindawi）至 \$1088 美元（PLOS）之间，另外 eLife 的成本数据更高（PP49）。

开放获取

29. 随着新的商业模式的出现，期刊出版变得更加多样化和更具潜在的竞争性。开放获取（OA）使得原创性研究成果通过网络免费获取，大多数论文在再使用和授权许可方面是免费的。有三种 OA 方式：开放获取出版（“金色 OA”，包括完全 OA 期刊和混合 OA 期刊）、延迟免费获取和自存储（“绿色 OA”）（P88）。
30. 在开放获取目录（DOAJ）中列有大约 10090 种（其中 7245 种为英文版）完全 OA 期刊。OA 期刊在某种程度上仍然显示出比非 OA 期刊更少地被 Scopus 或 Web of Science 等选择性的摘要和检索数据库收录，部分地反映出 OA 期刊的创办时间更新近以及在规模总体上比其他类型的期刊小（有一些很知名的期刊例外）。因此，在每年出版的 200 万篇论文中，OA 期刊所发表的论文比例大大低于非 OA 期刊。最新的估计是 OA 期刊发表的论文数占论文总数的约 12%（而 OA 期刊数量占期刊总量的 26%~29%），其中超过 5% 的论文通过出版商的网站延迟免费获取，另有 10%~12% 通过自存储版本的方式免费获取（P31；98）。
31. 金色 OA 有时被视为文章发表费（APC, article publication charge）商业模式的代名词，但严格来说仅仅是指期刊一经出版即立刻提供开放获取。然而，Scopus 收录的相当多部分的金色 OA 论不涉及 APC，但使用其他模式（如机构支持或赞助）。APC 模式本身也变得更为复杂，可变 APC（例如基于论文篇幅）、折扣、预付款和机构会员制度、混合性出版的补偿和捆绑销售安排、个人会员体系，等等（P91；93）。
32. 基于 APC 的金色 OA 有很多潜在的优点。它在规模上可与研究产出规模的增长相一致地发展，有潜在的系统范围的节省，并且再使用得到简化。资助机构通常会偿付发表费用，但即使有广泛的资助机构支持，有关大学内部的资助安排细节仍有待于充分地研制。还不清楚 OA 出版费用的市场在哪里：目前该费用低于论文出版的历史平均成本；大约 25% 的作者来自发

展中国家；只有 60%的研究人员有专门的研究经费；并且多数研究密集型大学仍然担心对他们项目预算的影响（P90；123）。

33. OA 出版导致了一种新型期刊的出现，即所谓的巨型期刊。以 *PLOS ONE* 为例，巨型期刊有三个特点：全文免费获取且出版费用相对较低；基于“合理性但不考虑科学意义”的快速“非选择性的”同行评议（即：论文的选择取决于研究工作的科学合理性，而不是基于对其影响力的主观判断、多取决于其对于某个特定研究群体的重要性或相关性）；更为广泛的学科范围。巨型期刊的数量在持续增长：表 10 列举了大约 50 个案例（P99）。
34. 研究资助机构在学术交流中起着越来越重要的作用。他们期望通过强调问责制和传播来衡量和提高研究投入的回报。这些因素促使他们支持并强制开放获取（以及相关的鲜有争议的数据共享政策）。这些政策也增加了诸如影响因子等计量指标的重要性（有些人认为具有弊端），并且最近正在创建研究评估服务的市场（P88）。
35. 绿色 OA 和知识库的作用仍然存在争议。机构知识库的作用或许比学科知识库（尤其是公共医学中心，PMC）的作用更低。缺乏独立可持续的商业模式意味着绿色 OA 并不能颠覆（订阅模式的）期刊。正反两方面的证据都有：PEER 项目发现 PEER 开放知识库中论文的可获取性对出版商网站的全文下载没有负面影响，但这与那些能够在更早创建且更知名的 arXiv 和 PMC 知识库中可持续获取其大量期刊内容的出版商的体验相矛盾。对 PEER 使用数据的研究还进一步证实能够通过知识库中绿色 OA 获取的期刊论文使用的半衰期更长，并且在不同学科之间半衰期的变化很大（结果显示尤其是对那些使用半衰期较长的学科，超过 6~12 个月的更长的保密期具有重要的意义）。绿色 OA 的支持者指出 STM 出版有继续盈利的能力，由于缺乏自知识库出现以来期刊停刊和新刊创办速率降低的数据，绿色 OA 对期刊影响较小的说法仍然盛行。许多出版商的商业直觉告诉他们相反的结论，但他们在接受诸如 NIH 这样的大型资助机构支持的项目课题组投稿时必须接受其发表的文章可以通过绿色 OA 获取，不过出版商的绿色 OA 政策已有一些收紧（P102）。

1. 学术交流

STM¹出版发生于包括正式形式（如期刊论文、书籍）和非正式形式（如会议讲演、预印本）的、更为广泛的学术交流系统。学术交流的供应链传统上一直被视为包括两个主要角色，这两个主要角色服务于以学者（作者和读者）为代表的学界和学者的资助者及其所属机构的需求，即：出版商（负责实施质量控制、生产和传播）和图书馆员（负责管理内容的获取和导航、长期保存，不过长期保存因电子出版的出现而正在发生改变）。在某些市场（如电子书、医疗、产业），资源积聚也发挥着重要并且可能是愈加重要的作用。然而，学术交流正在演变，并且研究基金资助机构正随着开放获取和相关进展的发展而成为该系统中最重要的一部分之一，并且，其他的角色正起着日益重要的作用（尤其是数据仓储及服务供应商）。

1.1 研究周期

学术交流所发挥的不同作用可以在研究周期的背景下理解（括号中是交流的作用）（图 1，Goble（2008）引自 Bargas）：

- 思想发现，生成假说（认识，文献回顾，非正式）
- 基金申请/获准资助（文献综述）
- 开展研究（认识）
- 传播成果（正式出版，非正式传播）

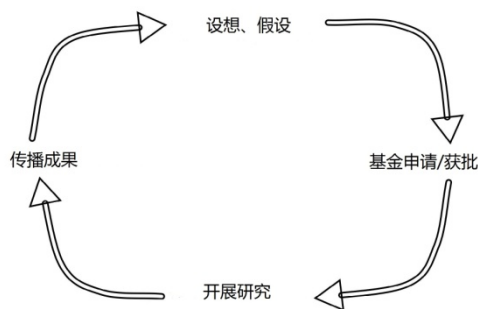


图 1 研究周期

¹ “STM”是科学、技术和医学(scientific, technical and medical)的缩写，但有不同的含义。它可能是一种出版模式，包括社会科学、艺术与人文；有时被用来表述科技期刊；也是出版商协会的名称（“STM”），即本报告的发起者。本报告中这三种用法均存在，所赋予的含义在其出现上下文语境中清晰明了。

1.2 学术交流的类型

如上所述，学术交流涵盖广泛的活动范围，包括会议讲演、非正式的研讨会、面谈或电话交谈、电子邮件交流、邮件群发、正式的期刊和专著出版、预印本、灰色文献和日益增长的社交媒体。学术交流的一种分类是基于公共交流或私人交流以及评议或非评议两方面，详见图 2。本报告主要关注正式的、期刊论文形式的书面交流。正式和非正式交流的区别有时比较模糊（例如，arXiv 知识库中作者未经评议的原稿正越来越多地被正式出版物引用，并且期刊论文变得越来越非正式地允许读者加注类似博客的评论），不过期刊论文在学术交流中的核心作用比以往任何时候都重要。

然而，我们也有兴趣了解学术交流可能会受到当前和未来的电子通讯方式的哪些影响。对于各类人际交往，可将其划分为三种基本模式：一对一，一对多，多对多（对于这方面争议更广泛的处理见参考 Inger & Gardner, 2013），这些可进一步划分为口头和书面交流。从学术交流类型的口头和书面交流方面看（如表 1 所示），可发现电子和基于网络渠道的引入在很多方面为旧的交流模式带来了新的方式（如网络出版物正在代替纸质出版物），但没有衍生全新的模式。维基(wiki)和社交媒体是例外，维基（及可视为维基衍生的、诸如论坛和专门的讨论/合作平台等类似的网络空间）提供了一个促进多对多地书面交流的实践方式，这种多对多地书面交流确实提供了离线世界某种程度上全然没有的方式。社交媒体，尤其是微博（Twitter），也提供了也可能在多类网络平台实时传播的实时、多对多的讨论。这种视角也许对于平衡某些持有技术中心论的“方案解决者”的观点有所帮助，该观点声称数字和网络技术的引入将会自动导致学术交流发生革命性改变（参见“作者的行为、认知和态度”部分）。

1.3 学术交流体系的变化

学术交流的过程经受着深刻变革的压力，这种压力主要来自技术和经济的驱动。与此同时，研究人员的潜在需求总体上没有大的变化（见“作者的行为，认知和态度”部分）。变化可被认为发生在以下三个方面（参见 Van Orsdel, 2008）。

- 出版市场的变化（例如：开放获取的新商业模式；（联盟）协议许可的新销售模式；全球化及新兴地区的增长）。
- 研究实施方式的变化（例如：网络的使用；数据密集和数据驱动科学的发展；研究的全球化）。
- 公共政策的变化（例如：研究资助机构强制自存储和数据共享；版权的变化）。

这些变化的细节和影响将在后面的章节中进一步讨论。

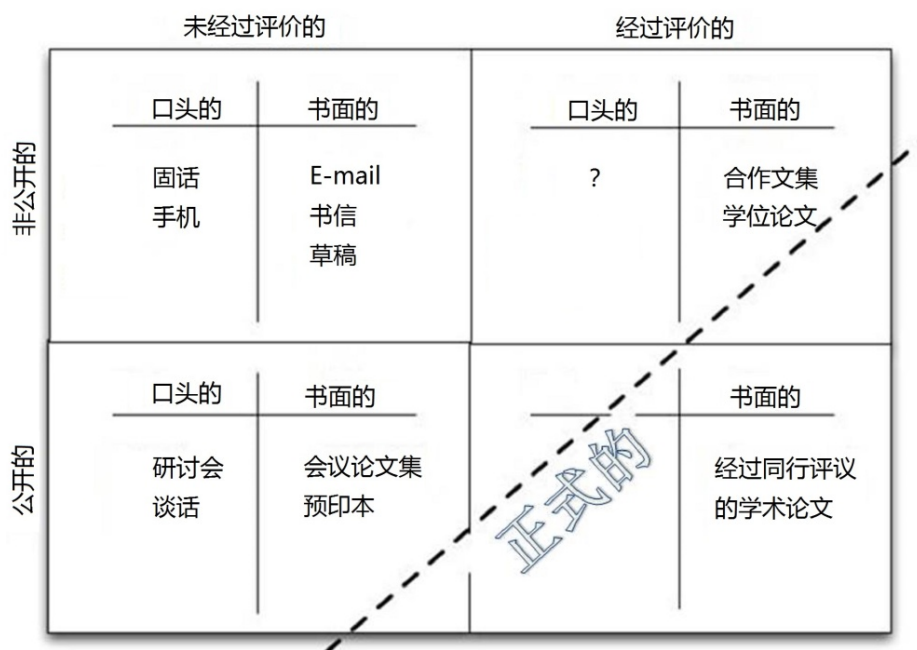


图 2 正式和非正式的学术交流类型

表 1 交流模式

模式	联络	旧方式	新方式
口头	一对一	面对面交谈	即时消息
		电话交谈	网络 (VOIP) 电话 视频电话会议
	一对多	演讲	即时消息
会议报告 电视/广播		网络视频	
书面	多对多	电话会议?	基于网络的会议
	一对一	信件	电子邮件
		一对多	印刷出版物
	多对多		无

2. 期刊

2.1 什么是期刊？

有一系列的出版物类型可被泛泛地称为期刊，如《自然》(*Nature*)，《核物理 B》(*Nuclear Physics B*)，《新科学家》(*New Scientist*)，对于非科研人员而言没有明确的分界线。本报告主要关注学术和科学文献，即：发表科研人员撰写的经同行评议的周期性出版物，而不是以新闻为主的杂志。

期刊在传统上一直被视为有四大功能：

- 登记：第三方建立的关于作者对某个思想的优先权和所有权的时间印记。
- 传播：通过期刊的品牌形象向目标读者传播研究发现。
- 认证：通过同行评议确保质量控制和奖励作者。
- 档案记录：为将来参考和引用保存固定的版本。

现今也许要增加第五个功能：导航，即基于海量的出版材料（和不断增加的相关材料，如数据集）为相关的工作提供筛选和标记。或者，这可以被看作是传播功能的一部分。

我们不厌其烦地重申这些基本认识是为将要讨论的新体系做铺垫，如开放知识库，它只具有部分功能，而不是上述所有的功能。

此外，值得注意的是，可以将这些功能看作会如同为读者服务一样地为作者服务。事实上，当作者将文章的权益免费转让给期刊出版商时，他们并没有“放弃”这些权益而是用来交换以上这些服务（及文字编辑、标记和语义富集等其他服务）。

2.2 期刊的出版周期

在期刊出版过程中，信息在不同参与者之间的移动通常称为“出版周期”，如图 3 所示。某个特定研究群体的作者所生产的研究信息，在以期刊论文形式回到作者所在研究群体的读者手中之前，要通过作者所选定的期刊编辑部传递至期刊的出版商，征订型高校图书馆（通常是经由代理商，虽然协议授权正在降低大型出版商的作用）。当然，在电子出版时代，读者也可平行于图书馆途径直接从出版商获取期刊论文（尤其是开放获取的论文），不过订购型期刊的获取仍然主要由图书馆管理。

作者通过出版传播研究成果的同时也建立自己的个人声誉和对成果的优先

权及所有权。期刊的第三方日期认证机制记录论文的收稿和录用日期，而期刊的声誉与论文和对作者的推广都相关。

期刊的编辑通常是独立的，他们多是由出版商任命并给予财政支持的相关领域的一流学者（最常见但不是普遍来自大学学术团体）。期刊编辑受理作者的稿件，判断稿件与期刊的相关性，并指派同行专家进行同行评议。

同行评议是对作者论点合理性、引文的权威性、结论的原创性等方面的方法性检查。虽然同行评议通常不能确定论文中数据是否正确，但提高了大多数论文的质量并为作者所欣赏。期刊编辑基于审稿人的意见作出是否发表的最终决定。有关同行评议将在以后的章节中进行更深入的探讨（见“同行评议”部分）。

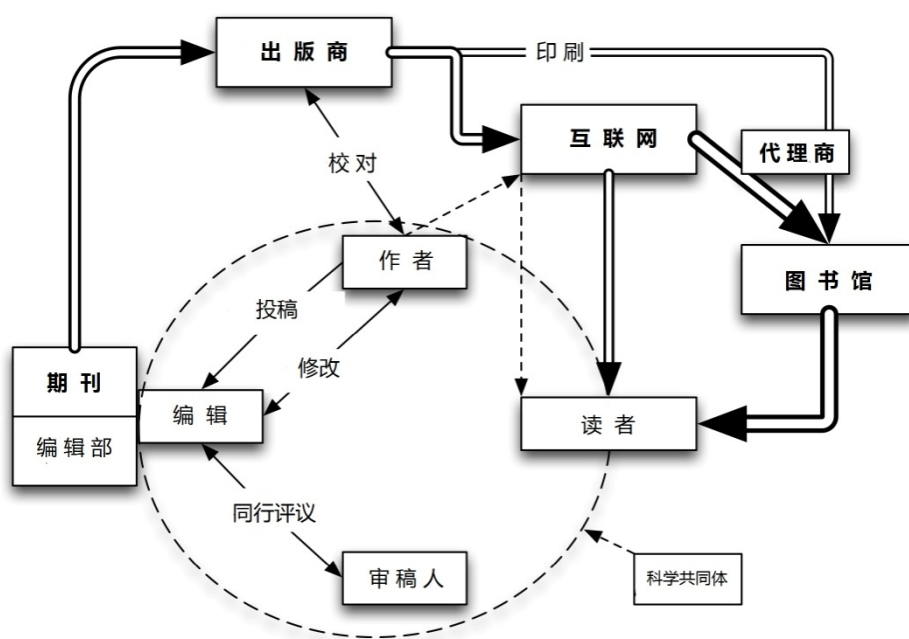


图3 出版周期

出版商的作用

出版商的作用通常与印刷者或生产者混淆，实际上它更广泛。期刊出版商的关键作用是甄别发现对路的市场定位以创办新期刊，或扩展（或关闭）现有的期刊。这一创业方面的特征既要满足学术界内部的对新期刊的需求（值得注意的是期刊出版商基于对一些学科发展前景的信任并支持这些领域的新期刊，因此促进了许多学科的诞生），又要实现令人满意的投资回报。作为企业，期刊出版商也需要具备以下能力：

- **制造商/电子服务提供商**——文字编辑、排版和标记，以及（在目前用户和市场仍然需要的情况下）印刷和装订期刊。
- **市场营销者**——吸引论文（作者），扩大读者群（开放获取期刊与订阅期刊同等重要）和新订户。
- **传播者**——出版商维护订阅运作系统以保证货物的按时交付，保持与订阅代理、图书馆员和学术界的联系。
- **电子存储**——电子期刊需要许多额外的技能，通常是数据库供应商、网站开发人员和计算机系统方面更为需要。

另一种评判出版商作用的方式是考量他们在哪些方面增值。STM 信息领域广泛（即包括但不限于期刊），可以考虑 STM 出版商的作用是否在流程的以下几个方面增值（改编自 Outsell, 2011）：

- **研究产出的分级和评估**：同行评议的好处之一是期刊感知质量的满意度，这种对质量的感知被广泛应用于对研究产出的评估（Ware, 2008）。
- **内容聚集**：虽然其他参与者（如 Google、PubMed）也参与其中，但目前出版商的聚集服务可提供广泛使用的服务。
- **证据提炼**：例如参考资料和超级综述（meta-reviews）。
- **建立标准和寻求共识**：大量的出版商主导的举措提高了 STM 内容的质量、可发现性和可使用性，包括 CrossCheck、CrossRef、CrossMark、ORCID 等等。
- **颗粒化、标注和语义富集**（包括分类和本体的发展）、内容的优先级、识别、规则的应用：这些方面的增值可能会变得越来越重要。
- **系统集成、数据结构和交换标准、内容维护、更新流程**：例如 SUSHI 标准、KBART 标准。
- **多来源内容的整合**：超越简单的聚合服务，例如，基于多个内容类型和来源建立复杂的循证医学服务，在医疗方面支持医生。
- **建立和监控行为变化**：例如，加强医学期刊的利益披露标准，有些期刊鼓励（或要求）研究数据的并行存储。
- **workflow 分析的发展及个体、部门、机构和地区实体的最佳实践标杆管理**：例如，支持研究评估的工具。

有关出版商作用的一个更详细的描述请参见博客文章“出版商所做的 82 件

事（2014 版）”（Anderson, 2014a），这实质上是上述同样功能的更细化的分解，但 Anderson 也强调需要有盈余产生的长期可持续发展模式（例如在新技术平台的再投资）。

Cliff Morgan 及其合作者综述了出版商在开放获取发展中的作用，指出类似的活动仍然是必要的，并指出自 2000 年以来出版商在在线出版技术方面累计投资了 35 亿美元（Morgan, Campbell, & Teleen, 2012）。

论文的版本

随着自存储的广泛采用，一个潜在的问题是论文有多个版本可被读者（和其他人，如知识库管理人员）获得。为建立论文在出版周期的各阶段的统一的名称，美国国家信息标准组织（NISO）与学术和专业出版商协会（ALPSP）合作推荐了一个使用名称（NISO, 2008）。NISO 推荐的术语：

- AO = 作者的原稿（Author's Original）
- SMUR = 提交的评审中的稿件（Submitted Manuscript Under Review）
- AM = 录用的稿件（Accepted Manuscript）
- P = 校样（Proof）
- VoR = 记录的版本（Version of Record）
- CVoR = 记录的修改版（Corrected Version of Record）。
- EVoR = 记录的增强版（Enhanced Version of Record）

上述术语是为了多种目的（如本报告中的很多地方）而建立的细化的结构，在讨论期刊出版时不是必须的。STM 协会与欧盟等机构讨论中使用的术语为阶段 1（作者的原稿）、阶段 2（录用的稿件）和阶段 3（任何版式的最终论文）。

预印本（pre-print）也用来指作者的原稿（有时指录用的稿件），后印本（post-print）是录用后的稿件。这些术语是过时的，因为他们含义模糊，容易产生潜在的混淆（例如，后印本肯定不是出现在印刷之后），但这并不妨碍这些术语广泛地继续被使用。

CrossRef 组织于 2012 年 4 月引入了 CrossMark 服务来识别记录的版本（Meyer, 2011）。用一个可见的风筝标记以方便读者识别，也有定义的元数据以方便搜索引擎等识别。CrossMark 不只是识别文章作为记录的版本，而且还提供有关出版流程（例如，同行评议）前的信息和出版后的事件，如勘误、订正和撤稿。

另一个潜在的问题是有可能出现采用更动态的文章注释的期刊。例如，期刊

Faculty1000Research 鼓励作者发布（多次）修订版本的文章，所有版本的文章被链接并且可独立地被引用。

2.3 销售渠道和模式

基于订阅或许可的期刊的市场大致可分为两大类购买者，即图书馆和个人（另见下述“开放获取期刊”部分）。虽然个人订户（个人订阅或会员式订阅）对于某些期刊而言很重要（例如杂志 / 混合型期刊，如《自然》或《科学》和某些学会期刊，尤其是医学会期刊），个人订户的期刊购买和使用多年来一直在下降，并且个人订购的价格通常有很高的折扣，期刊市场的大量收入主要来源于对图书馆的销售。

传统的图书馆销售与个人的期刊订阅形式是一样的，这种形式的市场份额正在下降，尤其是大型出版商更是如此，因为越来越多的期刊以捆绑形式销售到图书馆或图书馆联盟。

在印刷版继续的时代（见下文），大多数出版商提供三种模式的单个期刊订阅：只是印刷版、只是在线版、印刷版和在线版相结合。大多数出版商收取只订购在线版的费用比只订购印刷版的费用少，以及对纸版订户的在线获取收取一定的额外费用。

单篇文章的销售越来越普遍（虽然基数很小），提供单篇文章销售的出版商比例由 2003 年的 65% 增加至 2012 年的 83%（Inger & Gardner, 2013），并且相关的模式（如文章租赁和文章打包）越来越普遍。然而，更重要的是期刊的打包或集群销售。回溯档案（过刊）的销售也很重要，许多图书馆热衷于以一次性的价格（包括或不包括维护费）获取物理的文档在本地存储，同样的还有授权获取模式。

最后，销售模式的一个关键部分是有关“永久访问”，即订户在订阅期结束后访问以前订阅内容的权利。大多数出版商提供永久访问，但在某些情况下有额外的费用。大型出版商（91%）比小出版商（50%）更有可能提供永久访问（Inger & Gardner, 2013）。

订阅代理

订阅代理是销售渠道的一个重要组成部分：图书馆的订阅平均大约有 80% 是通过代理商进行的。代理商作为图书馆的代表，可使图书馆只与一、两个代理商打交道而不必与大量的期刊出版商发生关系，处理不同出版商各自不同的订购流程、条款及条件等。代理也通过聚集图书馆订单并转换成机器可读的数据、处理日常的续订等为出版商提供有价值的服务。STM 出版商给代理商的折扣优惠通

常低于其他许多行业，并且该折扣价一直在下降，从而使代理商可以通过向图书馆收取获取费用而盈利。代理商有着悠久的历史，第一家代理商成立于 1793 年（Everett & Son）。代理商协会（ASA）最近列出的名单上约有 25 个成员²，但代理商会员的数量在最近几年一直在下降（本报告 2012 年版中所报道的 ASA 会员是 30 个，2009 年版是 40 个），下降的主要原因是行业的兼并和收购及新成员的缺乏（或者更极端的是生意的失败，如 2014 年 Swets 的情况）。一个关键的原因是向电子出版的转移特别是联盟销售的上升使传统代理商的中介功能日益降低。因此，较大的订阅代理商正进行自身的重塑，如：转变为聚合商、出版商、分析服务的提供商等。有观点认为这种改变将更加有利于较大型的出版商（因此有利于加强出版商联盟），因为大型出版商撤回大量的业务（代之以直接向联盟销售），代理商的成本将加速摊派到剩余的小型出版商（Aspesi, 2014）。

内容捆绑

随着电子出版的兴起，单个期刊征订的销售已转为青睐于将总体的一部分捆绑销售。据统计（Cox & Cox, 2008），几乎所有（95%）大型出版商和大多数（75%）中等出版商都提供捆绑销售，尽管该比率对于小型出版商而言下降至 40%（原因很明显）。出版商越来越多地提供捆绑的内容，包括非期刊内容，特别是电子书、参考资料和数据集，并且这个趋势很可能会继续。小出版商更有可能参与诸如 ALPSP 学术期刊集群、BioOne 或 MUSE 项目等多出版商捆绑销售。2012 年研究型图书馆协会（ARL）实施的图书馆会员调查发现，图书馆采购的内容远超过 90%是来自较大型出版商的捆绑销售（Strieb & Blixrud, 2013）。然而，相比 2006 年而言捆绑的性质已有所改变，“全期刊”捆绑的比例已经比较低（归因于 2008/2009 经济衰退期的调整）。

ARL 的调查还发现大多数出版商的（捆绑）许可定价仍然基于历史印刷版（有时称之为“优先印刷版”）的模式，这与 Cox 在 2008 年的发现相似。在历史印刷版模式中，图书馆可获得对所有捆绑期刊电子版的获取，其价格反映于图书馆现存的纸版订阅情况（通常是延续的订阅），另加上对非订阅期刊的只获取电子版的充值式收费。这种充值收费模式（尤其是捆绑中包括出版商的所有期刊并且征订是面向图书馆联盟的）常被称为**大额交易（Big Deal）**。其他主要的定价模式包括：

- 基于使用的定价。第一次尝试开始于 2000 年中期，但没有获得很大的推动力。ARL 的调查发现几乎没有证据表明其会员在 2012 年采用基于使用的定价，这种现象也反映在有报道指出基于使用的定价仍然处于初生

² <http://www.subscription-agents.org/>

期并且只有很少(~10%)的出版商声称采用了这种模式(Inger & Gardner, 2013)。

- 基于机构规模分类的层次定价。Inger & Gardner (2013) 发现采用这种定价机制的普遍性仅次于采用基于历史印刷版的模式，最常见的对规模的界定是站点的数量。(类似 Carnegie 或 JISC 的分类机制不常用，因为这些分类只涉及大部分出版商的少部分内容)。
- 基于客户类型(如医院，学术机构，企业)的差异定价。
- 基于并发用户数的定价。对于数据库而言这种定价模式已存在多年。
- 对于捆绑的所有期刊的集成统一价格。

虽然定价模式的发展处于明显的停滞³，业内的讨论认为在未来的几年将有更多的出版商放弃基于历史印刷版的订价模式。

对于图书馆而言，一个关键的问题是出版商是否许可取消捆绑。Cox 发现只有 40%的出版商允许取消，有意思的是商业性出版商比非营利性出版商更可能允许取消(46%对 24%)。ARL 的调查表明，在经济低迷导致了图书馆预算受到严重制约的时候，出版商的政策实际上更加灵活。

图书馆联盟

以捆绑形式销售的期刊数量的增加与对图书馆联盟进行捆绑销售的重要性增加是相一致的(认识到这两个概念的不同很重要，有些出版商开展对图书馆联盟销售但不提供捆绑销售)。联盟的出现是为了提供高效的集中服务(例如：共享的图书馆管理系统、书目分类、馆际互借(ILL)、资源等等)和集中采购，通过与出版商谈判来提高图书馆的购买力，并且越来越多地利用捆绑的电子内容。联盟的数量一直在强劲增长：Ringgold 联盟在线目录⁴列举的 400 个联盟涵盖了 100 多个国家的 26500 余个个体图书馆⁵，其中约有 350 个联盟负责许可的内容。国际图书馆联盟联合会有约 200 个成员。联盟的规模和性质有很大的不同，从国家级联盟至小的地区性联盟，包括学术、医疗、公众、学校和政府图书馆。联盟

³ 值得注意的是，历史印刷版模式一直以来是一个务实的而不是保守的做法，因为优先印刷版在许多情况下是图书馆和出版商之间定价原则协议中的最后一点。更高级的数据库模式有优点也有缺点，而任何一方都不想要缺点。

⁴ <http://www.ringgold.com/cdo>

⁵ 本报告的上一版也指出这种增长，活跃的联盟由 2003 年的 164 个上升至 2008 年的 338 个，但自 2012 年以来变化相对较小。

⁶ <http://icolc.net>

所涵盖的个体图书馆总数多达 5000 多个。ARL 的调查指出 2012 年联盟的作用还保持在核心位置，61%~97%（取决于不同的出版商）的协议都是通过联盟签订的。根据 ALPSP 的一份报告（《学术期刊出版实践》，Cox & Cox, 2008; Inger & Gardner, 2013），大约一半的出版商（大型出版商有 90%）积极开展对联盟的市场销售。其中，大约有一半的出版商在基于个案谈判的基础上采用与捆绑销售相同的定价模式。联盟的交易通常（60%）是以 3 年为一周期，30%以 1 年为周期，10%以 2 年为周期，只有一半的出版商提供价格限制。取消的条款如同前述的捆绑销售。包括“作者权利”条款（通常包括作者在许可机构的自存储权力）和保密协议这样的新条款在捆绑中开始变得重要起来。

图书馆系统供应商

图书馆系统供应商⁷提供编目、企业资源规划及图书馆使用的链接解析器和其他接入系统。虽然他们的业务主要是面向图书馆而不是出版商，但他们是读者与出版商内容链条中一个重要的组成部分。出版商与系统供应商有关供应链的标准有诸如 ONIX 系列出版物⁸和 KBART（知识基础和相关工具）⁹。准确地识别机构对于出版商而言是十分重要的：Ringgold¹⁰的识别服务是领先的商业服务，其数据库中有超过 40 万个机构和财团，Datalon 于 2014 年推出免费的数据集 OrgRef¹¹。

供应商在“发现”工具方面进行了大量的投资，包括所谓的“网络规模的发现”，其中知名的产品有 EBSCO Discovery、Proquest Summon、Ex Libris Primo、OCLC WorldCat Discovery 等等。这些服务提供简化的界面(比较流行的是用户熟悉的 Google 界面)，可使用户通过一次搜索发现大量的图书馆收藏（包括摘要和引文数据库）和网络资源，快速提供结果，如：相关度排序、各方面结果浏览、内容建议、全文链接、以及各种社群和研究管理功能。此外，还有详细的机构使用的计量和报告。简化界面与强大结果相结合的流行可反映在 81%的图书馆将发现服务作为其默认搜索工具这一事实。图书馆的试验表明，发现服务的使用增加了用户的满意度并增加了图书馆订阅内容的使用（Somerville & Conrad, 2014; Outsell 2014e; Levine-Clark, McDonald & Price, 2014）。

所有这些服务，如同其他基于图书馆的搜索工具一样，都在与 Google Scholar

⁷ 相关概述及供应商目录见：<http://www.librarytechnology.org/>

⁸ <http://www.editeur.org/8/ONIX8/>

⁹ <http://www.uksg.org/kbart>

¹⁰ <http://www.openrfp.com/pages/identify.html>

¹¹ <http://www.orgref.org/web/about.htm>

竞争(并且在一定程度上也与 Microsoft Academic Search 有竞争), Google Scholar 除了提供特色搜索功能外还提供整合了馆藏、Scholar Metrics 的引文链接和其他特色的功能 (Inger & Gardner, 2012; Van Noorden 2014b)。确实, 正如 Somerville & Conrad (2014) 所述, “谷歌学术图书馆 (Google Scholar Library) 可以直接在 Google Scholar 的搜索页面保存文章, 在个人的我的图书馆 (MyLibrary) 页面基于话题组织内容并搜索全文, 设置了期望值增高的工作流集成解决方案”。Google Scholar 没有披露其涵盖期刊清单, 但估计其检索的学术文献介于 1000 万和 1600 万之间 (Khabisa & Giles, 2014; Orduña-Malea, Ayllón, Martín-Martín, & López-Cózar, 2014)。

2.4 期刊经济与市场规模

期刊经济与市场规模

据 Outsell 估计, 2013 年全球 STM (包括期刊、图书、技术信息和标准、数据库与工具、医学交流及其他相关领域) 的市场总规模为 252 亿美元 (Outsell, 2014c)¹²。预计 2017 年之前该市场总规模的年增幅约为 4%。

Outsell 曾估计, 在 STM 信息的总体市场中, 期刊的收入大约占 40%, 图书大约占 16% (Outsell, 2012c)。2013 年的市场还可以细分为科技信息 (122 亿美元) 和医学信息 (130 亿美元)。

期刊出版的收入主要来自学术图书馆的订阅 (占总收入的 68%~75%), 其次是企业用户的订阅 (15%~17%)、广告 (4%)、会员费和个人订阅 (3%), 以及各种形式的作者付费 (3%) (RIN, 2008)。由于广告的收入一直很弱, 因此广告收入的比例可能在下降 (尤其是在 2008/09 年经济衰退之后), 而订阅和许可证收入继续增长。

在各地区市场方面, Outsell 估计全球 STM 收入 (包括非期刊 STM 产品) 的约 55% 来自美国, 28% 来自欧洲、中东、非洲地区, 14% 来自亚太地区, 4% 来自其他各地 (主要是除美国以外的美洲地区) (Outsell, 2014c)。这些比例可能夸大了美国市场对期刊的重要性。

市场分析师 Simba 估计, STM 市场如果采用比 Outsell 给出的更狭义些的定义, 在 2011 年 211 亿美元的收入中, 期刊收入约 90 亿美元 (Simba, 2011)。Simba 估计 2013 的科学和技术 (即不包括医疗) 的市场为 107 亿美元, 较上一年度只增长了 0.2%, 但 2010~2012 年的年均复合增长为 2.3%。报告描述了正处在穷途

¹² 该项和其他市场规模的数字是所指当年的实际值, 即: 不更新为当前价值。

末路的科技图书市场,其销售下降了 4.2%,2012 年的销售额为 27 亿美元(Simba, 2013b)。

开放存取市场相对总体市场而言继续保持更为快速的增长,但收入额仍然很小。Outsell 估计 2013 年 OA 期刊的市场价值为 1.28 亿美元,并预计 2015 年将增长至 3.36 亿(Outsell, 2013)。Simba 更新的估计是 2014 年为 2.99 亿美元,相比全球 2013 年的 STM 期刊销售额(约为 2.42 亿美元)的增幅为 23%,相对 2012 年的增幅为 32%。Simba 预测 2011 至 2017 年之间 OA 的收入将增加 3 倍(达到 4.4 亿美元),而市场总体每年的增幅只有 1%~2%(Simba, 2014)。

据粗略估计,该行业全球雇用的人员大约为 11 万人,其中约 40%是在欧盟。2014 年 AAP 的专业与学术出版部门资助的一个独立调查估计美国的雇员总数超过 38000 人,有 350 多个出版商,工资成本为 23 亿美元(czujko & Chu, 2015)。此外,除了生产供应链环节的雇用人员以外,全球范围内估计还有 2 万~3 万全职员工是间接地从事 STM 产业工作(自由撰稿人、外部编辑,等)(据 Elsevier 估计)。

中国

尽管近几年增长放缓,从 STM 角度看全球经济最大的变化是中国的崛起。虽然中国已经成为世界第二大研究论文产出国(NSF, 2014),但其占全球 STM 市场份额远低于论文产出的份额。例如,Outsell 估计 2011 年度全球 STM 市场收入中中国的比例小于 5%,相比前文所述,美国的份额是 55%(Outsell, 2014c; Outsell, 2012e)。基于营业收入的大致划分是国际出版商占三分之二,地区性出版商占三分之一。

研究支出与 STM 市场份额之间产生这种差异的部分原因是中国研究基础设施正处于早期阶段。然而,另一个原因是有些出版商先采取低价格进入中国的市场,这个策略仍在继续压低市场价格。

尽管如此,2011 年度中国市场的增长为 10%,显著超过全球市场,并且 Outsell 估计 2011 至 2013 年中国 STM 市场的增幅在 10%至 11%之间。

图书和电子书

STM 图书市场作为一个整体近年来一直在紧缩,原因是尽管电子书的营业收入增长较为迅速,但总体体量要小得多,还不能抵消印刷版收入的下滑。Simba 估计医学书籍 2012 年全球市场下跌 2.5%至 29.6 亿美元,而科技书籍下跌 4.2%至 27 亿美元(Simba, 2013b; Simba, 2013a)。

2012 年 Outsell 的报告(Outsell, 2012d)估计 2011 年全球电子书的市场为 6.7

亿美元，约占 STM 图书市场的 17%。电子书的市场增长比 STM 整体市场的增长快得多，2011 年为 23%，而整体 STM 市场为 4.3%，并且一直以来比图书业整体的增长都更快，2008~2011 年的复合增长率为 33.7%，图书整体为 2.1%。

Outsell 发现科学和技术领域图书的市场吸纳率大于医学，后者的图书市场份额为 44%，而电子书的市场份额只有 35%。这似乎是自相矛盾的，因为医疗工作者在工作场所使用数字内容和移动设备在各类用户中是最高的。给出的原因是在科技领域出版商和集成商可以随现有的渠道出售捆绑内容，然而，虽然医疗从业者个人热衷于数字图书和内容，但机构采购安排的结构不利于医学电子书的批量购买，并且集成服务、循证医学、医疗点产品等方面存在着预算竞争。

参考性图书（及更小程度上的专著）是数字转换的先锋，出版商所报告的数字出版收入的绝大多数都来源于参考性图书。相比之下，教科书最难适应，数字版本的收入不到 10%。有两个原因：参考性图书更容易数字化（虽然编辑/生产流程必须重新设计，以便频繁的定期更新），而教科书需要更多的附加功能来支持学习和教学；此外，参考类图书的商业模式更为直接，而教科书出版商正在艰难地适应基于印刷和个人销售的模式向数字环境的转变，以及学生对数字教科书不同的反应。

基于 170 个出版商的调查分析，ALPSP 的《学术图书出版实践》报告提供了 2009 年度 STM 图书和电子书市场的详细图景（Cox & Cox, 2010）。虽然 2009 年以来的市场已经发生变化，尤其是与电子书相关的市场变化，但这份报告仍然是一个有用的信息来源。这些出版商代表了相当一部分的市场，包括每年出版超过 2.4 万种新书的出版商，共计涵盖 35 万种学术图书，其中包括参考资料、专著、教材、会议报告、专业手册、研究报告等。大部分（90%以上）出版商的出版是面向研究人员和研究生市场，约三分之二是针对大学生，并且约 40%为一般的参考性图书。与 2004 年 ALPSP 的研究相比，虽然电子书出版有了飞速的发展，但只有约三分之二的出版商出版电子书，并且电子书的收入占图书销售总收入的不足 10%。

图书和期刊的显著区别在于学者们本身更倾向于购买图书；例如，2012 年 Tenopir、Volentine 和 King 报道阅读的学术图书的最常见的来源是个人的存本（占 39%），远远超过通过图书馆的获取（26%），而期刊论文主要是从图书馆获取。

学术类图书的开放获取市场还处于萌芽状态，但对于某些领域如人文科学的开放获取具有潜在的意义，在这些领域，专著和其他学术著作仍然是重要的研究产出形式。有些举措和发展将在“开放获取图书”的章节中讨论。

数字教科书和相关的教育市场在商业模式上有相当大的创新，其中很多可能

会对现有的商业模式造成极大的破坏，包括免费增值模式（例如，基本的内容是在线免费，额外的服务需要收费，如更多功能的格式、打印、测试和课堂支持工具等，包括 FlatWorld Knowledge 和 Knowmia 等产品）；自然出版集团的《生物学原理》允许学生终身获取一期定期更新的在线课本；广告支持（例如 World Education University）；赠款资助（例如，Rice University 的 OpenStax College）；等等。开放教育资源（OERs）也一直在大规模地增长：最初由包括麻省理工学院等一流大学将课程材料提供在线免费获取¹³，其后有一波这方面的创业公司热潮，如 Coursera、Udacity、Udemy 和 Good Semester，并且将向大学提供的产品扩展到证书（著名的是 MITx 和 EDX）。领先的教育出版商，如著名的皮尔森（Pearson），正通过构建或提供端到端的服务而向另外的方向推进，不仅包括课本和教学内容，还有测试、在线学习环境、创建和上线他们自己的课程，并在不久的将来将根据英国法律的变化有权授予他们自己的学位（Pearson, 2012）。

学术交流体系的全球市场成本

2008 年剑桥经济政策协会的英国研究信息网（RIN）报告估计，期刊发表研究与交流的系统总成本为 1750 亿英镑，其中研究工作本身的成本 1160 亿英镑；出版、传播和论文获取为 250 亿英镑，阅读 340 亿英镑¹⁴。

250 亿英镑的出版费用中包括出版和图书馆成本；出版成本总额为 64 亿英镑：其中 37 亿英镑为确定发表第一稿的成本（包括 19 亿用于同行评议的非现金成本）27 亿英镑的可变成本和间接成本（包括出版商的盈余）。除去非现金的同行评议成本，出版和发行成本为 49 亿英镑，约占总成本的 3%。

（另见 *期刊出版的成本部分*。）

STM 市场的前景

虽然一直在恢复，但 2008/2009 年经济衰退的影响仍然存在，欧洲许多地区的增长停滞。因为试图控制预算赤字，许多发达国家公共部门支出很弱，从而导致在机构和图书馆预算方面不可避免地面临着预算削减的压力。在全球层面，国际货币基金组织（IME）已经降低了 2014 和 2015 年的全球增长预测，并警告说世界经济可能不会再恢复到金融危机前的增长速度（Elliot, 2014）。对 IME 来说关键的经济风险依然存在：对于未来的金融市场过于自满；俄罗斯与乌克兰之间及中东的紧张局势；欧元区三资探底的经济衰退可能会导致通货紧缩。其他风险包括量化宽松政策的退出及中国房地产市场（泡沫）的崩溃等。

¹³ 见 MIT 的 OpenCourseWare <http://ocw.mit.edu>

¹⁴ 2008 年的市值，没有计算当前的通胀。

由于对广告的依赖比较有限，STM 市场的周期性远不如许多信息市场，但广义的 STM 市场（如同 Outsell 界定的）在 2008 至 2013 年间的增长率只维持在 2.7%。Outsell 预计，在没有进一步的重大经济冲击（如欧元区三次探底的衰退及乌克兰或中东地区新的战争）的情况下，未来的增长率会更高，预计到 2016/2017 年达到 5% 以上并且 2013~2017 年的平均增幅为 4%（Outsell, 2014c）。然而，这些总体性的数据掩饰了不同的地区和产品细分市场的混合财富。STM 出版商的传统核心市场（在许多发达国家机构图书馆的期刊和图书销售）已经持平，仅仅维持着与通货膨胀相应的增长，并且可能会保持这种状态。市场的增长来自包括中国在内的新兴地区（中国的目标是从制造业经济向 2020 年的“创新驱动”转移）；同样，这些地区将继续成为大多数出版商收入增长的焦点地区。例如，Simba 的分析认为到 2017 年印度市场预计可与大多数的 G8 国家相提并论。

在增量方面，大型出版商报告文章投稿和在线使用有强劲增长，反映了对研究产品的持续增长的需求。就产品部分而言，亮点是电子书和在线的数据库、工具和其他服务。随着作为重要推动力的更广泛的移动应用程序加入到电子书，这可能成为保持增长的源泉。开放获取将仍然快速增长，尽管其占总体收入的比例不大。另外一个可能的增长来源是企业 and 中小企业部门：STM 出版商一直没有对这部分（制药公司和不同层次的技术部门除外）提供足够的服务，因为期刊和产业部门的需求之间不匹配。大型出版商如 Springer 和 Elsevier 似乎已经认识到这个机会并进行了相应的重组，设立有专门的销售和市场团队专注于企业部门、具体产品和门户网站¹⁵。

虽然只有有限的确凿证据，但似乎可能是经济衰退和相关的预算压力进一步推动了商业模式创新。尽管有一些批评的声音，但大交易将仍然是期刊（和越来越多图书）市场的主要形式，不过这是不断变化的，如未来几年新的定价模式有可能取代历史的定价方式，并且可能有更多的定价与使用的联动。

2.5 期刊和论文的数量及趋势

2014 年全球大约有 28100¹⁶ 种活跃的经同行评议的英文版学术期刊，全年共出版论文近 250 万篇（Plume & van Weijen, 2014）。图 4 显示了 2002~2012 年间乌利希期刊目录所列的活跃的、经同行评议的期刊数量增长情况，在此期间的期刊数量每年增长约为 2.5%。在本报告撰写的时候，CrossRef 数据库中包括了超

¹⁵ 如 Springer 的 R&D (<http://rd.springer.com>); Elsevier 的收购和扩张在线 (<http://www.elsevier.com/online-tools/knovel>)

¹⁶ 乌利希的网站上列有 28134 种正在出版的经同行评议的英文版学术期刊名录（据 2014 年 12 月 16 日的统计），如果包括非英文版期刊，则该数字为 34585。

过 710 万的数字对象唯一标识符 (DOI), 其中 550 万篇为来自超过 36000 种期刊的论文。更广泛地看, Google Scholar 检索的文档估计有 1 亿至 1.6 亿, 包括期刊论文、图书和灰色文献 (Khabisa & Giles, 2014; Orduña-Malea 等, 2014), 而 Web of Science 数据库收录的记录有 9000 万条。

只发表原始研究论文的期刊约占 95%, 其中包含一部分所谓的混合类期刊 (融合杂志和研究类期刊特点的发表广泛内容的学术期刊)。混合类期刊销售对象是个人和机构, 具有高发行量且有相当多广告收入, 这是纯研究类期刊所没有的 (Mabe, 2008)。最大的单一学科领域是生物医学, 有约 30% 的期刊, 艺术与人文学科的期刊占相对少数, 为 5%。

一个重要的期刊群是包含于汤森路透的《期刊引证报告》数据库的 2550 个出版商的 10900 种期刊, 其中自然科学 8700 种, 社会科学 3000 种, 这些期刊每年共计发表论文约 150 万篇。该期刊群之所以重要是因为它包含了最常被引用的期刊, 即 (至少基于这一测度的) 核心文献。汤森路透 (Thomson Reuter) 引文数据库收录期刊的规模总体上也远大于那些没有被收录的期刊 (被收录的期刊每年发表的论文数是 111 篇, 未被收录的期刊每年发表的论文数是 26 篇 (Björk, Roos, & Lauri, 2009), 美国科学基金会 (NSF 2014) 基于汤森路透数据库的分析表明这些期刊的年载文量由 1988 年的 111 篇上升至 2012 年的 168 篇)。另外一个摘要与索引数据库 Scopus 的覆盖范围更广, 收录了约 5000 个出版社的 22000 余种同行评议期刊。Scopus 包含的记录有 5300 万条并且 2013 年增加的记录有近 200 万条。

同行评议期刊的数量在过去三个多世纪以来每年以约 3.5% 的速度非常稳定地增长 (见图 5), 不过该增幅在战后的 1944~1978 时段确实稍大。

论文的数量在相似的时间尺度也以每年约 3% 的增幅增长。增长的原因很简单: 世界上科学研究人员数量的增长。图 6 绘制了论文和期刊数量伴随美国科研人员数量增加的情况。其他经合组织国家的类似数据可确认该结果 (来源: Elsevier)。

目前论文的增长可能远高于上述长期的趋势; Plume & van Weijen (2014) 报道当前论文总量以平均每年 6.3% 的增幅上升, 从 2003 年的 130 万篇增至 2013 年的 240 万篇。同样, 从 PubMed 检索的文章在 2003 至 2013 年的十年中每年增幅为 6.7%; Web of Knowledge 和 Scopus 收录的文章在 2010 至 2013 年间的平均增幅为 5%。然而, 美国国家科学基金会 (NSF, 2014) 报道 2001 至 2011 年的论文数量增幅只有 2.8%, 这个比较低的数字可能与采用的数据集较小有关。

最近的一项分析 (Bornmann & Mutz, 2014) 表明论文的数量增长有三个阶段,

每一阶段的增长率都是前一阶段的3倍：十八世纪中期之前的增幅不到1%到，二次世界大战期间的增幅为2%~3%，至2012年为8%~9%。

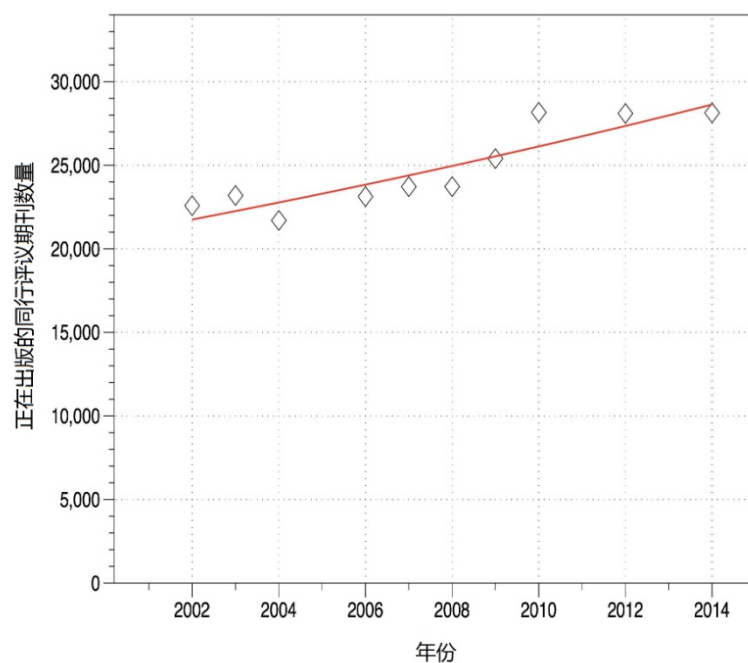


图 4 2002~2012 年乌利希指南收录的活跃的同行评议学术期刊的增长情况

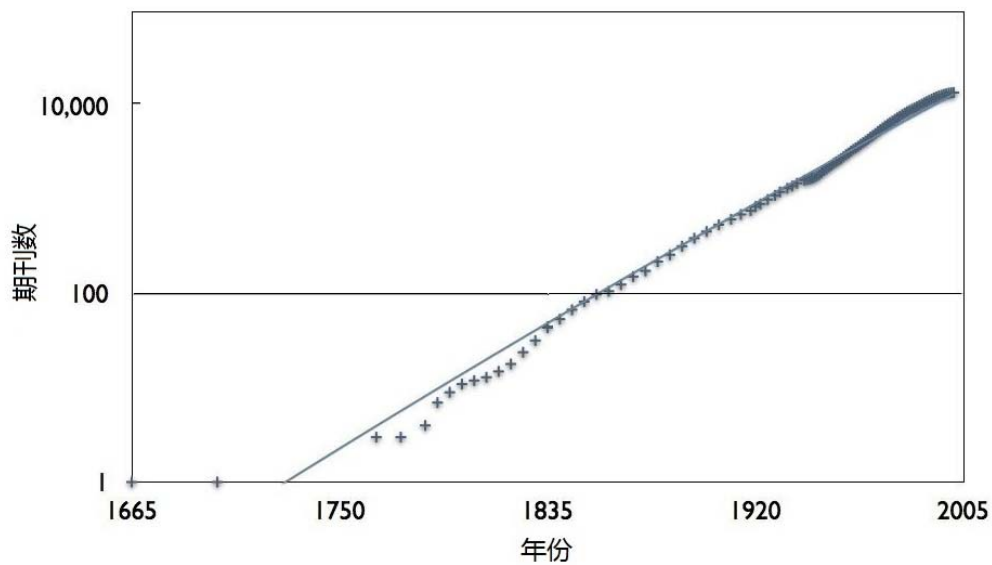


图 5 1665 年以来同行评议学术期刊的增长情况 (Mabe, 2003)

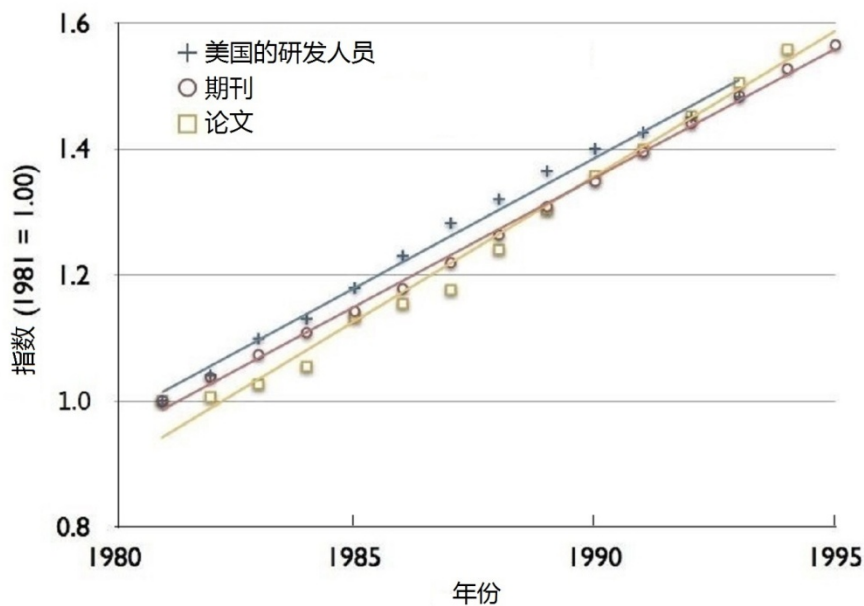


图 6 科研人员数量与期刊和论文数量的关系 (Mabe, 2004, 基于 ISI 和 NSF 数据统)

在线期刊

目前所有的 STM 期刊都可以在线获取，只有少数例外（例如：一些非常小的期刊和一些人文学科的期刊）。早在 2008 年，ALPSP 的学术出版实践报告 (Cox & Cox, 2008) 就已发现 96% 的科技与医学 (STM) 期刊和 87% 的艺术、人文与社会科学 (AHSS) 期刊可以通过电子版形式获取。与 2003 年 (STM 83%, AHSS 72%) 和 2005 年 (STM 93%, AHSS 84%) 的调查结果相比较，这个百分率呈稳定的增长。

最新的 ALPSP 报告 (Inger & Gardner, 2013) 给出了类似的数字，这表明在在线获取方面市场已经达到近于饱和的阶段，大量的出版商都已经有 90% 以上的内容可以在线获取。然而，过刊的在线获取是另一回事，有 70% 的出版商有 90% 以上的内容可以在线获取，20% 的出版商在线获取的内容不足 50%。（要牢记的是这个调查报告的数字是出版商而不是期刊：由于滞后的都是规模较小的出版商，因此对于期刊和文章而言当前内容和过刊内容的在线比例都明显更高。）

然而，只有为数甚少的期刊舍弃了现存的印刷版。主要的原因是还有一些市场需求，包括个人和学会会员需要、有些地区的机构订户等。维持纸版需求的因素包括：对于有些用户而言纸版有其优越性 (2014 年 Outsell 的调查显示，有 35% 的反馈表示阅读时更喜欢看纸版 (Outsell 2014b))，对数字格式长期保存的担忧，担心取消订阅或出版商消失事件影响数字内容的获取，某些广告客户谨慎向数字

格式切换，在某些地区有税收障碍，等等。在便携性和可读性方面，印刷版相对数字版的优势似乎可以被最新的平板（阅读器）减弱，并且移动格式也似乎给广告商提供了一些压倒性的优势。然而，自 2015 年开始，目录和行业讨论的迹象是，印刷版将在未来的几年将最终开始大量地从出版商的期刊名单中消失。

数字保存和持续获取方面的问题也被各种各样的项目所解决，包括 LOCKSS/CLOCKSS、Portico、国家图书馆项目等，并且图书馆人员和用户越来越习惯于基于大额交易和新创期刊（包括开放获取期刊）的纯在线期刊。此外，数字印刷技术（包括按需印刷）使得提供更低价格水平的印版需要经济可行。

图书是另外一种情况。如前所述（图书和电子书），2011 年电子书只占约 17% 的 STM 图书的收入。尽管预测的增幅比较高（尤其是参考书和学术专著），但是教科书可能需要更长的时间向数字化转移，虽然在这方面有很多创新之处。

开放获取期刊的数量

2012 年 12 月开放获取期刊目录（DOAJ）¹⁷ 所列的 OA 期刊为 10091 种（其中 7245 种为英文版），在自上一版 STM 报告发布以来的约 2 年时间增加了 1976 种。不是所有加入 DOAJ 的期刊都采用完全的同行评议期刊（尽管所有的期刊都通过编辑、编委或审稿等途径实施了某种形式的质量控制）¹⁸。乌利希目录列出了 7111 种同行评议的 OA 期刊，大约占收录的同行评议期刊总数的 25%。

主要的摘要与索引（A&I）数据库中收录 OA 期刊的比例低于 Ulrich 的比例，考虑到更高的收录要求和 OA 期刊总体上相对年轻，所以这不足为奇。Scopus 涵盖了 22000 种同行评议期刊，其中 2390 种或 13% 是 OA 期刊，而 Web of Science 涵盖了 726 种（比率约为 9%）OA 期刊。

开放获取论文的数量

统计开放获取期刊的数量有其挑战性（如：如何筛除掠夺性期刊），并且由于期刊的规模差别很大（例如，从每年出版 20 篇文章的季刊到 2014 年发表了 31883 篇文章的 *PLOS ONE* 这类巨刊），更好地测度研究群体吸纳开放获取的统计是论文数量，包括绝对数量和占论文总数的比例。

然而，计算开放获取文章因为其定义的问题及方法和度量的挑战而复杂化，

¹⁷ <http://www.doaj.org/>

¹⁸ DOAJ 在 2012 年迁移到一个新的平台，即 IS4OA，该平台于 2014 年重新发布。重新发布后引进了新的更严格的选择标准，要求申请表格并声明遵守标准（这在部分程度上消除了“掠夺性”期刊和其他低质量的期刊）。至 2014 年中期，只有一个很少部分（几百种）DOAJ 所列的期刊通过这种方式的重新认证。

不同的研究者对 OA 文章的分类有所不同，有时因为理想化的原因使得其他不同的估计很难或不可能相互比较。概括地说，有以下 4 类文章的统计（见后文的开放获取的章节）：

- 金色 OA：发表于纯 OA 期刊的论文（不论是否支付论文出版费）；有些研究将混合性 OA 包含在这一类。
- 混合性 OA：在订阅性期刊上发表的论文一经出版立即公开，通常是支付论文出版费（APC）的结果。
- 延迟性 OA：在订阅性期刊上发表的论文经过一段保密期后公开。
- 绿色 OA：论文不同版本的副本在机构知识库等各类知识库和作者的网页可获取；通常在论文发表后有一段保密期；可能有多个储存的版本并且存档于不同的储库。
- 其他：严格来说这不是 OA，只是免费获取的论文，有些是合法的（如提升可获得性），有些是不合法的（即发布的内容侵犯版权）。搜索 OA 论文的自动化工具或机器人可能无意地（或某些情况下故意地）将这类文章计入。

绿色 OA 的最大挑战为：如同其定义和重复数据删除的挑战，由于回溯的文章可能在任何时候补充，因此，不能及时统计固定的结果。

在方法方面，文章的统计可以基于精心设计的检索系统（如 Web of Science 或 Scopus）或使用专门的搜索引擎机器人。前者只针对检索系统收录的期刊发表的金色 OA 文章，从而低估了文章总数（虽然不一定改变比例）。后者不能正确地识别和分类可公开获取的文章，研究者试图通过尽可能多的样本或时间的手工验证来评估自动发现结果的可靠性。

表 2 选择了一些针对上述问题的研究结果。由英国 BIS 实施的 Elsevier 的研究结果比其他的结果略低，但其研究设计似乎更稳健，较低的结果可能是更严格的定义和更严格的排除。由欧盟委员会进行的 Science-Metrix 调查采用了不寻常的定义，因而其结果很难与其他结果比较：与大多数研究不同的是，其“绿色”分类不包括学科仓储（如 PubMed Central）、其他集成网站（如 citeseerx），作者的个人网页等；“其他”包括这些网站以及混合、延迟、及非法发布的副本（他们称之为“Robin Hood or Rogue OA”（罗宾或流氓 OA））。

有关金色 OA 的发展历史见图 7（来源于 2012 年发表于 *Nature* 的一篇文章（Van Noorden, 2012a）），该图总结了 2003 至 2011 年估计的金色 OA 文章的比率。图中所使用的数据源于不同的数据库，包括 Web of Science 和 Scopus 数据

库。图 8 显示的是基于 2014 年早期一组估测的更长的趋势，(Archambault 等, 2014)。

2011 年的数值 12% 是基于 Mikael Laakso 和 Bo-Christer Björk 对 Scopus 数据库中 OA 论文比例的估计 (其后的报道见 Laakso & Björk, 2012)，指论文能够立即通过出版商的网站获取 (即完全或混合 OA)，另有 5% 指可以通过延迟获取。

不同学科论文的可获取性相差很大，如图 9 (Björk 等, 2010) 和图 10 (数据源于 (Archambault 等, 2014)) 所示。生物医学领域采纳金色 OA 的比率仍然是最高的，该学科支持 OA 的基金和支持 OA 的基金资助机构更多。例如，使用 PubMed 搜索工具表明，2013 年发表且被 PubMed 收录的论文 (在 2014 年后期) 可免费获取全文的论文比例为 34%，最抗拒金色 OA 的 STEM (科学、技术、工程、医学) 领域是工程及物理和天文学，以及人文学科的哲学和神学及经济学和商学。其原因是多种多样的，例如，物理与天文学及经济学的预印本文化比较昌盛，这可能会降低对金色 OA 的需求。

表 2 部分研究报道的 OA 论文份额 (详见正文中说明)

	Björk 等, 2010	Gargouri, Larivière, Gingras, Carr & Harnad, 2012	Laakso & Björk, 2012	Elsevier, 2013	Archambault 等, 2014
金色 (总体)	5.3%		11.0%	10.2%	12.1%
金色, 需 APC				5.5%	
混合	2.0%		0.7%	0.5%	
金色, 无 APC (补贴)				4.2%	
延迟免费获取 (开放档案)	1.2%		5.2%	1.0%	
绿色 (总体)	11.9%			11.4%	5.9%
绿色, 预印本				6.4%	
绿色, 接受稿件 (mss)				5.0%	
其他 OA					30.9
所有 OA	20.4%	21.0%		~23%	46.9%

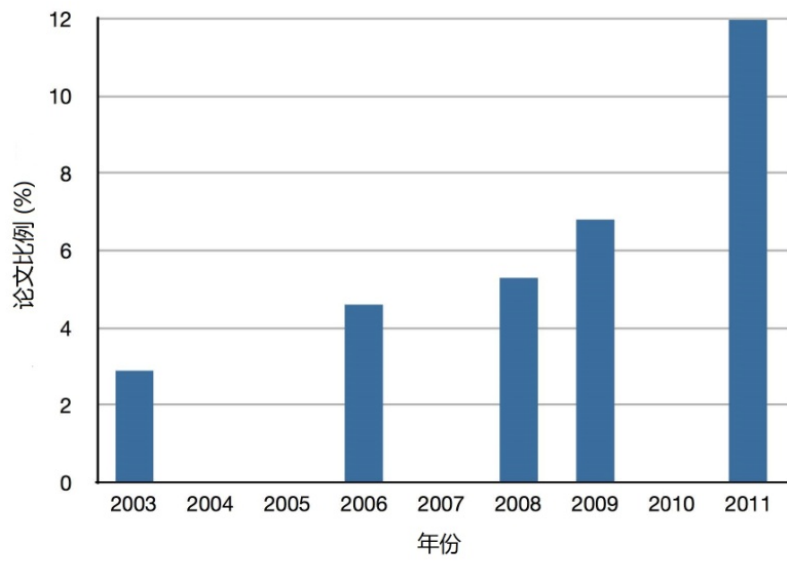


图 7 以金色 OA 方式发表的论文比率的增长 (Van Noorden, 2012a)

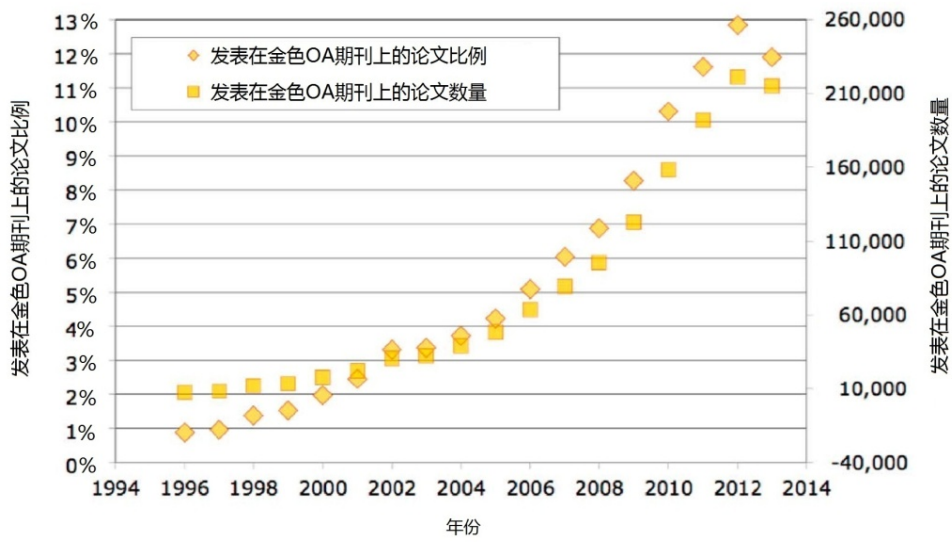


图 8 1996~2013 年金色 OA 的增长 (Archambault 2014)

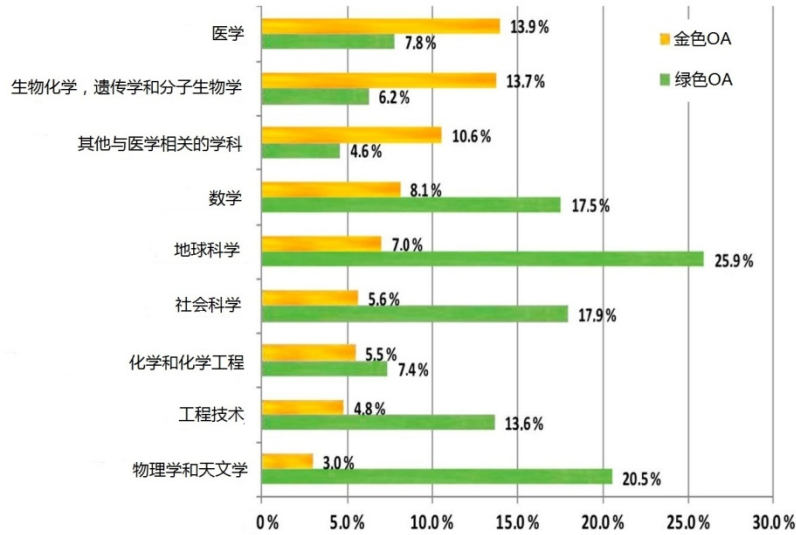


图 9 不同学科的 OA 可获取性。(Björk 等, 2010) doi:10.1371/journal.pone.0011273.g004 修改

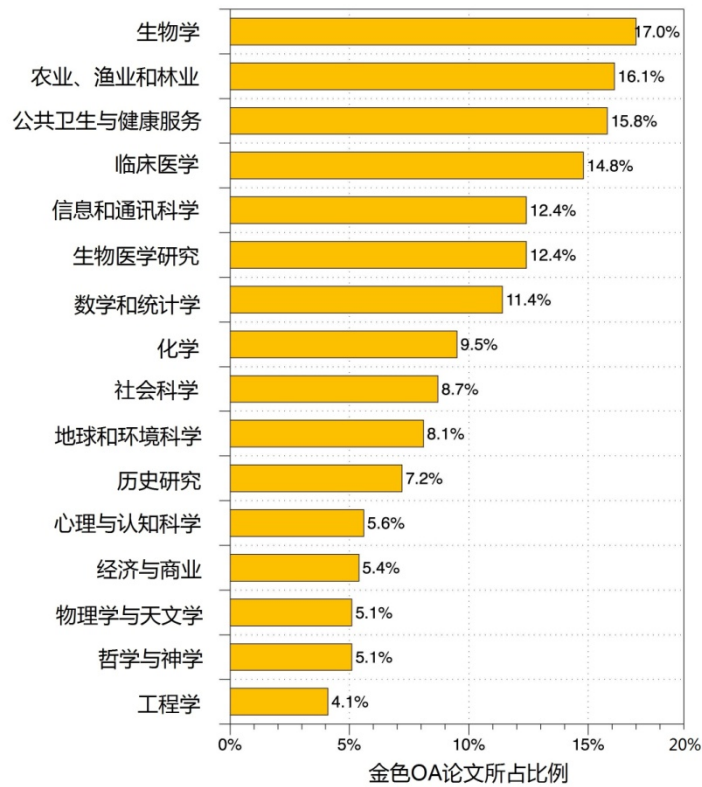


图 10 不同学科的 OA 可获取性。(数据源于 (Archambault 等, 2014), 表 II; 由于作者对绿色 OA 定义与其他研究不可比, 因此排除)

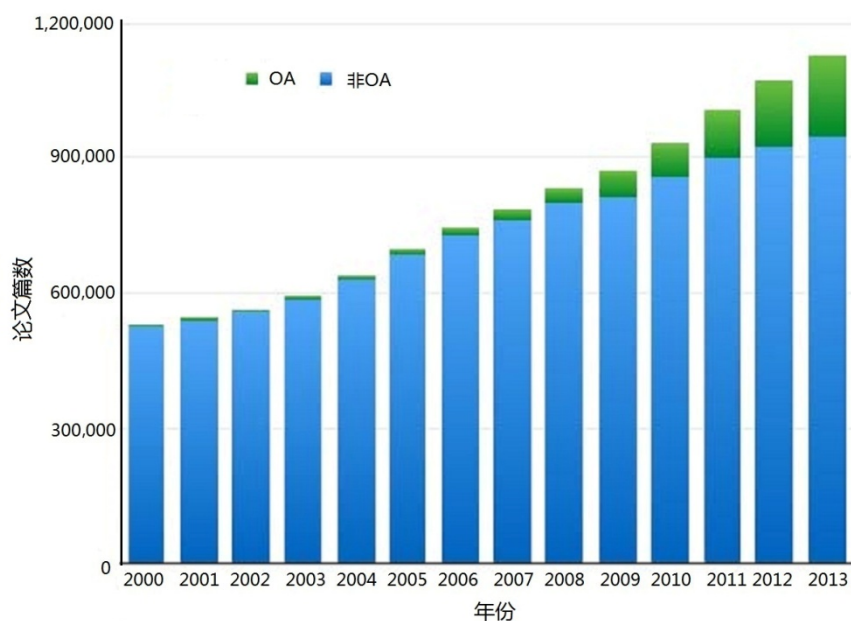


图 11 PubMed Central 中 OA 内容的增长

2.6 全球科学产出的趋势

研究与发展 (R&D) 的经费支出

正如我们所见，论文的数量与研究人员的数量密切相关，研究人员又反过来与研究发展的经费投入量有紧密的联系。

全球研发支出的增长长期以来一直快于全球 GDP 的增长，从 1996 年的 5220 亿美元上升至 2009 年的 1.3 万亿美元 (NSF, 2012)，估计 2014 年约为 1.6 亿美元 (Battelle, 2013)。这些支出大部分 (92%) 发生在世界上三大经济区，北美、欧盟和亚洲。美国以 4650 亿美元 (占 GDP 的比率为 2.8%，远高于全球平均值) 的最大数额迄今为止远比其他单个国家多，占全球 R&D 份额的 34%，虽然这个位置第一次开始受到挑战。

尽管如此，2008/09 年的衰退对全球 R&D 的影响是显著的：2008~2012 年 OECD (经济合作与发展组织) 国家的年增长率仅为 2001~2008 年的一半 (经合组织 2014)。

政府将把研发的支出视为创新、增长和国际竞争力的关键¹⁹。2010 年世界各

¹⁹ 对政府而言 R&D 支出似乎是一个非常好的投资：私体的回报估计平均约 25%~30%，社会回报率通常是其 2~3 倍 (Department for Business, Innovation & Skills, 2014)。

国的研发投入占国民生产总值的平均比例约为 1.7%，虽然（不足为奇）各国所占比例的变化范围相差很大，石油富有的沙特阿拉伯为 0.04%，印度为 0.8%，加拿大为 2%，瑞典为 3.7%，OECD 国家平均为 2.4%。有关研发支出的增加趋势将长期继续下去：虽然美国 1950 年代设定的目标是占 GDP 的 1%，其目前的支出正接近 2.9%，并且许多国家（包括欧盟作为一个整体）将目标设定为占 GDP 的 3%（OECD, 2011; UNESCO, 2010）。

中国研发支出的增长尤为显著，由 1996 年占 GDP 的 0.6% 上升到 2009 年的 1.7%，增长了 3 倍，并且同期 GDP 的年均复合增长率为 12%。更新的统计显示 2008 至 2012 年间中国的 R&D 支出翻了一番（Rowlands & Nicholas, 2005）。下文将对中国研发投资的策略及影响进行更详细的讨论。其他新兴国家也在大力增加研发支出：巴西计划在 2012 年之前将研发支出占 GDP 的比例增至 1.5%，并在 2020 以前年达到 2%（BIS, 2011）。

尽管研究成果和研究人员数量有正相关性，不同国家的研究产出率仍存在实质性的差异。例如，英国研究人员的人均论文产出数量和被引用次数、篇均论文被引用次数在世界前五国家（美国、中国、日本、德国）位居首位（Royal Society, 2011; Elsevier, 2013）。

经济衰退对美国 and 欧盟的影响大于中国、巴西和印度，这使得后者的研究支出占全球研究支出的比例增长更快。由于持续的衰退效果和刺激消费方案的终止，预计 2012 年全球研发支出的增幅会略微减缓（5.2%），低于 2011 年的 6.5%，2013 年更是降低至 2.7%。Battelle 预测 2014 年会恢复到 3.9%（Battelle, 2013）。

不同学科的 R&D 支出

学术研究支出的总体趋势是远离物质科学趋向生命科学。例如，在美国，生命科学得到了大约 60% 的资助（其中约 55% 是医学研究），并且是美国唯一的获学术资助份额有较大上升的学科领域（增长 6 个百分点）。

主要地区的优先研究重点根据其经济需要和政治压力的不同而变化。美国和欧盟主要集中在生物科学和医学，而日本的论文分布在生物科学、医学、化学和物理。新兴经济体的研究重点一直以来更集中于经济增长和基础设施建设，例如，中国的投资组合目前主要为化学，物理和工程，虽然其 2006~2020 年的 15 年规划为聚焦于能源、水资源和环境保护方面。

产业部门的作用

研发支出主要由产业部门资助：美国约为 66%，欧盟为 54%（最低为英国的 45%，最高为德国的 70%），中国、新加坡和台湾地区在 60% 到 64% 之间。产业

部门在研发实施中所占的比率甚至更高,例如在美国该比率超过 70%(NSF, 2012; Battelle, 2011)。这些对于出版是比较重要的,因为研究论文主要来源于学术作者。

在被统计的研究支出中,大多数用作应用技术的研发而非基础科学研究。在美国,对基础科学的研究支出约占 R&D 总支出的 18%,然而,有些令人吃惊的是,美国基础研究投入占 R&D 支出的比率近 50 年来翻了一番。目前几乎全部的基础研究都由学术机构实施,虽然产业部门和政府部门的研究人员曾经做的更多——“贝尔实验室们”大量涌现诺贝尔奖获得者(最后一轮的统计是 13 位)的年代已经一去不复返了。因此,美国产业部门的创新支撑比以前更加依赖学术机构对基础科学的研究。

知识密集型产业的增长

NSFC 最新的科学与工程指标(NSF, 2014)强调 2008/09 年经济衰退以来全球与科技相关的经济活动的图谱与以往有何不同。总体趋势仍然为知识密集型经济增长的重要性之一。1997 至 2012 年间,知识和技术密集型(KTI)产业在发达经济体的份额由 29%增长至 32%。在发达国家,该行业的增长在中国最为明显,其高技术制造业在 2003 至 2012 年间增长超过 5 倍,致使其全球市场份额从 8%攀升到 2012 年的 24%。

研究人员的数量

有关研究人员的数量一直没有全面且被广泛接受的确定数据,部分原因是当研究人员离开学界时,对其身份的定义就变得十分困难;部分原因是不同国家该数据的统计方法不同。

最新可获得的 OECD 的统计是 2011 年所报道的 840 万名研究人员报告(全日制的研究人员为 630 万名),涵盖了 OECD 以及一些非 OECD 国家(如中国和俄罗斯),但不包括其他一些重要国家(如印度,巴西)。2010 年该数字的增幅为 7.5%,表明了经济衰退以来的反弹。2000 至 2011 年间的平均年增长率为 4.2% (人数)和 3.2% (FTE-全职雇员),表明更大的兼职工作的趋势(OECD n.d.)。

最新的可获得的 UNESCO (联合国教科文组织)²⁰的数据是 2007 年的 570 万人,相比 2002 年的 570 万人复合年均增长率为 4.4%(UNESCO, 2010)。同时,Elsevier 公司最近提供给英国政府的报告中给出的数字较低,估计 2011 年的数字是 673 万人(Elsevier, 2013)。(爱思唯尔公司提供的数据较低是因为他们使用了弗拉斯卡蒂手册(Frascati Manual)中有关研究人员的定义,该定义比联合国教

²⁰ 新版的 UNESCO 科学报告将于 2015 年 11 月出版,见 <http://www.unesco.org/new/en/natural-sciences/science-technology/prospective-studies/unescoscience-report/>

科文组织对“科学家和工程师”的定义更为苛刻，例如，中国的研究人员数量基于弗拉斯卡蒂手册的定义由 160 万下降至 110 万。)OECD 估算每 1000 个在职人员中有 7.5 个为研究人员或工程师(欧盟占 7.5%，美国占 9%，日本占比小于 10%)。世界银行给出的数字更高，为 890 万研究人员。

不论采用何种定义，全球研究人员的数量在长时段都维持着每年 4%到 5%的增长（虽然伴随着经济衰退的短时段下降，最近的一次是 2009 年）。增长主要来源于新兴国家的驱动，亚洲领先国家的年增幅高达 8%到 12%，相比之下 G8 国家约为 2.9%，美国和欧盟只有 1%。最快速的增长是韩国（1995 至 2006 年期间人数翻番）和中国（2005 至 2008 年间增至 3 倍）。因此，中国在研究人员数量上将很快超过美国和欧盟；同样，韩国、台湾地区、中国和新加坡的研究人员总数占全球研究人员总数的比例已经从 2003 年的 16%增至 2007 年的 31%(Royal Society, 2011)。

地区变化

新兴经济体持续的世界平均值的 R&D 支出的增长已经大大地改变了世界研究的格局。正如麦肯锡顾问表述，到目前为止世界经济重心在 2000 年到 2010 年间发生最为快速的转变，同期全球研究分布的图谱也在快速变化。自从第二次世界大战以来，美国的领先地位首次受到中国的挑战。例如，2011 和 2001 间美国和欧盟的全球研发份额大幅下降，美国的份额从 37%下降到 30%，欧盟从 26%下降到 22%。同时，东亚、东南亚和南亚经济体的全球份额从 25%上升至 34% (NSF, 2014)。

根据预测，中国将于 2016 年左右取代美国成为全世界最大的经济体。自 2005 年以来，中国的 R&D 支出已经增长了 3 倍达到 700 亿欧元，且中国计划至 2015 年将 R&D 支出提升至占其 GDP 的 2.2%，并设置 2020 年的目标为 2.5%（优先研究领域为能源、水资源和环境保护）。在过去 10 年中，中国的 R&D 支出年均增长超过 20%。按照目前的趋势，中国将于 2020s 年代早期 (Battelle, 2013; NSF, 2012) 甚至更早 (OECD, 2014) 在研究支出方面超越美国。

在预算增长的支持下，中国正致力于加强更高教育机构和研究院的研究能力。“211 工程”旨在将一些重点大学建设为世界级大学。113 个 211 大学培养了全国的 80%的博士研究生，拥有 96%的国家重点实验室，获得了 70%的科研基金资助。同样，“985 工程”旨在建设世界知名大学。在参加该项目的 39 所大学中有两所（清华大学、北京大学）目标是跻身世界最好的大学，另有 8 所的目标是成为世界一流大学，其余的目标是成为中国最好、世界著名的大学 (Outsell, 2012e)。

可以预测中国对研究支出的大幅增长所带来的结果。目前中国已经取代英国

成为世界上每年发表论文数量第二的国家，目前的论文份额介于 11%（NSF, 2014)和 17%（Elsevier, 2013），并且 2020 年以前有望超越美国（图 12 和图 13 分别为 2008~2012 年及 2001~2011 年的论文产出趋势）。

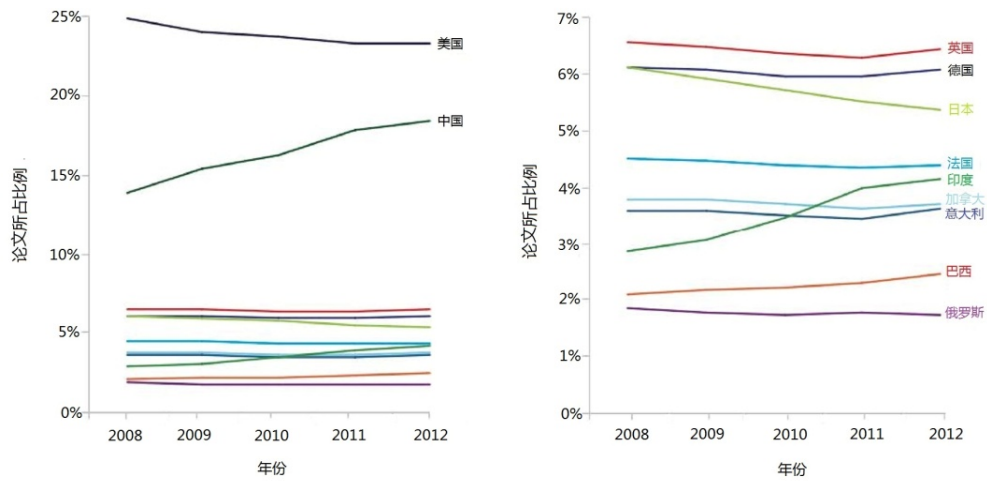


图 12 世界论文的份额（Elsevier, 2013）

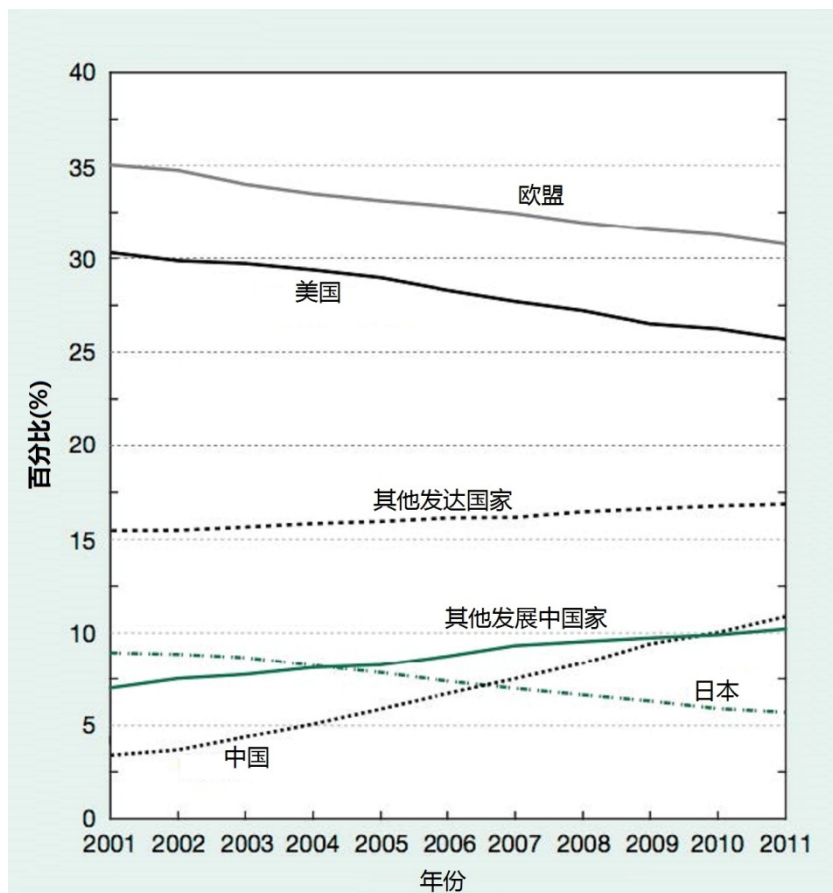


图 13 论文的全球份额（NSF, 2014）

中国不仅在研究产出数量方面的增长，质量也一直稳定增长，尽管质量仍然还远落后于美国和欧盟。目前中国占世界总引用的份额接近世界平均水平（Thomson Reuters, 2014），这尤其体现在计算机科学，其相对引证指数（引用的平均数与全球所有的引用均分到各国家的期望值的比例）为 1.3，稍高于美国。2002 至 2012 年间中国占有最高被引前 1% 的论文增加了 6 倍（NSF, 2014），目前正接近世界平均值。2013 年度自然出版指数（NPI）的编辑预言中国“正加快步伐在未来的 2 至 3 年成为亚太地区对 NPI 的最大贡献者”（Nature Editors, 2014）。

中国并非唯一的论文产出份额增长的国家。在 2003 年到 2012 年间，金砖四国（俄罗斯、巴西、印度、中国）中除俄罗斯的份额从 3.0% 降为 2.1% 以外，其他三国的论文份额都有增长。拉美国家的论文产出份额从 2002 年的 3.8% 升至 2008 年的 4.9%。目前美国仍然维持着论文产出份额世界第一的位置，但其份额（百分比）下降最快，从 33% 降至 27.8%（Thomson Reuters, 2014; UNESCO, 2010）。巴西、印度和韩国的经济状况也有可能 在 2020 年代早期超越日本和法国。

发展中国家的研究优先领域和新生势力与发达国家的传统优势有很大的不同。例如，英国和美国都侧重于生物医学、临床医学及地球科学和空间科学，而中国（其研究侧重更专注于某些特定领域）则主要侧重于物理学、化学、数学和工程科学。

对联合国教科文组织（UNESCO）来说，这些变化引发了“知识贡献模式的结构突破从而促进了全球经济水平的增长”。换言之，国家不再需要从源头上通过国家性的研发来建立其知识基础，但发展中国家可以在世界知识宝库的基础上，采用其他国家正在使用的技术来建立其知识基础，并且这样做的风险很低。同时，地理界限对于研究和创新来说，既关联越来越小，又比以往任何时候都更加重要。（译者注：这句话的意思是当前的研究和创新已不能为地缘所缚，各自为阵，而应该突破地理界线，分享知识的储备和技术的突破。）

研究人员的流动性

Elsevier 认为，在全球化研究的世界，20 世纪 50 年代的“人才流失”观点应被更微妙的概念“人才循环”所取代。这种观点认为，研究人员在国外时所获得的技能和构成的网络在他们终将回国时使其本国受益，并且即使他们不返回本国仍可作为侨民发挥作用（Albanese, 2009）。Elsevier 的研究表明研究人员处于高度流动，但流动性因国家而异，以英国和加拿大的流动性最大，“久坐不动”者（1996~2012 年没有在本国以外发表成果）比例最低，为 27%，而日本为 60%，中国为 71%。英国有一个特别流动研究群体，英国有近 72% 活跃的研究人员当下所发表的文章标署的是非英国机构（Elsevier, 2013）。

合作及合著

研究继续变得更加国际化和更加合作性，其驱动因素包括：超越单一机构的知识分享和技能分享的科技优势；通讯和航空成本的降低；信息技术使用的增加；鼓励国际合作的国家政策及冷战的结束；研究生“国外留学”项目等等。

当今的合作已司空见惯，体现在每篇论文中平均作者数量和机构数量的增加及国际合作论文比例的增加。在 1988 年，只有 8% 的论文为国际合著，但截至 2009 年这一数字已经高达 23%，并且在主要科技地区这一比例在 27%~42% 之间（NSF, 2012）。英国皇家学会引用基于一个不同的数据集估测，相比 15 年前的 25%，如今国际合作论文数量占总量的比例已经达 35%（Royal Society, 2011）。图 15 显示了国际合作论文比例的变化趋势。有趣的是，并非所有的国家都呈增长趋势，中国大陆和台湾地区大致相对稳定或甚至下降。土耳其和伊朗的国际合作论文比例（图中没有显示）也稍有下降。这可能反映了这些快速发展地区的研究机构的新生性。

总体而言，美国每篇论文平均作者数从 1990 年的 3.2 人增加到了 2012 年的 8 人（NSF, 2014），汤森路透数据库中的全部论文的篇均作者数从 2007 年的 3.8 人上升至 2011 年的 4.5 人（ScienceWatch, 2012）。图 4 显示了 20 世纪下半叶论文篇均作者数的增长情况，图 16 显示了不同学科论文作者数的不同增长情况，其中合作者最多且增长最快的是物理学和天文学，而合作者最少的是数学和社会科学。该图反映的另一个趋势是合著论文的比例在 1990 年至 2010 年间由 42% 增加至 67%。更新近的趋势是超过 50 人合作的论文数量增加，甚至超过 1000 位作者（超多作者，hyperauthorship）合作的论文数量也有所增加，主要是在国际高能物理领域的合作。1981 年，美国科学信息研究所（ISI）收录的论文中最高作者数为 118 人，而 2011 年这一数字是 3179 人。这一趋势引发了人们对作者属性的讨论，有人认为在此类情况下应该区分“作者”和“贡献者”的概念（另见“数据引证”章节）。

国际合作也在增长：1997 至 2012 年间，世界总体的国际合作论文由 16% 上升至 25。美国是一个特别重要的国际合作伙伴。2009 年，有 43% 的国际合作论文都包含了至少一位美国作者（NSF, 2012）。此外，由于受到与 G7 伙伴（美、英、法、德、日、意、加）合作的影响，金砖四国之间的学术合作非常少。英国，法国和德国都有很高的国际合作，国际合作论文超过 50%（Elsevier, 2013）。然而，中国在不断增长的国际合作的总趋势方面是例外，在过去的 10 年中其论文快速增长，国际合作论文保持在相对稳定的 27%（NSF, 2014）。

国际合作模式随不同学科而有所差异。在 STEM 学科，天文学是最具国际化

的学科，而心理学、化学、社会科学合作率很低（图 16）（NSF, 2014）。

对于研究人员而言，合作的一个明显好处就是能增加引用（在一定程度上增加了使用）。论文的篇均被引用次数会随着合作国家的增加而上升（即：在领衔作者国家的基础上增加）；增加 5 个国家的论文所获得的被引次数是非国际合作论文的 3 倍（Royal Society, 2011）。合作的影响程度对于不同国家有所不同，但发展中国家趋向于受益更大，可能是因为受益于与发达国家的更成熟团队的合作。例如，中国一流机构以外的国际合作论文的被引用次数是非国际合作论文的 3.1 倍（Elsevier, 2011）。

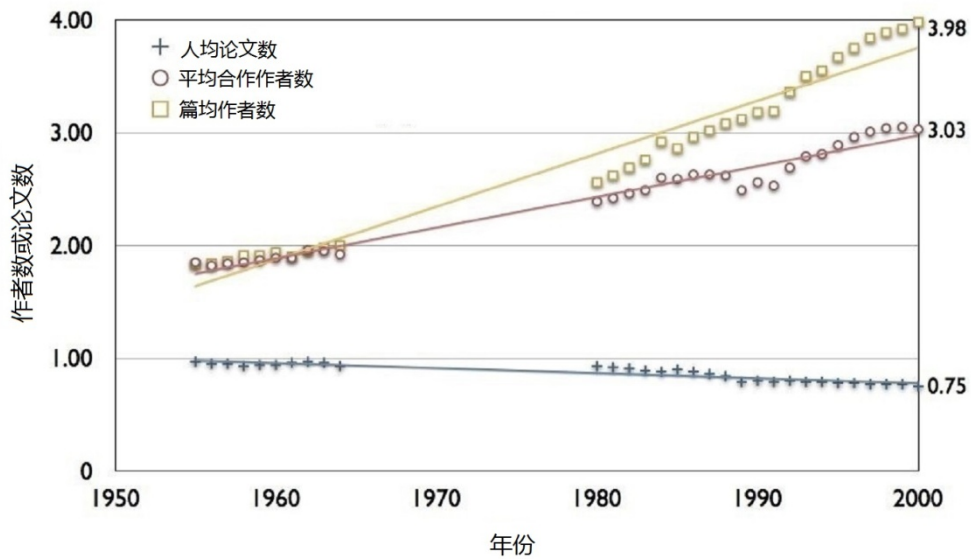


图 14 1954 至 2000 年论文合著的趋势图(据 Mabe & Amin 2002, 数据来自汤森路透的 SCI)

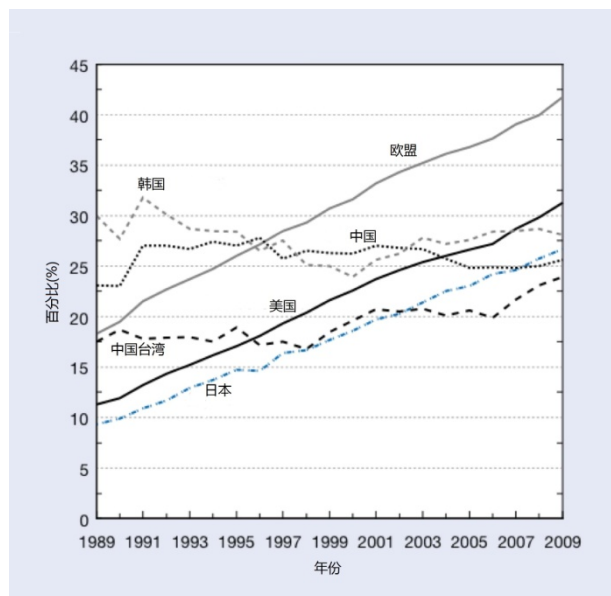


图 15 1989 至 2009 年某些地区/国家/经济体的国际合作论文的比例（据 NSF，2012）

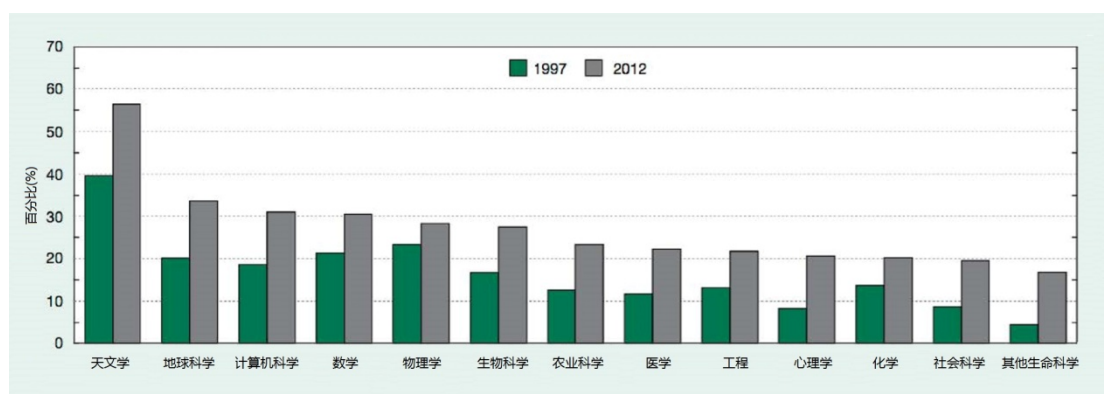


图 16 不同学科合著论文的增长：基于 SCI 和 SSCI 收录的 1997 和 2012 年的全球论文统计（NSF，2014）

2.7 作者和读者

作者

全球活跃研究人员的数量依据不同的定义在 670 万到 890 万之间不等（参见“研究人员的数量”章节）。然而，作者的数目不同决不是因为所有这些人要在给定的某年中发布一篇文章。例如，Plume & van Weijen (2014) 报道 2013 年共有 416 万个体作者发表了 240 万篇论文。（作者总数有 1000 万，因为每篇文章平均有 4.2 个个体作者。）这些数字自 2003 年稳步增长，当时大约有 210 万个体作者发表的 130 万篇文章。Scopus 数据库拥有约 150 万个作者标识符。

期刊论文主要由学术界撰写。据 King & Tenopir 报道，美国虽然只有 10%到 20%的科学家被大学雇佣，但他们发表了全美 75%的论文(King & Tenopir, 2004)。

Tenopir & King 后续的研究揭示，美国只有 15%到 20%科研人员作为作者在同行评议期刊发表过论文。这一作者与读者不对称的估算被 Mabe & Amin 证实：他们认为 UNESCO 当时估算的全球的 500 万至 600 万研究人员中，只有大约 100 万人(约 18%)多次发表过文章，并且有约 250 万作者在近 5 年发表过论文(Mabe & Amin, 2002)。

最近的一项研究调查了最具产出力的作者，定义为纳入研究的 16 年（1996~2011 年）中每年至少发表 1 篇文章的作者。调查发现了 1520 万位发表过文章的科学家中只有 150608 位（或不到 1%）每年发表一篇论文。并且，这个活跃的核心产出了 42%的论文和 87%的高被引论文。许多这些多产的科学家可能是在其团队产出的论文中署名的实验室或研究小组的负责人（Ioannidis，

Boyack & Klavans, 2014)。

读者

还需要有一个标准来区分核心的、活跃的研究人员群体和更广泛的期刊阅读群体，后者的数量可能非常庞大。很多期刊阅读者阅读论文的频率不高且不太深入，研究生和本科生都属于这个阅读者群体。目前似乎并没有可以准确估计这个群体数量的有效方式，Elsevier 通过统计 ScienceDirect 用户使用情况的一个内部研究估计，全球的期刊读者总数大约在 1000 万到 1500 万之间。最近科学社交网络 Academia.edu 报道有超过 1600 万的注册用户，这表明上述的数字可能被低估了（见“科学社交网络章”节）。

论文作者和读者群体的重合情况如图 17 所示。这个重合的程度在不同的学科间有很大不同：在专精的学科如理论物理这样的纯科学领域，读者和作者几乎近 100%重叠；在类似如护理和医学等行业领域，读者的数量往往是作者数量的数倍。

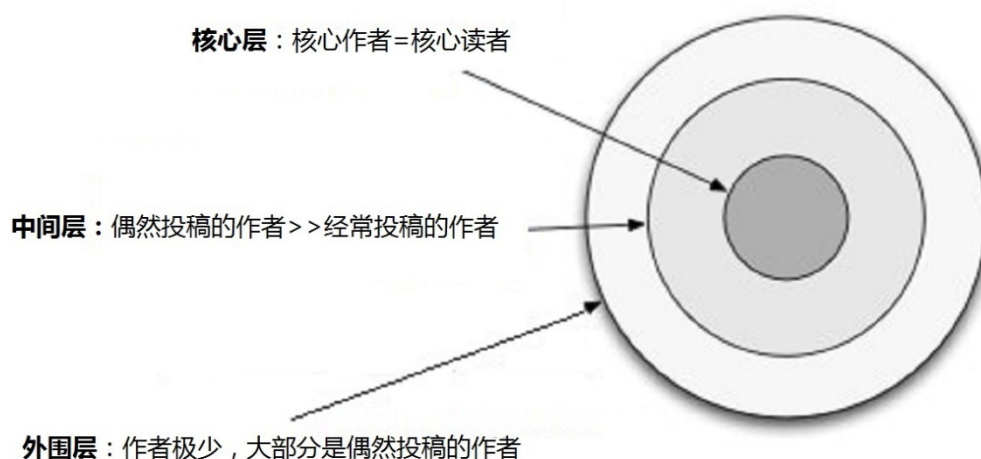


图 12 论文作者和读者群体的重合。在全球约 850 万 R&D 工作者中（据 UNESCO 数据），每年大约有 400 万人发表论文（Plume & van Weijen, 2014）

人们曾认为科学论文一般很少被阅读，这个误解源于 Garvey and Griffith 最初有关文献阅读调查时使用的有缺陷的研究方案 (King, Tenopir & Clarke, 2006)。电子出版使得论文使用情况可以得到精准的统计，即可以统计论文的下下载次数。尽管并非每次下载都对应着一次完整的阅读，每年全球从论文在出版商网站下载的次数约 25 亿次（据 STM 非正式的调查结果），另有 4 亿次从诸如知识库的其他站点下载。在英国的大学，2006~2007 年间有 1.02 亿次的论文全文下载，平均

每个注册的图书馆用户下载了 47 篇,并且下载次数每年以 30%的速度增长(RIN, 2009b)。2005 年的一项研究显示,学会期刊《儿科学》(*Pediatrics*)中的论文平均被阅读了 14500 次(King 等, 2006)。

PEER 项目有关论文使用的研究(CIBER Research, 2012a)发现,研究论文有 99%几乎每 6 个月就会被从相关的出版商网站下载至少 1 次,且有相当多数(74%)是从 PEER 知识库下载的。正如作者指出,“学术文献正被仔细阅读”。

顺便提及的是,平均每篇科学论文需要作者花费 90~100 小时准备(King & Tenopir, 2004),需要 2 至 3 位审稿人花费 3 至 6 小时评审(Tenopir, 2000; Ware & Monkman, 2008)。

2.8 出版商

据估计,全球有大约 5000 到 10000 个期刊出版商:在 Scopus 数据库统计中,有 22000 种期刊来自 5000 多个出版商,并且,在“长尾”一端显示约有 10000 种不包含在 Scopus 数据库的同行评议期刊是以一社一刊形式出版。

主要的英语语种的商业和专业学会(ALPSP, SSP 和 STM)的会员包括大多数的大型出版商及少部分的较具全球性的出版商。据 Morris (2006) 统计,ALPSP (商业和专业学会期刊出版商学会)在 2006 年共有 657 个出版商会员,他们共出版 11550 种期刊,大约占当时期刊总数的 50%。在这些出版商中有 477 (73%) 个出版商和 2334 (20%) 种期刊是非营利性的。早期的乌利希国际期刊指南调查显示有大约一半的期刊来自非营利机构。这个差异表明乌利希在统计中运用了更广泛的数据。Elsevier 基于 Thomson Reuters 期刊引证数据库统计揭示了不同种类的出版商的论文产出比例:商业出版商(包括学会出版物)为 64%,学会出版者为 30%,大学出版社为 4%,其他出版者为 2%。

出版商出版的期刊的分布极不平衡。一方面,有 95%甚至更多的出版商只出版过 1 或 2 种期刊,另一方面,出版期刊数量前 100 位的出版商拥有期刊总量的 67%。前 5 名的出版商出版了近 35%的期刊,其中四大出版商(Elsevier、Springer、Wiley-Blackwell 以及 Taylor & Francis)各出版 2000 种以上的期刊(表 3)。在“长尾”的一端,有很多只出版 1 或 2 种期刊的机构,其中很多机构自己都不认为他们自己是“出版者”(例如:学术或政府研究部门)(Morris, 2007)。

2.9 同行评议

同行评议是学术交流尤其是期刊的根基。同行评议是一个将某领域作者的稿件在出版之前交给同一领域的专家仔细审查的过程。(这种方法也用于研究申请书的评估。)期刊审稿过程各不相同,一般都是 2 到 3 个审稿人把评议报告交给

有最终决定权的期刊编辑。所有 STM 期刊的稿件平均录用率约为 50%。

表 3 出版期刊数量位于前 10 位的大型出版商

出版商	期刊数量	累积占期刊总数的百分比
Springer (不含 NPG)	2987	10.6%
Elsevier	2500	19.5%
Wiley	2388	28.0%
Taylor & Francis	2105	35.5%
SAGE	750	38.1%
Wolters Kluwer (包括 Medknow)	672	40.5%
Hindawi	438	42.1%
CUP	350	43.3%
OUP	362	44.6%
NPG (包括 Frontiers)	178	45.2%

尽管同行评议有一些缺陷（例如存在潜在的偏见——见“对同行评议的批评”章节），但这种方式在学术界仍然被广泛使用。例如，出版研究联合会（PRC）调查显示有 93% 的人认为同行评议是有必要的（Ware & Monkman, 2008; Sense About Science, 2009）。尽管学术界整体广泛认同同行评议，但越来越多的作者认为该方法需要改进，尤其是在评议时间过长和评议人存在偏见方面。相比 1993 年至 2005 年间的调查，Mulligan & Mabe 在 2011 年发现研究人员对同行评议的态度基本维持不变：同行评议依旧被高度重视，并且有很大一部分人愿意继续承担评议工作。

最近由 CIBER 实施的一个大规模调查（Nicholas 等, 2015）继续显示出对同行评议的强烈支持。该调查发现同行评议的关键性益处是，对等审查的关键利益被认为是起到提供信任的核心支柱的作用，尽管对文章的改善也得到珍视。对同行评议的最大批评是时间漫长、不进行充分干预的放手编辑、轻度同行评议——研究人员珍视有价值的详细且强有力的评议、以及变化较大的评审质量。研究人员（尤其是年轻人）愿意使用非同行评审的材料，但远不太可能引用它们：被视为正式的活动需要有经同行评审的内容。同行评议对于期刊的选择也仍然重要，并且影响因子仍然很重要。社交媒体和开放获取没有被视作改变对同行评议的重要替代物：研究人员已经从以印刷为基础的体系转向数字化体系，但这并没有显著改变人们决定信任什么的方式。

同行评议的有益之处

有大量的支持同行评议的观点。同行评议可被视为提高研究质量的保障过程

（明显不同于在发表前提高稿件的质量的过程）。尽管有些人认为提高研究质量是同行评议的目的之一，但这为同行评议设置了一个很高的门槛，且目前很少有证据显示同行评议在这方面的有效性（Jefferson, Rudin, Brodneý Folse & Davidoff, 2007）。

另一方面，研究人员支持同行评议是因为他们认为这样做可以提高发表的论文质量。PRC 的调查显示，压倒多数（90%）的研究人员认为同行评议的主要作用就是提高发表论文的质量，并且有同样多的人认为这种做法提高了他们本人上一次所发表论文的质量。Mulligan 和 Mabe 报告了相似的结果，尽管这种认同的程度因不同的学科领域而略有变化。

同行评议也能充当过滤器以惠及读者。对于职业研究人员，这种最重要筛选的意义不仅在于论文经过了同行评议这一事实，而且还在于这种筛选通过感知质量为期刊的分级提供了基础：同行评议的过程其实是将更好的文章发表到更好或最合适期刊的过程。尽管还在热烈争议这样的分级是否为最有效的文献筛选方式，但到目前为止同行评议对于作者和管理者仍然是一个重要的信号（Tenopir, 2010; Ware, 2011）。

同行评议还可作为认可的一个凭证，例如，可以用来区分可信的经过同行评议的科学和未经同行评议的材料。相比科研人员，这个区分对于外行读者和期刊从业人员可能更加重要。然而，“掠夺性期刊”和类似期刊的增长，也许已经增加了区分正确履行同行评议的期刊和论文与那些伪装同行评议的期刊和论文的那些需要（见下文“同行评议的校验”）。

对于同行评议的批评

对于同行评议当然不是没有批评。主要的反对声音认为它不够有效、不可靠、检错能力不足、受偏见的影响太大（尤其是在单盲评审方式的情况下）、为评审提供了不端行为的空间、评议时间长而不必要地耽搁发表时间（Ware, 2011）。补救的方法包括有争议的开放评议（见下文）——被认为可以同时提高评议的公平性和质量；级联评议（cascade review）——旨在减少低效并加快发表速度；发表后评议——最为激烈的方式（“先发表再筛选”模式），通过论文发表后实施评议来加速发表。

同行评议的种类

目前有两种主要类型的同行评议在广泛使用，单盲评议（审稿人知道作者的身份但作者不知道审稿人的身份）和双盲评议（审稿人和作者双方互不知道对方的身份）。单盲评审比双盲评审明显更加普遍（例如，PRC 调查显示有 84% 的作者经历过单盲评审，但只有 44% 经历过双盲评审），但目前学术界有相当多的呼

声支持双盲评审，估计是回应单盲评审中审稿人潜在的偏见。目前双盲评审在人文和社会科学领域期刊比（基础）科学领域期刊更加常见，临床医学类期刊则介于二者之间。然而，这种情况也许正开始发生改变，因为有些领衔期刊如 *Conservation Biology* 和某些 *Nature* 研究类期刊正采用双盲评议(Cressey, 2014)。

双盲评审最根本的缺陷是它很难对审稿人隐藏作者的身份。大多数作者通常都更多地引用自己先前的成果，而且稿件中的研究主题和写作风格也可能会把作者的身份暴露给知识渊博的同行们。另外，期刊使用双盲评审也许会招致明显更多的编辑成本，以在评审前消除稿件中作者的身份线索。

开放同行评议

一个更加新颖的解决对单盲评审非议的处理方式是开放式同行评议：在这种模式中，作者和审稿人互相知道对方的身份，并且审稿人姓名（可选）和他们的评议报告也与论文一起发表。开放式评议的赞同者认为这样更加公平，因为某人对他人工作的重要评估不应秘密进行。同时，他们认为公开身份有助于审稿人避免随意和粗鲁的评论从而更好地完成工作。

历史上，开放式同行评议比两种标准的评议方式（单盲和双盲）更加少见（PRC 调查中有 22%的作者表示他们曾经历过开放式同行评议）。受调查的作者对这个方法表示出有限的支持，并且似乎不愿参与（例如在《自然》的开放式评议试验，Campbell, 2006）。最重要的原因可能是审稿人比较介意因负面评审而可能产生的后果。

尽管有上述的担心，对开放式评审的支持显示仍在增长：随论文一起发表的审稿意见（署名或不署名）也正变得更加普遍。比较著名的例证包括《英国医学杂志》(*BMJ*)、生物医学(BMC)期刊群、欧洲地球物理学会(EGU)的期刊群、欧洲分子生物学组织(EMBO)的期刊群等，这些都是开放评审有益处的很好案例(Pulverer, 2010)。

发表后评审

近年来，电子出版技术使得不同形式的开放式同行评议有所发展，所有读者，不限于编辑遴选的审稿人，都可以在论文发表后对论文进行评论，甚至对论文进行基于数字分级的评级。发表后评议可针对无论是否经过传统同行评议的论文。这样做的好处是能够将广泛的意见作为参考（“大众的智慧”），并且使评议处于一个更加活跃和透明的过程。知名的案例是 *PLOS ONE* 期刊，尽管该刊所发表的文章只有不到 10%有评论。对于发表前的开放评议，到目前为止学界不情愿参与大多的试验。除了类似于对出版前开放评议的担忧外，学者们还表示他们没有足够的时间针对已发表的论文撰写大量的评审意见。

然而，最近的发展表明这个概念的最好时期正在临近：

- NIH 创建了 PubMed Commons，这是一个允许评论 PubMed 数据库中 2200 万篇文章的试点服务。评论仅限于 PubMed 的作者（在该领域是一个很宽泛的范畴），并且用户可以评论和分级其他评论的有用性。在生物医学领域 PubMed 的中心地位赋予了该方法很高的信誉。
- 开放获取出版商的 Frontiers (Macmillan Science & Education (麦克米兰科学与教育) 所属的一部分) 已经开发出一种创新的评价系统，该系统采用算法提供发表后研究的评价，使读者间接地参与对某篇文章的排名。
- PLOS 已试验开放评估，允许用户（在私人测试版）对文章进行四个维度的分级：兴趣水平、文章的意义、研究的质量、写作表达的清晰度。这些都是典型的传统上出版前评议的问题，这大概不是巧合。
- 目前声称有超过 1600 万的注册用户的 Academia.edu 宣布收购启动 Plasmyd，以期整合其技术以支持出版的同行评议。

与此密切相关的是，有越来越多的兴趣关注于一篇文章潜在影响的聚合“信号”，包括发表后文章的评论数（在杂志的网站和其他网站），也可以作为对影响因子的补充（见“论文层面的计量和替代计量”章节）。

级联同行评议

级联同行评议在评审过程方面与普通同行评议有所不同。这种方式是通过（经作者本人同意）将论文和先前的评审意见一同转给新的期刊，从而避免了每次论文遭到拒绝后投到另一期刊所必需的重复评审。这种方式最初被开放获取出版商 BioMed Central 采用，随后成为 *PLOS ONE* 类巨型期刊的特点（虽然这种评审方式在 *PLOS ONE* 的论文中从来也没占多少比例）。对级联同行评议持积极态度的是神经科学同行评议联盟的期刊，它们同意接受联盟内其他期刊的审稿意见，尽管这些加盟的期刊属于不同的出版商。实际是该联盟只形成了很小的影响，只有很少一部分稿件转投。

此外，eLife, BioMed Central, 公共科学图书馆 (the Public Library of Science) 和 EMBO (the European Molecular Biology Organisation) 于 2013 年宣布成立一个新的同行评议联盟 (Clarke, 2013)。

便携式同行评审

在级联式评审中，同行评议由期刊以正常的方式进行，如果稿件被拒绝，评审意见可能伴随着稿件投稿到一份新的期刊。

一个更激进的想法是作者在向期刊投稿前其本人寻求同行评议。这也许可作为旨在传统方式的投稿前提高稿件的预投稿过程，或者甚至是参与期刊可能同意接纳的（如果不是必须被约束）完全的“便携式”的审稿。

有两个机构提供便携式的同行评议服务，各自有显著不同的商业模式。rubriq²¹以 500~650 美元的费用为作者提供同行评议服务。Rubriq 希望作者可以承担这一费用以作为提升论文质量的成本，或者开放获取期刊在 APC 中扣除此费用以反映提交的评议的价值。Peerage of Science²²提供一个供期刊独立评议的平台，出版商可以在该平台搜寻潜在的投稿。它不收取作者的费用，但每篇发表的文章要收取出版商约 400 美元以涵盖其成本。

同行评议滥用和不当行为

有些研究人员已经利用同行评议系统中的漏洞来操纵同行评议以便对其有利。2014 年 *Nature* 的一篇文章详细介绍了有些作者为达到此目的，如何使用多个通用的电子邮件地址（例如 Gmail）在在线审稿平台建立虚拟的和实际的研究人员的欺诈账户（eLife, 2014），然后在投稿时建议他们本人（虚拟的欺诈账户）作为稿件的审稿人。在过去的 2 年中，期刊至少有 6 次因同行评议操纵而被迫撤稿 110 余篇文章。在一个主要的情况中，SAGE 出版社基于 14 个月的调查发现了约 130 个可疑的审稿人帐户并撤回了 60 篇论文。（另见“出版道德”章节。）

奖赏审稿人

期刊文章的同行评议历来被视为研究人员职业责的一部分，费几乎从来不支付审稿费（一个罕见的例外是 CDL 新近创办的巨型期刊 *Collabra*；也有实物形式的支付，如免收投稿费、免收或打折 APC、提供有限地获取基于征订的资源等）。

有关研究人员和出版商的日常经验的调查表明，对此类费用的需求是很有小的（虽然反感情绪可能认为出版商的利润很大但不支付同行评议费用）。然而，确实出现了对审稿工作进行更正式认可的需求。目前采用盲审的期刊对此认可的方式通常是采用年度声明的形式列出其审稿人以示感谢。研究人员可以并且确实有人在其简历中列出审稿活动。

通过认可奖赏审稿的更直接方式正在出现：

- Publons 提供一种审稿人可以发布其同行评议历史的在线服务，然后根据他

²¹ <http://www.rubriq.com>

²² <https://www.peerageofscience.org>

们选择来展示（比如在简历中展示）。在本报告撰写的时候，Publons 已经记录了 59000 个审稿人的近 30000 条审稿意见，覆盖 2800 期刊。Van Noorden (2014d) 访谈一些更多产的用户。

- 出版商正在试验更直接的方法。例如，Elsevier 举办一个公开竞争以提高同行评议；目前正在实施的是获胜者入选建议的“审稿人徽章和奖励计划”。审稿人可以在社交媒体页面展示徽章（通过 Mozilla OpenBadges 生成）。在第二阶段的“审稿人认可”平台已经为约 40 种期刊开发。一旦为这些期刊中的某份期刊完成一位审稿人，审稿人就能获得该平台的一个个人主页的链接，以显示他们的审稿活动（van Rossum, 2014）。

同行评议所花费的时间

同行评议不可避免地要花费时间。不同学科要求的审稿时间各不相同，高速发展的生命科学发表快速，期刊要求审稿时间常以周为单位（或更短）；而数学、人文学科和社会科学的论文则可以用更长的时间（数月或更长）审稿。PRC 的调查显示平均的审稿时间约为 3 个月。一般来讲，作者认为 30 天或更短的审稿时间是可以令人满意的，但如果审稿时间超过 3 个月，满意度就会直线下降，只有不到 10%的作者对超过 6 个月的审稿时间满意。

学术界从事同行评议活动的程度可以通过花费在评议方面的时间来反映。PRC 的调查显示，审稿人在每次审稿中所花费时间的中位数为 5 个小时（平均数为 9 个小时），且平均每年评审 8 篇论文。然而，大多数的稿件评审是由一部分高产出的审稿人完成的，这些人的审稿量差不多是平均值的两倍。

全球同行评议的成本相当大，尽管大多数都是非现金成本：RIN 的一项报告估计，每年的同行评议成本是 19 亿英镑，相当于每篇论文 1200 英镑（RIN, 2008）。霍顿报告（Houghton Report）得出的数值更高，认为每篇论文平均成本是 1400 英镑（Houghton 等, 2009）。这些数字为全部成本，包含学术界从事评议所花费的时间。根据 PEER 报告，出版商在管理同行评议方面花掉的平均费用（只算薪金和直接费用，除去管理、设备、系统等方面的费用）为平均每份投稿 250 美元（Wallace, 2012）。

出版商在同行评议中的作用

出版商在同行评议中最基本的作用就是构建并支持一个期刊及任命并支持期刊的编辑和编辑部。从组织管理的层面来讲，出版商的作用是组织协调工作进程，近年来也起到开发期刊采编网络平台支持评议过程的作用。在线投稿系统目前已经很普遍：Inger & Gardner (2013) 报道只有 5%的出版商没有同行评议系统（由 2008 年调查的 35%急剧下降）。大多数出版商倾向于选用市面上的三大系

统之一, Editorial Manager(Aries)、eJournal Press 以及 Manuscript Central(Thomson Reuters)。

一项汤森路透基于其在线投稿系统 ScholarOne 的数据 (涵盖 2012 年的 365 个出版商的 4200 种期刊) 分析表明, 稿件提交数量从 2005 年的 31.7 万篇增加到 2010 年的 100 万篇²³。(这些数据未对使用该投稿系统的期刊数量的增加而导致的稿件上涨作出调整)。在这段时间内, 稿件的录用率从 41% 下降到了 37%。从稿件上交到最终接受间的时间从 2005 年的 65 天下降到了 2010 年的 59 天, 不过从投稿到第一次决定的时间维持在 40~41 天 (Morris, 2009)。

出版商使用同行评议操作的规模可由领先的在线投稿系统的吞吐量来说明。至 2013 年 ScholarOne 系统每年处理的原始稿件提交总量为 160 万篇 (包括重新递交的稿件则为 220 万篇)。其竞争对手 Editorial Manager 每年处理的稿件为 200 万篇, 涵盖 250 多家出版社的 5800 种期刊。

在线投稿系统的使用缩短了同行评议所耗费的总体时间, 并减少了一些相关的直接成本 (例如纸张打印费用和邮寄费用), 但这些开支往往会转变成管理费用 (软件、硬件及培训费用)。启动全电子化的流程也带来一些另外的好处, 包括如下几个方面:

- 更快捷的发表: 系统能生成一经录用即可发表的全链接的同行评议后的作者最终稿件。
- 更高效的生产: 系统能自动执行“飞行前检测”, 例如在投稿阶段检查图片的分辨率。
- 为审稿人和编辑提供支持: 自动链接作者稿件中的参考文献, 有助于编辑遴选审稿人并帮助审稿人评审稿件。有些出版商还为编辑提供文献检索系统的查询, 以便评估和遴选审稿人。更新的基于文本挖掘的人工智能系统还可以整合在线投稿系统以帮助识别审稿人。
- 查核学术不端: CrossCheck 等检测系统可以将提交的稿件与已发表的论文及与网上的文章进行对比 (见“出版伦理”章节)。
- 集成电子商务: OA 的文章处理费、或版面费或彩版费等可以使用出版商自己的系统或类如 RightsLink 的第三方的插件管理 (版权结算中心, CCC)。
- 投稿或录用时收集的元数据可用于与其他服务生成集成; 例如, 依赖于

²³ 至 2013 年原始稿件提交的数量为 160 万篇, 包括重新递交的稿件则为 220 万篇 (据 Thomson Reuters)。

FundRef 数据采集的 *CHORUS*。

同行评议认证

如果文章的读者无法知道他所读的这篇文章是否经过了同行评议，在没有很好地了解某一个领域期刊的情况下参照什么标准来判断文章质量呢？外行读者通常不会有这方面的知识，即使是专家级的研究人员遇到非本领域的文章也是如此，并且采用良好同行评议的期刊也不总是（在文章层面）清楚地标注文章是否经过了评议。为了解决这个问题，一些倡议提出文章和/或期刊层面的同行评议过程认证。

有一个这样的认证是 PRE（同行评议评估，Peer Review Evaluation）²⁴，“一套旨在支持和加强同行评议过程的服务”。它的第一个服务是 PRE-val，为最终用户验证内容是否通过同行评议过程并提供有关过程质量评估的信息。此外，文章水平的服务 preSCORE 也被初步提出。

同理，BioMed Central 的医学编辑呼吁采用“风筝标记”来标识经过专业人士同行评议过的研究论文（Patel, 2014）。

2.10 阅读方式

大学教职人员人均每年阅读的论文数量正在逐年稳步增长，如图 18 所示（Tenopir, 2007; Tenopir, King, Edwards & Wu, 2009）。其他来源的数据也给出了相似的估计，认为大学教职人员大约每年阅读量为人均 250~270 篇论文，而在大学就职的科学家人均每年阅读的论文数量只是这个数字的一半（King & Tenopir, 2004），不同学科间有很大的差异（见“学科间差异”部分）。最近 Tenopir 对英国的研究显示，大约每个学者每月阅读 39 篇论著，其中包含 22 篇论文，7 本书和 10 个其他类型的出版物（Tenopir 等, 2012），累积为每年约 448 小时的阅读时间（相当于 56 个每天 8 小时）。

2008 年的一项国际性调查（Tenopir, Mays & Wu, 2011）发现，科学研究人员表示他们每周阅读学术内容的时间介于 12.3 小时（健康科学）和 15.3 小时（生命科学）之间；而社会科学的研究人员则称每周花了（有点难以置信？）25.9 小时（虽然总体上并没有阅读更多的文章）。

阅读的广度也随时间的推移而增加：在 1977 年 Drexel 的科学家平均每人每年阅读 13 种期刊，现在这个数字是那时的两倍。

²⁴ <http://pre-val.org>

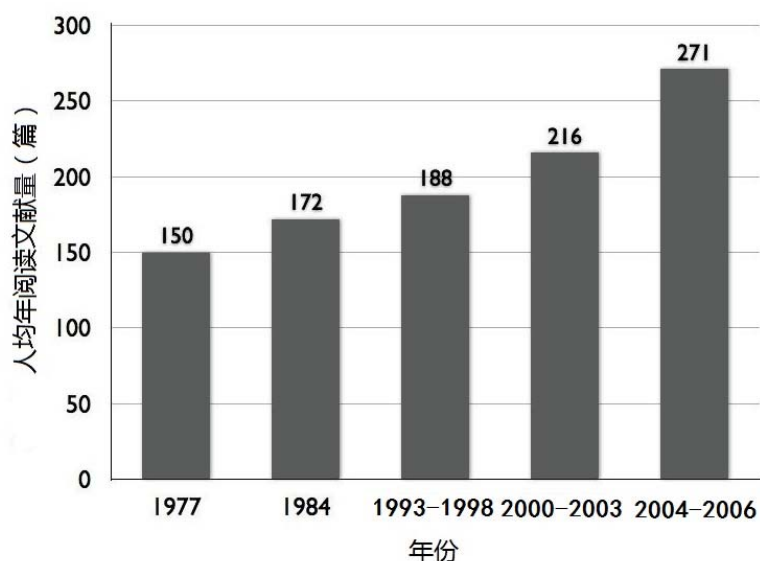


图 18 大学教职人员每年人均阅读的论文数量 (Tenopir, 2007)

在 1977 年到上世纪 90 年代中期, 研究员平均阅读每篇期刊论文的时间维持在 45~50 分钟, 但在其后这个时间下降到略多于 30 分钟 (Renear & Palmer, 2009)。健康科学的研究人员每篇文章花费的阅读时间最少, 为 23 分钟左右, 相比其他科学和社会科学研究人员为 30 分钟 (Tenopir 等, 2011)。这还是在期刊中论文的平均长度大幅增加的情况下 (从 1975 年平均每篇 7.4 页增加到 2001 年的平均每篇 12.4 页)。

CIBER 研究小组受 RIN 资助的调查得出了一个合理的解释 (Nicholas & Clark, 2012)。在分析出版商的记录后, 他们发现很少有用户在学术网站花费长时间阅读电子论文。一般浏览时间都很短, 而且大多只浏览 1 至 3 页, 一半的浏览者不会再回来阅读同一篇论文。有研究显示, 只有 40% 的研究人员表示全部读完了他们阅读的最近一篇对自己“重要”的论文。用户一般会把论文下载下来留作以后的参考, 但追踪调查发现至少有一半的文章在下载后没有被阅读 (这还可能是乐观的估计)。CIBER 的作者们表示研究人员在电子环境中已经由纵向阅读转换为了横向阅读, 这意味着他们在不同文章之间快速地扫读 (“跳跃式, 翻阅式、浏览式”), 而非深度阅读某一篇文章。虽然作者们指出有很多现代阅读环境因素导致这种阅读习惯 (过多的文章、更少的筛选时间、更多的工作量、多任务情形的正常化、社交媒介使得我们适应接收快节奏信息) 的形成, 他们也指出研究人员可能总是选择式阅读和片断式阅读, 深度阅读作为常态只是一个传说。

Renear 和 Palmer (2009) 讨论了阅读策略和技术创新 (“策略阅读”), 这可

帮助读者从更多的论文中提取信息而花费在每篇论文的时间更少。这种方式中有相当多的聚焦技术，包括语义网络技术（如分类学和本体论），文本和数据挖掘技术，以及新计量方法的使用。下文将讨论（见“学术交流的新进展”部分）。

论文的获取和查找

学者们使用多种方法来定位论文，如图 19 所示，图 20 显示了更详细的近期数据。在这些数据中可以明显发现在线搜索的重要性增加以及浏览的重要性减少（很多出版商的网页浏览记录中都表明有大约 60% 的阅读来自谷歌搜索引擎）。询问同事仍是一个重要的查找论文的办法，虽然排名位于浏览和搜索之后。

在上世纪 70 年代到 90 年代间，阅读论文的来源基本完成了从个人订货到图书馆提供获取的转变。

读者在网上获取和查询期刊内容的方式对出版商和图书馆人员造成了影响。Inger 和 Gardner 在 2012 年通过对引用搜索、核心期刊浏览和主题搜索的调查（Inger & Gardner, 2012, 对 2005 年和 2008 年调查的修订），得到如下结果：

- 读者比以往任何时候更有可能直接从论文或摘要进入期刊，而不是从期刊的主页（更不用说出版商的主页）的导航找到论文。这当然部分地归因于搜索引擎（特别是 Google 和 Google Scholar）对学术内容的定位，但调查中值得注意的是读者所采用途径的多样性。专业书目数据库依然是读者搜索特定主题文章的最常用的选择，仍然领先于网络搜索引擎。学术搜索引擎（Google Scholar、Microsoft Academic Search）似乎在 2012 年超越了一般性搜索引擎。
- 读者非常看重期刊网站的目录提醒服务（期刊内容更新提醒是读者得知新论文的最主要方式），并且十分重视将期刊网站的主页作为获取最新论文的地方，但并不看重个性化和搜索功能（也许是因为他们习惯于使用外部搜索引擎搜索多个期刊/出版商）。RSS（聚合内容）提醒仍然是小众工具但其使用量在 2005 至 2008 年有大幅增长。
- 学科之间有一些显著的差异，例如，人文和教育方面的研究人员比物理学和天文学方面的研究人员在有可能使用图书馆网页搜索文章。
- 地区差异可能也很重要：例如，亚洲研究人员更可能直接搜索文章，而不是从期刊提醒或论文的引文中获取文章。
- 图书馆的在线目录和网页开始受到通用搜索引擎增长的影响，但仍然保持着作为查询起点的重要性，尤其是用于主题和后续引用的搜索。图书馆控制的网络空间具有链接到图书馆购买内容的优势，且满足图书馆的选择标准。图书馆部署的链接解析器和网络层面的发现技术进一步加强了其重要性。

- Inger 的报告指出，尽管出版商了解读者很少使用个性化功能，但迫于来自编委会和学会会员的压力，他们依然保留了这方面的功能。

“Y 一代”研究调查了出生于 1982~1994 年间的博士生关于信息查找方式（JISC & British Library, 2012）。结果显示电子期刊是所有学科的主要搜索来源。尽管这些博士生被视为复杂的信息搜索者和复杂来源的用户，他们却比老一辈研究人员更加习惯于在没有在线文章的时候将就着使用论文摘要。尽管他们总体上来讲是新信息技术的积极使用者，但他们在自己的研究中依然十分谨慎地使用最新的网络技术，且只在该方法可以很方便的结合自己的工作的工作的情况下才使用，对社交媒体不太热衷。

Newman & Sack（2013）的一个采访研究提供了一些关于阅读和信息发现行为的定量数据的有用的定性背景。举例来说，下述的建议在通过提供易于使用的工具来帮助研究人员保持了解知识进展方面仍然有很多要发展的：

“大多数受访者没有一个系统的策略来保持跟进最新的知识。[...] 受访者严重依赖于被引用的文献、同事的建议、或者是一小部分熟悉的期刊。只有少数人获得文摘和索引数据库的提醒并辅以重要期刊的提醒。有不少人表示对于寻找最新信息时缺乏技巧而失望。

“有一位计算机科学部门的受访者表示，‘对于不能做足够的工作以保持了解最新的信息，我不断地有内疚感。’”

Ithaka S+R/JISC 于 2012 年对英国学者的调查证实上述图景，受访者称保持了解最新研究进展的首要方式是参加研讨会，其次是其他学者建议的材料，浏览最新出版的重要期刊，浏览目录。电子发现工具（推荐系统，保存关键字提醒）的重要性要小的多，社交媒体的重要性最小（Ithaka S+R, JISC & RLUK, 2013）。

就阅读地点而言，阅读模式正在慢慢改变：在一项 2005 年的调查中有相当一部分对象（22%）都表示更喜欢在舒适的家中浏览网上资料，这一比例在医学研究人员中最高，达到了 29%（Mabe & Mulligan, 2011）。

2.11 学科差异

值得注意的是以上描述的一般性情况隐藏了不同学科间在出版情况及阅读和使用学术材料方面的一些重要的差异。

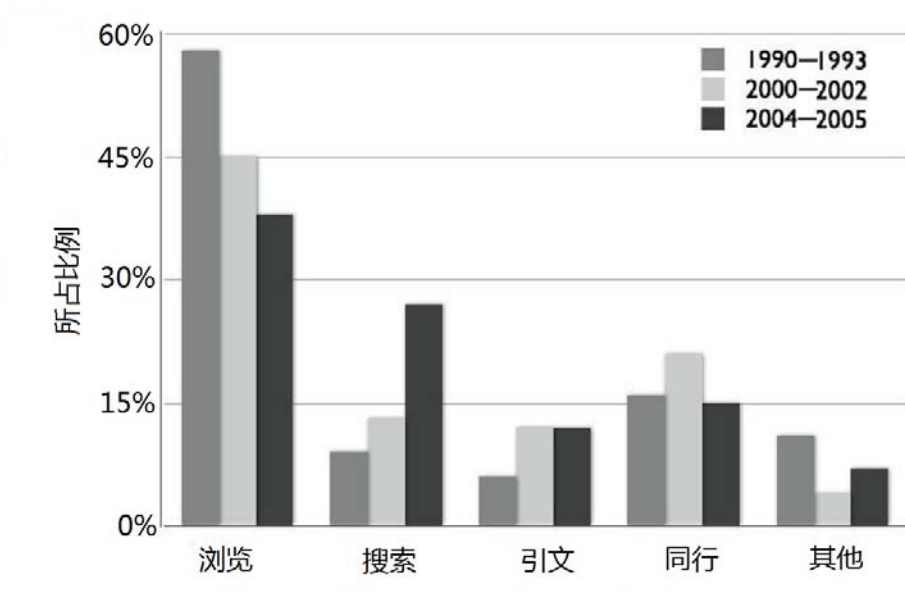


图 19 大学教职人员查询论文的方式 (Tenopir, 2007)

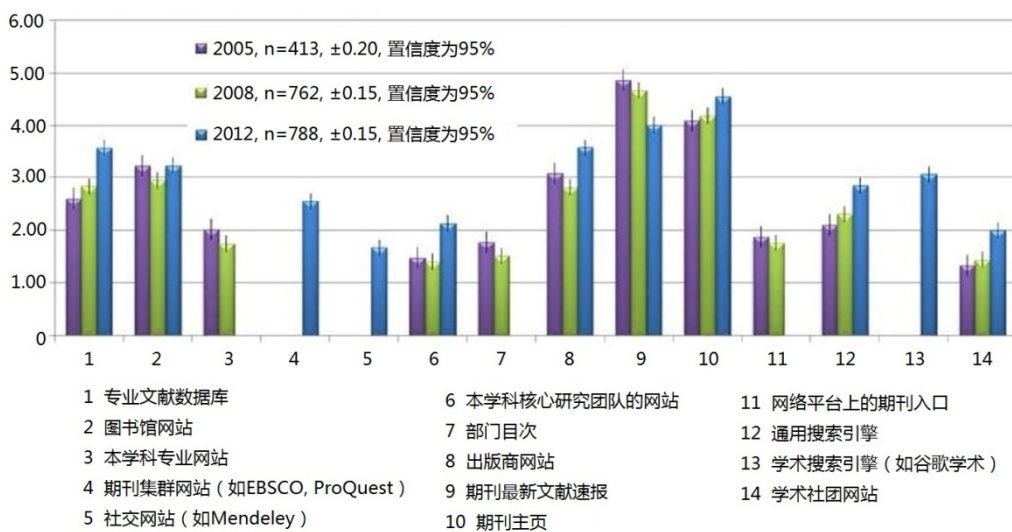


图 20 发现最新文章起始点的趋势：2005 至 2012 年间的趋势

(Inger & Gardner, 2012)

例如，虽然《期刊引用报告》(JCR) 收录的期刊平均每年发表 120 篇论文²⁵，科技期刊年均发表文章数量为 140 篇左右，远高于社会科学和人文学科学期刊的

²⁵ 严格地说，这是指“可引用的论文类型”，即：学术论著包括但不限于论文、综述和进展性文章。数据来源于汤森路透的产品《期刊引证报告》。

年均不足 45 篇。这也可部分地解释为什么科技期刊的价格往往远高于人文及社会科学期刊的价格。

2005 年英国 JISC 有关学科差异的报告 (Mabe & Mulligan, 2011) 是基于对英国学术界的调查,但这个报告的结果不见得不适用于更广的范围。该研究发现:

- 不同学科群的论文产出数量有很大不同。“硬”科学(物理、生物医学和工程)中发表的论文数量最多,为人均每 3 年约 7.5 篇,而社会科学其次(人均每 3 年 5 篇),艺术和人文科学最少(人均每 3 年不到 3 篇)。
- 论文作者的合作度也有显著差异并且与产出数量的模式相似。生物医学作者最倾向于合著(85%的受访者表示其产出的 75%甚至更多的论文是合著的),物理学和工程科学其次,其后是社会科学,而艺术和人文科学最少(76%的反馈者称他们有 25%甚至更少的文章是合著的)。
- 众所周知,期刊在科学技术及医药领域的重要性远超其在人文及艺术中的重要性(在后者的领域中书和专著更加重要)。不过,报告称随着期刊论文在人文艺术领域的重要性增加,这个差别正在减小。一个可能的原因是研究评价中对(高影响因子)期刊中所发表论文的强调。
- 不同学科所需求的论文年份高峰值变化很大。人文学科需求的论文年份峰值约为 20 年前,化学、工程科学和医学是 10 年前,计算机科学、生命科学和信息科学是 5 年前。

2005 年 JISC 报告指出研究人员阅读(和编写)书籍相比于阅读期刊论文可能在下降,后来的 RIN 研究证明了这一点,并发现研究人员也对这一现象表示关注。目前还不清楚究竟是因为图书馆预算缩减导致图书供应量下降,还是在线论文的增加,或许仅仅是由于研究人员缺少时间,但文献计量分析证实期刊论文和其他出版形式的产出中对书的引用显著减少(RIN, 2009a)。另一项研究证实该变化中评估活动产生了此压力(Adams & Gurney, 2014; 见“研究评估对研究者行为的影响”)。

有一系列精彩的信息使用的案例分析,深入研究了不同学科间(包括生命科学、人文科学、物理科学)的研究者如何发现、获取、分析、管理和传播信息(RIN, 2009c; RIN, 2011d; RIN, 2012)。由于研究结果十分丰富和详细,因此在此很难总结;但这些研究结果表明,即便是在同一学科的研究人员也不会采用一个单一的“工作流”。

在某些学科,期刊“认证”功能的重要性远低于其他学科,这体现在某些学科中人们愿意接受预印本(未经同行评议的作者的原稿)代替最终发表的版本。认

证在理论性学科和大型实验性学科（高能物理学、理论物理学、数学、计算机科学）中不那么重要，这些学科中合著非常多并且领域之小意味着每个研究人员的研究质量能被同行个人了解，但是，期刊“认证”功能在中-小实验领域（生命科学、化学、地质学等）更为重要。值得注意的是，从单纯的人数上来讲，后者拥有全球研究人员数量的绝大部分。

不同学科间的论文阅读和查询手段有着很大的差别。例如，医学研究人员人均阅读论文数为人文学者的三倍（见图 21）。这些数据反映出期刊论文在这些领域中的相对重要程度和组成阅读的“本性”，并且也反映了解释诸如医学这类主要由执业人员组成的学科的难度。图 22 以浏览中重要性的显著差异为例揭示了读者寻找论文的不同方式。

不同学科的作者对同行评议有着明显不同的态度。总体上，PRC 调查显示物理学和工程学科的作者认为同行评议更加有效，他们比人文学科和社会科学作者对当前的评议制度更加满意。双盲同行评议在人文和社会科学（94%的作者经历过双盲评议）比在物理学和工程科学中（31%）更加常见，并且人文和社会科学的作者比其他学科的作者优先双盲评议远甚于单盲评议。

不同学科的研究人员对开放获取的态度也有明显的差异。这些差异有些反映了资助结构（如在人文科学和数学领域缺乏外部研究资助），而其他差异反映出研究群体长期规范（如早于开放获取的预印本文化）。更多细节参见[开放获取](#)章节。

然而，不同学科领域在某些方面显示出没有（或很小）的不同：

- JISC 研究发现，英国在期刊和学术资源获取方面，不同学科间差异很小；后来的 RIN 研究证实了该结论（RIN, 2011a），但在产业部门的学科领域中有所不同（见“研究人员对期刊的获取”部分）。
- 所有学科的作者都称职业发展和同行交流是发表论文的最重要原因。

2.12 引用和影响因子

引用是科技论文的一个重要的部分，它可以帮助作者通过引用先前的工作在不重复详细解释的情况下建立自己的论点；它还可以十分有效地指引读者阅读相关的文献（调查显示这是读者浏览文献的最常用方法，见 Inger & Gardner, 2012）。电子期刊额外地允许超链接，即链接到引用的论文，该功能也可得到索引和检索服务的支持。

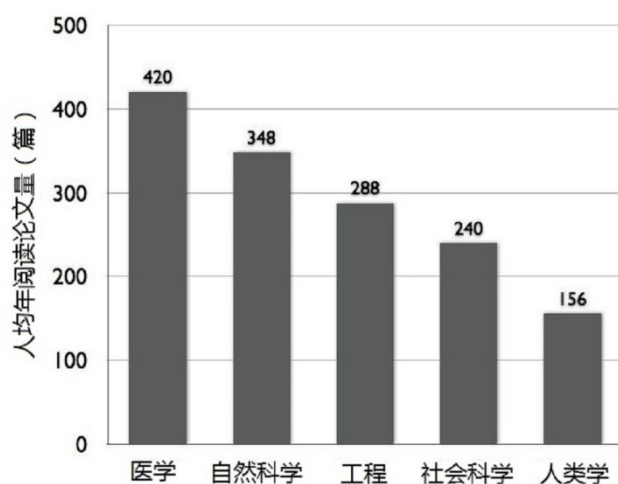


图 21 平均每个大学教职人员每年阅读的论文数量 (Tenopir, 2007)

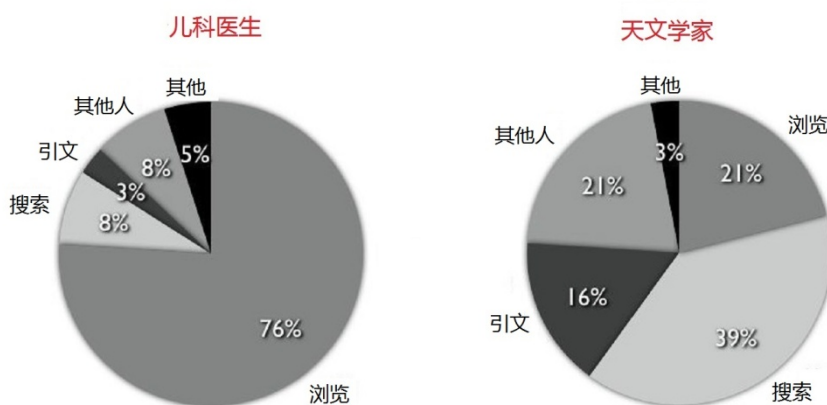


图 22 不同学科间查找论文方式的不同 (Tenopir, 2007)

引证膨胀

引文数量的增加超过出版物的增加。相比 1999/2003 的 5 年, 在 2004/2008 的 5 年间出版物的数量增加了 33%, 而引用数量增加了 55%。图 23 显示出 1992~2012 年每篇论文中平均引文数的增长趋势, 所有国家的平均数值由 1992 年的 1.7 增加至 2012 年的 2.5 (NSF, 2014)。增加的因素可能是文献数量的增长 (有更多可以被引用的材料)、合著数量的增长, 以及近来参考文献列表越来越长的趋势 (Elsevier, 2011)。

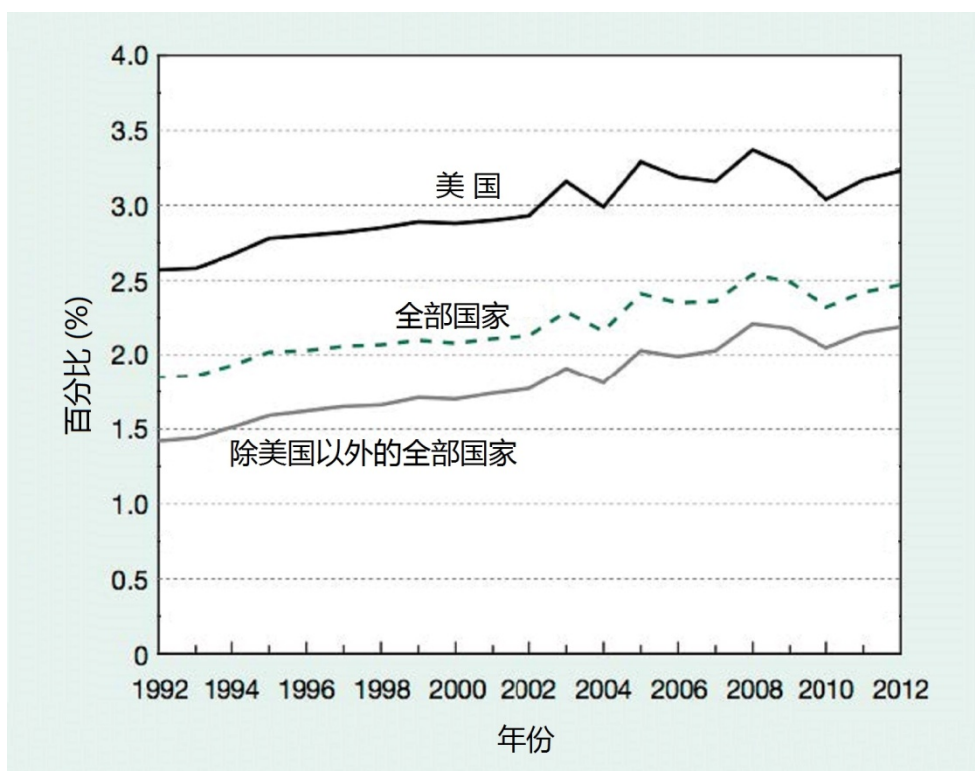


图 23 引证膨胀：基于作者国别的每篇论文中平均引文数的增长 (NSF, 2014)

引用的国际趋势

在论文出版方面，全球化导致全球范围内不同地区引用份额的变化。表 4 显示了 2000 年到 2010 年间的变化：在这段时间中美国和日本的份额在减少，而中国和其他亚洲国家的份额在增加。图 24 中的近期数据证实了这种趋势的继续 (Elsevier, 2013)²⁶。

科学研究的国际化发展体现在作者越来越多地引用作者本国以外论文。如同国际合著（见“合作和合著”部分），除中国以外的所有主要科技国家的国际引文在过去的二十年稳步增长。1992 年中国科技论文中的 69% 文献来自中国以外，2012 年这一比例已下降到 49%，这表明中国论文产出的增加正在中国境内被使用 (NSF, 2014)。

²⁶ 图中的数据与表 4 中 NSF 数据的不同在很大程度上可能是由于使用了不同的数据集：Elsevier 使用的是包含了约 22000 期刊的 Scopus 数据库，NSF 使用的是包含 5087 种期刊的 Web of Science。

表 2 各国科技论文的引用份额（科技与工程指标-2012, NSF, 2012）

地区/国家	2000 (%)	2010 (%)
美国	44.8	36.4
欧盟	33.3	32.8
中国	0.9	6.0
日本	7.1	5.7
亚洲 8 国	1.8	5.3

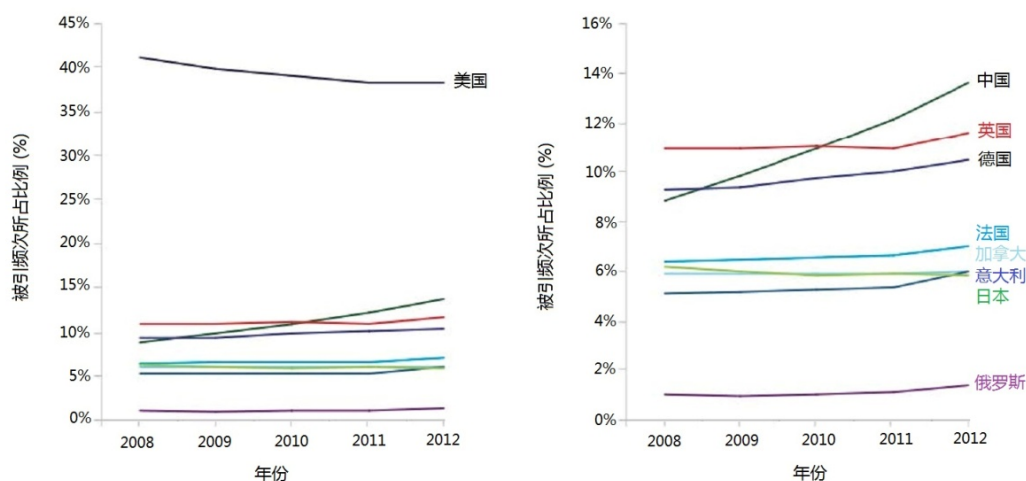


图 24 世界引文的份额（Elsevier, 2013）

引证与影响力计量

一篇论文被引用的次数常常可以用来测度其影响力并扩展到测度其质量。使用引用数作为影响力和质量的替代已经以影响因子形式延伸到期刊。一份期刊的影响因子是对其在某段时间内平均每篇论文被引用频率的衡量。（官方定义是：影响因子是一份期刊中论文在出版后的第二年和第三年被引用的次数，除以在同时期中所有可被引用论文的篇数）。

虽然使用引用数据（尤其是期刊层面的影响因子）评价一个研究人员或者一个研究部门的研究产出的做法很普遍，但却受到越来越多的批评。发表于同一期刊的所有论文的质量都相差不多的假设也与研究数据相悖：一个偏态分布显示期刊中一般有 15%的论文占有 50%的被引用量，而有 90%的被引用数量出自 50%

的论文 (Seglen, 1992)。也就是说, 一份期刊中被引用次数较多一半论文是较低一半论文的 9 倍。人们对这影响因子的不满导致了替代计量方法的发展(见后文), 虽然这种方法目前还处于发展初期。

引用的分布遵循广泛发现的帕累托模式, 约 80%的引用来自约 20%的文章。例如, Scopus 引用数据显示 2008 年论文在 2008~2012 年的被引用显示几乎就是这个结果, 而有 32%的论文其间未被引用 (Elsevier, 2013)。

在尺度的另一端是使用最高被引 1%的论文来测度国家和机构的影响力。图 25 显示, 尽管美国和欧盟仍然保持不变, 但中国的份额在 2002 和 2012 年间稳步增长。自然出版指数 (Nature Publishing Index) 显示了类似的趋势 (Nature Editors, 2014)。

平均影响因子在不同的学科间显示出了很大的差异, 主要原因是不同学科论文合著的平均水平变化很大。因此, 数学的合著度为 1.25, 平均影响因子为 0.5, 而生物的合著度和影响因子都在 4 左右。相比于应用和专业技术学科, 基础和纯研究学科倾向于有更高的影响因子。这种学科间的差异很大, 以致于某个学科中影响因子最高期刊的影响因子有可能比另一学科中影响因子最低期刊的影响因子还要小 (论文来源归一化指数(SNIP)考虑了这点, 见“其他计量测度”部分)。与学科间差异相关的是多作者的问题, 每个学科中的论文平均作者数都不同 (见“学科间差异”部分)。鉴于作者倾向于引用其本人的工作, 因此这种变化反映在变化的层次。因此, 各学科之间引用情况差异非常大, 跨学科比较可以用学科权重调整后的引用指标来比较, 虽然这种方法很少被使用。

使用影响因子作为质量测度的另一个问题是该数值是一个统计平均值, 有着统计上的波动。这个波动对小型期刊尤为重要 (因为小型的样本意味关有更大的统计波动)。对于一个平均规模的期刊 (约每年 115 篇文章), 不同年份 $\pm 22\%$ 的影响因子波动并不显著, 但对于更小的期刊 (小于每年 35 篇), 这个幅度是 $\pm 40\%$ 。一个影响因子为 1.50, 每年发表 140 篇论文的期刊与另一个相似规模、影响因子 1.24 的期刊并没有很大的差别。因此处罚一个在小于某一特定影响因子的期刊中发表论文的作者是不明智的, 假定一份平均规模期刊的影响因子是 2.0, 影响因子可以没有显著性差异地变化于 1.50 至 2.25。有关这个问题的更完整讨论见 Collins & Tabak (2014)。

一个有趣的问题是, 发表在开放获取期刊的论文和作者把发表在传统期刊的论文放到个人储库中, 是否会有更多的被引用次数。这将在之后有关开放获取的章节中讨论 (见“开放获取的引用优势”部分)。

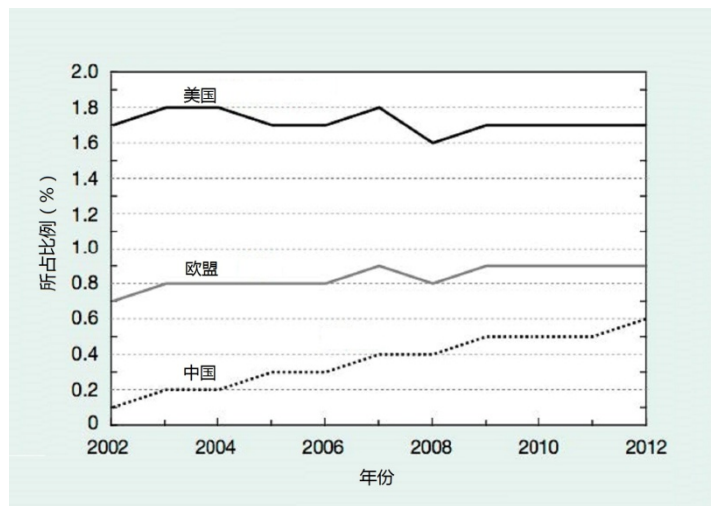


图 25 2002~2012 年美国、欧盟和中国被引用位于前 1% 的文章的世界份额 (NSF, 2014)

研究评估对研究者行为的影响

古德哈特的法则²⁷认为，当指标成为目标时，它就不再是一个好的测度。换句话说，指标不能再真正地反映原始变量，但越来越多地最大化测度机构或个人的效果，而这样做也可能以不理想的方式改变行为。

有明确的证据表明，研究评估活动已经改变了研究者的行为，如 REF（英国的卓越研究框架）或 ERA（澳大利亚的卓越研究）。例如，Adams & Gurney (2014) 分析了英国的数据显示，研究人员将优先产出被认为是领域内核心产品的期刊论文，他们倾向于选择高影响力的期刊，并且将稿件投递给此类期刊[为评价的目我]，即使文章得不到很好的引用，并且，有时甚至不是研究论文。这些作者们指出，这是因为他们认为期刊总体高影响力的品牌是一个更好的真正卓越证据的代理“信号”。后续的 RAE 周期所观察到的投稿行为导致了向期刊论文的持续集中，导致这种相对的增长体现在工程科学学术会议论文和社会科学学术专著向期刊论文的转变。

其他文献计量方法

在影响因子方法有缺陷的情况下，其他起到补充或替代作用的文献计量方法被提出。一些比较知名的方法如下：

- 每篇论文的源标准化影响 (SNIP)：通过 Scopus 数据库的基于某学科领域总引证次数的加权引证来测度特定语境的引证影响。

²⁷ http://en.wikipedia.org/wiki/Goodhart's_law

- SCImago 期刊等级 (SJR)²⁸ 是可免费获取的期刊水平的度量, 是一种声誉度量, 基于并非所有的引用都是相同的观点, 也就是说, 引证是根据施引期刊的威望加权的。
- 谷歌提供学者的度量²⁹, 是基于 h 指数的免费的期刊水平的影响度量。
- **即时指标**, 能够衡量一份期刊中的论文在发表多久后会被引用。
- **被引半衰期**: 能够衡量一份期刊发表的论文多长时间后还在被引用。
- **h 指数**: 定义是: 如果一个作者有着 h 的指数, 便说明在他的 N_p 篇论文中有 h 篇每篇有着至少 h 次的被引用, 而其余的论文 (N_p-h) 每篇最多不超过 h 次引用。这个指标旨在用来测度学者个体 (而非期刊) 的科学产出的质量和可持续性。
- **特征因子(eigenfactor)**: Google 使用一种类似网页排名的网络理论算法, 通过基于期刊被其他高影响期刊的引用来测度期刊的影响。

实际上还有很多其他计量方法。洛斯阿拉莫斯的 MESUR 团队比较了 39 个科学影响的测评方法 (Bollen, de Sompel, Hagberg & Chute, 2009), 他们用统计将不同的方法大致归结为声望和流行两个维度, 得出了影响因子计量的层面“也许算不上是‘科学影响’的理念的核心”。诸如使用关联度的基于使用的计量方法实际上可能是更好的共识性方法。当然, 应当注意的是使用和引用所测度的是不同的事物。

实际情况中, 影响因子的使用范围非常广泛, 以至于即便有更好的测度方法出现, 它也难以被取代, 尤其是当那些计量方法比较复杂时。当然, 考虑多种计量方式比只依赖于一种计量方式更加明智。

论文层面的计量和替代计量学

替代计量运动起始于一些对影响因子 (或影响因子滥用) 的不满: 首先, 期刊影响因子被用于衡量单篇论文的质量, 尽管上文列举了一些对其批评; 其次, 引用测度只是影响力的一个很小的方面; 再次, 引用 (即便是论文层面的测度) 是一个慢且滞后的指示。为了克服这些不足, “替代计量”运动³⁰ 提出了一系列基于引证和下载的补充计量方法, 以建立更完美的影响力图谱 (Priem, 2010)。替

²⁸ <http://www.scimagojr.com>

²⁹ <http://googlescholar.blogspot.co.uk/2014/06/2014-scholar-metrics-released.html>

³⁰ 不能与替代计量学项目和应用 (<http://altmetric.com>) 混淆, 后者是一种基于数字科学开发的、在论文网页收集和展示替代计量数据的工具。

代计量方法在很大程度上依赖于社交媒体和社交工具，且包含了推特消息、博客文章、书签数据（如 CiteULike, Mendeley）以及新闻媒体和论文层面的评论、批注和评级。

已涌现大量的工具和服务支持替代计量的跟踪、报告和可视化，包括 Altmetric, PlumX（EBSCO 于 2014 年早期开发），PLOS Impact Explorer（基于 Altmetric），PageCritic，等等（当今的可能不完整的清单见 <http://altmetrics.org/tools/>）。

随着追踪文章或作者层面影响的兴趣的增长，作者可能越来越多地寻求最大化地扩大他们的影响（或者至少是他们的替代指标分值）。有二个方面的举措旨在支持这方面：

- Kudos³¹ 旨在帮助作者扩大他们研究论著的读者群并增加引用，其举措是通过包括撰写外行能读懂的总结和有效地利用社会媒体的结构化过程。它对研究人员是免费的，但收取出版商相关服务的费用。
- ImpactStory³² 让研究人员创建反映他们研究产出的在线的个人资料（论文、数据集、会议报告、软件），并跟踪和显示其替代指标的影响。收取研究人员的订阅费是 60 美元/年。

有一些社交媒介活动可以预测引用的初步指标，不过证据还不是很强（如 Eysenbach, 2011）。对将社交媒体提及、论文层面评论和评级等作为影响力评估的主要批评是除时效性和大众化之外还测度了什么。有吸引人眼球和特别题名（尤其是包含关于性的）的文章似乎可以与有着真正深远意义的论文一样在互联网世界获得很大反响。

同样值得铭记的是，对于大多数论文而言，在论文层面的引用和使用通常是处于很低的水平，这个数字是如此之低以至于试图把他们变成一个有意义的判别指标会有可能被计数所困扰：大多数数据都在统计噪声的水平并且彼此间是很难区分的。

引用的类型和贡献者的角色

目前引用的是一个生硬的工具：作者引用的实际意图不明显：同意、不同意，等等。为提高引用的价值，有建议以结构化的方式划分引用的“类型”，这主要对文本和数据挖掘的应用及研究网络的可视化有好处。有一个倡议是 CiTO（引

³¹ <https://www.growkudos.com>

³² <https://impactstory.org>

文类型本体 (Citation Typing Ontology (Shotton, 2010; Shotton, 2009))。然而, 作者采用这样一个结构化过程的前景看起来确实还很遥远。(另见“开放注记”部分)

类似地, 大多数学科领域的论文有多个作者是很正常的, 但不同作者的作用可能会有很大的不同。贡献者的作用可能包括研究思路、方法、调查、数据分析和统计、写作, 等等。这些作用可以在论文的致谢部分阐述 (特别是医学期刊), 但表述是非结构化且是不一致的。为了解决这个问题, 编辑、期刊和出版社正在促进标准的分类学的发展, 以描述可用于 STM 期刊贡献者作用 (Allen, Scott, Brand, Hlava & Altman, 2014; Meadows, 2014)。

关于研究评价的旧金山宣言 (DORA)

研究人员及期刊和出版商对研究评价的不满在 2012 年底发布的旧金山宣言 (DORA) 中作出了明显的表达 (American Society for Cell Biology 等, 2012)。宣言指出, 研究结果是多种多样的, 声明了反对研究评估中使用影响因子。其关键的建议是不使用期刊的计量指标作为以研究评价为目的的论文质量的替代品, 宣言也给出版商和指标提供者提出了很多建议:

- 大大降低晋升中对期刊影响因子的强调;
- 提供论文计量层面的指标, 以鼓励从期刊层面指标的转移;
- 解除对研究论文中参考文献再使用的, 并使它们可免费使用;
- 解除或减少有关研究论文中参考文献数量的限制;
- 用来计算所有指标的供数据和方法要公开、透明;
- 基于不受限制使用的许可提供数据, 并在可能的情况下提供数据的计算获取。

在本报告撰写时, 该宣言已有 12000 多个人和约 550 个机构签署。

引证行为的变化

除了上述的发展趋势, Google Scholar 最近发表的两篇论文提供的证据表明目前作者比以前更多地引用早期的论文, 并且高被引论文更有可能出现在非顶尖的期刊中 (Acharya 等, 2014; Verstak 等, 2014)。作者推测, 早先的不知名期刊内容的在线可获得性和日益增长的易发现性 (例如, 通过搜索引擎或其他发现工具) 对于这两种情况都发挥了作用。

专利的引文

专利对 STM 论文的引用有时作为超出学术界更广泛影响的另一种测度。专利中的引用通常要比科学文献中的引用老得多，主要是因为专利授予的延迟；例如，NSF 的分析着眼于滞后 5 年的 11 年的窗口。在美国，引用学术文献的专利比例从 2003 年的 12% 增加到 2012 年的 15% (NSF, 2014)。大多数被引用的文章分为三个领域：生物科学 (48%)，医学科学 (23%) 和化学 (11%)。

使用与期刊使用因子

据估计，2010 年全球论文从出版商网站的下载次数在 11 亿至（见表 3）至 25 亿之间（根据一个非正式的 STM 调查统计），并且另有约 4 亿是从诸如知识库的其他网站下载的。

表 5 2010 年不同国家的论文下载情况 (Elsevier, 2011)

国家	论文下载(百万)	占全球总量的比例(%)
全球总量	1065	100.0
美国	327	30.7
中国	105	9.9
英国	100	9.4
德国	70	6.6
日本	62	5.8

有人认为论文的下下载次数也许比引用能更好地测度论文的广泛影响（如上文提到的，大部分科学家都并非论文作者），这在临床医学期刊和其他有着很多执业读者的期刊中尤为明显。

英国的 Serials Group 开展了基于下载的统计数据计算“使用因子”是否可行的研究。该报告发布于 2007 年中期，认为创建有意义的期刊使用因子是可行的，且图书馆和出版界有着对这种做法的支持。UKSG 和 COUNTER 随后委托 CIBER 进行更多的细致调查，并在 2011 年发表 (CIBER Research Ltd, 2011)，在 2012 年 3 月，UKSG/COUNTER 发布了“COUNTER 有关使用因子的实践规范：草案稿 1” (COUNTER, 2012)

这个规范草案将出版和使用周期规定为连续的 2 年：即 2009/2010 年的使用因子是基于 2009/2010 年间对 2009/2010 发表的论文的使用情况。使用因子：期刊 (UFJ) 定义为“一组有序的全文使用数据的中位数”；之所以使用中位数而非平均数是因为这个数据集合是高度偏斜的，大部分都在较低的使用水平，很少有被使用很多次的。每年这个数据都会以整数形式报告（更高的精确度意义不大，

因为方差之大说明有很多统计噪音干扰)。UFJ 将整合从稿件录用阶段到出版中的文章,且会参考多个平台的使用情况,反映出文章使用的多种来源。可能计算出二种期刊的使用因子(UFJ):出版商使用因子(基于出版商兼容 COUNTER 平台的全文使用)和综合利用因子(来自一组兼容 COUNTER 平台的总体使用量)。

CIBER 发现使用情况在不同的文献种类和版本间有很大不同。因此他们估计有两个版本的使用因子来源:一个是基于除编委会名单、订阅信息和许可细节以外的所有文章种类,另一种只基于学术内容(短篇交流、研究论文、综述性论文)。

CIBER 发现在所提出的使用因子和基于引用的测度(如影响因子)间有着很小的相关性。由于两者测度的是不同的事物(反映出了读者和作者分别的选择),这个结果并不出乎意料。高被引论文确实趋向于有着高下载次数,但反之并不必须成立,尤其是在有大量执业人员的学科。引用和下载频率在时间上也有不同:在文章发表之后的数月中文章下载量就会形成高峰,而引用则是多达 2~3 年。

目前产生的共识似乎是下载次数(作为阅读的替代数)是一个对引用数据的潜在有用的补充,但它并不应该取代后者,因为两者反映的是论文“使用”的不同方面。下载和阅读论文在研究设计和论文写作的早期有更重要的意义,而引用则趋向于出现在这个过程的结果。期刊层面的使用因子被图书馆应用于获取配置并且可能对于作者选择拟投稿的期刊有用,但在许多情况下论文水平的计量与上述讨论的原因更为相关。

2.13 期刊出版的成本

因为对系列出版危机和开放获取的争论,因此有关期刊出版成本的认识对出版商和广大学术界都有重要意义。

剑桥经济政策协会(Cambridge Economic Policy Associates)于 2008 年发布的 RIN 报告详细研究了期刊出版过程中的成本,包括图书馆获取的供应成本及学术界同行评议、查找和阅读论文所产生的非现金成本。这个报告提供了对期刊成本的很可靠的估计,CEPA 在其后的报告中更新了他们的估计(RIN, 2011c),给出了 2010 年每篇期刊论文的平均生产成本(纸质+电子)为 3095 英镑,由以下各项组成:

- 第一版成本(与发行数量无关的成本,例如同行评议的管理费用、编辑加工费用、排版费用等): 1261 英镑。
- 可变成本(印刷、纸张、发行): 581 英镑。
- 间接成本(工作人员和管理费用): 666 英镑。

- 盈余：586 英磅。

值得注意的是 RIN 的数据中包含了盈余项，因此该成本是基于买家视角而非商家视角。考虑各组成的比例，与 Campbell & Wates 计算的 Wiley-Blackwell 期刊的平均数值就大致相似了（Campbell & Wates, 2009）。

PEER 项目报告称每篇投稿的平均同行评议管理成本在 250 美元左右，且录用稿件的平均生产成本为每篇 170 至 400 美元之间（这些数值都指只包含工资和报酬，不包含管理费用、基础设施和系统建设费等）（Wallace, 2012）。

重要的是要记住这些数字是平均值。第一版成本的数据差别尤其很大，取决于期刊的类型。早先 RIN/EPS 主报告（EPS, 2006）引用文献中的数值从 350 美元到 2000 美元不等，但是 2008 年的 RIN 报告引用的数据范围很小。对于低退稿率的期刊，RIN 的作者们给出的数据是 1670 英磅，高退稿率的期刊则为 4091 英磅。对于热门的综合性期刊（《科学》、《自然》等），RIN 的数据为 4116 英磅，虽然有其他的估算认为这一数值为 10000 美元甚至更高。

RIN 还估计，商业出版商的间接成本为每篇文章 705 英磅，而学会出版商为 428 英磅。我们还没有看到任何其他系统、完整的数据可证实这点。

期刊价格除了抵消出版商的这些成本以外，大多情况下还包含了盈利的部分（对于商业出版商）和盈余（对于非赢利性出版商）。盈利是再投资和创新的主要源泉。在这方面，学会常常将期刊出版带来的盈余用于支持其他活动，如会议和讨论会、差旅和研究资助、科普教育等（Baldwin, 2004; Thorn, Morris & Fraser, 2009）。据 RIN 估计，2008 年的收入中平均盈利或盈余为 18%，相当于每篇论文的盈利有 517 英磅（这些数据在 2011 年的报告中没有更新），并且商业出版商（642 英磅）和学会出版商（315 英磅）间有所差异，这至少部分地反映了他们之间不同的税收地位和实际的盈利能力（非营利期刊不支付公司税务，因此用税后利润和盈余比较比用税前利润更为合理）。

纯电子出版的成本节约

出版商认为转向纯电子出版有可能节约 10%到 20%的成本。RIN（2008）估计，如果全球 90%的期刊都改为在线出版，全球范围内的出版成本可节约 10.8 亿英磅。这种节约的最大部分不是源于出版商而是来自于图书馆（不再需要管理、装订和保存纸版本等），出版物和发行成本的减少相当于全部出版成本的 7%。去掉盈余或盈利后，这一数值相当于出版商成本的 9%，略低于出版商的估计。

开放获取和可能的成本节约

开放获取对成本节约的潜在影已经讨论了很多（如见“开放获取”章节）。

然而，纯开放获取期刊出版商的出现使得我们可以从其财务报表获得一些平均的论文成本的证据：

- PLOS 的 2013/14 年度报告显示其总成本（包括管理费用）为 2960 万美元，相对发表的约 34000 篇论文，每篇论文的平均费用为 1088 美元（不包括 871 美元的管理费用）。结合低成本的 *PLOS ONE* 与高成本的选择性（高退稿率）期刊，表明 *PLOS ONE* 的平均成本更低（PLOS, 2014）。
- *eLife* 第一个全年运行的财务报表显示其成本为 264.4 万英镑，相当于平均每篇文章 8370 英镑（以英镑/美元 = 1.67 计，约合 14000 美元；2013 年发表的论文为 316 篇）。虽然 *eLife* 是一份高退稿率的期刊并且其编辑人员的费用昂贵，这个数字仍应谨慎对待，因为它没有透露有多少比例的费用是一次性的，或蓄意加大了必要的费用以为建设未来的规模需要（eLife, 2014）。2014 年的出版量超过了一倍达 636 篇，这似乎可能显著地降低了平均成本。
- 在成本天平的另一端是 *Nature* 报导的 Hindawi 出版社，该出版社 2012 年出版了 22000 篇文章，平均每篇文章的成本为 290 美元（Van Noorden, 2013）。Hindawi 采用低成本的出版模式，并且位于相对较低工资的地域（埃及）。

Nature 的同一篇文章中还援引了其他出版商所声明的成本，尽管不被所公布的账目支持：

- 国家科学院院刊（*Proceedings of the National Academy of Sciences*）估计他们的平均每篇发表的文章成本为 3700 美元；
- 援引 *Nature* 的主编是其内部成本为每篇文章 20000~30000 英镑；
- PeerJ 说他们的平均成本是每篇“低至数百美元”。

另一方面，现有的大型出版商大量节省可能不易被发现：金融分析公司 Bernstein Research 估计，全面过渡到开放获取将节省订阅型出版商的基础成本的 10% 至 12%（Aspesi, 2012）。

期刊定价

期刊定价一直是许多讨论和争议的源头，并且人们认为已经很高的价格和其快速增长一直是推动开放获取议程的因素之一。期刊定价增长的速率的确超过了通货膨胀率，例如，研究图书馆协会（ARL）发表的统计数据显示 1986 年到 2011 年间期刊开支年均增长 6.7%，而同期美国 CPI 年均增长为 2.9%（ARL, 2011）。

历史上期刊价格上涨的原因一直有变化，包括（King & Alvarado-Albertorio, 2008）：论文发表总量的增加导致了平均每种期刊中论文发表数量的增加，同时

论文的篇幅也有所增加，从而导致期刊规模的增大；版面费和彩版费的恢复；“新期刊”效应（学术的增长导致了新领域的增加，进而导致了新期刊的出现；新期刊一般在其定位的领域发行量很小（至少在初期是这样的），且多在经济运转方面不够有效，并因此导致更高的订阅价格）；增加的特殊要求和特点；纸质向电子期刊的转化；出版者提高定价来弥补下降的订阅数量和汇率的影响；当然，还有成本膨胀（尤其是薪资和纸张成本），这方面在过去的二十多年来年均增幅约为3%。

总之，1990年到2000年间的期刊平均价格增长的原因有很多的因素，其中论文数量的自然增长（3%）和价格通胀（3%）是最重要的，其次是向电子出版转型的成本，新期刊的专业化及损耗（价格螺旋）和汇率波动的影响（约为1%）。

期刊危机的出现不仅是因为价格方面的压力，还来自研究预算的增加（转化为更多的论文产出）持续超过图书馆的预算增加。例如，在2004年到2008年间，英国大学的总支出实际上涨了22%，而图书馆在“信息内容”方面支出只上涨了15%（RIN, 2011b）。在美国，大学对图书馆的资金投入比重到2009年已是连续14年下降，并首次低于2%（ARL, 2011）。部分原因是效率的提高（如捆绑销售和协议销售，其他共享服务，图书馆编目和文献服务的外包，以及工作人员的缩减），但这也反映了图书馆未能在总预算增长中维持其份额。

捆绑销售和协议许可对价格的影响

出版订阅价格的统计数据越来越具有欺骗性，因为这些数据并不能代表图书馆的实际支出，原因是由于电子传输效率和多期刊许可的增长（一定程度上致使美国研究图书馆协会（ARL）和拉夫堡大学图书馆和信息统计部门（LISU）都在他们的年度统计中停止了记录订阅数量）。

一个正在被越来越多地使用的期刊定价方法是计算每次下载的成本。每次下载论文的平均成本已有大幅下降，部分原因是学者们正越来越习惯于使用电子文献，还有部分原因是“大额交易”（Big Deal）和类似的协议许可提供了大量的相对廉价的额外内容。LISU在他们2005年的年度报告中称这些事实一定程度上导致了当前英国期刊订阅的平均成本在过去5年中（至2003/2004）下降了23%（Creaser, Maynard & White, 2006, 第133页）。这个下降还在持续，英国学术机构每次下载的平均成本从2004年的1.19英镑下降到2008年的0.70英镑，下降幅度为41%（RIN, 2011b）。

这种情况也在2012年的一份报告中有所说明（Gantz, 2012; Gantz, 2013），该报告挑战了上述引用的ARL统计的解释。该报告认为，1990至2010年之间图书馆在连续出版物方面的支出确实增加了三倍，但ARL图书馆馆藏通过新的

订购和已订购产品的内容扩张也增加了三倍。因此，平均每种期刊的成本与 1990 年相同。如此看来 ARL 所报道的明显的 6 倍增长是不正确，但这个是基于印刷版的价格表，而图书馆现在购买的是电子内容包。这说明 1990 至 2000 年所获取期刊的平均成本在增加，然后 2010 年又下降到 1990 年的水平。

2.14 作者的行为、认知和态度

有很多对作者的行为、认知和态度的研究，其中有两个开拓性的成果以其（当时）很大的国际规模（4000~6000 余份的调查者反馈）和严格的研究设计而脱颖而出：CIBER（伦敦大学学院的一部分）进行了两个调查，结果于 2004 年和 2005 年发表（Rowlands, Nicholas & Huntingdon, 2004; Rowlands & Nicholas, 2005），其中一个调查为爱思唯尔委托 CIBER 与 NOP 在 2005 年合作进行（Mabe, 2006; Mabe & Mulligan, 2011）。RIN 和 Harley 后来的研究延伸并扩大了这些发现（RIN, 2009a; Harley 等, 2010）；并且更新近的工作记录了作者对开放获取的不断演进的态度（Taylor & Francis, 2014; Nature Publishing Group, 2014）。

在“新期刊发表模型：对资深研究人员的国际性调查”中，Rowlands 和 Nicholas 报道了第二次 CIBER 调查，这次调查收到了 5113 个资深期刊作者的回复。他们有关开放获取方面的一些发现在一定程度上已经有所变化（例如，大多数作者都认为向开放获取的大规模转型会损害学术出版，但目前这至少也是英国政府的政策——见“开放获取”章节），但一些观点仍然不变：

- 同行评议的至关重要性被重申。
- 资深作者和研究人员认为论文下载数量比传统的引用对于测度研究的有用性更加可信。
- CIBER 发现作者对机构知识库了解很少，而且有证据显示有相当一部分人（38%）不愿意使用机构知识库。虽然有一些特别的例外，这种情况时至今日还是如此（Wallace, 2012）。

爱思唯尔/CIBER/NOP 在 2005 年的调查使用了类似 CIBER 调查的方法——有 6344 个在线问卷的反馈——并且还有 70 个后续的电话深入访谈。其中一些至今仍被认为有效的重要发现如下：

- 虽然表面上期刊出版最重要的理由是传播研究结果，内在驱动力是基金资助和作者的职业发展。除了“建立优先地位”和“他人认同”的重要性有所增加，该图景与一个早在 1993 年的研究结论（Coles, 1993）相似。1993 年到 2005 年间向电子出版的转变并没有明显改变作者的动机。
- 研究人员对研究资助机构的态度模糊：63%的人认为资助机构对开展何种研

究有着过大的权力。即便研究人员感受在高影响力期刊发表的压力很大，资助机构并未对论文发表在何种期刊上做出规定。（这个调查在资助机构实施强制论文存放知识库之前，因此它无法调查研究人员对该规定的观点）。

- 作者对究竟在知名期刊发表还是在专业性期刊发表论文持不同态度。
- 同行评议的重要性被再次强调（另见“同行评议”章节）。
- 大多数的作者（60%）认为出版商增加了他们论文的价值；但 17%认为没有，计算机科学领域（26%）和数学领域（22%）这样认为的人更多。
- 研究人员对在 10 年前（即在电子期刊出现之前）发表的论文有多的需求。

发表的动机

研究人员对学术交流的基本需求在过去的 20 年里进行了研究，这种基本需求根据他们角色的不同而有所不同，即：是作为一个作者或还是一个读者。作者的核心需求被视为首先是报道思想；在交流中感觉安全；（为实证对象）说服读者其结果是普适性的，并且产生于规范的（科学）方法的制定；使其观点被同行接受；把他们的观点报道给正确的读者；使他们的理念得到理解；并使自己的工作拥有一个永久性的公共记录（Mabe, 2012）。

至于发表的具体动机，2005 年调查所报告的最重要的动机是在“传播”（73%），“促进我的职业”和“未来的资助”是关键次要动机。与 1993 年的一个类似的研究比较，这三个动机或排序的变化很小，但次要的动机，“认可”和“建立首发权”已明显增加，尤其是后者（Mulligan & Mabe, 2011）。

期刊的选择

多项调查显示，影响作者选择期刊的主要因素是期刊的质量、相关性和出版速度（按此顺序）。这些观点一直以来非常稳定。例如，源于“爱思唯尔的作者反馈项目”（Mabe & Mulligan, 2011）的 10 年有价值的数据分析比较了 2002 年和 2009 年的数据（包含了接近 100,000 位研究人员的反馈），并且显示质量、相关性和出版速度仍然是最重要的因素，且排序相同。该整体图景得到 2012 年对英国学界的一项调查证实（Ithaka S+R 等, 2013），并在 2014 年自然出版集团（NPG）的调查中再次证实，NPG 报道的五大因素是期刊的声誉、相关性、同行评议的质量、影响因子、第一次决定的速度（Nature Publishing Group, 2014）。

作者对开放获取的认知和态度

有兴趣的是开放获取状态是否影响到作者对期刊的选择。最近的调查表明，对于大多数作者而言上述三个因素仍然是主要的，但 OA 状态在次要因素中的重

要性正在显现。例如，最近 NPG 的一项调查发现，相比 90%~96% 的表示基于相关或质量因素而言，少数（37%）科研人员表示将立即 OA 作为一个非常或重要的因素（Nature Publishing Group 2014）。对于那些选择了 OA 期刊的人来说，最常见的理由是该期刊只提供 OA，也就是说，他们因为其他一幅幅理由选择了该期刊。然而，选择一份 OA 期刊的次常见原因是，他们认为研究在出版后可立即可以获取。有趣的是，资助者和机构的强制并不是选择 OA 出版的重要原因，最重要的原因是认为研究应该是免费获取的，其次的信念 OA 出版物能被更广泛地阅读。

OA 的状态也可能是期刊选择中的一个消极因素，至少在目前 OA 涉及出版费：对英国学界的伊萨卡 2012 调查发现，期刊选择中的第四个最重要的因素是“期刊允许学者免费发表文章，无需支付版面费或文章的费用”，该项因素的重要性在出版速度之前（Ithaka S+R 等, 2013）。

有几次大规模的调查都探究了作者对 OA 的态度，包括 NPG 的作者见解调查和 Taylor & Francis 的 OA 调查。这两次调查在 2013 和 2014 年都开展了，可用来估计一些态度上的变化（Nature Publishing Group, 2014; Taylor & Francis, 2014）。这些调查是互补的，NPG 更强自然科学而 T&F 更强调人文和社会科学。这些调查包含了丰富的细节，一些亮点如下：

- OA 的有利之处似乎被告普遍接受：只有 11% 反馈认为 OA 没有什么好处；
- 绝大多数研究人员认为 OA 能够使他们的研究得到更广泛的传播（81%）和更高的显示度（75%），这些信念在 2013 和 2014 之间有所加强；
- 约一半的研究人员认为 OA 出版物的发表速度要快于订阅期刊（虽然目前还不清楚事实上是否如此）；
- 研究人员 OA 期刊是否有更多的被引方面有分化：29% 同意，31% 不同意，39% 中立。然而，相比 2013 年，同意的更多，不同意的更少。
- CC-BY 协议对于相当一部分的作者没有吸引力：T&F 调查的 65% 受访者认为，不事先征得他们的同意而将他们的工作用于商业收益的目的是不可接受的。当问及对不同的 OA 协议的倾向时，最普遍的选择是 CC-BY-NC-ND，超过 CC-BY-NC 和 CC-BY-ND，远远超过易被最少优先选择的 CC-BY。11% 的反馈者将 CC-BY 列为最多或次多的选择，而 53% 的反馈者将 CC-BY-NC-ND 列为最多或次多的选择（然而，T&F 的受访者偏重于人文和社会科学）。
- 对于 PLOS ONE 式“合理但没有重要性”的同行评议的支持可能正在衰退，

2013 和 2014 年间支持水平正在下降。

- 将论文存放于仓储是个人使研究成果免费获取的责任，向其他研究人员索取论文也是如此。
- 相反，不存放论文于仓储的主要原因是缺乏对出版商政策了解及缺乏可获取的时间。
- 严格（快速）的同行评议是作者期望支付发表费所回报的最重要的服务，紧随其后的快速出版。
- 展望未来，作者认为期刊仍将是主要的出版物出口，标定高质量的研究，但研究论文将有很大一部分只发表于与期刊共存的学科或机构知识库。

对同行评议的态度

研究人员在上述调查中一致表示支持同行评议，并在调查中表示致力于探索同行评议（详见“同行评议”章节）。Mabe 的纵向数据显示，对同行评议的态度在 2002~2009 年期间没有显著的变化（Mabe & Mulligan, 2011）。

对社交媒体和“science 2.0”的态度

RIN 有关研究人员对 Web 2.0 和社交媒体的使用与态度的研究反映出研究人员对现有手段和工具的持续喜好（RIN, 2010）。

一个加州大学伯克利分校的研究（Harley 等, 2010）也发现研究人员仍然专注于传统的正式出版，并且对基于网络形式的学术交流新模式持谨慎态度。研究人员在一个研究周期的不同阶段中使用一系列的交流方法，且这些方法在不同的学科有所不同，生物学研究成果输出的形式变化最小（即：主要是学术期刊）。他们发现“没有证据显示‘精通技术’的年轻研究生、博士后或助理教授反对传统的出版方式”，并且“一旦新生学者进入某个行业——不论是研究生、博士后还是助理教授——他们就会采纳这些行为、规则和他们导师的建议，以此来提升他们的职业生涯”。事实上，只有成熟的研究人员才有资本尝试新的方法。（早先加州的一个研究提出了相似的结论，认为年长的教职人员比年轻者对创新更为开放，更愿意尝试和参与新的倡议，并且发现人文和艺术领域比其他领域对改变有更大的兴趣（加州大学，2007））。Harley 研究确认了一些需要关注的问题，包括：对同行评议方法和时间的再评估；新的可以适应不同长度、富媒体和超链接的出版方式；以及对管理和保存新的数字研究方法和产出的支持（自然语言处理的成分、可视化、复杂的分布式数据库，GIS，等等）。

然而，在过去的几年中，对包括 Academia.edu, ResearchGate 及 Mendeley 等

科学社交网络的注册用户数已经迅速发长,表明研究人员可能更愿意为了专业目的使用社交媒体或网络(详见“科学社交网络”部分)。

另见 *社交媒体* 章节。

2.15 出版道德

人们越来越意识到期刊出版中需要更高(或至少是更透明)的道德标准来关注和解决诸如利益冲突、代笔作者、论文挂名、引用圈、同行评议操纵、著作权纠纷、数据造假、学术欺骗、不道德实验和抄袭剽窃等问题。大量的批评都针对生物医学期刊与医药行业之间的利益关系,但这些问题绝不是这个领域所特有的。

在线投稿系统的采用使得期刊系统化地收集诸如竞争利益声明、伦理同意书等方面的数据变得更加容易。有关领域的期刊中随论文发表这样的声明已经越来越常见。

近年来人们对快速增长的撤稿数量越来越关注,撤稿数量已经从 2000s 早期的每年 30 篇增加到 2011 年的 400 篇以上,而同期论文数量只增长了 44%(Van Noorden, 2011)。即使如此,这个数字只达到了全部论文数量的 0.02%,虽然调查中有 1%~2% 的科学家承认了他们曾经至少有一次编造、伪造或修改数据或结果。似乎撤稿数量的增加是有积极影响的,因为人们对这方面问题更加关注且检测的方法更好,这些可能都超过不端行为本身的增长幅度。撤稿而导致的一个问题是作者们仍然倾向于继续引用被撤回的论文,对 CrossMark 的采用应当有助于遏止这种情况,或至少提醒跟踪引文的读者。

出版道德委员会

出版道德协会(COPE)³³成立于 1997 年,旨在为出版商和科技期刊编辑提供了一个可以讨论有关投稿与出版中学术诚信问题的平台。它有 9000 余个会员,其中大多数都是科学期刊的编辑。他们每个季度都要开会,并为会员提供审查工具来衡量对最佳行为指南的遵守程度。所有的 COPE 会员都被要求遵守行为规范(Code of Conduct)和期刊编辑最佳操作指南(Best Practice Guidelines for Journal Editors),这些规范和指南的最新版本发布于 2011 年(COPE, 2011)。

其他与出版道德有关的组织

国际医药期刊编辑协会(ICMJE)³⁴为医药出版的伦理制定了详细的准则(很多规则在其他领域也适用),包括作者身份和贡献、研究数据共享(包括临床实

³³ <http://publicationethics.org/>

³⁴ <http://www.icmje.org>

验数据)、编辑行为、同行评议、利益冲突、隐私和保密、研究中人权和动物的保护等。ICMJE 建议(以前称为生物医学期刊投稿的统一要求)相当于一个特设的标准被广泛遵守(ICMJE, 2013)。

世界医药期刊编辑协会(WAME)³⁵也对伦理道德方面的问题有所阐释,并出版了同行评议医学期刊有关利益冲突的政策声明(WAME, 2009)。

撤稿监察日志定期记载有关论文撤稿及其引发的相关问题。日志的作者们要求期刊采用一个透明度指数来说明相关事项,如:期刊的同行评议政策,是否使用剽窃检测软件,处理错误或不当行为指控的机制,更正和撤稿是否符合 ICMJE 和 COPE 指南等(Marcus & Oransky, 2012)。

CrossCheck 和其他自动检测工具

CrossCheck³⁶是 CrossRef 组织专门为学术期刊开发的用于检测抄袭的工具。尽管有很多软件能够将稿件文字与网络上的文本进行比对,但这类服务在检测科学期刊的投稿时并不实用,因为这种软件不能获取科技文献的数据库内容。CrossCheck 通过创建一个包含 STM 合作数据库(参与的出版商联合创建)并联合商业抄袭检测软件(iThenticate)解决了这个问题,软件的使用者可以比较已提交的稿件和已出版的文献。这个软件会自动给出不同作品之间的相似度,但最终的判断一篇文章是否是涉嫌抄袭、重复发表或其他一些非恶意的原因则要由人来判定。

其他检测学术不端的工具包括基于图像编辑软件的筛选来检测对照片或影像的操纵、数据检查(数字偏好分析可以检测伪造的数据)。

arXiv 仓储有自己专用的软件来筛选有潜在抄袭的投稿。2014 年的一个研究调查了整个数据库中从 1991 年中期到 2012 年的全部 757000 篇文章的潜在抄袭情况。文字的重用是相当普遍的:过滤掉综述和合法的引用后,大约每 16 位作者中就有 1 位从自己先前发表的作品中复制长短语和句子。大约每 1000 位投稿作者就有 1 位从其他人已发表的文章中复制相当于一段价值的文字而不加以引用。也许最有趣的发现是,重复使用已经发表作品的文字越多,这篇作品被引用的频率越低(Citron & Ginsparg, 2014)。

2.16 版权和许可

一个被大部分体系内参与者认同公平的强大的版权(或者更广义的知识产权)

³⁵ <http://www.wame.org>

³⁶ <http://www.crossref.org/crosscheck.html>

制度被认为是商业内容和传媒产业的前提条件，期刊出版（包括开放获取）也不例外。对于通过订阅获取的期刊来说，作者或者把版权让渡给出版商（在保留一定的规定的权益的情况下）或者许可出版商使用另外一套界定的权益（目前约三分之二的大型出版商倾向于这种版权选择的许可（Inger & Gardner, 2013））；以上两种情况下的结果都是一样的，即允许出版商的商业使用权利，而作者可以享受服务（同行评议、编辑加工、荣誉等）。对开放获取期刊来说，作者通常保留版权并在知识共享许可协议或类似协议（见下文）前提下允许使用和再使用，但有附加条件，如取决于版权的作者归属。然而，原则上开放获取也可能存在于传统版权制度下。

版权和其他的知识产权法律（如专利法）试图在赋予创造者垄断性权益（为了鼓励创造和创新）与不限制内容获取而使受益者更广泛这二者之间建立平衡。这个平衡可能需要不断评估，例如要及时了解技术的发展。数字化转型已经给传统的基于文档的复制和完整性控制的版权制度带来了许多挑战——一个单一的数字文档可以服务世界并且其本质上并非完全不变。

版权改革

最近英国和欧盟有关版权的综述（Hargreaves 报告和随后的政府咨询³⁷、信息学会的版权³⁸及后续的欧洲版权协议³⁹等多个项目）涵盖了任何制度下都相关的数字环境所提出的问题：

- 数字版权除外。版权除外是从公共利益判断允许的特殊情况下，免除一些一般性的版权限制。其管理需要遵循伯尔尼国际条约的三步检验法：除外必须被限制在该特殊情况下；它不应干扰对该作品的正常利用；并且没有不合理地损害版权所有者的合法权益。
- 评估的除外包括：图书馆需要的备份（如用于替换损坏的档案原件或用新形式代替过时的旧形式）；支持盲人和视觉缺陷者；图书馆间互借；在图书馆内获取印刷版资料的数字化内容；教学课程包资料；孤儿作品。
- “孤儿作品”：指使用者无法获知版权或无法联系版权所有者的作品。这种作品风险除外合法开发，因为版权遵守者可能宁愿不使用也不冒险侵权。为了防止这种现象发生，使用者在“仔细搜索”以识别版权所有者之后，孤儿作品允许除外。

³⁷ <http://is.gd/3il5QW>

³⁸ http://ec.europa.eu/internal_market/copyright/copyright-infso/index_en.htm

³⁹ <http://ec.europa.eu/licences-for-europe-dialogue/>

- “商业脱离”作品是指作品仍然受版权保护，但通过正常的商业渠道不再能获取（例如，脱离印版的图书是出版商不打算重印或发行电子书版）。各国政府已提出或赋予公共/国家图书馆和文化机构的数字化权利使这些作品可以获取。

英国的版权变化

在英国，Hargreaves 提出的大量与出版商相关的重要建议至少都已经部分地实施了（Hargreaves, 2011; Intellectual Property Office, 2014）：

- 允许文本和数据挖掘（TDM）的版权除外已经实现，尽管这正是 STM 发展中一个活跃的领域（专责委员会希望看到出版商建立可用且负担得起的许可方案），允许用户“复制作品‘用于文本和数据分析’”，前提是非商业性的研究，并且复制本伴随“足够的承认”（在切实可行的情况下）。（另见下文“文本和数据挖掘”部分）；
- 2013 年 3 月英国政府宣布资助版权中心和数字版权交易中心⁴⁰，这是 Hargreaves 所已提出的更容易获得版权许可的方案，包括“孤儿作品”（那些原始版权持有者无法追查）。它旨在提供有关版权的信息资源；不同权利和其所有者的登记注册点；以及许可交易的平台；
- 格式转换的版权除外：2014 年 10 月 1 日生效，虽然有来自各方面权利持有人的反对。该除外的只限于个人的复制，但是，学术图书馆和英国图书馆都希望有这样的权利以便于长期保存；
- Hargreaves 还建议版权除外不能被协议推翻。相应于新的版权除外，这项也已实施（引证，模仿和私人复制/格式转换）。

欧盟版权协商与发展

如上所述，欧盟已就信息社会的版权方面可能的改变和欧洲方案的后续协议进行了磋商，并于 2013 年 11 月线束。其后是更进一步的有关欧盟版权规则的公众咨询审查，下一步的报告和建议于 2014 年 7 月发布⁴¹。

欧盟的进程类似于英国的 Hargreaves 报告，包括重要的 STM 问题，如文本和数据挖掘，但也包括诸如越界获取和服务的便携性等事项，以及用户生成的内容和微许可。该委员会正在起草拟于 2014 年底发布的白皮书《欧盟创造与创新

⁴⁰ 见：政府拨付 15 万英以启动版权中心, <https://www.gov.uk/government/news/government-gives-150-000-funding-to-kick-start-copyright-hub>

⁴¹ http://ec.europa.eu/internal_market/consultations/2013/copyright-rules/index_en.htm

的版权政策》(A copyright policy for Creativity and Innovation in the European Union),其中阐述了其建议的变化,内部稿在2014年7月流出(Baker & McKenzie, 2014)。

委员会所谓的进一步举措与 STM 出版商关联较小,包括非商业作品、私人复制税,纸版阅读障碍人员的版权获取、视听作品的在线传播等等。

美国及其他地区

在美国也有着对版权改革需要的激烈争论。版权的登记旨在通过立法解决孤儿作品问题、改进有关图书馆使用的公平使用规则、促进商业上事可获取作品的大规模数字化 (Samuelson, 2012)。最后一项的出现是由于谷歌图书解决方案的失败,并且也涉及“非商业”的作品

然而,改革议程实质上更为广泛。美国版权注册师 Maria Pallante 于 2013 年 3 月在国会作证阐述了需要改革广泛清单,包括:“明确版权除外的范围,为图书馆和档案馆修改除外和限制,解决孤儿作品问题,容纳印版障碍人士,向教育机构提供指导,在适当的情况下除外附带的副本,更新强制的规定,提供法定赔偿指南,回顾 DMCA 的功效,帮助小的版权声明,改革音乐市场,更新有线和卫星传输的框架,鼓励新的许可制度,完善版权登记和备案制度。”

法国已经通过了法案允许其国家图书馆数字化诸如失去商业化的作品。

对版权的理解和认识

值得注意的是,很多关于 STM 版权方面的争论都发生在当前人们对版权的无知和误解以及现行体制下的版权获取方式。例如,一篇中国 2009 年发表的论文调查了作者们对自己文章的权益的认识,并将其与出版商实际允许作者拥有的权益作对比 (Morris, 2009)。这个调查发现作者们低估了他们对自己作品的出版前版本的权益(如自存储、在课程中使用、将副本提供给同事阅览),但高估了出版商政策中有关允许他们使用正式出版版本的范围。尤其是很多作者认为他们可以自存储正式发表的版本,但很少有出版商允许这种做法。这个研究的结论是出版商并没有有效地与作者交流他们的版权政策。

这个显示版权和相关使用及再使用权益被误解的情形在其他对学术界的研究中也常常见到,甚至对于图书馆人员也是如此。例如, RIN 报告中有关文献获取差异性的一项研究中发现,(研究人员)不太明确权益许可的范围,尤其是预约权益的内容,特别是有关网络资源的获取许可 (RIN, 2011a);此外,缺乏对版权的认识也被认为是作者在知识存储时犹豫的原因之一。最近对作者的调查证实有关版权了解和保留权利的混乱仍持续存在。

示范协议

示范和样本协议已由一些组织机构制订，这些组织机构包括出版商组织、中介机构和采购机构。这类协议的使用是可取的，主要有二方面原因：其一是简化了交易和市场的运作；其二是因为协议通常代表了有关各方进行实质性的协商和妥协之后的“最佳实践”。案例包括：

- IFLA 许可原则：这些实际上不是一个示范协议，而是一套管理图书馆和出版商之间合同的原则。最初于 2001 年起草，最近的更新是 2014 年 5 月。
<http://www.ifla.org/publications/ifla-icensingprinciples-2001>
- licensingModels：一套电子资源协议，最初产生于 1999 年，为主要订阅代理的协作制订，其后由 John Cox Associates 扩展（该网站现在由 Ringgold 管理）。协议包括学术图书馆、学术联盟、企业图书馆、公共图书馆、电子书，30/60 天免费试用。<http://www.licensingmodels.org>
- P-D-R 模式协议 ALPSP，STM 和 Pharma Documentation Ring（医药文献环）制订，覆盖出版商和制药公司之间的协议条款。2012 个更新包括一条新的对文本和数据挖掘的权利进行指导的条款。
<HTTP://www.p-d-r.com/content/publications/>
- 文本和数据挖掘（TDM）：STM 制订的示范协议涵盖征订内容的 TDM 及以前未征订内容的 TDM。<HTTP://www.stm-assoc.org/text-and-data-mining-stm-statement-sample-licence/>
- STM 开放获取协议：见下文的开放获取协议
- JISC 示范协议的应用只限于 JISC 协议，包括 NESLi2，以及档案、数据库和 SHEDI 的次级协议。http://www.jisc-collections.ac.uk/model_licence
- 在美国没有相当于国家的采购，但示范协议包括 LIBLICENSE 示范协议，以及那些由各大财团制订的协议。
<http://liblicense.crl.edu/licensing-information/model-license/>

SERU

SERU（共享电子资源的理解）建议实践是 NISO 的一个最佳实践。它提供了一种供图书馆和出版商商定的可供选择的协议，主要为北美市场制订（和使用）。SERU 声明表达了：内容提供商、订阅机构和授权用户的共识；内容的性质；材料的使用和不当使用；隐私和保密；在线性能和服务条款；存档和永久使用。好处是通过省略双边许可证而简化了电子资源的采购。

SERU 于 2008 年在电子期刊中最初采用，目前的版本是 2012 年更新的，涵盖范围更广的内容，包括电子书。出版商、图书馆和联盟都愿意注册加入 SERU（在其网站上可获取），但这并不要求他们未来的订单必须使用它（NISO SERU Standing Committee, 2012）。

开放获取协议

对于 OA 期刊而言，根据协议所发表的文章允许用户在特定情况下获取、复制和重复使用。从作者的角度来看，典型的安排是为他们保留版权，但要与发行商签署许可协议，允许后者在规定的 OA 协议下发表其作品，但其他安排也是有可能的。

OA 期刊最常使用的协议是知识共享（Creative Commons, CC）协议。最新版本（V4.0）于 2013 年 11 月推出；相比 V3.0 的主要改进是进一步国际化，提高与其他协议互通，更持久的未来发展的期待，以及对数据、科学和教育的具体要求。

CC 有时被描述为“保留部分权利”（相对于“保留所有权利”的版权声明）；与把材料放置在公共领域（即放弃权利）的原则完全不同。CC 有多种类型：

- CC-BY 允许用户最大自由地再利用内容：实质上允许所有的复制和再利用，前提是承认作者（著作权人），包括衍生作品的创意、出于商业用途的再利用。该协议是大多数 OA 倡导者所倾向的，包括开放获取学术出版者协会（OASPA）；
- CC-BY-NC 除了不允许商业目的的再利用外（没有事先获得许可与标准的版权），与 CC-BY 是同样的。许多 OA 倡导者倾向于 CC-BY 协议，认为商业是一个模糊的术语，并认为允许公共资助研究的商业开发是为了公众的利益；
- CC-BY-NC-ND 附加了排除衍生作品的创作。OASPA 不允许其成员使用该协议，因为它认为学术研究中基于以前工作的衍生使用是基本的；
- CC-BY-SA：“共享”骑士需要那些创造衍生作品附加到共享协议。该协议（可能令人吃惊地）受到大多数 OA 倡导者反对。例如，OASPA 不允许其成员使用，因为共享文章中的材料分布只能是组合的且与其他共享的内容重新分布。

STM 于 2014 年 8 月发布了新的开放获取模式协议⁴²，其目的是作为 CC 协议的补充。该协议的设计为包括最新的发展，如多语言获取、文本挖掘、以及商

⁴² <http://www.stm-assoc.org/open-access-licensing/>

业利用的具体情形（涵盖与 OA 内容相关的付费广告）。

协议没有很好地得到 OA 倡导者和活动家的认可；一个联盟（包括资助者、机构、出版商、策展人和公共资源的用户）发表声明呼吁 STM 撤回该协议（获取研究、科学与教育的组织机构的全球联盟呼吁 STM 撤销 2014 年新版示范协议）。STM 的回应声明指出（STM, 2014），对该问题有多个观点，包括它自己的成员：有些倾向于基于单一协议规范化的效率，而另一些则倾向于为作者提供选择和选项，这也许反映出特别的关注。这最终很可能是一个由市场决定的问题。

文本和数据挖掘权益

文本和数据挖掘（TDM）被认为是使用 STM 内容的一种重要的且持续增长的方式。“学术交流的新进展”中对其有更详细的讨论，但也值得在这个版权章节中作为一个词条，因为权益问题正在热烈讨论并且有更多的问题加入。

在撰写本报告时，在没有出版商同意的情况下 STM 期刊允许对其数据进行挖掘是不常见的，且大多数出版商（OA 出版商除外）没有明确的可公开获取的政策，而是基于个案处理每个请求（Smit & van der Graaf, 2011）。对于打算挖掘一定数量文献的研究者而言，要求逐一联络出版商无疑是一个沉重的负担⁴³

在对这个问题的一个平行的公开回应中，Hargreaves 报告为 TDM 提出了一个版权除外的建议；这已经在英国实施，但没有解决大多数研究人员的问题，因为这个问题是全球范围内的，并且这样做只能挖掘已经授权许可的内容，而一个更普遍的问题是要挖掘授权和未经授权的内容。

一个更好的改进方法是在更一般的情况下可以是一个全面的协议过程，涵盖多个出版商。2012 年采纳的一个小而重要的一步是示范协议条款涵盖 TDM，STM、ALPSP 和 P-D-R（医药文献环）已同意该条款⁴⁴。

出版商已发声明承诺促进 TDM 的非商业性使用（STM, 2013a）。在这方面，STM 已制订示范协议条款，可以添加到现有的出版商-图书馆协议中以支持规定术语的 TDM（STM, 2012）。

一些更雄心勃勃的跨行业协作的举措已经出现，特别是那些由 CrossRef、CCC 和 PLS 引领的举措，这将在下节的文本和数据挖掘中讨论。

这个问题也与开放获取有密切关系。比如，英国研究理事会 2012 年的获取

⁴³ 例如，自 2000 年以来，PubMed 中发表 1000 篇以上论文的出版商有 587 个，这显然对于大多数人来说是难以协调的。

⁴⁴ <http://is.gd/UXnRMI>

取政策要求作者存储于知识库中的论文允许 TDM，还要求在期刊中发表的开放获取论文允许 TDM。有些出版商将他们 OA 论文从 CC-BY-NC 协议改为 CC-BY 以适应这些和其他方面的压力。

机器可读和嵌入式许可

一个潜在可行的应对孤儿作品和数字内容用户的获取权力误解的方法是将许可以一个可被机器识别的格式嵌入到信息资源中。这在一些特定种类的媒体文件中已经被使用，尤其是在线销售的音乐和视频。在这些领域，这种方法经常与数字版权管理（DRM）有关，但也并非必须：许可能够简单地声明作品所属，并明确指明允许的后继使用和许可条件。

2011 至 2013 年链接内容联盟（Linked Content Coalition）⁴⁵ 建立了一个权利数据网络以提供“能够完全互操作和互链接的基于标准通信基础设施的框架”。这可包括整个供应链以及最终用户，并可能涉及到所有的媒体类型，不仅仅是 STM。权利数据网络将是“一个权威的链接的数据，其中的权利数据网络中所有实体都有标准，解析的标识符；这些标识符以标准的方式链接；并且标识符和链接的管理遵照确保适当授权下的登记程序，有合法利益的当事各方能确认利益是正确和公开认可的。”

2.17 长期保存

在印刷界，长期保存显然更是图书馆（而非出版商）的责任。保存的成功需要纸张的耐久度（无酸）、多份分散的保存和保存机构的长久存在。

对于电子期刊，事情远没有如此简单。最基本的问题是长期电子保存的问题还没有被完全解决：即使保存二进制数据看起来可行（定期在旧存储器老化的时候转移到新的存储中），问题却在于这些数据在未来可能无法被识别，例如，当相应的硬件/操作系统无法使用的时候。一个不那么根本但仍然重要的实际问题是大多数电子期刊是通过出版商的服务器访问的；图书馆本身并不拥有这些期刊用于永久保存，并且不能指望出版商在未来的任意一天都会继续存在。这种不能被证实的长期保存的解决方案是阻碍图书馆员全部转向为电子订阅的原因之一。

技术问题也在研究项目中涉及，例如，在 Koninklijke Bibliotheek（荷兰的国家图书馆）、数字管理中心、英国的大英图书馆及其他地区的项目中。然而，在现阶段所面临的技术至少是机构组织与技术方面是同样的。

当前的主要解决方案如下：

⁴⁵ <http://www.linkedcontentcoalition.org/>

- 国家图书服务：最早的且最知名的这类服务是 Koninklijke Bibliotheek 的 e-Depot⁴⁶。它的电子存档服务面向全世界的出版商，且被很多主要出版商采用，例如 Elsevier、Springer、Wiley-Blackwell、Taylor & Francis、OUP 和 Sage。e-Depot 还为荷兰的知识库提供存档服务。
- LOCKSS（多份副本保证文档安全）⁴⁷。正如这个方法的名字，它基于冗余的原则工作，这与保存多份期刊会更加安全的原理相似。基于斯坦福的 LOCKSS 系统允许图书馆收集并储存有特别许可的订阅内容（超过 500 个出版商都许可他们的内容储存在 LOCKSS 系统中）。这个软件允许每个图书馆的服务器与其他的图书馆服务器不断进行比较并修补任何损坏。
- CLOCKSS（控制 LOCKSS）⁴⁸ 是一个基于 LOCKSS 技术的、有约 200 个学术出版商和 750 个研究性图书馆的可持续合作组织，。
- Portico 是一个非营利性的为学术内容提供储存服务的机构⁴⁹，在发展成为一个独立组织之前是一个 JSTOR 项目。它提供了一个电子期刊和电子书籍（和其他电子产品）的永久可控存档，图书馆可以受益于其对某些界定的触发事件所造成损失的保护（如出版物不能从出版商或其他来源获取）。它还为取消订阅后的获取提供便利。2014 年末期有大约 920 个参与的图书馆，涵盖超过 20,600 种期刊和 340,000 本电子书籍的 274 个出版商，共有 7.05 亿个文档。
- 永久获取联盟（APA）⁵⁰ 旨在通过信息交换、合作和特定的项目建设，为科学信息永久从一个可持续发展的组织的基础设施中获取建立一个共享的愿景和框架。一个相关的组织，APARSEN（欧洲科学的 APA 记录）是一个支持研究有关数字信息和数据长期可获取和可使用的障碍的优秀网络。

根据一份 ALPSP 的报告（Inger & Gardner, 2013），Portico，其次是 LOCKSS / CLOCKSS，是大中型出版商最受欢迎的选择。调查中的所有大型出版商都有某种存储的安排，但几乎五分之一（18%）的小出版商没有。

⁴⁶ <http://www.kb.nl/en/organisation/research-expertise/long-term-usability-of-digital-resources/information-for-international-publishers>

⁴⁷ <http://www.lockss.org>

⁴⁸ <http://www.clockss.org/clockss/Home>

⁴⁹ <http://www.portico.org>

⁵⁰ <http://www.alliancepermanentaccess.eu>

2.18 转换准则

UKSG 转换操作准则⁵¹是出版商之间对于期刊转换而形成的最佳实践自愿声明,旨在减少图书管理员和终端用户之间潜在的中断。它规定了转出和接收出版商的作用和责任,且涵盖诸如对先前订阅内容的永久获取权、数字内容和征订目录的转换、利益各方之间的交流,以及期刊 URL 和 DOI 的转换。目前这个准则的版本为 3.0,已经有约 50 个出版商表示认同,包括所有大型期刊出版商。

除了保持该准则,转换工作组保持着提醒服务(包括通知数据库,表格和目录清单),并且提供非正式的建议。有关该准则更多的背景和演化详见 Phillpotts 及其同事的阐述(Phillpotts, Devenport & Mitchell, 2015)。

2.19 研究人员对期刊的获取

科技期刊在线版的发展使得对科技文献的获取大幅增加,且平均每次使用的成本大幅下降。这很大程度上是因为电子出版发行的低边际成本使得出版商可以为一系列期刊(达到且包含出版商的所有作品)提供在线获取,与机构早期的印本订阅总体费用相比较,这只增加一个相对较低的许可费用。在需求方面,图书馆已经形成联盟来增强他们在与出版商协商电子许可时的购买力,这也导致了他们的读者可以获取更多的期刊。

统计显示,上世纪 90 年代后期随着电子期刊的引入,每个图书馆获得的期刊数量有了大幅增长,且对每种期刊的平均支出下降了。例如,ARL 的统计(ARL, 2011)显示每个 ARL 图书馆的期刊购买数量在 1990s 间有所下降,在 2001 年达到了低点的 13682 种,但随后大幅增长至 2011 年的 68375 种(并非所有的都是经同行评议的期刊),与此同时,平均每种期刊的价格从 2000 年的顶峰持续下降。相似的,平均每个英国高等教育机构当前订阅期刊的数量在 1994/1995 到 2004/2005 的 10 年间翻番,从 2900 种到 7200 种(Creaser 等, 2006)。SCONUL 的数据显示了在英国获取有相似的增长,澳大利亚的数据也呈类似的趋势。

研究信息网络(RIN)发布的两份报告《电子期刊:使用、价值和影响》(RIN, 2009b; 2011b)描述了英国高等教育机构中有关协议许可获取的巨大影响。例如,全文下载在 2003/2004 至 2006/2007 年间增加了两倍以上,达 1.02 亿篇次,并且到 2008 年一直是保持年均增幅 20%以上,同时获取的费用在 2008 年下降到每篇 0.70 英镑(在研究密集型机构为 0.65 英镑)。该研究发现大学在电子期刊方面的支出费用与下载数量间呈正相关,并且期刊的使用和支出费用与研究成果产出有很强的正相关性,与机构的规模大小无关。

⁵¹ <http://www.uksg.org/transfer>

当前获取的程度

评估当前学术期刊获取的程度对于政府部门和其他政策制定者来说是一个核心问题，但这方面的调查在方法上都有或多或少的不足(Meadows, Campbell & Webster, 2012)。近期政府部门的咨询报告尤其印证了这个现象(OSTP, 2012; European Commission, 2012b; European Commission, 2012a); 公平地讲, 这些调查是明确的咨询行为而非市场研究, 但如果认为这些结果具有代表性或广泛性就会出现危险。

这些方法上的差别和不足使得不同的调查之间难以比较和解释。2011 年末 CIBER 代表 RIN 进行的调查(RIN, 2011a) 分析了 20,000 个调查中的 2645 个反馈(反馈率为 13.2%)。这个调查再次印证期刊论文的重要性(会议论文次之)。在大学, 有 93%的人表示研究论文的获取容易或相关容易, 72%的人表示获取情况在过去的 5 年有改善。这个发现与早先使用相似方法的研究结果相似, 似乎表明获取方式表面上看没有什么问题。

类似的, Outsell 为澳大利亚 Go8 图书馆组织开展的研究分析了来自 13807 个澳大利亚研究人员的 1,175 个反馈(反馈率为 8.5%)。调查发现 91%的反馈者都表示对信息资源的获取能很好或足够满足他们的需求。

当 CIBER 调查的受访者被问及他们最希望看到哪些方面的资源获取得到改善时, 很大一部分人(大学中有 39%)都认为期刊论文是他们的首要选择。

在欧洲委员会的调查中, 主要的受访者是图书馆管理员, 有 84%的人不同意或强烈不同意“欧洲的科技出版物没有获取方面的问题”的论断。OSTP 咨询的受访者还主张更强有力的政府强制和集中知识库, 以改善获取。

怎样来调和这些立场呢? 首先, RIN 的作者发现, 大部分文献的“容易”获取对许多研究人员而言并不够。虽然大学获取的程度总体上来说不错, 但其他很多地方获取并不容易, 尤其是产业部门和其他群体, 如没有学术图书馆获取途径的独立专业其他人员(Ware, 2009)。

更普遍的是, 以前一直版权除外的东西可能不再能满足目前的需要。据 Meadows 推测, 因为研究人员知道几乎所有的期刊论文都有电子版本可获取, 因此当他们无法获取某些特定资源的时候会很沮丧并且会表达不满。另一个因素可能是通过搜索引擎增加了研究论文的显示度和发现有容易度, 并增加了使用这些去寻找学术内容。

Finch 报告(Finch Working Group, 2012)注意到, 学术界和研究密集型公司的大多数研究人员都比以往任何时候能够获取到大量期刊, 但是他们想要更多:

“每年产出的可在线免费获取和使用的文章有近 200 万篇,此外还有以前的出版物;并且能够使用最新的工具和服务来分析、组织和管理所发现的内容,因此,搜索新知识的效率更高。”

获取的障碍

获取的障碍是一个重要事项: RIN 调查结果指出:“信息障碍可以导致显著的无效劳动及研究人员和知识工作者失去机会”。相似的, Finch 报告认为提高获取有助于提高透明度、开放性和责任性、以及公众的参与度,密切研究与创新之间的联系,促进经济增长,提高研究过程中的效率,增加研究投入的回报。

以上讨论中所有调查和咨询中最常被提及的获取障碍是费用:在 EC 和 OSTP 的咨询调查中,期刊高额订价和缩减的图书馆预算均被 85%以上的受访者提及。RIN 调查也发现最常见的障碍是研究人员不得不为获取的内容付费:认为期刊获取很重要的受访者中大多数都认为他们在当前的策略中没有得到足够的获取。除了高订阅价格外, RIN 的受访者还觉得单篇论文的收费过高。

虽然价格障碍是最重要的,但它们并非这些(以及以前的)调查所指出的唯一问题。其他障碍包括:对现有可获取的资源缺乏关注,繁琐的购买程序,数字出版的增值税,格式和信息技术问题(包括数字版权管理问题),有可获取内容的图书馆会员的缺乏,以及作者或出版商的权益与对内容使用渴望之间的矛盾。

对英国学界的 Ithaka 调查发现,当想要一个图书馆没有收藏的资源时,表示在网上找一个免费版本的受访者比例最高,比例第二高的是表示放弃,这二者都高于使用图书馆的馆际互借和文献提供服务的比例(Ithaka S+R 等, 2013)。

在这个背景下 Finch 组织在英国成立,并发表了一个简报调查增强获取的方式。它所提倡的方法主要集中在将开放获取扩展到更长的时段(见“开放获取”章节),但这些建议中也包含了若干个旨在过渡到开放获取的扩大短期获取的建议(RCUK 政策):增加资助国家许可的扩展和合理化的覆盖,为公共图书馆提供大部分期刊的预约获取(见“公众获取”),发展诸如中央和地方政府、志愿部门和商业部门的许可证。

中小型企业(SMEs)

公共政策对中小型企业(SMEs)获取科学文献的关照有所增加。中小型企业被视为是创新和就业的一个源泉,因此在全球经济衰退的情况下尤其有着重要的作用。中小型企业还不是期刊出版商的核心市场,因为他们通常并不订阅期刊,但他们从图书馆、数据库和文献支持服务中获取期刊。出版研究协会的一份调查(Ware, 2009)发现英国高技术中小型企业中的人比大公司中的人更加看重信息,

并阅读更多的论文。在认为信息很重要的群体中，71%的人认为他们有着不错的获取状况，60%认为状况比5年前有所改善。然而，报告发现一半以上的人有时难以获取一篇论文，并且也举出了一些为提高获取而有可能采取的方案：单篇付费获取可以通过面向公司的更合适的付费机制和降低价格而变得更加简单；高等教育机构的期刊许可能够包括在线版本并为地区性企业预约获取；此外，可探索创建更全面、更集中的国家许可。Finch 组织对其中的一些方案付出了实践，即使也注意到从事研发工作的中小型企业比例非常小。

自上述的2009年调查以来，对这一问题的进一步研究相对较少。Houghton、Swan 和 Brown (2011) 研究了丹麦的中小企业在获取方面的情况，调查了获取和使用水平、是否有任何的获取障碍、获取困难或间隙、涉及获取研究成果的成本和收益等，发现对学术研究的获取可实质性的收益。在过去的3年中，如果没有对学界的获取，开发或引入的27%的产品和19%的工艺被推迟或放弃，同时新产品对年度销售额贡献的平均值为46%。大约有一半的受调查者认为研究文章非常或特别重要，也有类似的比例(55%)表示了获取研究论文的困难。最广泛使用的访问非开放获取材料的方法是个人订阅和内部图书馆或信息服务。公共图书馆、图书馆间的借阅和按次付费阅读(PPV)很少被使用。

公共获取方案也许能帮助中小企业，虽然预约获取(如英国获取研究试点计划)没有在中国的研究中得到高度认可(见“公共访问”部分)。

在过去的几年里，已经推出了新的服务以获取期刊论文，其价格低于出版商的全文“按次付费阅读”，供应商包括 DeepDyve、ProQuest Udini、ReadCube Access 以及 Infotrieve (CCC, 版权结算中心)。获取或者受时间限制(文章出租)，或者受功能限制(如不能打印和本地存储)，商业模式包括一次性费用或针对集团或公司的月度或年度计划。同理，Reprints Desk 的 Article Galaxy Widget 允许用户搜索文章并寻找成本最低的获取选择。

发展中国家的获取

在各类调查中，富裕的英语国家(美国、加拿大、英国、澳大利亚)的获取情况最好，小的欧洲国家和中东地区次之，亚洲不出意料地在世界上排名最低。

有不少方案旨在为发展中国家的研究人员提供免费的或高折扣的科技文献资源。

Research4Life 项目⁵²是联合国机构、STM 出版商、大学及大学图书馆、慈善机构和科技伙伴的合作项目。该伙伴的目标是帮助联合国2015年实现8个千

⁵² <http://www.research4life.org>

年发展目标中的 6 个，减少工业化国家和发展中国家之间的科学知识差距。目前有 4 个方案，为 109 个发展中国家的约 7700 个机构提供免费或低成本获取来自 180 个国家超过 30000 种期刊和其他全文资源：

- HINARI，2002 年 1 月与世界卫生组织（WHO）联合启动，为发展中国家的地方性非赢利组织提供免费的或低价的在线获取，包括主要期刊、全文数据库和其他生物医学和相关社会科学资源。
- AGORA，2003 年 10 月由联合国粮食及农业组织与主要出版商联合建立，提供基于数字图书馆的来自 70 个出版商的 3000 余种期刊的在线获取，涉及食品、农业、环境科学和相关的社会科学。
- OARE（环境研究的在线获取），以联合国环境项目作为合作伙伴于 2006 年末建立，提供超过 3900 种期刊的环境科学方面文献的获取。涉及学科包括环境化学、经济、法律和政策，以及其他的环境学科，例如植物学、保护生物学、生态学和动物学。
- ARDI（发展与创新研究的获取），于 2009 年与世界知识产权组织合作建立，并于 2011 年加入了 Research4Life，旨在促进发展中国家和最不发达国家融入全球知识经济体系。

这些项目为最贫困国家（基于人均国民生产总值）提供免费获取和费用极低的获取（通常每个研究机构花费大约 1000 美元就能获取所有信息）。

其他的项目包括：

- HighWire 出版社为发展中国家提供近 500 种高质量期刊的免费获取⁵³，基于一个可以识别用户获取站点位置的软件。
- 有些出版商独立提供相似的项目，如英国皇家化学学会、美国国家科学院出版社。
- INASP 的 PERI 计划于 2013 年结束，但 INASP 通过其加强研究与知识体系计划⁵⁴继续支持获取研究成果。
- eIFL（图书馆电子信息）⁵⁵合作伙伴是图书馆和图书馆联盟建立的能力，倡导获取知识，鼓励知识共享，以及启动创新型图书馆服务的试验方案。

⁵³ <http://highwire.stanford.edu/lists/devecon.dtl>

⁵⁴ <http://www.inasp.info/en/work/what-we-do/programmes/srks/>

⁵⁵ <http://www.eifl.net>

发展中国家在获取和使用文献方面的问题不仅仅局限于支付的困难。Research4Life、INASP 和 eIFL 都认识到存在更广泛的问题,并多方面提供培训、推广和支持、宣传、带宽拓展。此外,还有对作者的支持,例如,通过 INASP 的作者帮助项目⁵⁶。

也有担心提供西方期刊的免费获取(或相应地提供开放获取费用的减免)可能会导致限制本地的新生出版产业的预料不到的后果(Dickson, 2012),很多这类项目都在密切关注这方面的影响。

⁵⁶ <http://www.authoraid.info/>

3. 开放获取

开放获取（OA）是指内容（特别是指期刊论文，尽管对其他形式的研究产出 OA 的兴趣也在逐渐增加，包括专著、会议论文集等）可供在线获取，免费，并且基本不受版权或者许可限制，也不会有技术限制或其他获取障碍（例如数字版权管理或要求注册获取）⁵⁷。

因此，严格来说，OA 是单篇文章的属性，而不是期刊的属性。不同方式的 OA 可从以下三个方面考虑：开放什么（What），何时开放（When），如何开放（How）。

“开放什么”可以从以下 3 个阶段区分：

- 第一阶段：未被期刊考虑的作者的未评审初稿，通常叫做预印本（特别是在物理学界）。“作者原稿”（author's original）是 NISO 版本中倾向使用的术语（见 *文章的版本*）。
- 第二阶段：作者的经过同行评议并已被期刊录用、即将出版的论文版本，文章的内容已经根据同行专家的建议作了修改（“已录用的稿件”）。
- 第三阶段：最终发表、可从期刊网站获取的可供引用的版本（“记录版本”）。

用户可以获得哪些再使用权是一个有争议的问题，从本报告上一版本发布至今，这一问题获得了很多的关注，一些基金资助机构要求遵循 CC-BY 协议，然而很大比例的作者不愿放弃商业再使用等权利（见 *开放获取许可*）。另外一个因素是文本和数据挖掘（TDM）日益重要，不过这更为复杂，因为 TDM 方案需要跨越 OA 和非 OA 的资源（见 *文本和数据挖掘*）。

在 OA 时间（when）方面有三个选择：（正式）出版之前；出版的同时；出版一段时间之后（“保密”期）。“如何”（how）开放在很大程度上是一个商业模式问题。

基于这个框架可以区分当前采用 OA 的主要类型：

- **全开放（金色路线）**：在第三阶段文章出版时立即开放，利用“翻卷”（供方）商业模式或者赞助模式。
- **延时开放**：第三阶段，但是文章发表后延时开放，基于订阅的商业模式。

⁵⁷ 见 <http://www.earlham.edu/~peters/fos/overview.htm>

- **自存储**（绿色路线）：在第二阶段时，或者立即开放或者延时开放；没有独立的商业模式。

每种方式都有不同的变异，下一节将简要讨论并且了解当前进展情况。

3.1 开放获取驱动力

促使 OA 发展的驱动力有研究资助者和决策者的政策和干预，OA 出版的逐渐发展、成熟，以及能够为作者提供更多可靠 OA 期刊的商业活动。

研究资助机构的政策很可能是建立 OA 环境的最重要因素。值得关注的里程碑事件包括：

- 2012 年 6 月英国 Finch Group 报告出版(Finch Working Group 2012)。随后英国政府接受了他们的建议，标志着明确的支持学术论文 OA 的政策，表示“受到公共基金资助的研究成果应该能够被免费、不受版权限制地获取，这个原则是强制性的，从根本上是无可辩驳的”。
- 维康基金会（Wellcome Trust）的紧缩政策（2012 年 6 月）⁵⁸，特别引入了对于不遵守原则情况的制裁和对 CC-BY 协议的推动。
- 英国研究委员会（UK Research Councils）新的统一政策（2012 年 7 月宣布，2013 年 4 月引进），这一政策进一步发展了之前的 Finch 建议(RCUK 2012)；另见下文的 *RCUK 政策*。
- 欧盟最新的一项长达 7 年研究计划，Horizon 2020，在 2014 年开始实施，其包含了欧盟 800 亿欧元的资助，该计划对作者的要求更严格，目标在 2016 年前开放 60%的资助项目(de Vrieze 2012)。政策支持通过 APCs（文章出版费用）报销的方式进行金色 OA；此外，所有的文章必须在发表之前保存在一个开放的知识库中，并在 6 个月以内（或者对 HSS 来说 12 个月）开放获取。
- 在美国，围绕国立卫生研究院（NIH）的强制以及其可能的扩大有争论。两个相互竞争的法案 Research Works Act (RWA)和 Federal Research Public Access Act (FRPAA)被取消，而第三个法案，the America COMPETES Bill 在 2011 年通过，法案要求科学技术政策局（OSTP）协调全联邦的资助机构的获取政策。因此 OSTP 在 2013 年发出通知要求较大的资助机构制定提供公众获取的计划。见下文的 *科学技术政策局*

⁵⁸ <http://www.wellcome.ac.uk/News/Media-office/Press-releases/2012/WTVM055745.htm>

OA 出版业的发展与成熟体现在 OA 期刊数量的增长（见 *OA 期刊和论文的数量*）。Björk (Björk 2011) 这样描述其发展，90 年代为自愿模式（通常由个体的学者发起）；经过一段历史悠久期刊的浪潮，特别是学会期刊或者地区性期刊比如拉美期刊，这些期刊在文章出版电子版时将其 OA；之后从 2002 年开始将选择性 OA 作为一种商业模式，最初是市场新入者实施，然后是现有的商业和非商业出版者实施。一些著名的出版商例如 Nature Publishing Group 和 AAAS，特别是非商业性机构例如 OUP、the Royal Society 和许多领先的学术团体等采用这种模式，帮助作者建立了这种模式的可信度。尽管许多作者不愿在没有影响因子的新期刊上发表文章，许多 OA 期刊还是生存了下来并有了可信的影响因子(如 Björk & Solomon, 2012a)。

政策焦点转向黄金模式

在 2012 年左右，一些政策制定者和资助者改变思路，倾向于支持金色 OA 模式。这种情况在英国尤其显著，Finch 报道不仅建议被纳税人资助的研究成果应该开放获取，而且建议应首先考虑以金色模式开放，资助资金应该包括出版费用在内。该团队（和英国政府）似乎认为，在为 OA 提供可持续发展的商业模式方面，金色模式比绿色模式更具有优势，另外也避免了可能给英国工业带来破坏的风险。2012 年 RCUK 也认可了金色模式的优势，并且为提供文章出版费用 (APC) 资金提供了一个资助方法（大学整体补助金）。然而一些研究密集型大学（例如英国罗素大学集团）仍然担心潜在的成本影响。

近期这股倾向于金色模式的势头似乎已经放缓，尽管 HSBC 一份关于学术出版方面 OA 政策的分析指出 (Graham 2013)，基本上所有的重要研究资助者支持 APC 报销的金色 OA 模式，但是很少表达出对金色模式的偏好，并且许多资助者积极地推动绿色 OA 更甚于金色 OA。

3.2 开放获取的商业模式

随着 OA 的规模化发展和趋于成熟，并且扩张到研究经费并不同等重要的实验科学领域，一种单一的商业模式显然不能满足所有的状况。然而基本的 APC 模式依然很重要，在这个基础之上出现了不同模式，同时还有日益增长的 non-APC 模式（见表 6）。

全开放获取（金色 OA）

在全 OA 中，最终的文章能够在出版时即时在线获取，出版商通过某种商业模式，使出版费用获得支付，而用户不需要支付使用费用，主要有两种形式：

- 立即全开放：期刊所有内容能够在出版时免费即时获取。

表 6 OA 的商业模式 (有关 OA 期刊的社论的定价方法另见 Björk & Solomon (2012b))

模式/策略	描述	实例
文章出版费用 (APC)	征收出版和相关服务费用, 常见不同的折扣和减免。	广泛应用
页面和其他出版费用	在 APC 基础上附加费用, 例如长度超过规定限制, 包含彩版、多媒体等	Science Advances (AAAS); PhysRevX; 一些有彩版本收费标准的混合期刊
投稿费	论文投稿费, 无论同行评审结果是否通过均不退回 拒稿率高的期刊可行	少见: 如 Hereditas; JMIR (投稿费意外地常见于基于订阅模式的期刊)
预付款	以折扣价格批量支付 APC	Taylor & Francis
机构会员	一系列相关模型, 例如以机构为基础的优惠、预付款、捆绑销售、补偿等	BMC; PLOS
抵消	针对特殊机构或联盟用户的覆盖所有订阅和 OA 费用	IOP; RSC; JISC “Total Cost of Ownership”; Austria FWF
捆绑销售	包含订阅、许可费加上 OA 出版费或机构会员制在内的组合费用	Springer
关联订阅的以机构为基础的优惠	为订阅同一出版商其他期刊的机构内作者提供折扣的 APC	OUP
个人会员	个人购买会员的一次性费用 (分层次); 所有的共同作者必须是会员 (达到最大数量); 会员需要参与同行评议来保持良好的信誉	PeerJ
免费增值	对一个基本的在线版本实行 OA, 加上附加服务费用(打印、增强电子格式等)	JMIR; OECD (data/books); Knowledge Unlatched (books)
第三方支付 APC	通常作为永久性模式过渡期的支持, 由协会、机构、基金会等支付优惠的文章出版费用	Some BMC transfers-in; MedKnow; Versita (De Gruyter Open); eLife (at present)
赞助 (无 APC)	赞助商承担不采取 APC 收费的费用, 赞助商包括: 协会、机构、研究组织、基金会、研究资助者, 也可能包括志愿者	Eurosurveillance (ECDC); Clinical Phytoscience (Springer); Asia & the Pacific Policy Studies(Wiley); Journal of Law and the Biosciences (OUP/Duke/Harvard/Stanford)
图书馆合作伙伴补贴	为 OA 出版集体资金创建的新图书馆联盟	Open Library of the Humanities SCOAP3, Knowledge Unlatched (monographs)

- 混合(或选择性)开放: 期刊部分内容能够即时获取。期刊给作者提供选择, 作为支付费用的回报, 可以将他们的文章 OA 或者存放在订阅访问的期刊中。

最知名的 OA 出版模式是“作者方支付”模式, 即由作者(或者其研究资助者或所在机构)支付出版费用。全开放和混合开放期刊均使用这种模式。许多全开放和混合开放期刊也提供“机构会员”付费模式, 即负责支付的机构的成员能够享受较低(或免费)的出版费用。

这种方式有优势, 不仅能够衡量科研产出的增加幅度, 还能使学术内容广泛获取, 为出版商提供了一个商业模式。然而进一步广泛应用显然还存在一些阻碍, 将在下文中讨论(见[过渡和可持续问题](#))。

混合期刊

许多已经建立的订阅型期刊正在试用 OA 模式, 混合型开放方式为这些期刊的试用提供了一个相对低风险的方式, 让市场(即作者或其资助者)决定 OA 的价值。几乎所有商业性的和非营利性的主要期刊出版者都提供混合模式。然而迄今为止 OA 市场使用率却非常小(仅有 1%~2%, 另见 [OA 文章的数量](#))。

Bird 报道过她所工作过的 OUP 期刊和其他几个出版商的混合开放的市场使用率(Bird 2010)。总的使用率是 6%, 不同学科差异较大, 人文和社会科学 2%, 医学 4%, 数学 6%, 生命科学 10%。一些生命科学期刊更高, 比如 *Human Molecular Genetics* (人类分子遗传学) 18%, *Bioinformatics* (生物信息学) 30%。相比报道的其他出版社, OUP 的市场使用率更高。Bird 推测这可能是因为 OUP 为订阅期刊的机构的作者提供 50% 的优惠(通常大多数作者都来自这些机构)。她引用的其他出版社的使用率如下: Nature Publishing Group: STM 期刊 5%, 还有其他较高使用率的期刊比如 *EMBO* 11%; Wiley-Blackwell 调查了其他出版社的网站, 总体使用率为 1%~2%, 但是有些期刊高达 20%, Wiley-Blackwell 自身的使用率总体来说则非常低, 不过在 2008 年有两个生物医学期刊到达 10%~20%, 在 Bird 文章发表后出现了一个突出的例子, *Nature Communications*, 其选择性加入率超过 40%(每篇文章的费用为 5000 美元), 不过该刊现在已经转变为全开放模式。

一些图书馆员和资助者仍然对这种模式持有一定的怀疑态度, 他们关心这种混合开放模式是否会导致更低的订阅费用(所谓的“双重收费”问题)。应用这种模式的出版商表示他们将在制定订阅价格之前把 OA 费用的影响列入考虑范围之内; 费用的降低能够使所有的订阅者获益, 不仅仅包括作者出资将自己文章

OA 的机构。(另见抵销)。

抵消

为了应对针对双重收费的担心和来自大型购买联盟的压力,一份由 Wellcome 和其他机构委任的报告推荐了三种可能的途径(Björk & Solomon, 2014), 以使混合开放市场更透明, 更具有竞争力。这些途径包括: 按照标价偿还 APCs; 有限价的分级 APCs; 资助者仅偿还部分 APC, 由大学来弥补差额。

另外, Jisc Collections 设计了一种称为学术交流的“总所有权成本”: 共同管理订阅和 APC 费用(Lawson2015)。一些出版商已经同意对此进行试验。可供考虑的选择有: 出版商提供信用值来应对将来订阅被取消时的 APC; 或者相反的, 提供信用值来应对 APC 已支付时的将来订阅费用; 或者未来 APCs 的适度额外收费与订阅的捆绑。

市场中也出现了类似的抵消模式, 即 OA 费用由订阅费用抵消, 或者将两者作为单一收入管理。例如 RSC 首创的“Gold for Gold”模式⁵⁹, 英国物理学会出版社使用的抵消方案(Jump 2014a)。

文章出版费用

表 7 列出了 2014~2015 年度一些主要学会和商业性出版社的出版费用。全 OA 期刊的费用主要分布在 1000~5000 美元之间, 除了 Hindawi 均价为 600 美元, 对于案例报告、短篇报道和没有大量经费的研究领域文章, 有较低的收费 400~1000 美元。混合型的 APC 相对更贵一些, 分布较窄, 基本在 \$3000 左右。另外还有一系列的折扣、捆绑销售和抵消模式(表 6), 使情况更为复杂。一些期刊会征收附加费用(例如基于篇幅、更快速的同行评议、不同的许可授权等)。

表 7 中给出的费用范围反映在研究资助机构 Wellcome Trust 支付的平均 APC 是 2365 美元, 尽管其中包括很大比例的混合 APCs。然而, 期刊所有文章 OA 的平均 APC 要低一些, 一份 2014 年基于 Scopus 收录的 OA 期刊的调查显示(Björk & Solomon, 2014), 全 OA 期刊的平均 APC 在 1418 美元(纯 OA 出版商)和 2097 美元(订阅出版商)之间, 而混合开放的平均费用为 2727 美元。这份调查还显示相比混合开放期刊, 全开放期刊之间的 APC 差别较大; 在全开放期刊中, APCs 与影响因子相关, 混合开放期刊则并非如此。

这些平均 APC 与同一作者之前调查的结果相比偏高。一种解释是, 之前的研究利用 Scopus 而非 DOAJ 作为样本来源, 因此都是已经确立的(至少 2 年)、

⁵⁹ <http://www.rsc.org/Publishing/librarians/GoldforGold.asp>

更专业出版的期刊。前一份调查参考了 1370 本 OA 期刊，在 2010 年总共发表文章 100687 篇(Björk & Solomon, 2012)。作者总结出总平均费用为 906 美元，不同学科的变化情况为：生物医学约为 1100 美元，技术与工程为 530 美元，艺术与人文为 240 美元。在这份调查中，APC 的分布有两个峰，较高峰值为 1600~1800 美元，较低的峰值大约在 600~800 美元，对应于不同类型的出版商。这表明，随着 OA 的扩大，市场平均 APC 在逐渐提高，原因是相比那些收费低的小规模、兼职类出版社，“专业的”出版商有更大的扩大市场的能力。

ALPSP 学术期刊出版实践报告还提供了有关平均 APC 和其分布的信息(Inger & Gardner, 2013)：出版商的全 OA 期刊 APC 的中值为 1350 美元，混合开放期刊为 2500 美元。作者认为，与之前的版本相比，全 OA 期刊的平均 APCs 有所下降，而混合开放期刊则大部分保持在 3000 美元左右。

表 7 部分全 OA 和混合 OA 期刊出版费用。 各类优惠（学会会员、订阅/“会员”机构、低收入国家等）未包括在内。零 APC 由第三方资助，不含促销（来源：出版商网站，2014 年 12 月，£/\$=1.6, €/£=1.3）

期刊/出版商	全 OA/混合 OA	APC, 全 OA(US\$)	APC, 混合 OA(US\$)
American Institute of Physics	Full/Hybrid	1350~2200	1500~2500
American Physical Society	Full/Hybrid	1700	1700~2700
BioMed Central	Full	1015~2650 (中值2075) (不含0)	-
BMJ Group	Full/Hybrid	2160~2720	3120~4800
Cambridge University Press	Full/Hybrid	600~1600(STM)	2700(STM) 1500~2700(HSS)
Elsevier	Full/Hybrid	500~5000	3000(most)
F1000 Research	Full	1000	-
Hindawi	Full	300~2250 (中值 600)	-
Nature Publishing Group	Full/Hybrid	1350~5200	3975(CCBYNCNC) 4400(CCBY)
Oxford University Press	Full/Hybrid	0~2270	1600~4000
PLOS	Full	1350(PLOS ONE) 2250~2900(其他)	-
Royal Society(London)	Full/Hybrid	2160	2880
Springer(见上述 BMC)	Full/Hybrid	1070(SpringPlus) 0~1700	3000
Wiley-Blackwell	Full/Hybrid	0; 800~4500	3000(大多数)

2014年5月，另外一项基于 DOAJ 数据库的调查报道了一组非零 APC 期刊样本的平均 APC 为 1221 美元（中值为 1145 美元）(Morrison & et al, 2014, p.9)，这项调查还报道了一份 APC 优惠和附加费的长列表（p.11）。

资助者将向最大程度允许再使用许可的转移作为支付 APC 的条件也将产生影响。传统的期刊收入有很多渠道，不仅仅是订阅费，很大一部分来源于二次利用的版权费（通过诸如 CCC 和 CLA 的集体管理组织支付）、商业在使用或再版（特别是制药工业用于针对医师的药品促销的再版）。Wellcome Trust 和 RCUK 等资助机构要求全 OA 费用只能向使用 Creative Commons CC-BY 许可的出版商支付费用，因此，根本上说这减少了制药工业再版和复制渠道的收入。一种可能的方法是 APCs 随许可条件而改变，对于那些可能减少出版商收入的要制定较高的价格（最近 Nature Publishing Group 利用这种模式，CC-BY 许可比 CC-BY-NC-ND 许可高出大约 10% 的收费，尽管随后他们将 CC-BY 采纳为默认许可）。

尽管如此，这里显示的费用大多比之前报道的篇均成本要低（基于诸如(RIN 2008)），这可能是混合 APCs 要高于纯 OA 期刊的原因之一。

为了不将低收入国家和缺少经费的作者排除在外，大部分 OA 期刊会对这些作者免收费。因此在设定 APC 或者根据标价 APCs 计算市场规模时，需要对免除或优惠的费用比例进行补贴。例如，PLOS 的年度报告中假定免除的收费总计占 APC 收入的 8% 左右(PLOS 2014)。

“转型”期刊

将期刊从订阅转变为 OA 的过程被称作“转型”商业模式。”Open Access Directory”⁶⁰ 列出了 196 个例子（遗憾的是并不完全精确），表 8 提供了一个更具选择性的实例列表。

当出版商要确定候选的“转型”期刊时，需要考虑的因素有：适度的订阅收入；相比订阅量，更期望作者数量的持续增长；较高的拒稿率；对作者有吸引力；该学科有可获得（和可用的）OA 资助；现存混合型文章的数量；当期收入与已发表文章的比率(Jones 2014a)。

混合内容期刊

在另一种混合商业模式中，期刊中的学术论文能够在发表后即时获取，但是对于一些“增值”内容需要订阅，比如受邀的综述文章、新闻等。BMJ 实行这种

⁶⁰ http://oad.simmons.edu/oadwiki/Journals_that_converted_from_TA_to_OA

模式。OA 出版商 BioMed Central 一小部分期刊也采用这种模式。

Non-APC 模式

并非所有的 OA 期刊都需要付出版费：开放获取指南（DOAJ）列出的期刊中大约有一半没有列出作者的费用。这些期刊使用各种资助模式，包括补助金、会员订阅、赞助/广告、商业再版、分类广告、订阅印刷版、义务劳动、主办机构（有意或无意地）以实物补贴或支持。然而，DOAJ 中多数期刊没有列出版费的事实可能会引发误解，因为大部分发表在 OA 期刊的文章确实是收费的(如 Dallmeier-Tiessen, S et al 2010)。尽管如此，Non-APC 的金色 OA 的规模依然很大(见表 2; Elsevier, 2013)。

被赞助或者补助的 OA 期刊不收取作者或读者的费用，这种期刊在大型和商业型的出版社以及研究机构、学会等越来越常见。在某些情况下这可能是在期刊创办时吸引作者的过渡性措施(如 eLife)，不过在其他情况下只是不打算引进 APCs。

表 8 “转型”期刊举例

期刊	出版商	出版年	转型年	IF	APC (美元)
Nucleic Acids Research	OUP	1974	2005	8.8	1450
Acta Veterinaria Scandinavica	BMC	1959	2006	1.38	2350
Genetics Selection Evolution	BMC	1960	2009	3.75	1745
Evolutionary Applications	Wiley	2008	2012	4.57	1950
Aging Cell	Wiley	2002	2014	5.94	2800
Cancer Science	Wiley	1907	2014	3.53	2250
Influenza & Other Resp. Viruses	Wiley	2007	2014	1.89	2500
Journal of Diabetes Investigation	Wiley	2010	2014	1.50	3000
Developmental Cognitive Neuroscience	Elsevier	2011	2014	3.71	1500
Stem Cell Research	Elsevier	2007	2014	3.91	1800
Int J infectious Diseases	Elsevier	1996	2014	2.33	1750
Epidemics	Elsevier	2009	2014	2.38	1800
EJC Supplements	Elsevier	2003	2014	-	3000
“Central European Journal of” series (x8)	De Gruyter	~2009	2015	0.43	tba
Nature Communications	NPG	2010	2015	10.74	5200
Chemical Science	RSC	2011	2015	8.60	tba

Non-APC 模式在人文与社会科学领域特别受欢迎，这一领域的研究资金要

比实验科学低很多(Edwards 2014), 尤其历史学家已经呼吁抵制 APC 模式(如 Mandler 2014)。

图书馆合作伙伴的补贴

图书馆合作伙伴的补贴(也叫做联盟 OA)是一种 non-APC 模式, 该模式演化为旨在包括为 OA 出版集资的特设图书馆联盟的创立。这是一个有很少实例的新模式, 当不同出版社共享其核心概念时, 在细节上有相当大的差异。

- 人文科学开放图书馆 (Open Library of the Humanities) 正在使用图书馆合作补贴模式, 为他们新的、受 PLOS 启发的人文科学巨型期刊平台筹集资金, 每个图书馆的指标性年费约为 925 美元, 假设有 200 个图书馆参与, 每年有 250 篇文章⁶¹。
- Knowledge Unlatched (专著): 这个相当复杂的模型创造了一个市场平台, 在这个平台上, 参与图书馆的成员可以选择参加由参与出版社提议的专著板块。出版商向图书馆收取费用, 将基本的在线版本(如简单的 HTML 版式)上传网络, 参与的图书馆成员可以免费获取。每个图书馆需要支付的费用随着参与图书馆的增加而降低。出版商保留出售打印版和更高增值版本的权利, 比如电子书版本和数字增强版。然而, 尽管这个模式很复杂, KU 看起来依然获得了牵引, 并在 2014 年发布了其第一个试点板块的题目⁶²。
- 由 CERN 和 SCOAP3 领导的高能物理联盟可能也被归类为图书馆合作伙伴的补贴模式。在这种情况下, 经过一个征求图书馆的承诺和后续的招标的过程后, 可以完成学科领域中几乎全部的主要期刊从订阅向 OA 的转化。
- arXiv 预印服务拥有三重资金: 大致上, 三分之一来自 Simon Foundation, 三分之一(以实物偿付)来自康奈尔图书馆, 另外三分之一来自大量使用 arXiv 的机构的集资。

这种模式存在一个尚未解决的难题是“吃白食者”问题: 无论是否参与集资, 图书馆(和其主顾)均能获得使用权, Knowledge Unlatched 已经在很认真地考虑这个问题, 尝试引入激励措施(例如提供印版优惠)以避免这个问题。无论图书馆有多希望支持 OA, 如果没有一定的限制因素(例如多年期合同), 削减补贴预算(能够免费获取资源的条件下)一定比订阅同等价值的资源容易得多。

⁶¹ <https://www.openlibhums.org/about/library-partnership-subsidy-lps-flyer/>

⁶² KU 寻求最低 200 个图书馆支持其发布首个试点板块。在 2014 年 2 月的最后期限之前, 超过 250 个图书馆签署加入。到 10 月, 大约有 13000 个下载记录, 每本书每周平均 40 次。

K|N Consultants 在一份白皮书中提出一种类似的模式，设想建立一个中央管理机构和独立的审查小组，来监督用于出版的补助金经费，机构的资金则由高校、科研机构、基金会和捐助者筹集(K|N Consultants 2014)。

3.3 OA 期刊的类型

除了以商业模式加以分类，描述主要类型的 OA 期刊也许有所帮助，如表 9 所示。

巨型期刊

OA 市场发展最快的部分是“巨型期刊”，这种由 PLOS ONE 开辟的出版模式被证实取得了巨大的成功，代表着对学术期刊模式极少数的创新之一，有着显著的广泛影响。这种模式由三个关键部分组成：较低价格的完全 OA；基于“可靠性而非重要性”的快速同行评议（例如，选择文章时，看重科学是推导出来的，而非对一个特定领域影响力、重要性或相关性的主观判断），再有保证直接了当的同行评议政策（避免补充实验和重新投稿的可能性）；非常宽的主题范围（实质上受到作者投稿意愿和期刊寻找审稿人能力的限制）。此外，这种模式已经与级联同行评议模式相关联（尽管在实践中，根据文章提交的数量上来看，从来没有对 PLOS ONE 有太大的影响），这种期刊促进了快速出版，在一定程度上是更简化的同行评议的结果（它在发展的同时，努力缩短其发表时间，使其比同领域其他领先的期刊更快）。

PLOS ONE 巨型期刊模式的成功引发其他出版商大范围的效仿，表 10 列出了一些实例。其他出版商采用了该模式的要素（范围广，快速出版，低价 OA），不过保留了更为传统的选择性同行评审流程：例如 *Physical Review X* (APS); *Open Biology* (Royal Society); *Cell Reports* (Elsevier); *Nature Communications* (NPG); *Science Advances* (AAAS)。

然而在 2014 年，PLOS ONE 之前持续增长的势头被打断：在年底，月产量比 2013 年 12 月高峰期下降 25%，尽管总发文量比 2013 年仅仅下降了不到 5%，但还是促使各方对其原因的猜测(Davis 2014)。

大期刊模式是否是一个完全的新模式或者是品牌延伸的化身，亦或者是学科领域期刊最后的手段，现在仍然还不清楚。随着 PLOS ONE 吸引其核心的生物医学领域之外的更多文章，这可能会影响到实际的平均影响因子，从而降低其潜在吸引力。还不清楚“全领域”的期刊是否能够在长时间内保持稳定。如果不能，它是否恢复为更像是传统学科期刊的一种集群，或者进化为一种更新的基于文章或者数据的学术交流模式。这些因素成为改造期刊出版的基本动力的核心。

表 9 OA 期刊的类型

期刊模式/策略	商业模式	描述	实例
巨型期刊	APCs 机构会员制 个人会员制	范围广，客观的同行评审，费用低，容量大	PLOS ONE; Optics Express; BMJ Open; Scientific Reports (NPG); SAGE Open
大学科范围的选择性期刊	APCs 图书馆合作补贴	范围广，与巨型期刊类似，不过采取传统的选择性同行评议	Open Biology (Royal Society); Nature Communications (NPG); Open Library of the Humanities
主题 OA 期刊	APCs 页面/出版费 投稿费 第三方承担 APCs 赞助(non-APC 模式)	OA 而非订阅的“标准”研究期刊 越来越多的新标准发布(如 Springer 自 2011 年以来的所有发布)	不属于以上两类的大部分 OA 期刊
级联评审/次级期刊	APCs 共享关联期刊 APC	通常定位在测试 OA 并保护旗舰期刊 通常的学会策略 次级学科中增长的案例研究	J Nutrition Science (CUP); Physiological Reports (Wiley/ Physoc/APS)
出版商和多重团体合作	APCs APC 与相关合作伙伴共享	及被拒文章的多个学会伙伴期刊 合作期刊共享 APCs	Wiley Open Access journals –e.g. Ecology and Evolution; Immunity, Inflammation and Disease; Energy Science & Engineering; 等等。
转型期刊	参考主题 OA 期刊	期刊从订阅转向 OA 模式	见表 8 的开放名录所列的 196 个案例 (并非全部准确) http:// is.gd/456h5v
混合期刊	APCs	任何类型的订阅期刊，在文章层面有选择性的付费 OA	由大型出版社出版的生物医学领域的大部分期刊
OA 会议系列	OA，基于卷和服务向相关机构收费(相当于 APC)	出版一系列的会议录 可能包括对机构的服务，例如使用在线追踪系统 价格低，半自动化处理	IOP 出版 – J Phys Conference Series

表 10 巨型 OA 期刊的创刊时间和出版量(注: 广泛学科范围的选择性期刊, 比如 *Science Advances*, *Nature Communications*, *Palgrave Communications*, *eLife* 等不包括在内)

期刊	创刊年份	总出版量(截止 2014 年)
<i>Optics Express(OA)</i>	1997	27290
<i>Zootaxa</i>	2001	15713
<i>PLOS ONE</i>	2006	114211
<i>Ecosphere</i> (Ecological Society of American)	2010	526
<i>mBio</i> (American Society of Microbiology)	2010	1018
<i>FEBS Open Bio</i>	2011	255
<i>AIP Advance</i>	2011	1390
<i>BMJ Open</i>	2011	2987
<i>SAGE Open</i>	2011	822
<i>Qscience Connect</i>	2011	91
<i>G3</i> (the Genetics Society of America)	2011	719
<i>Scientific Reports</i> (Nature)	2011	7325
<i>Springer Plus</i>	2012	1392
<i>Cureus</i>	2012	232
<i>The Scientific World Journal</i> (Hindawi)	2012	4914
<i>F1000 Research</i>	2012	707
<i>Biology Open</i> (Company of Biologists)	2012	451
<i>PeerJ</i>	2013	736
<i>SAGE Open Medicine</i>	2013	93
<i>CMAJ Open</i> (Canadian Medical Association)	2013	66
<i>BMJ Open Respiratory Research</i>	2013	36
<i>Open Heart</i> (BMJ)	2013	74
<i>IEEE Access</i>	2013	186
<i>Journal of Engineering</i> (IET)	2013	122
<i>Royal Society Open Science</i>	2014	52
<i>BMJ Open Diabetes Research & Care</i>	2014	32
<i>Elementa</i> (BioONE)	2014	29
<i>Frontiers in Medicine</i> (Frontiers/NPG)	2014	32
<i>Frontiers in Surgery</i> (Frontiers/NPG)	2014	46
<i>Open Library of the Humanities</i>	2014	-
<i>Cogent Economics & Finance</i> (Cogent/T&F)	2014	56
其他 Cogent 系列(T&F), 计划共 15 种	2015	-
<i>Collabra</i> (U California Press)	2015	-
[暂未命名](Elsevier)	2015	-

3.4 延迟 OA

在这种模式中，期刊将其内容在一定时间段以后开放使其免费获取，通常为 6 到 12 个月，在某些情况下为 24 个月。越来越多的期刊（特别是生命科学和生物医学领域）开始采取这种延迟 OA 策略。最著名的是以生命科学为主的、采用 HighWire 系统的、遵循 DC 原则的学协会出版社群体。尽管该集团可免费获取的文章已不再单独呈现，但 HighWire 平台目前还是拥有了 250 万可免费获取的文章，其中大部分都是从延迟 OA 期刊中获取的⁶³。

这种商业模式取决于保密时间要足够长，不至于危害到期刊的订阅销售；更多细节将在下面章节中讨论（见*过渡和可持续问题*）。

出版商通常选择采用这种模式的期刊都不希望其影响到销售的领域，例如那些增长迅速、竞争力较强的研究领域。

Laakso and Björk 在 2012 年发布的一份调查中列出了 492 种使用该模式的期刊，在 2011 年共出版文章 111312 篇，大约有 78% 的文章在 12 个月内实现 OA，85% 在 24 个月内开放。延迟 OA 的期刊平均引用率要比封闭订阅期刊高出 2 倍，反映出这些期刊中很多都是各领域中领先的期刊。作者做出总结：延迟 OA 期刊构成了开放可利用的学术期刊文献的重要组成部分，在纯文章的数量包括很大比例的影响力期刊方面都是如此(Laakso & Björk, 2013)。

3.5 通过自存档实现 OA（绿色 OA）

OA 的绿色路线是指通过自存档，使第二阶段的已接收稿件能够立即或者延迟获取。自存档没有独立的商业模式，因此必须依赖于第二阶段的文章版本可免费获取并不影响第三阶段实施（例如期刊订阅）的假设之上。这个假设将在下文中讨论（见*过渡和可持续问题*）。

作者（或其代表）将文章存储在一个开放的知识库或者其他开放的网络空间上。资源库可以是由作者所在机构（通常是大学）运行的机构知识库或者以主题为基础的中央知识库（例如生物医学领域的 PubMed Central）。另一种非常重要的方式是文章被作者存放在他们所谓的主页上（通常是机构服务器上非结构化的空间）。

这三种主要路线构成的可利用的“绿色”文章的比例并不能精确的衡量。例如，Björk, Laakso, Welling, & Paetau（2014）报道的三份调查估计大体的不同的比例如下：

⁶³ <http://highwire.stanford.edu/lists/freeart.dtl>

- 主页等：27%~49%（“绿色”文章）
- 机构知识库：19%~44%
- 主题知识库：29%~43%

造成评价结果差异的一个原因是调查时间窗口的选择。例如，据估计超过一半的绿色期刊要在发表至少一年以后才可以上传（这实际上也降低了 OA 期刊读者的受益）(Björk 2014)。

Taylor & Francis 在 2014 年的调查中显示 23%的作者表示他们将最近出版的文章储存在机构知识库中，有另外 23%的作者储存在个人主页或部门网站上，12%的作者存在学科知识库中，52%的作者没有储存。除去没有储存的作者，大致相当于机构知识库和个人/部门主页各占 40%，学科知识库占 20%(Taylor & Francis 2014)。然而这项数据不太可能精确地反映可获取文章的比例分布，因为有许多作者也许不知道有些文章被出版商直接储存了。

如果需要更多细节可以查看 Björk et al., 2014 发表的综述。

OpenDOAR 网站⁶⁴将储存划分为以下几个类别（数据截止到 2014 年 12 月）

- 机构：2257（83%）
- 学科：296（11%）
- 整合：98（4%）
- 政府：77（3%）

OA 知识库注册(ROAR)报道了总数为 3914 的知识库，其中 723 个在美国，249 个在英国（数据截止到 2014 年 12 月）；2603 个是机构知识库，259 个为“研究交叉机构知识库”（与 OpenDOAR 的学科分类相似）。

对于机构和主题知识库宽泛的分类掩盖了在范围、功能和费用方面的巨大差异。例如，将 arXiv, RePEc 和 PubMed Central 归为同类是有误导性的。arXiv 主要包含作者已经被接收的文章；RePEc 实际上是一个有超过 1600 个知识库的索引服务；而 PMC 是一个高度集中的数据库。PMC 的确可以被称作一个“真正的电子图书馆”：它的功能包括将不同的输入格式转化为规整的 XML，对文本转换为电子版本过程中发生的结构、内容和一致性错误进行修正，为待下载的文章提供转换处理打印清晰的 PDF 版本。

⁶⁴ <http://www.opendoar.org>

知识库包含很多种内容类型，不仅仅是期刊论文：学位论文、书籍章节、工作报告、会议论文等也非常常见（见图 26）。

现在绿色文章席卷机构和主题知识库及个人主页，但是有最大收集量，最高显示度和利用率的依旧是主题知识库(如 Björk et al., 2014), (Björk et al., 2010); 另见入库的网页排名, <http://repositories.webometrics.info>)。Romary & Armbruster (Romary & Armbruster, 2009)认为中央资源库（未必是主题的）有巨大的优势：第一，在推动存储上，资助者的强制比机构更加有效，中央知识库享受单一的基础设施和大量知识库的最好服务，增加了资料收集的价值；第二，他们的调查显示机构知识库运营起来很麻烦，相比中央知识库，不太可能实现一个高水平的服务。然而，机构知识库的基础设施建设已经被加强，他们正在变成大学图书馆期待的服务之一，大部分机构现在都有一个正在使用的知识库(Björk et al., 2014)。

最近几年知识库数量大幅度增长，主要是机构知识库的发展，OpenDOAR 在 2006 年记录了 800 个，到 2014 年增长到 2260 个。它们联合记录总数超过 1450 万，年增长率大约在 35%，当然这些记录并不都是全文，但是比例是在增加的。

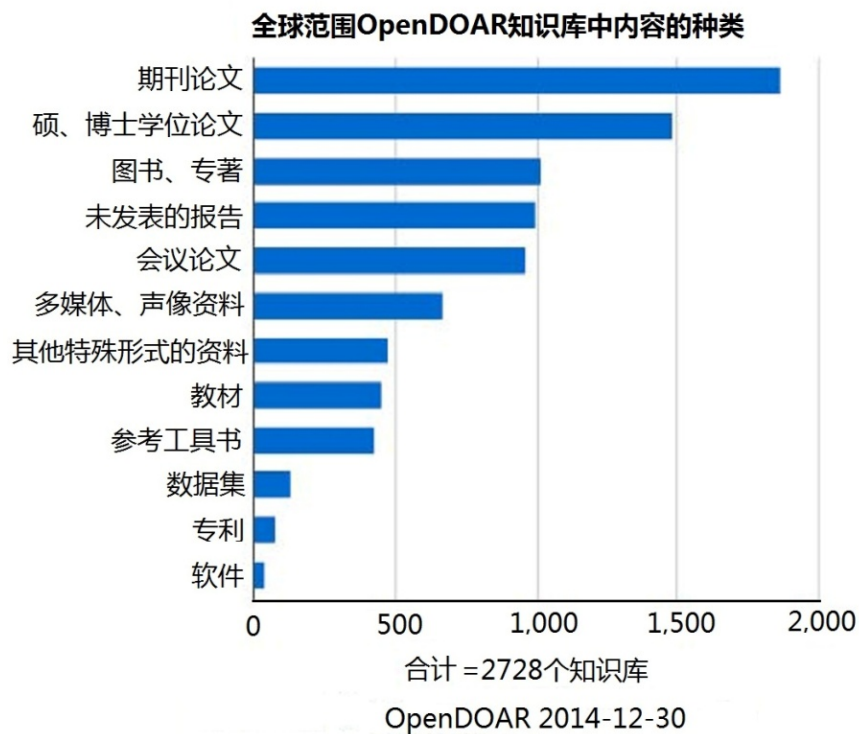


图 26 资源内容类型类型（来源：OpenDOAR）

PEER 基线给出了机构知识库增加的几个原因如下(Fry et al., 2009):

- 对学术型出版物的逐渐开放使用
- 增强的显示度（还有可能的使用和引用）
- 展示科研机构研究成果
- 越来越多的可利用的公众基金（英国的 JISC，欧洲的 DRIVER 项目基金）
- 教育部门越来越激烈的竞争

也许不出所料，新的主题知识库的启动将会越来越少（见下文的 *预印本知识库*）。2014 年 Pinfield 曾阐述过知识库的发展(Pinfield et al., 2014)。

机构知识库

机构知识库是一个收集和保存机构（特别是科研机构）知识成果的电子版本的在线数据库。

对一所大学来说，机构知识库的内容包括研究性期刊论文（例如作者的原稿和录用稿）、学位论文的电子版本，也可能包括常规学术活动中产生的数字资产，例如管理文件、课程笔记或者学习目标。

建立机构知识库的两个主要目标是：

- 为机构的研究成果产出提供 OA
- 储存和维护其他机构的数字资产，包括未公开出版的或容易丢失的（灰色）文献（例如学位论文和技术报告等）。

高校也可通过展示他们的科研成果产出而获益。

机构知识库起源于 21 世纪初期，始于 2002 年在 MIT 创建的 DSpace 和稍早些时候 Eprints software 在 Southampton 的发展。

机构知识库软件使用了一种能让 OAster 或者 Google Scholar 等特殊搜索引擎获得论文元数据的技术标准。尽管这种分散性的搜索没有集中化数据库（例如使用可控性关键词或分类法的 PubMed）那样强大，但是不管文章存在哪个机构知识库，用户都能相对轻松地找到感兴趣的文章。

尽管他们提供服务的复杂性有很大差异，但是机构知识库的数量一直在迅速增加。

作者在他们的机构知识库储存的文章的数量也在增加，最开始比较缓慢，不

过总的储存量现在以每年大约 35% 的速度增加。许多机构知识库（也许不包括荷兰）依然未被作者充分利用(e.g. see Björk et al., 2014; Salo 2008; Albanese 2009), 不过也有一些明显的例外（昆士兰理工大学知识库已经提供了超过一千万次下载）。（Eprints 在 2012 年列出的 1885 个机构知识库的文章总数量为 980 万，平均值为 5220，不过这些包含所有类型的记录，包括从其他来源输入的书籍记录。在一小部分成功的大型机构数据库和一长串小知识库共存的情况下，文章数量呈现偏态分布。）目前来看，大多数的作者对于自存档的潜在好处要么毫无所知要么漠不关心（见 Wallace 2012, 以及下文的 *PEER 项目*）。Stevan Harnad 估计通过倡导和劝说的方式使得作者自愿存储文章电子版会有一个上限，大约能够达到合格文章数量的 15%；采取机构强制的手段能提高存储率。

机构知识库的未来并不确定，在两派之间存在持续争议，一些人将机构知识库主要作为大学数字基础设施的一部分，或许在管理灰色文献、研究数据和其他机构内容中起到重要的作用；另一些（例如加利福尼亚大学的 eScholarship 知识库）则主要看重其在学术交流和出版方面起到的作用(Albanese 2009), UK Finch Group 更加赞同前一种观点(Finch Working Group 2012)。（另见：图书馆出版，数据密集型科学）

以主题为基础的知识库

以主题为主的中央知识库比机构知识库出现的时间更久远，Björk 在 2013 年回顾了主题知识库的状态，他认为主题知识库首次出现时迎合了强大的市场需求，之后由于互联网搜索引擎的发展、机构知识库的迅速增加和期刊出版商 OA 政策的紧缩，主题知识库的发展似乎受到了限制(Björk 2014)。然而领先的主题知识库确实看起来很健康，如下文案例所示。

arXiv

arXiv 是最早的主题知识库之一，由 Paul Ginsparg 于 1991 年在 Los Alamos 创立，现由 Cornell 图书馆管理，arXiv⁶⁵（在时间上早于万维网）设计目的是提高分享高能物理界文章预印本的现行惯例的效率和有效性。也许是因为它建立在已有的“预印文化”之上，也因为高能物理学家是电子网络的早期采用者，这个团体积极热情地采用 arXiv，以至于这个领域几乎所有的文章都是自存档的，最起码也有作者的原稿。虽然综合覆盖领域比较小，但现在 arXiv 已经把它覆盖领域扩大到其他的（并非所有的）物理学、数学、计算机科学和定量生物学等领域。当前 arXiv 拥有 1 百万份预印本。（见下文 *OA 的最近进展* 中有关 arXiv 资助模式的讨论。）

⁶⁵ <http://www.arxiv.org>

随着 arXiv 的发展(图 27),它的寄主机机构(现在是 Cornell,最开始是 LANL)努力在为资金需求努力。在 2012 年 8 月, arXiv 宣布了一项新的覆盖 2013~2017 年的资助模式⁶⁶, 资金由三种收入来源组成: 由 Cornell 图书馆提供的现款和实物支持; Simons Foundation 提供的资金拨款; 成员机构的集资等, 即: 高能物理的机构自愿同意为其支出做出贡献。Cornell 希望每年通过成员机构筹集 33 万美元(占总运行成本的 36%), 126 个机构平均每个每年支付 1500~3000 美元(以机构的大小区分费用高低)。2014 年项目共花费 88.6 万美元, 可计算出每篇新增增加的文章少于 10 美元 (Van Noorden 2014a)。

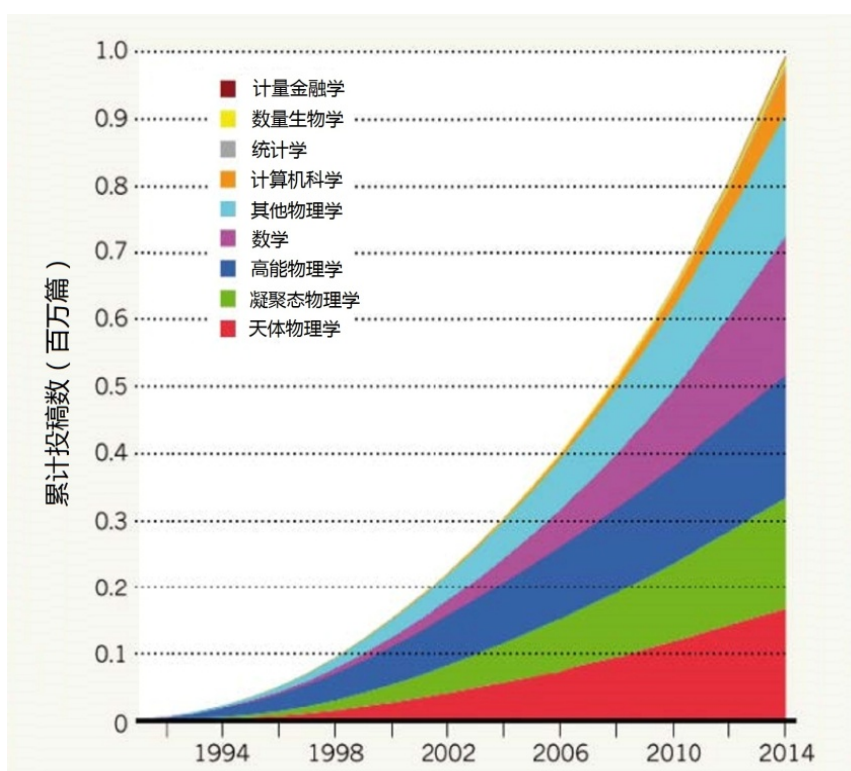


图 27 arXiv 的发展，物理和数学依然是最重要的学科

(来源: Nature News, Van Noorden 2014a)

RePEc

RePEc (经济学研究论文)⁶⁷ 是另一个早期的的知识库, 也是建立在一种已经存在的经济学界文化之上, 这种文化指的是分享还未发表的被称为工作手稿的文章。RePEc 现在拥有来自 1800 本期刊和 3800 份工作手稿系列的共计 140 万份

⁶⁶ 见 <http://arxiv.org/help/support>

⁶⁷ <http://repec.org>

研究资料，它有以下几点与 arXiv 有所区别：它是一个分散的（基于志愿者的）文献数据库而非一个集中的知识库，整合了 1600 个档案的内容；不包括全文文献，期刊论文的记录仅包括摘要和文献目录信息，不过许多文献都有全文链接，包括链接到出版社网站的完整版。还有一点区别是，出版社与 RePEc 合作来储存他们期刊论文的文献记录。在很多方面，RePEc 更像是一个免费的文献数据库而非知识库，促进了各种各样利用其数据的服务的产生。

PubMed Central

出版商现在非常感兴趣的一个基于主题的知识库是 PubMed Central(PMC)。PMC 是美国国立卫生研究院 (NIH) 的一个项目，而非发源于学科团体自身的自愿行动。其创建于初期的包括 Medline 在内的文献数据库 PubMed，并加入了文献全文。PMC 是 NIH 和其他生物医学研究基金资助机构指定的受资助研究人员的仓储库。PMC 得到了许多出版商的支持，他们自愿地代表他们的作者将其稿件版本（第二阶段）或者在某些情况下的全文（第三阶段）存放在 PMC，这些文章能即时开放（全开放期刊）或者经过一段保密期后开放（延迟开放期刊）。PMC 还与出版商合作，将以前出版的纸质内容数字化，这些文章必须免费获取。从 2004 年开始，PMC 就开始接收作者手稿归档来支持上面讨论的 NIH 资助政策。在撰写本报告的时间点 PMC 拥有 330 万篇研究文章，其中 94.5 万篇属于 OA 范畴（其他的可免费获取但不是 PMC 所采用意义上的 OA，在 Creative Commons 许可下允许重新传播和再利用）。

Europe PubMed Central⁶⁸ 基于 PubMed Central 建立并增加了一些附加服务和功能(McEntyre et al., 2011)。这是 PMC 国际合作的一部分，此外还有 PMC Canada。

新预印本知识库

2013 年生命科学领域有两个预印本档案库启动，传统的观点认为这个领域（不像物理或者金融）没有优先预印本文化（例如 NPG 曾在 2007 年建立了 Nature Precedings，在 2012 年就结束了），此举多少有些对抗传统观点的意思。

PeerJ Preprints 是 PeerJ 生态系统的一部分，支持作者从早期草稿到同行评审再到出版及其以后的工作流程。定价执行 PeerJ 的免费增值模式：作者要存储必须成为会员，免费会员每年只能提交一份预印本，要获得更多的使用权需要成为付费会员。经过 2013 年中的缓慢启动过程，2014 年的提交率明显上升，一年共储存 729 份预印本。

BioRxiv 在 2013 年底建立，比 PeerJ 稍晚，不过在文章提交数量上早已超过

⁶⁸ <http://europepmc.org>

它，在 2014 年存储了 918 份预印本。从名字上可以看出，它模仿了物理学界的 arXiv。它提供了一些附加功能，例如公众评论、补充信息或者链接到外部数据库等。

生命科学界其他可以相比的服务，比如 *F1000Research* 期刊，在提交论文时立即发布，结合了预印服务和公开同行评议。Figshare 数据档案库拥有一套自由的存储政策，一份调查显示一些用户把它当作分享预印本的平台。

其他类型的知识库

SSRN

SSRN，社会科学研究网⁶⁹，可以归类为主题知识库，不过它还拥有很多不同点，使其成为令人关注的案例研究。SSRN 在 1992 年建立，这个时间反映出经济学界长期存在的预印本文化（类似于物理学界）。尽管依然依赖于志愿者的工作（例如大约 1000 位志愿者担当咨询编辑、编辑和网络主管等），但从组织上来说，现在它是一个预算超过 1 百万美元的公司。它当前拥有 58 万篇文章的摘要和 48 万篇全文文章。迄今为止它的文献全文的 PDF 下载已经超过 8000 万次，现在每月大约有 100 万次下载。

SSRN 在其领域中与大部分出版商合作，使其能够在储存能力之外提供索引服务（类似于 PubMed 和 PubMed Central 的各自角色）。

SciELO

SciELO（科技在线电子图书馆）不是一个传统的知识库，而是一个文献数据库和 OA 期刊数字图书馆。SciELO 的模式用于发展中国家的合作电子出版。SciELO 于 1997 年创建于巴西，目前在 12 个国家运行，还有 3 个国家正在“发展”中。在 2014 年中的时候，它已经拥有 1161 份 OA 期刊，涵盖将近 50 万篇文章。SciELO 在 2013 年宣布与 Thomson Reuters 达成协议，将 SciELO 引文索引整合到 Web of Science 中。

Redalyc

Redalyc⁷⁰ (Red de Revistas Científicas de América Latina y El Caribe, España y Portugal) 是一个文献数据库和 OA 期刊的汇集，专门收集拉丁美洲地区的学术成果。创立于 2002 年，现在涵盖 930 本期刊和 36.5 万篇文章。它的服务包括文献计量指标、社交科学网络、期刊汇集和使用计量。

⁶⁹ <http://www.ssrn.com/>

⁷⁰ <http://www.redalyc.org/>

其他 OA 整合和索引服务

大量提高 OA 内容的发掘和利用的服务已经发展起来，表 11 列出了一些实例的概览

表 11 OA 整合和索引

服务	创立时间	描述	收录	URL
BASE(Bielefeld 学术搜索引擎)	2009	使用 OAI-PMH 协议的 OA 网络资源搜索引擎，全文和元数据搜索	超过 3200 处来源的 650 万多的文档	http://www.basesearch.net/about/en/
DOAJ	2002 ; 2013	针对于 DOAJ 注册期刊范围内的文章层面的搜索	来自 6032 本期刊的 180 万篇文章（共 10135 本期刊）	http://doaj.org
OAister	2002	通过 OAI-PMH 获取的 OA 档案的目录，通过 OCLC's WorldCat 可搜索到的	来自 1500 多位贡献者的超过 3000 万份记录	http://www.oclc.org/oaister.en.html
OpenDOAR	2006	允许通过谷歌自定义搜索查找在 OpenDOAR 注册的知识库	1000 万~1500 万?	http://www.opendoar.org/search.php
Paperity	2014	启动目标是做一个“多学科的同行评议 OA 期刊和论文（全 OA 和混合 OA）的聚合器”，使用基于 Redex 的专有网络爬虫/采集	来自 2200 本期刊的大于 39 万篇开放文章	http://paperity.org/
ScienceOpen	2014	结合预印本知识库、开放同行评议、出版和聚集的多功能平台	大于 140 万篇文章	https://www.scienceopen.com
Scilit	2013	来自 OA 出版商 MDPI 的试验性 OA 搜索引擎	大约 185 万篇文章	http://www.scilit.net/

自存档的政策和要求

在2004年，美国国立卫生研究院出台了一个政策来鼓励受资助的研究人员在知识库PubMed Central中储存被接收的稿件。然而遵循这项自愿政策的人特别少（<5%），因此NIH改变政策，强制要求研究者进行存储，从2008年4月生效。NIH新政策允许作者在出版后最多延迟12个月进行储存。

尽管并非第一次，NIH还是由于其研究经费预算（300亿美元）的规模引来了

众多关注。类似的政策现在非常普遍，SHERPA/Juliet网站⁷¹列出了140项研究资金(至2014年12月)，其中109项包含存储政策(从鼓励到要求OA存储各有不同)，包括所有的英国研究委员会、惠康基金会、霍华德-休斯医学研究所、欧洲研究委员会、德国科学基金会、德国弗劳恩霍夫研究院、澳大利亚研究理事会。(因为SHERPA/Juliet统计重点围绕UK，所以可能低估全球范围内的政策数量。)保密期为6到12个月，或者在某些情况下，在遵守出版政策的基础上尽快存储。

除了研究资助者以外，一些承担机构也采取了类似的政策。Eprints/ROARMAP网站⁷²在2014年12月记录了415个完整机构和65个附属机构的政策命令。另一份机构和资助者要求的名录MELIBEA⁷³列出了349个机构和150个资助者的政策，明确表示采取强制的机构包括Harvard、MIT、UCL、ETH Zurich、Fraunhofer-Gesellschaft和University of California。

这项政策最开始的影响非常小：一般来说作者没有自存档的积极性(例如，见PEER项目调查的讨论)，在缺乏监管和强制的情况下，一般不会优先考虑OA，尤其是对于科研机构的要求是如此，不过甚至连高知名度资助机构的要求至今也没有大范围地遵循：2012年中期统计NIH为75%，惠康基金会为55%(这两种情况都包含出版商自身重要的协助)。然而这种情况由于资助者更加强调执行强制要求的优先性而得到改变：例如，惠康基金会在2012年宣布加紧OA政策，包括对不遵守存储政策的研究者的制裁⁷⁴。尽管如此，这项政策的有效沟通也需要一定的时间：例如，在2014年Taylor & Francis作者调查中，仅有30%的调查对象称他们理解RCUK政策，许多人“似乎不确定这项政策是否适用于他们，超过半数(55%)的研究者并不能确定他们将来的文章是否需要遵循政策来发表”(Taylor & Francis 2014)。

最近较有影响力的要求作者存储文章的政策来自英国研究委员会(RCUK，见下文)、比尔和梅林达·盖茨基金会(Bill and Melinda Gates Foundation)⁷⁵(以其严格的要求而知名：文章出版后必须立即实行OA，没有保密期，利用CC-BY或等效的许可)。

在美国，科学技术政策局一份备忘录要求联邦政府保证公共资金资助的研究

⁷¹ <http://www.sherpa.ac.uk/juliet/>

⁷² <http://www.eprints.org/openaccess/policysignup/>

⁷³ <http://www.accesoabierto.net/politicas/default.php>. 除了列出政策外，MELIBEA还计算了一个计量，即“估计的OA百分比”

⁷⁴ <http://www.wellcome.ac.uk/News/Media-office/Press-releases/2012/WTVM055745.htm>

⁷⁵ <http://www.gatesfoundation.org/How-We-Work/General-Information/Open-Access-Policy>

成果能够OA，这看起来扩展了NIH政策在其他领域的影响，尽管该政策正在进一步推进中。

科学技术政策局备忘录

在美国，OA发展的一个关键点是白宫科技政策办公室在2013年2月发布了一份关于联邦政府资助研究公众获取的备忘录(OSTP 2013)。其中明确指出所有的研究预算超过1亿美元的代理机构必须将其研究成果——尤其是，“所有发表在同行评议的学术型出版物上的，基于联邦政府直接资助的研究项目的成果”——必须在发表之后最多12个月内实现免费获取。

这份备忘录没有明确如何去实现这些目标，而是要求代理机构制定计划来满足这些要求。于是出现了4种类型的解决方案：代理机构可以创建他们自己的专用知识库；NIH的PubMed Central知识库可以扩充或者分支出一个需付费的内部代理服务机构；SHARE，一个由图书馆管理的系统，基于联系高等教育机构的基础设施（包括机构知识库）；CHORUS，由出版商主导，利用现存的出版平台和技术。

CHORUS

CHORUS—美国开放研究交流中心，是由一批出版商和服务提供商创建的非营利公私合营的公司，为基金资助公司提供使其满足OSTP要求的服务⁷⁶（现在由一个非营利组织CHOR公司管理）。

CHORUS充当一个信息桥梁的角色，能直接在出版商平台上链接到基于联邦政府资助研究项目的开放期刊论文。CHORUS提供5个核心功能：识别、发现、获取、保存、使文章符合政策要求。其依靠于一套由CrossRef管理的标准方法FundRef来判断已发表的学术研究的资金来源。当文章被出版商接收时，这套系统能识别出由联邦政府代理机构资助的论文并将其通过CHORUS开放。因此这套系统不需要新的重大基础设施，只需要在原有的出版商平台上提供信息和访问界面，而且发展和管理费用大部分落在出版商头上。除了获取层界面之外，CHORUS允许代理机构为他们的内容创建搜索门户，为所有的利益相关者提供“仪表盘”来监视公共获取的遵守情况。它将提供开放APIs，允许任何人对联邦资助的研究创建概览和搜索工具。CHORUS建议通过CrossRef的TDM工具和LicenseRef项目来支持文本和数据挖掘（见*文本和数据挖掘*）。

CHORUS还通过与Portico的合作来提供保存功能（见*长期保存*），这样一来，联邦资助论文长期的公众获取将不再依赖于CHORUS或者出版商的持久存在。

⁷⁶ <http://www.chorusaccess.org/>

考虑加入CHORUS的出版商必须决定是否开放他们稿件的最终版本或者录用的稿件。

在撰写本文的时候，美国能源部已经选择了CHORUS作为遵守OSTP要求的方法。最大的代理机构NIH则致力于开发自己的PubMed Central平台，而其余的代理机构还未公开他们的计划。

有OA评论家指出，相比CHORUS，PubMed Central具有更强大的功能（包括全文搜索和复杂的接口和发现工具），并视其为一种保护价值和出版商平台优势的方式（如Eisen 2013）。

SHARE

SHARE⁷⁷（科研共享访问生态系统）是另一种可供选择的方案，由美国研究图书馆协会（ARL）、美国大学联合会（AAU）、美国公立和Land-grant大学协会（APLU）倡议合作，斯隆基金会提供启动资金。SHARE并不主要专注于满足OSTP的要求，而是更加广泛地强调对科研成果的保存、获取、再利用的需要。

服务框架计划由四层构成：通知服务；注册；搜索；挖掘和再利用。COAS（开放科学中心）正在开发一种通知服务的原型：它将借助于大范围的信息来源（包括机构和主题知识库、出版商、CrossRef等）同时使用推送协议和获取服务来搜集关于“开放事件”的信息（例如文章的发表和研究数据的发布）。这些信息将作为一套元数据提供给利益相关者，例如资金代理机构、研究室、机构和学科知识库。像CHORUS一样，SHARE将不会储存研究成果副本，而是采取对内容的注册，使其之后能支持搜索界面。

SHARE将不仅提供期刊论文的获取，而且打算涵盖在其负责范围内的研究数据(Lynch 2014)。

RCUK Policy

英国研究委员会政策⁷⁸在2012年提出，自2013年4月生效，它创建于2005年的一份早期政策之上，明显受到Finch Report影响。与大部分其他资助政策不同，这项政策认可金色OA模式的优势大于绿色OA模式，其他所有的都同等且应有基金资助。这项政策的总结如图28所示。

为了支付文章的处理费用，高校和有资格的科研机构将收到一份总的补助款。RCUK负责依据经验审查其执行政策的过程；高校汇报的第一年(2013/14)执行政

⁷⁷ <http://www.arl.org/focus-areas/shared-access-research-ecosystem-share>

⁷⁸ <http://www.rcuk.ac.uk/research/openaccess/policy/>

策的符合率在35%（Hull）和97%~89%（LSE; Huddersfield）间变化⁷⁹。

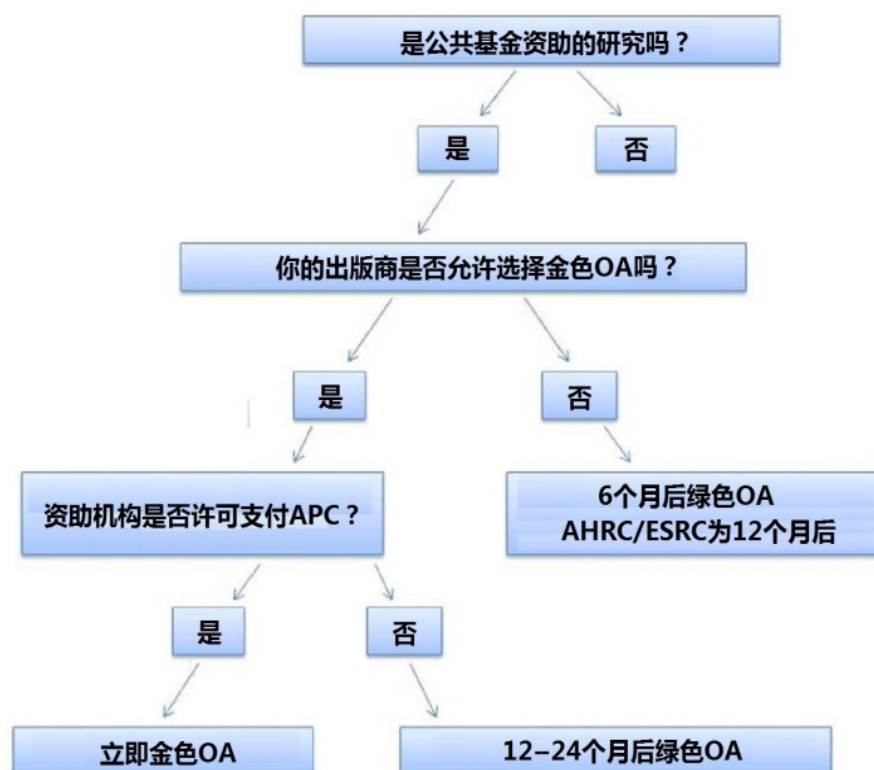


图28 RCUK OA 政策总结

英国科研机构为了遵守RCUK和（更严格的）HEFCE政策⁸⁰而引起产生的费用由研究委员会估算，2013/14年的费用为920万英镑，其中金色APC仅占不到10%（见图29的分类）。从高额的管理费用可以明显看出在系统范围内还没有有效的支付和APC追踪系统⁸¹。Jisc在对RCUK OA政策的评论中提出了类似的观点：“工作流和信息流意味着需要在利益相关者和系统之间加入新的成分，许多部分不能有效地结合起来。存储系统、技术标准的缺乏、不同需求之间的冲突等意味着能

⁷⁹ <http://e2eoa.org/2014/10/01/open-access-reports-to-rcuk-collected-here/>

⁸⁰ HEFCE 要求所有材料经验收存放于仓储，以便有资格将来提交 REF 评估—
<http://www.hefce.ac.uk/whatwedo/rsrch/rinfrastruct/oa/policy/>

⁸¹ 出版商的内部系统正变得更有效，例如通过诸如 CCC's RightsLink (www.rightslink.com/) 的外包服务，但这并不一定影响到付款人的最终效率。另一个中介是开放存取的关键（OAK），其系统的目的旨在提高机构和出版商的效率（<https://www.openaccesskey.com>）

够巩固一个可扩展性的APC环境的数据经常缺乏、不完整、错误、不明确或者是被卡在某些地方。”(Jisc 2014)



图29 RCUK OA政策的税务执行费用

出版商的自存档政策

大部分出版商有相当开明的政策，允许作者将他们自己的文章存储在网络，尽管这些政策最初在储存不够系统化的条件下被引进。随着存储和搜索工具越来越系统化，存储政策在个人主页（更自由的政策）、机构知识库和主题知识库（更严格的要求，反映出对订阅更大威胁）区分开来。政策还将根据特殊资助者的要求而改变；例如极少数的出版商不允许在PubMed Central上存储。（另见科学社交网络）

由SHERPA/RoMEO项目维护的一份出版商政策的数据库⁸²，包含了1766个出

⁸² <http://www.sherpa.ac.uk/romeo/>

版商（截止2014年12月）：

- 34%允许存储作者的原稿和录用稿
- 34%允许存储录用稿
- 7%允许储存作者原稿
- 25%不正式支持存储

因此有75%的出版商允许某些形式的存储。因为较大的出版商通常允许某些形式的存档，所以自存档期刊的比例将会更高一些。

一些出版商还允许作者存储出版的最终版本，这种模式比较少见（根据Inger & Gardner 2013, 表7, 大约三分之一的出版商允许将最终版本文章存储在机构或主题知识库，在小型出版社更常见一些，大型出版社显然不太可能在2005年之前允许这种存储）。一些出版社会添加附文，例如要求一个从存档稿件到出版商最终在线版本的链接。考虑到保护订阅，出版商一般会有一个保密时段（例如出版后一段时间内不允许自存档）。

知识库成本

有大量的关于引进和管理知识库成本的报道。最原始的机构知识库之一，MIT的DSpace估算了它的年度运行成本为28.5万美元（职员225美元；运营成本2.5万美元；3.5万美元）(MIT 2003)。一份对ARL的调查(Bailey 2006)发现知识库启动资金在8000美元到180万美元范围内变化，平均值为182550美元，中值为45000美元。持续运营的预算在8600~50万美元的范围内变化，平均值为113543美元，中值为41750美元。

Houghton估计的高等教育机构知识库年费用为100000英镑（包括高级管理层在政策和宣传活动方面的时间因素）(Houghton et al., 2009)。除此以外，学者存储他们文章所花的时间成本估计为每次存储的成本为10英镑，将英国作为一个整体的成本为160万英镑（全球为1500英镑），这样算起来每篇文章平均总成本为20英镑。

一份2007年关于美国机构知识库的报道发现(Rieh, Markey, St Jean, Yakel, & Kim, 2007)，几乎所有情况下知识库资金均来源于图书馆，并且没有额外的预算供给（即：资金出自日常的图书馆运行费用）。报告没有提供预算总金额，不过提供了按类型分类的支出细节：职员约为37%，供应商费用38%，硬件约10%，软件约2.5%，硬件/软件维护和备份约12.5%。

最近，PEER项目发现很难获取机构知识库的启动和运行费用的数据，原因

是对平台启动的投资、软件升级和知识库维护的花费被看作是沉没成本，没有分别核算，费用覆盖多个部门。项目能够估算出技术人员支持的成本；它报道了每篇参考文献的成本为2~50欧元，每篇全文文章的成本为2.5~53.2欧元。成本范围分布较广反映知识库的规模的有效扩展，也就是较低的成本指的是较大的知识库，反之亦然。

英国知识库支持网络⁸³（2006~2013）采用了直观的硬件成本为2000~150000英镑，并建议20000英镑的启动资金应该可以处理50~100 000篇文章。该网站列出了其他领域的启动和运行成本（主要是工时），不过并没有迹象表明这些花费可能所处的水平。

大型的学科知识库总体来说自然要更加昂贵，不过规模经济可能导致较低的个体成本。例如康奈尔大学图书馆估算的（高度自动化的）arXiv 2014年运行成本为88.6万美元，平均每篇新文章存储花费不足10美元。美国国立卫生研究院曾估算出管理自存档政策的花费是400万美元⁸⁴，每年大约有9万篇文章，因此每篇文章大约为44美元。然而这只是PMC总花费的一小部分，反映的仅仅是收集、处理、将NIH资助的稿件转化为PMC存档模式的费用。

文章的多个版本

自存档的广泛使用带来的一个潜在问题是读者（或其他人员，例如知识库管理员）将获取到文章的多个版本。

作者一般将存储其原稿或者已录用的稿件，或者在某些情况下两者都存储（较少数出版商允许存储最终版本(version of record)）。大多数资助机构和机构要求至少将已录用稿件存档。作者很有可能将不同版本的文章存储到多个知识库中（例如一个机构知识库和一个中央知识库）。

大型的知识库（包含机构和主题知识库）正在与出版商合作，提供从存储版本到最终版本的链接。CrossMark服务能够从其他版本中区分出补体最终版本，将显示出其价值。

3.6 其他的 OA 衍生模式

Willinsky (Willinsky 2003)识别出9种不同的OA衍生模式。除去上面列出的和自存档路线，他包含了“双重模式”（打印订阅加OA在线版本）；“人均”（基于国家人均收入将文章对国家开放—见以上关于发展中国家使获取的讨论）；“摘要”

⁸³ <http://www.rsp.ac.uk>

⁸⁴ http://publicaccess.nih.gov/Collins_reply_to_Pitts121611.pdf

（开放获取期刊目录和摘要—大多数出版商提供）；“合作”（机构会员支持OA期刊—与图书馆合作伙伴模式非常类似，在表6中列出）。

一种不太常见的混合开放获取的衍生模式是学会会员提交的文章发表在该学会的期刊上可立即实现OA⁸⁵。

最后一个可能被提及的衍生模式是一种虚假的OA，许多调查显示学者将OA与由他们的机构提供的免费在线访问使用相混淆(如 Biosciences Federation 2008)。根据对作者调查的反馈意见，OA期刊可能会受到这种混淆现状的影响。

3.7 SCOAP3

SCOAP3⁸⁶（国际高能物理开放出版资助联盟）欧洲核子研究委员会CERN发起的一个雄心勃勃的计划，目的是将发表高能物理学（HEP）论文的所有期刊转化为可持续的OA形式。在高能物理领域，每年大约有5000~7000篇文章发表，其中80%集中在来自4个出版社的6本期刊中。实际上几乎所有这些文章的原稿和/或最终稿都会在出版之前出现在arXiv上，因此这些期刊正在失去（或者已经失去了）传播功能。剩余的关键功能被认为是高质量的同行评议和担当“记录保管者”的角色。SCOAP3估计的高能物理领域全球范围内期刊出版费用大约为1300万美元（5000—7000篇文章，每篇2000美元）。

该理念是构建一个联盟，包含国家级HEP资助单位、图书馆和同意为该项目作出贡献的联盟机构（通过重新调整订阅的方式），这些都基于每个国家贡献的HEP文章比例。SCOAP3利用这些资金使出版商出版同样的期刊，不过在新的OA模式下，集中资金替代了需要向作者收取的费用。

SCOAP3在2012年9月经过第一次招标确定了来自7个出版社的12本期刊参与⁸⁷。这些期刊在2011年发表了6600篇文章，大多数都是高质量同行评议的文献。项目会存在一些遗漏，特别是美国物理学会的*Physical Review Letters*（物理评论快报）的投标因价格原因被拒绝。

SCOAP知识库在2014年初建立⁸⁸，至2014年底已经有4300篇文章存入。

由SCOAP3资助的文章将实现永久OA，遵循CC-BY协议，同时出版商会相

⁸⁵ 这方面的一个例子是美国植物生物学会的期刊*Plant Physiology*，见
<http://www.plantphysiol.org/cgi/content/full/142/1/5>

⁸⁶ <http://scoap3.org/>

⁸⁷ <http://scoap3.org/news/news94.html>

⁸⁸ <http://repo.scoap3.org>

应地降低订阅费用。

SCOAP3建议他们的项目可以看作其他领域的试点，HEP是一个相当不寻常的领域，很高比例的文章集中在很少数量的期刊上并且有很大一部分早已通过自存档实现了OA。天体物理学、核子物理以及其他一些理论物理有类似的特点，不过还难以确定该模式是否适用于诸如生物医学这样的更加多样化的出版生态。

3.8 学术书籍的开放获取

开放获取运动最初的焦点集中在获取期刊的研究性论文。此后，人们对期刊论文以外其他类型的文献越来越感兴趣，如教育资源和学术书籍，特别是学术专著。Finch 报告（Finch Working Group, 2012）建议相关的各方应该共同努力，在学术专著的开放获取出版方面进行更多的尝试。

OAPEN 图书馆是一个免费获取人文与社会科学相关学术书籍的在线图书馆和出版平台，它是从早期欧洲网络计划中的开放获取出版发展而来的⁸⁹。欧洲网络计划是一个在人文与社会科学领域发展和实施学术书籍的可持续开放获取出版模式的协作项目，起初是由欧盟共同出资的 eContentplus 项目的一部分。OAPEN 在 2012 年 4 月创建了《开放获取图书目录》⁹⁰；截至 2014 年 12 月，目录中已经包含 79 家出版商和超过 2482 种 OA 图书，并且每年以 40% 的速率增加。DOAB 许多出版商都是大学出版社，但是也有一些是商业型出版商，包括 Bloomsbury Academic, Brill, De Gruyter, Palgrave Macmillan, Springer, Taylor & Francis, 和 Ubiquity Press 等。

OAPEN-UK⁹¹ 是 2010~2015 年运行的独立研究项目。它是由 Jisc/AHRC 资助的搜集证据以帮助利益相关方针对人文社科学术专著开放获取做出决策的合作研究项目。

人文社科领域专著可行的开放获取模式不外乎以下两种：延迟 OA，或在提供诸如印刷版、电子书阅读器版本、增强电子书版等高附加值的收费版本的同时提供一个基本的电子版本可供在线免费获取（Milloy 2013; Ferwerda 2014）。由于高额的图书出版费用以及缺少足够的资助，APC 模式非常少见，但是也有一些正在运行（例如 Palgrave Open, Ubiquity Press）。在自然科学领域，一些出版商⁹²采用作者支付出版费的金色 OA 模式出版合著的专著，书中每个章节与期刊中

⁸⁹ <http://www.oapen.org/>

⁹⁰ <http://www.doabooks.org/>

⁹¹ <http://oapen-uk.jiscebooks.org/>

⁹² 例如，Springer (<http://www.springeropen.com/books>); InTech (<http://www.intechopen.com>)

的文章是等价的。

最近探索出了两种创新模式：**Knowledge Unlatched** 图书馆的合作伙伴补贴模式（详见前述“开放获取商业模式”部分中的*图书馆的合作伙伴补贴*）；众筹模式，在这种模式下，由出版商设定一个目标价格，达到该价格后就会开放提供专著的电子版本（例如，开放图书出版商和 **Unglue.it** 的合作商 **De Gruyter**）。

2015年初，加州大学出版社（UCP）建立了一个新的开放获取图书的平台，**Luminos**。它的商业模式将文章处理费用（APC）、图书馆合作补贴以及印刷版销售组合在一起⁹³。基础的出版费用为15000美元，作者需要支付其中的7500美元（有减免）（with waivers available），剩余的由自愿加入图书馆的会员费（每家图书馆每年1000美元）、纸质书的销售收入以及UCP获得的补贴补齐。

3.9 公共获取

为研究提供公共获取的政治优势为，它至少是被 **Finch** 报告承认的，而 **Finch** 报告是美国政府制定政策的主要驱动力（详见“科技政策办公室备忘录”）。在英国，通过公共图书馆不经预约可直接访问的一个试点项目“**Access to Research**”从2013年开始启动，这个项目受到24个顶尖出版商的支持，通过英国80%的图书馆服务提供超过1000万篇文章的获取⁹⁴。一些出版商，尤其是美国的物理学会，也独立提供类似的公共图书馆获取服务。

3.10 全系统和经济视角研究

随着政策制定者对开放获取兴趣的不断增长，目前已经有许多人尝试研究OA对经济的影响，包括对学术交流的全系统影响以及更为广泛（更有争议）的经济影响。

RIN/CEPA 研究

如前面所述，2008年 **RIN** 报告（**RIN, 2008**）估计期刊出版和发行的成本为49亿英镑（不包括非现金的同行评审费用），而出版业和图书馆业的总成本是250亿英镑。然后作者建立了模型将90%的文章转变为需作者付费的OA出版系统，并评估其影响，他们估计这将使整个系统节省成本约5.6亿英镑，其中出版商和图书馆节省的成本几乎相等（与此同时，如果全球出版业向纯电子出版（**electronic-only**）转换，将节省成本约10亿英镑）。图书馆因此将会节省约29亿英镑的订阅费用，但这项节省与实际收取作者的费用等同而相互抵消。对机构

⁹³ <http://www.luminosoa.org/why-oa.php>

⁹⁴ <http://www.accesstoresearch.org.uk>

整体而言将出现成本和收益的不平衡：研究密集型机构将支付比图书馆订阅费更多的出版费，而相反的情况会出现在其他机构。另外，节约的费用不包括出版商、基金会和机构用于管理作者付费的管理成本。

Houghton 报告

在 2009 年 JISC(英国联合信息系统委员会)的报告中(Houghton 等, 2009)，经济学家 John Houghton 估计，下一年英国开放获取出版在全系统范围可节省 2.12 亿英镑，扣除作者支付的 1.72 亿英镑的费用，净节省 0.41 亿英镑（这个数据与 RIN 报告中估计全球范围可节省 5.60 亿英镑的数字大致一致）。其中最大的一部分（1.06 亿英镑）来源于研究过程中的节省，包括研究人员花费在搜索、寻找及获取授权的时间，通过使用更好的开放获取获得更快的同行评审，并且由于由于查找文献更加便捷，减少了作者花费在写作上的时间。据 Houghton 的介绍，基金资助机构很满意将一部分研究经费用于支持开放获取，因为在研究过程中节省的支出会超过支持 OA 的费用。

出版商对以上观点提出异议，认为上述分析低估了当前订阅系统的效率和英国科研人员享有获取的水平；并且，许多成本节省的假设将取决于世界上其他国家是否也采用作者付费或自存储的出版模式。Houghton 模型中使用的许多数字是不准确的估计，而不是源于行业的统计数据。

除了全系统的成本节省外，Houghton 还估计英国公共研发部门通过增加获取可增加价值约 1.70 亿英镑的经济回报。很明显，这个数字更是由推测得到的，因为对当前的获取水平和增加获取的边缘回报率假设难以检测。

迈向开放之路

2011 年英国一个“迈向开放之路”(RIN, 2011c)的研究试图解决以上两项研究的局限性，他们没有采用静态假想经济模型，即假设现行出版模式已完全转换为 OA 模式，而是关注由现行的出版模式动态过渡到开放获取（分别建立金色 OA 和绿色 OA 的模型）和其他途径（如增加许可和交易式获取）共存的多样化模式的过程中可能出现的各种问题。其关键结论是 OA 途径多样化为感兴趣的政策制定者推动文献的开放获取提供了最大的潜力。虽然绿色 OA 能够增加获取，但它有损害现有的基于订阅的出版体系的危险，从而可能破坏出版体系自身的可持续性。从长远来看，金色 OA 是首选的 OA 路线，它具有较好的可持续性，有更高的透明度、较低的准入门槛及更高的收益/成本比的潜力；并且，在平均 APC 不是太高的情况下，金色 OA 可为英国的公共财政和大学节约成本（该研究采用 1450 英镑作为较低 APC 的情况，并且估计了平均 APC 的最高阈值，如果金色 OA 的 APC 高于 2000 英镑则金色 OA 的性价比不佳）。

其他开放获取的经济学研究

其他的一些研究也尝试评估开放获取的经济学影响，包括如下内容：

- Houghton 拓展了他针对丹麦和荷兰的初始报告（包括英国），并且概述了对各个国家的分析比较（Houghton 2009）。
- Houghton 后来也使用了相同的方法来模拟联邦研究成果公众获取提案（FRPAA）在美国的影响（Houghton 2010）。
- 2011 年，日本工业标准委员会（JISC）研究了 OA 为高等教育和学术研究领域的私立部门带来的潜在效益（Parsons, Willis, Holland, 2011）。
- 基于 ESRC 资助的访谈，Dagmara Weckowska 探索了 OA 促进经济革新的潜能（Weckowska 2014）。

经济学研究的局限性

由于研究方法的限制，上述所有研究的发现存在以下局限性：

- 非现金方面：例如，因获取改善而节省研究人员时间的估计是有问题的，因为这种节省无法用金钱来计量，只可假设转化为更高的工作效率。即使这个估算可能是合理的，由于信息技术的采用，也很难辨别由 OA 带来的经济生产力的增长。
- 几个关键变量存在很大的不确定性。
- 特别是 Houghton 和 2011 年 RIN 的报告⁹⁵中经济多元化效应所产生的大量数据可能掩盖其他效应，并且都是基于未经检验的假设。
- 没有一种方法可真实地模拟尽可能多种多样的 OA 实施方案，如不同的国家甚至同一个国家的不同机构之间可能有不同的 OA 政策和实施方式。

正如 Rick Anderson（Anderson 2014c）所描述的那样，一个可能存在的更深一步的问题是，过分的鼓吹、顽固的保守以及理想化的思维往往淹没对开放获取优点的理性认识。

3.11 开放获取的其他进展

eLife 和 *PeerJ*

开放获取出版规模的增长及被主流接受的事实似乎激起了新一轮的探索和

⁹⁵ 这两项研究中使用的是 Solow-Swan 增长模式。

创新(Van Noorden 2012b)。PeerJ 就是一个很好的例子,它在 2012 年末由前 PLOS 和 Mendeley 的工作人员创办,并得到 Tim O'Reilly 风险投资基金支持(SAGE 后来也有注资)。PeerJ 提出了作者会员制的论文发表模式。终生会员费为 99~299 美元,缴纳 299 美元会费的作者可无限发表文章。有多个作者的文章中的每个作者(最多为 12 位)都必须成为交费会员;PubMed 中收录论文的平均作者数为 5~6 个,所以最终累计的发表费可能接近巨型期刊的费用(每篇约 1350 美元)而不是 99 美元。鉴于研究者们在不同的论文中合作署名的变化,该模式类似内置的病毒式营销。在这一点上,PeerJ 的意义不在于其对出版市场的影响(到目前为止 PeerJ 只发表了 730 篇文章,还没有其他的出版商采用其出版模式),而是代表了可信赖的出版专业人员和风险资金在学术出版领域尝试根本性创新的意愿。

另一个完全不同的方式是以 eLife⁹⁶ 的创办为代表(也是在 2012 年下半年创办),该刊由 3 个研究资助机构创办,分别是 Howard Hughes 医学研究所、Max Planck 学会和 Wellcome 基金会。创办 eLife 的初衷就是给那些国际顶尖期刊(*Cell*, *Nature*, *Science*) 树立一个开放获取的竞争对手,促进学术交流的创新。eLife 可能会也可能不会采用常规的金色 OA 出版模式(创办初期的一段时间将对作者免费);其意义在于研究资助者能否直接参与出版,这是前所未有的。另外如果抛开其他一切不谈,eLife 推翻了那些以为开放获取几乎没有成本的人的观念:它的年度报告显示在 2013 年每篇文章出版的平均费用(第一个出版全年)是 14000 美元,尽管由于产出的加倍 2014 年的单篇出版费用看起来又大幅降低(eLife 2014)。

再利用的权利

对于资助者,更加重视开放获取文章相关的许可和再利用权利的趋势将继续下去,同时这种现象已经在研究人员当中普遍持续存在。

现在许多基金会和机构强制要求受资助研究所产出的论文不仅能够被在线免费获取,而且要使用知识共享协议(Creative Commons)中最为宽松的一款许可协议——“仅保留署名权”的许可协议(CC-BY),为 OA 论文的再传播和再利用提供最大的便利(RCUK, 2012)。原因之一是人们对文本和数据挖掘的兴趣不断增长(见“文本和数据挖掘”部分)。“How Open Is It?”指南中阐明了该观点,根据权利类别呈现了开放图谱,如读者权利、再利用权、机器可读性等(SPARC, PLOS, 2013)。

为了回应这些规定,一些出版商由“保留署名权-非商业性使用”(CC-BY-NC)

⁹⁶ <http://www.elifesciences.org/>

许可协议转换到“仅保留署名权”(CC-BY)许可协议作为其对开放获取文章的默认。放弃“非商业性”的限制使出版商放弃如前所叙的任何再利用的商业性收入,如为制药公司重印和其他权利收入(医学科技期刊的重要收入来源);一些出版商通过向选择 CC-BY-NC 许可收取较 CC-BY 许可更多的费用来回应相关规定。

另一方面,对作者以及他们在有可选择的条件下行为的调查(一些出版商允许作者自主选择许可协议),显示相当一部分作者更倾向于 CC-BY-NC 许可,甚至相对于 CC-BY 许可协议更倾向于传统的向出版社进行版权转让方式(Taylor, Francis, 2014)。这种现象主要表现在人文和一些社会科学领域。

“掠夺性出版商”

开放获取出版的声誉在一些评论家的眼中已被所谓“掠夺性出版商”的出现玷污(Beall, 2012)。这些出版商被指控利用进入 OA 出版的低门槛的优势来开办大量的期刊,并且大规模群发电子邮件给出版市场的作者,甚至有时不公开出版(全部)成本,直至文章被接收后才公开,同时也会列举一些未征得其同意的编委的名单,或者剥夺研究者想要发表或不发表文章的意愿。覆盖范围小于掠夺性出版商的另一个现象是“被劫持”期刊,即通过模仿合法期刊创建一个欺诈性网站来吸引投稿并骗取文章处理费用(Jalalian, Mahboobi, 2014)。

通过对研究者的调查发现,品牌认知度是少数人不选择开放获取期刊投稿的一个原因(NPG, 2014; Frass, Cross, Gardner, 2014)。合法的开放获取出版商通过建立开放获取学术出版者协会⁹⁷(OASPA)来应对公众对开放获取出版商的低信任度,OASPA 要求其成员遵守行为准则并提供投诉处理服务。开放获取期刊目录(DOAJ)也清理了其数据库中不符合类似于 OASPA 成员规范的期刊和出版商,经过排查,发现了至少 900 种可疑期刊(Anderson, 2014b)。然而令人担忧的是,这些简单而主观的标准将一些期刊和机构被误认为具有掠夺性。

开放获取按钮

开放获取按钮⁹⁸是一类基于浏览器的书签工具,它鼓励由于收费壁垒而无法获得文章内容的用户使用。它具有三种功能:首先,它从用户处收集信息来构建存取差距图。第二,它给用户间提供机会传播自己的存取经验(例如通过Twitter)。最后,它尝试向用户建议同一篇文章的多个版本或是在开放资源库中的一个版本,如果无法获取该文章除了需付费获取的其他版本,它将向用户推荐类似主题的其他可免费获取的文章。它还能联系那些被标记文章的作者,并建议他们提供一个

⁹⁷ <http://oaspa.org>

⁹⁸ <https://www.openaccessbutton.org/>

可被免费获取的文章版本，如果这个请求得到那些作者的回应，那些需要文章的人便会得到通知。因此，开放获取按钮主要是一个鼓励开放存取的倡议活动，而不是一个用于获取所需内容的主要渠道。（参见“科学社交网络”部分）。

3.12 转型和可持续发展问题

学术界和出版市场的行动表明目前开放获取争论的焦点已经转移到了什么才是 OA 可持续发展必需的（而不是 OA 是否可能实现），以及如何管理 OA 出版转型的问题（当然这里有分类重叠的问题）。金色 OA 正在快速发展，但目前仍然只占市场的一小部分（约占市场文章的 10%，且只占市场收入的约 2.5%），目前的问题是如何继续有效扩大其规模。

关键的问题是：

- 什么会对出版的经济效益产生影响：经济回报是否足以继续吸引当前的出版商，或者结构调整可能会造成什么影响？
- 是否有相同的模式适用于所有学科领域的开放获取出版，或适用于所有类型的期刊（一种“放之四海而皆准”的问题）？
- 如何做好过渡期间的资助工作？
- 就 OA 快速发展的现状，如何构建资助机制，这种机制对学术界和出版商的影响如何？
- 采用不同 OA 方式的影响（包括对不同政府、采取不同政策的资助机构和学术机构，以及不同学科的文化准则等的影响）是什么？
- 地缘政治的影响是什么：地缘政治的变化将如何影响那些欠发达经济体和新兴经济体的研究人员？

金色 OA：基于文章发表收费的可持续开放获取模式

公众普遍认同在适当的条件下基于 APC 的金色 OA 可以提供一种可行的商业模式，这种模式既能在经济上实现自我可持续发展又能提供比传统订阅模式更广泛的经济收益和可获取性。这个结论已反映在很多关于政策导向的研究中（RIN 2011c; Finch Working Group 2012），同时在新 OA 出版商的盈利能力（Ithaka S+R 2011; PLOS 2012; 2014）和对现有商业出版商创办 OA 期刊的行为

中 (RIN, 2011c; Finch Working Group, 2012) 得以体现, 并得到 STM 成员签署的支持可持续 OA 声明的认可⁹⁹。

在忽略学科差异和其他复杂因素的前提下, 对于金色 OA 模式的可持续性发展, 作者和他们的资助方愿意支付的费用需要大于总成本 (包括可进行再投资的盈余和资金成本)。当前生产一篇文章的平均成本估计为 2364 英镑 (或 3800 美元) (RIN 2011c)。正如上所述, 多项研究估计收取的 APC 平均为 906~2097 美元, 根据选取的样本不同而波动 (Björk & Solomon, 2012c; 2014; Morrison & et al, 2014)。表 7 显示多数一流出版商收取 APC 的范围在 1300~5000 美元之间 (然而 Hindawi 的 APC 平均为 600 美元), 并且由 Wellcome 基金会在 2010 年末支付的平均 APC 为 2365 美元 (已充分考虑各种 APC)。无论选择以上哪个数字, 均低于目前已知的每篇文章的平均成本和收益。

因此, 若要做到可持续发展, 要么提高 APC 的价格, 要么降低成本, 或者两者兼具。

从经济学角度来说, 开放获取模式中有争议的优势之一是更高的价格透明度及因此产生的价格竞争, 如果这是真的 (市场的发展确实表明价格被新进入者作为一个竞争的元素使用), 将难以实现更高的定价。正如 2009 年 Outsell 报告所示, 在这些情况下, 大量开放获取出版取代订阅期刊, 会由于订阅收入的下降导致 STM 出版市场规模的萎缩 (Outsell, 2009)。基于他们的假设, 出版物市场将缩小为原来的一半 (在报告中假设出版物市场已全部转化为 OA)¹⁰⁰。

此外, PLOS 的结果清楚地表明, 如果 OA 出版在规模足够大和低成本的方式下, 收取超过 1350 美元或更少的 APC 是能够盈利的 (在 PLOS 的所有期刊中, 每篇文章的平均成本为 1088 美元)。由于较高的退稿率和更优质的编辑服务, 一些编辑成本比 *PLOS ONE* 高的期刊, 就只能收取更高的 APC: 美国科学院院刊 (*PNAS*) 预计会收取 3700 美元的 APC (Van Noorden 2013); *EMBO* 将收取不少于 2000 美元以支付期刊的成本 (Pulverer 2014); *Nature* 的出版商预估其内部成本在 10000 美金以上 (Jha 2012); *eLife* 在 2013 年平均每篇文章的成本为 14000 美元 (其第一个出版全年), 2014 年由于发文量翻倍, 成本有望大幅下降。

假设市场的动态是价格不断走低, 一些业内人士担心可能有意想不到的不良后果。第一, 降低成本的压力可能会导致 OA 出版走捷径及质量下降 (即“逐底

⁹⁹ <http://www.stm-assoc.org/publishers-support-sustainable-open-access/>

¹⁰⁰ 假定论文产出数量随着研发投入的增加而增长, 这份报告模拟了不同层次的 OA 采纳和定价增长模式, 结果表明在某些特定情况下, 出版市场可以恢复到预 OA 的水平, 但前提条件是 APC 必须高于目前市场的定价。

竞争”，Anderson 2012）。对于某些出版类型，市场的选择是低成本和低服务质量（*PLOS ONE* 的成长就印证了这一点），但这种方式不适合低录用率并且刊登大量非研究性文章的期刊，也不适合对使期刊更优秀的新型工具日益增长的需求。

第二，收入和微薄利润的压力可能使编辑或出版商降低文章的录用标准，以刊登更多的文章。可以说订阅模式下也存在相同或相似的压力（因为出版商可以通过扩大期刊容量来提高订阅价格），这两种情况下的答案无疑是相同的：期刊的标准越低就越无法吸引好的作者或编委会成员，期刊将越办越差。

到目前为止不同学科采取金色 OA 出版的情况存在很大差异，采用最多的是生物医学领域，最少的是人文科学、数学和（或许更令人吃惊）化学（Björk et al., 2010; Archambault et al., 2014）。因此大多数（但不是全部）创新性的出版活动也集中在生物医学领域。

生物医学领域采取金色 OA 出版的有利因素包括大量的研究资助和已经设定规则的资助机构。随着政府的政策逐渐转向将所有的学术产出开放获取，产生了一些问题：那些外部科研经费较少的学科（如数学、人文科学）和众多没有资助的作者（只有 60% 作者有经费支持）将如何采用 OA 出版？答案似乎是：大学或其他雇主将从单位事业经费中划拨 APC（如英国研究理事会（RCUK）的新政策批准了一大笔经费用作 APC 基金）；然而，这里又出现了其他诸如由谁来决定及如何支配有限的出版基金的问题（Crotty 2012）。

更广泛采纳金色 OA 的一个主要障碍是过渡期资金资助的问题。对于独立的机构而言，如果采纳有利于金色 OA 的（国家或地方）政策就会因为支付 APC 而加大其运营成本，因为与此同时，他们还需要继续支付订阅期刊的费用（Swan & Houghton, 2012）。这个问题在一个率先于世界其他国家推行金色 OA 政策的国家层面也同样存在。这些问题均在《开放获取之路》报告中给出了模型（RIN 2011c），并且在 Finch 报告中进行了讨论，该报告建议英国政府在过渡期间每年提供额外的 3800 万英镑，再加上 500 万英镑的一次性补助，用于承担过渡期间的额外支出。但不足为奇的是，这些建议都未被政府采纳（虽然政府提供了额外的一次性 1000 万英镑补贴）。正如一些人已指出的那样，从全球范围内看，这些过渡成本将是一笔非常大的支出。

看来在全球范围内，出版模式过渡期间的资金资助来源于政府和科研基金资助机构是不太可能的，资助机构和出版商正在探索以市场为基础的解决方案。一种可行的方案是抵偿和捆绑（见补册部分）。

过渡期间的另一个问题是金色 OA 的经济状况体现在非现金节省（例如，研究者的时间）和不确定经济的多重影响。在经济紧缩的环境下，与执行出版转型

过程中的现金成本相比，获得的非现金收益将无足轻重。

开放获取可能导致市场萎缩的另外一种方式是其对版权收入的影响（假设采用仅保留署名权的 CC-BY 协议会使收入减少）以及对企业订户的影响（目前企业订户的贡献大约占期刊收入的 15%~17%，但其对论文产出的贡献只有一小部分（约 5%））。如果期刊的经济来源转向收取 APC 的话，这应该不成问题，但有人指出，这实际上会将企业的花费转嫁到大学，是一个令人不愉快的转变。

同样，APC 定价时需要考虑作者申请减免的因素（如针对发展中国家的作者）。这也代表了一种转移或补贴，虽然很少有人会反对。据牛津大学出版社(OUP)报道，其出版的约 70 种期刊中，这种减免率一直稳定在 6%~7%；PLOS 的年度报告显示，减免率稳定在 7%~8%。来自中等收入国家的研究人员可能反而会对要求减免感到不舒服，有一项调查显示，这些国家的作者用自有资金支付 APC 的比例更高。

还有其他一些有关开放获取对新兴和欠发达经济体影响的担忧。一是对当地出版商的影响：如果开放获取出版成为一种常态，作者可能放弃向本国无力减免 APC 的期刊投稿（这些期刊的大多数作者都符合减免条件），而倾向于投稿给能够提供 APC 减免的西方期刊。因此，从长远看，除了在最贫困的国家，减免的体系似乎是不可取的（Dickson 2012）。

非 APC 的金色 OA 模式的可持续性是很难建立模型的。非 APC 的金色 OA 并不是学术交流环境中一个不重要的或完全低质量的部分：基于对 Scopus 数据库的分析得出一个令人信服的估计，即没有 APC 的金色 OA 文章（4.2%）水平与有 APC 文章（5.5%）相当，同时几乎所有的大型期刊出版商都有这两类案例。

现今的这些模式（很大程度上是依赖于补贴和赞助）是否能够被衡量还是未知的，而一些新型的模式（例如，图书馆合作补贴或众筹）还没有经过证实或者有着无法弥补的缺陷（例如免费搭便车的现象）。

基于 APC 的金色 OA 对于很有名望的顶级期刊，如 *Nature*、*Science* 这些依赖于昂贵的编辑质量控制流程的期刊，并不是一个好的模式，因为这意味着必须要收取很高的 APC。*eLife* 的创立就是为了推翻这种假设，然而，尽管它已经成功地吸引了高质量的文章，但在它明确自己的长期商业模式之前都无法定论。

有些期刊（包括发行量大的普通期刊）印刷广告的减少是一个问题，虽然广告收入仅占期刊总收入的 4%。这个问题的产生应更多地归咎于期刊的数字化而不是开放获取；另有一些迹象显示，平板电脑版本对广告商更具有吸引力。

被 *EMBO* 以及其他一些退稿率高的期刊所采用的一种方式建立向其他期

刊转移稿件的系统（见“级联同行评议”部分），这样可以减少一些 APC。对于退稿率高的期刊来说另一种可行的模式是引入投稿费制度（Pulverer 2014），但这种模式也有严重的缺陷：出版商们不喜欢在一个充满竞争的市场中承担稿件流失的风险，并且资助机构也不赞成这一做法（因为他们主要是赞助非出版机构）。

目前仍未讨论的一个问题是：如今的政治思潮，尤其是公众对互联网的态度，是否会导致付费出版不可持续。潜在的威胁来源于如下观点：在线的内容应该是免费的，内容的分享是网络的首要功能，知识产权的概念已经过时，公共基金资助自动等同于公共获取。Michael Mabe 已经在他的一本书中讨论了这些问题，并得出结论，只要现存版权的框架可以继续维持，并且政客们了解其中的风险，这场战斗还不能说失败（Mabe 2012）。

混合（或选择性）开放获取

混合模式的开放获取最初的提出是作为订阅型期刊向开放获取转型的一个低风险路径，以避免一次性的转型所带来的风险。当然，有些出版商选择混合模式并不是期望在短期内过渡到 OA，而是利用它来获得资助，并为作者提供一个遵循基金资助机构政策的途径。

在实践中全球性采纳混合 OA 模式的比例一直非常低（约 1%~2%），但也有例外（见 *混合期刊* 部分）。

基金会和机构担心“双重收费”现象会一直存在。新的商业模式将订购与 APC 成本捆绑或者允许补贴，这也许会让混合 OA 在未来扮演更重要的角色（见 *补贴* 部分）。

延迟的开放获取

延迟开放获取期刊是期刊在一段设定的保密期后再提供对内容的免费访问（虽然不是通常意义的开放获取）。Laakso 和同事们将延迟开放获取描述为一种被忽视的具有高影响力的可以公开获取科学文献的方式，它能为读者免费提供 5% 的（不包括当年发表的）期刊文章（Laakso & Björk, 2013）。

延迟开放获取商业模式的可行性在于图书馆愿意继续订阅期刊，即使期刊（过刊）的大部分内容已经免费提供。这里需要考虑的两个（相关）关键因素是保密期长短和学科领域。此处所涉及的问题与自存储相同，除了延迟 OA 的保密期完全处于出版商的控制下，我们将在以后的内容中详述。

3.13 自存储对期刊的影响

出版商一直担心期刊论文普遍的自存储可能带来的影响。按照常理假设，如果强制要求导致高度的自存储，图书馆（可能长期处在预算紧张的压力下）会越来越多地选择依靠自存储的版本，而不是从出版商那里订阅。

2006 年秘密情报处（SIS）为出版研究联合会所做的一项研究一定程度上支持了以上假设（Beckett & Inger, 2006）。这项研究调查了图书馆员的采购偏好，并得出图书馆员都倾向于用 OA 资源代替订阅的结论，前提是这些资源是经过同行评议的（正如所有基金会和机构所要求的）并且保密期不是很长。最后也是关键的一点：对于保密期为 12 或 24 个月的 OA 版本来说，图书馆员更有可能倾向于订阅版本，但 6 个月或少于 6 个月的保密期对图书馆员的选择倾向几乎没有影响。然而这只是对图书馆员的调查；大量的调查研究（包括 PEER 项目研究）揭示了研究人员对 OA 或订阅版的选择倾向，至少反映了他们对研究出版周期的一些阶段的看法。

PEER 项目

假定使用期刊自存储对图书馆员决定是否订阅很重要，并且对期刊的市场定价也很重要，自存储是否会导致出版商网站文章下载量的减少也成为一个问题。欧盟的一个项目 PEER（欧洲研究生态与出版项目）在 2008 至 2012 年间共进行了 45 个月的调查研究，涉及整个欧盟的 12 家出版商和 6 个知识库，该项目在自存储对开放知识库的影响方面进行了最全面和最详细的研究¹⁰¹。该项目的研究结果涵盖了广泛的议题，其中包括对出版商和知识库经济和行为学的研究（其中的一些内容已在本报告的其他部分阐述），尤其是项目报告中关于自存储使用的研究为知识库对期刊的影响提供了最相关的数据。

关于自存储使用的研究（企业工程技术研究咨询公司的研究（CIBER Research 2012b）是一项随机抽样调查，旨在比较在有和没有 PEER 网络知识库存储的情况下出版商网站文章的下载情况。这项研究的主要发现是，出版商网站文章的下载量不但没有减少，从出版商的网站下载的文章在 PEER 知识库的显示度还有了适度的增加。下载量总体增加了 11.4%（95%的置信区间为 7.5%~15.5%）。研究者们认为可能的解释是，PEER 知识库高质量的元数据和一项自由的搜索引擎机器人索引政策提高了数字化文章在搜索引擎中的显示度（在很多情况下对出版商网络平台的元数据的摄取得到增强和拓展）。调查的结论是，没有实验证据支持 PEER 知识库会对出版商网络平台文章的下载产生消极影响的断言，并认为“知识库与出版商”的二元对立是一个错误的假设，另外“在复杂的学术生态系统

¹⁰¹ <http://www.peerproject.eu/>

中，高显示度才是王道，并且通用搜索引擎越来越成为关键工具”。需要加以考虑的另一个因素是 PEER 项目中的每个期刊都能够选择一个适合自己的保密期。这意味着现实世界中可能存在的“一刀切”标准在该项目没有得到反映。

这个结论可能有片面性：首先，PEER 项目的研究者竭力强调他们的研究结果仅针对 PEER 知识库，不具有典型性，其中的原因多种多样。例如，PEER 研究者发现出版商的突出贡献在物理学领域有显著的统计学意义，但这一发现同与 arXiv 知识库涵盖的物理学科出版商的经验（非设为对照）正好相反，而 arXiv 不仅在物理学领域更为知名，而且更关键的是，它基本包含了某些物理学二级学科期刊的全部内容，虽然大部分论文都是第一阶段或预印本内容版本。其次，即使该研究结果具有普遍意义，由知识库促进的自存储使用的增加不一定意味着面临预算缩减的图书馆在订阅时会优先选择那些可以在知识库中找到的期刊。例如，2012 年全球学术和专业学会出版商协会（ALPSP）和英国出版商协会（PA）对图书馆员进行了一项调查，询问他们如果期刊的大部分内容在 6 个月的保密期后可以免费获取，是否还会继续订阅期刊（Bennett 2012），34% 的受访者表示他们会取消订阅一些自然科学的期刊，10% 表示他们会取消所有的此类期刊；对于人文社科类的期刊，42% 表示会取消其中一部分期刊，另有 23% 会取消全部期刊。

争论的关键问题为是否允许保密期的存在以及保密期的时长。出版商认为缩短或取消保密期，如像基金资助机构强制规定的那样，将会给期刊的订阅带来更大的风险，然而 OA 的拥护者则认为缺少相关证据证明此观点。出版商还认为不能为不同的学科设定统一的保密期，因为不同学科期刊的使用模式有很大的不同。PEER 项目关于使用的研究也提供了文章终身使用规律的数据。图 30（取自企业工程技术研究咨询公司的研究，CIBER Research, 2012a）显示了不同学科领域的文章出版后的累积下载量。生命科学、医学和物理科学领域文章的使用自出版后的第 56 个月开始进入平台期，而人文社会科学的则是在 80 个月才开始进入平台期。然而，对订阅用户而言关键的问题不是文章被使用时间曲线的整体长度而是其形状，如研究人员更关心在出版后第 6、12 或 24 个月时获取全文的重要性（或价值）。事实上，无论证据如何，政策制定者（包括英国研究理事会和欧盟）越来越倾向于选择缩短保密期（见“自存储政策与规定”部分）。除了使用数据的研究，PEER 项目还有另外两个主要的研究主题，即关于行为学 and 经济学方面的调查（Wallace 2012）。

有关行为学方面的主要发现可能与行为学项目本身无关，而是作者完全没有兴趣将文章自存储。最初的计划是推广一半文章由出版商直接存储、另一半由作者存储。尽管发送了近 12000 封邀请函，由作者存储的论文仅有 170 篇。这可能与项目的实验性质有关，并且作者以前不知道 PEER 项目的存在。但是以前也有

相关证据显示，一些研究者认为自己发表在期刊上的文章能否被 OA 获取不是他们分内的事。确实，通过自存储使文章 OA 的作者只占少数。总体来看，PEER 行为学项目的研究表明，“学术研究者对学术交流系统的态度、认知和行为都比较保守，他们不期望现有的传播和出版方式发生根本性的变化”。对知识库，研究人员并不一定抱有负面态度，但却相当谨慎，并且不相信可以从改变他们自身的行为中获益。

PEER 项目的经济学研究结果中新的信息较少。研究证实了同行评议确实有实际存在的成本并且规模经济很小，另外估计出版商每篇投稿的同行评议的平均成本为 250 美元（仅包括薪金和税费，不包括管理费用、基础设施、系统建设费等）。去除同行评议费，估计每篇论文的平均生产成本为 170~400 美元（不包括管理费等）。出版商的网络平台年度维修费用在 170~400,000 美元之间，另外平台初建和开发的费用一般需要数十万美元。

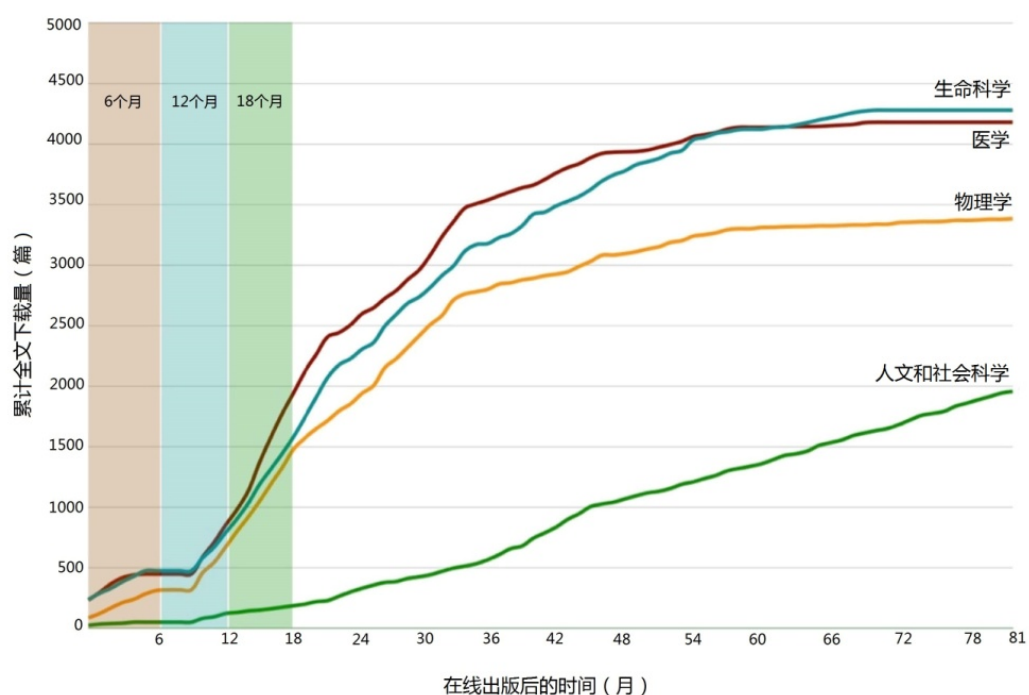


图 30 不同时间不同学科领域出版商网站的全文下载情况 (CIBER Research, 2012a)

3.14 开放获取对论文使用的影响

对使用（下载）的影响

现在有大量的关于开放（或免费）获取对使用行为影响的研究，包括下载和引用，可以回溯到至少 15 年前，这可能是开放获取文献中最多内容之一¹⁰²。（现在一个类似的问题，即关于研究数据公开带来的影响也正被研究，见“数据引用”部分。）

由于缺乏适当的方法学来控制其他因素的影响，使用数据的研究存在一些缺陷（包括早期的视图效果和选择性偏差，见下文），人们普遍认为可以免费获取文章的下载次数应明显高于不可免费获取的文章。例如，对比 2007 年刊登在 *Journal of Experimental Biology* 上的选择性 OA 文章和非 OA 文章，发现 OA 文章全文的下载量较非 OA 文章约高 40%（Bird, 2010）。

更有力的证据来自 2011 年的随机对照的试验，该试验发现，HTML 格式 OA 文章的下载几乎翻倍，PDF 格式 OA 文章的下载增加了 62%（Davis 2011; Davis & Walters, 2011）。

一部分下载量的增加可能不是由于（人们）对内容的使用：牛津大学出版社出版的 *Nucleic Acids Research* 在转换为 OA 期刊后文章的下载量增加于一倍，但其中的大部分下载是搜索引擎的贡献，只有 7%~8%是因为对其内容的使用（Bird 2008）。然而，由于 *Nucleic Acids Research* 是非常成熟的一流期刊（订阅量很大），它作为一个研究性专业出版物不大可能有其相关研究群体之外的读者感兴趣。

随着研究人员行为的改变，便捷的无处不在的可获取性也可能改变“使用”的含义。例如，2008 年企业工程技术研究咨询公司（CIBER）和其他人的调查结果显示，用户的“内核浏览器（power browse）”通过最初点击的文章列表（通常来源于搜索结果）形成，在保留一些文献供以后的研读和使用的同时，会略读并摒弃很多文献。Outsell（2009）指出这意味着虽然这些文章对出版商来说价值都是一样的，但对研究者来说，它们的价值可能从零（立即丢弃）到非常有用来分布。

开放获取引用优势

大量的研究都针对开放获取如何影响到文章的引用（Lawrence, 2001）。一般的假设是，与仅能从订阅途径获取的文章相比，可开放获取的文章将被更多地

¹⁰² 例如，OpCit Project 的在线文献目录是这一领域的文献资源最丰富的数据库（OpCit n.d.）

使用，并因此被更多（或更早）地引用。然而，由于文章所得到的全部引用实际来源于其他学者，引用总数量的增加只可能在期刊所在领域内的大部分活跃的研究者不能获取到该期刊的文献时成立。

大多数研究表明，相对非自存储文章，自存储的文章确实能够获得 200%~700% 的引用率，导致这种现象的原因分为 3 种效应：（1）先行阅读效应，存储文章也可能是因为可获取的时间更长而得到更多的引用；（2）选择倾向效应，作者更有可能将自己的优秀文章存储，或者更优秀的作者更有可能进行自存储；（3）开放获取效应，文章的高被引纯粹是因为文章的开放获取(Kurtz et al., 2005)。合作者的数量和地域也影响到引用率，但通常是不可控的。

2010 年的一个文献计量学研究(Gargouri et al., 2010)试图控制选择性偏差，他们将强制实行（因此可以很好地观察）自存储政策机构的自存储文章与其他机构（作者选择自存储的比例低）的自存储文章进行了比较，发现开放获取状态与引用数量之间存在很强的统计学相关性，这种相关性在高被引文章中最为明显。

Craig 和他的同事们在一篇文献综述中表明，当时最为严谨的研究结果（Moed 2007，涉及凝聚态物理研究）证实了先行阅读效应对引用差别的影响主要是由选择性偏差造成的，但是没有证据支持开放获取效应的影响（Craig et al., 2007）。然而，不同学科的引用模式不同，因此，其他学科中开放获取效应是否有效仍有待考证。

Davis（2011）报道一个随机对照试验的结果，发现开放获取文章在出版后的第一年中下载量显著增加，拥有更广泛的读者，但与对照组订阅获取的文章相比，并未获得更早、更多的被引用率。在随后的一篇综述里，Davis 和 Walters（2011）对现有的证据进行评估，并得出结论：OA 对文章被引用率的影响是“不清晰”的。最近的研究显示大量的引用优势只是由未能充分控制混杂变量的人为因素产生的，并且他发现 Craig 及其同事的结论在后续的工作中被证实。通过统计学或方法学控制混杂变量后，几乎没有证据显示开放获取状态独立地影响被引用次数。

在 Björk 和 Solomon 进行的另一项研究中，比较了 OA 期刊和订阅期刊的引用率（Björk & Solomon, 2012a）。他们发现订阅期刊的整体引用率高出约 30%，但在排除了创刊时间、学科和出版地的影响因素后，这些差异大多消失了。收取 APC 的 OA 期刊的引用次数明显高于那些无 APC 的 OA 期刊。在医学领域，在过去 10 年里创办的 OA 期刊得到的引用率与同期创办的订阅期刊大致相同。

最近对经济学和商学领域期刊的一项研究明确地显示：“过去的研究中出现的重要的影响因素是由于未控制文章的质量而人为造成的，一旦将固定的影响因

素设为对照，之前的差异因素就会消失”（McCabe & Snyder, 2014）。然而，研究者们发现，整体聚合效应的缺失掩盖了在线资源之间的差异：在这个方面，西文过刊全文库（JSTOR）系统中的期刊引用率提高了 10%。

相反的，由同一批作者开展的另一项研究（McCabe & Snyder, 2013）发现：“在我们调查的样本中，从订阅期刊转向开放获取期刊可以使文章的引用率平均提高 8%。获益主要集中于顶级期刊。事实上，在我们的研究中，开放获取会让排名低的期刊的引用率显著降低，因此我们推断，开放获取会加剧文章争取读者关注上的竞争，从而分化出期刊中的失败者和获胜者。”

Nature Communications 是一个混合 OA 期刊（现在已转向完全开放获取），多于 40% 的文章已选择了开放获取。因此可以在同一个期刊中比较使用开放获取和不使用开放获取的效果。自然出版集团（NPG）委托信息网络研究（RIN）出具的报告显示：开放获取文章的下载量有显著的增长（与前面许多研究结果一致）。据它报道，就算在研究中未控制选择性偏差、绿色 OA，或者作者数量和位置的影响，OA 文章的引用率“似乎看起来”也会有小幅的增长（RIN, 2014）。

具体是什么因素造成的影响还是未知的，但是目前最充分的证据表明，OA 文章总体上可能不会被一直引用，但由于能被较早地看到以及选择偏差效应，它们能被更快地引用。在一些特定的情况下 OA 期刊的引用率确实增加了：例如，在一些特定平台、领域内，或是在高排名的期刊中。

引用效应是否存在以及它有多大的影响可能越来越成为学术界关注的热点，因为开放获取的文献比例正在稳定增长。引用增加的效应就是开放获取的众多益处之一，它同时影响着作者。例如，在 2014 年 Taylor 和 Francis 的调查显示，29% 的作者认为 OA 期刊的引用率高于订阅期刊的，而在 2013 年的调查中只有 25% 的作者这么认为（Taylor & Francis 2014）。

4. 学术交流新进展

技术的发展正在促进研究行为和交流方面的深刻变革,或者说是为这种变革创造了机会,这两方面的变革都可能对期刊出版产生影响。

技术进步速度的加快,对这份涵盖技术发展趋势的报告每3年一次的更新行为成了挑战。虽然如此,我们编写本报告时仍然达成共识:尽管在其相对重要性方面确实有争议,但这份报告中所讨论的技术变革趋势对学术出版有重要的作用。

其他相关更新更频繁的信息,建议参考STM未来实验室委员会的报告¹⁰³。最新的报告(STM Future Lab Committee, 2014)确定了2014~2016年的3个主要议题:“机器是新的阅读器”(丰富的内容、文本和数据挖掘、智能化文章等);“对作者的回馈”(更好的作者/研究人员经验、扩大的同行评审、新的写作工具、出版自动化等);“改变游戏的新播放器”(社交媒体、新启动器、自助出版、开放写作工具等)。

值得注意的是,STM未来实验室委员会的报告在过去两年提到的主要发展趋势,不仅不会马上过时,而且已经融合到当今的形势中:

- 2013年的核心主题是:“混合阅读器体验”(未来的PDF等),“开放获取实施”(开放课件、开放ID、开放文献数据等),“从认同到扩大”(数据处理和分析、替代计量学、新的计量学)。
- 2012年的核心主题是:“从内容的可发现性到内容的可操作性”,换句话说,技术的焦点从如何使内容更容易被搜索和发现(如平台架构、语义富集、搜索引擎优化等),转换为如何使内容更有用、更互动、更可用和再利用(如应用程序接口(API)技术、数据集成、数据和文本挖掘、语义网络技术、生产和工作流程工具)。本报告确定了三个主要议题:API平台、研究数据及身份管理,这是提高内容可操作性的关键。

2012年一篇关于STM出版平台发展的综述(Outsell, 2012a)指出了以下几个关键主题:在日益开放的信息时代明确出版商平台的作用;用户体验重要性的日益增加;可发现性;超越简单的个性化;社交媒体和网络。以上主题的讨论基于如下的技术发展背景:移动设备、语义富集、搜索工具、应用程序接口¹⁰⁴、电子商务和货币化、集合和整合、功能性(活跃)内容、分析。尽管这份报告中

¹⁰³ <http://www.stm-assoc.org/standards-technology/committee/>

¹⁰⁴ Application Programming Interface, http://en.wikipedia.org/wiki/Application_Programming_Interface

的一些细节不可避免有些陈旧，但大部分主题仍是重要的。

Outsell 对 STM 出版平台供应商市场的评价 (Outsell, 2014d) 揭示了一些最新的发展趋势：对数据可发现性和可获取性更多的需求，重点是在研究者身上而不是研究内容的“容器”上，对信息而非格式——随之出现的相应驱动——用户行为和工作流程的需求，商业模式的定制化服务，以及对标准化的需求。

以上各方面的观点和主题将在下面的章节中进行阐述。

4.1 “科学 2.0”或“开放科学”

尽管“科学 2.0”这个词作为表述一种新兴的数字化方法（通常是以网络为基础）来研究已经被非正式地用了至少十年 (Waldrop, 2008)，最近欧洲政策制定者将它用在更为正式的领域。欧洲委员会召开了一次公共的研讨会，题目为“科学 2.0”：科学转型 (European Commission, 2014)，将科学 2.0 定义为科学研究及其组织的一个基于数字技术的系统性变化，受科学团体全球化及解决“Grand Challenge 倡议”(由比尔及梅琳达·盖茨基金会资助)中涉及问题的需求所推动¹⁰⁵。

报告指出科学 2.0 包含三个明显趋势：研究产出的显著增加，科学数据密集化（见后面阐述），科学家数量和科学受众的双重增加（两者都包括专业和非专业（民间科学）群体）。

科学 2.0 的政策利益一部分体现在对社会和经济有争议的影响方面，包括科学对社会需求更具响应性，更大的公开性能够提升科学的信任度，由创新率增加带来的经济利益（比如，中小型企业可以开放获取研究数据），利用大数据技术、文本和数据挖掘等产生新型的服务和生产力（尽管本报告指出的上述大部分说法目前依据还不够充分）。

从出版的角度来看，欧洲委员会报告指出，和加快出版以及重组现存的同行评议体系一样，开放获取为科学 2.0 的一个重要部分。需要改变的相关领域还包括由于缺少可获得数据而对研究结果再现形成的挑战，对研究的评价和激励措施的改进，比如对开放数据的方法采用与发表文章相同的标准进行奖励。

近年来科学 2.0 被重新命名为开放科学，由于科学 2.0 被主张的范围过于广泛，使得其不被研究团体支持。开放科学还有一个优势，他似乎不会回到现在已经过时的网络 2.0 时代。

¹⁰⁵ <http://grandchallenges.org/>

4.2 FORCE 11 和“科学转型”

研究团体已经率先提出了一系列倡议（不仅仅是在出版领域），举例如下：

- FORCE 11 是一个致力于促进学术交流的团体工作的网络，除了每两年举办一次口碑不错的会议（原先叫做“不止是 PDF”）¹⁰⁶，它于 2011 年的声明中提供了对新机会和新挑战的实证分析。它着重关注作为 STM 产品制造者的研究人员对工具的需求，为顾客身份的研究人员提高产品的质量、声誉管理的工具及服务（FORCE 11, 2011）。FORCE 11 已经融入到后面讨论的一系列具体问题中（见数据引用部分）。
- “科学转型”项目与欧洲委员会提出的“科学 2.0”讨论会具有许多相同的主题，它已经在荷兰具有重要的影响力，但在其他地方的影响没那么大¹⁰⁷。

4.3 出版平台和应用程序接口（API）

众所周知，大量的搜索行为不是从出版商的网站开始的（比如多达 60% 的网页链接是来源于搜索引擎）。鉴于此，出版商的平台在研究者的研究工作中扮演什么样的角色呢？如果研究者对期刊和出版商一无所知，只是想尽可能快地进入出版商的网站找到并下载 PDF 文件（CIBER 2008），那么出版商是否应该将网站设计成包含（智能）内容的（智能）储库，且拥有最大限度的开放网络可搜索性和开放应用程序接口，使得终端用户无论选择什么途径都能够以最快的速度得到他们想获得的资源呢？亦或是出版商是否应该投资语义富集技术，提升交流，通过开发或集成工作流程工具创建一个丰富的、高效的网络环境？事实上，无论是开发跳进/弹出功能强大的浏览器还是让处于深度探索阶段的研究人员能得到更好的互动体验，这些行为都已得到了出版商的支持。

未来 STM 出版平台的一个关键的技术特征是具有开放的应用程序接口（此处“开放”是指接口的技术参数是免费提供的，而不是指内容）。从战略角度考虑其原因，是平台的价值将越来越多地取决于可互操作性（例如，可以从多个来源整合内容，可以整合和共享数据，可以增加平台的功能性，并允许用户从其选择的切入点或使用的工作流工具获取他们所需的内容）。更具战术性的是，当代应用程序接口的发展将有助于出版商开发更多新产品，提供更快捷的服务，制定内部的工作流程处理并更轻松地管理它们，以及更容易地支持多台设备。

¹⁰⁶ 例如 <https://www.force11.org/meetings/force2015>

¹⁰⁷ <http://www.scienceintransition.nl/english>

4.4 社交媒体

社交媒体和网络（有时称为 Web 2.0）具有促进正式和非正式的学术交流的潜能，但是其影响迄今为止还是有限的。它们的影响正在逐渐地扩大，尽管与学术交流的传统途径相比仍具有一定的局限性。

目前已有许多关于科研人员使用社交媒体的研究。RIN 的题目为“如果你创建了它，人们会来吗？”的报告中发现，科研人员使用社交媒体的比例较低，一般低于 15%（RIN, 2010）。只有一小部分研究人员（大约为 5%）使用社交媒体来发布他们的科研结果和工作进展。RIN 认为，阻碍社交媒体更多地被使用的原因是人们不清楚其潜在的价值：由于社交媒体的使用需要投入相当大的成本，而且没有明显并立竿见影的好处，研究人员更倾向于使用他们熟悉且信赖的学术交流渠道或服务。服务系统的快速发展和扩散意味着社交媒体很难被追踪或评估其潜在的价值，并且服务系统的扩散趋势意味着每种社交媒体都缺乏用户需要的集群效应。社交媒体使用的第二个障碍是关于其质量和信任度：研究人员不愿使用非同行评议或缺乏公众认可的新型学术交流形式。另外，与人们一贯的观点不同的是，使用社交媒体的人口因素差别很小，其中包括年龄在内。RIN 的总体结论是，很少有证据表明，Web 2.0 这种由开放式研究团体提倡的学术交流模式会在中短期内给学术交流带来根本性的改变。

其他的研究也发现了类似的结果（Ithaka S+R, 2010; Procter 等, 2010; RIN 2009a）。更有趣的是，David Crotty 深入研究了被生物学家广泛使用的 Web 2.0 网络工具，并分析了它们未能取得进一步成功的原因。研究发现它们未被更广泛使用的主要原因是：缺乏时间，缺乏激励，缺乏归属感，缺乏用户的集群效应，行为惯性，以及不符合科学文化的不适当的工具等（Crotty 2008; 2010）。

这些社交媒体除了被主动使用的频率相对较低（比如展示内容），研究者也很少被动使用他们作为搜索信息和增加认知的资源。比如说，Ithaka 发现很少有受访者会把博客和社交媒体作为对他们科学研究比较重要的工具，尤其是看其他科学家的博客或社交媒体对一些研究者来说是迄今为止追踪其科学领域最新研究进展的最不重要的方式（Ithaka S+R 2013）。有趣的是，越来越多的科学家（尽管比例依然较低，大约为 15%）表示他们会使用社交媒体来分享其研究成果，却很少使用其追踪最新研究进展（大约 5%），可能是由于他们将社交媒体更多地作为一种在学术界扩大影响的工具，而不是与同行交流的工具。

Web 2.0 技术可以被用来当作同行评议的补充手段，例如在文章发表后允许读者对文章进行评论和打分。尽管经过尝试（比如说对期刊文章进行评价和评论），这个工具的使用率仍旧很低，同时还存在关于使用这些技术测试的是什么样一

些亟待解决的问题 (Harley 等, 2010)。尽管如此, 这个领域新的技术仍会继续出现并可能会大有作为 (见发表后同行评议部分)。

由于社交媒体在普通人群中的使用已成为强劲的潮流, 许多人相信社交媒体将逐渐演变成为学术交流的主渠道, 尤其是它们被更加紧密地整合到个人电脑和移动操作系统。科学社交网站已经发展得十分迅速 (见以下内容)。有一些迹象表明 Twitter 可能会成为预测高被引文章的主要工具 (Eysenbach, 2011)。日益普及的文章层面的计量指标也许能够提高研究人员使用 Twitter 或博客来讨论或宣传发表期刊文章的意识, 并因此带来一些正面的反馈。与仅设置一个“喜欢”按钮相比, 更多地融入了社交特征的服务系统 (如 Mendeley) 可以更自然地建立社交行为。因此, 总体上这些例子可以让我们相信社交媒体至少在科学内容发现和分享方面发挥了一些作用。

维基百科

维基百科 (Wikipedia) 不仅是当今最知名的由普通用户创作的多用途的百科全书, 而且已有越来越多的研究人员和学者使用它查询关键信息以外的内容, 尽管始终有对其某些内容的质量存在质疑的声音。Wikipedia 已经启动了一些关联项目 (“WikiProjects”), 旨在提高某些特定学科文章的数量和质量。

虽然一些研究者们对 Wikipedia 感兴趣, 但它不太可能对学术交流的核心领域产生太大影响。更有意义的是一些利用 Wiki 平台的核心功能进行研究或达到其他学术目的的特定项目。也许其中最令人兴奋的基于 Wiki 的项目是允许学术团体创建和维护共享数据库和资源。例如, WikiPathways, 是采用标准的 Wiki 软件创建了一个“由科学界建立并服务于科学界的专注于管理生物信号通路”的网站; 再如 OpenWetWare, 是用以促进生物学和生物工程研究界信息共享的软件。

作为一项相关的服务, Wiki 数据在包括 Wikipedia 的 Wiki 媒体姐妹项目中主要用于储存结构数据。一个旨在建设作为中央枢纽来连接开放研究数据的 Wiki 数据提议正在进行, 该提议暂时被命名为“用于研究的 Wiki 数据”, 由柏林自然博物馆和 Wiki 媒体德国公司合作实施¹⁰⁸。这项提议成功与否还有待进一步观察。

学术出版商还迟迟没有采纳 Wiki 作为出版平台, 这很可能是因为 Wiki 系统依赖于开放、可编辑和可重复使用的内容, 而这种运作模式不容易产生利润¹⁰⁹。

¹⁰⁸ <http://blog.wikimedia.de/2014/12/05/wikidata-for-research-a-grant-proposal-that-anyone-can-edit/>

¹⁰⁹ 本报告上一版中提及的两个项目 (Elsevier 的 WiserWiki 和 SciTopics) 目前已被停止, 又创办了一些新的项目, 例如国际水质协会出版社 (IWA) 的 WaterWiki, 它将正式发表的文献内容与学会整合的内容结合起来, <http://www.iwawaterwiki.org/xwiki/bin/view/Main/WebHome>

RNA Biology 期刊自 2008 年起要求有关 RNA 家族研究的文章作者向 Wikipedia 的“RNA 家族”条目提交论文的初稿,希望 Wikipedia 网站能够成为收集有关 RNA 家族信息的重要窗口。但这一政策没有得到其他期刊的普遍效仿。

科学社交网络

自从本报告的前一版刊出后,科学社交网络取得了迅猛的发展。在 2008 年左右创建的三个主要社交网络系统是 Academia.edu¹¹⁰ (据报道已有超过 1600 万注册用户)、Mendeley (将近 350 万用户,且在 2013 年 4 月被 Elsevier 收购)和 ResearchGate (超过 500 万用户)。第四大网络系统是在 2011 年建立的 Colwiz,迄今为止已有 26 万用户。人们对网络系统的熟知度看起来相对较高,尤其是在 STM 领域的 ResearchGate 以及在人文社科领域的 Academia (Van Noorden, 2014c)。

用户文档的数量也是非常庞大的: Mendeley 报道它的用户们已经上传了超过 4.7 亿份的文件; ResearchGate 报道可以通过其平台下载获得 1400 万份文件。

研究人员利用这些平台以达到多种目的,最被大家认可的普遍动机就是为了建立在线的简介以使他们自己和他们的研究成果被更多人发现,其他比较受欢迎的用法就是发布文章内容、发现相关的研究者、追踪数据以及发现新的或由他人推荐的研究性文章 (Van Noorden, 2014c)。目前,一项看起来似乎不那么重要的活动便是直接互动和讨论;这些平台看起来更像是为研究者收集资料提高自己的声望并使资料被更多人发现的工具,且这些平台可以成为使研究者更多地接触工作流程工具和服务系统的利器,而不是用作社交互动的交流工具。

一些分享期刊文章导致侵犯版权的社交网络的使用使得出版商开始担心版权问题,以至于在一些情况下他们会发出撤回文章的通知 (Economist 2014)。需要制定一些行业的规范来解决这些问题,例如利益相关者可以自愿地使用一系列的条例以规范在社交网络上分享文章的行为,也可以制定版权许可协议。

参考文献管理软件 (比如 Endnote (Thomson Reuters)、Flow (Proquest)、Pages (Springer)、Zotero 等) 同样允许用户将他们的研究图书馆分享给其他人,但是比较有特色的是这些分享是内在的一对一或者一对小部分人群,或是对能够分享内容用户人数有着明确且强制性的限制。

4.5 移动接入和应用程序

在更强监管、问责和由不断增长的数据导致的超负荷环境中,各行各业的专家都在压力越来越大的条件下从事更加复杂的任务。毫无疑问,在这种情况下,

¹¹⁰ 尽管是“edu”的域名,但他是由风险投资支持的商业性创业项目。

通过移动设备获取信息、工具和服务具有创造巨大效益的潜力。

即使以互联网时代的标准评判，移动计算机设备在大众人群中的普及速度一直以来都是极为迅速的。2011 年售出的智能手机数量已经超过了个人电脑的数量，在 2013 年平板电脑销售的数量超过了个人电脑。移动设备的使用数目在 2013 年也开始全面领先于个人电脑，手机和平板电脑的使用者超过 20 亿，而个人电脑仅仅有 15 亿 (Meeker, 2014; Blodget & Cocotas, 2012)。不足为奇的是，全球移动网络通信的发展也十分迅猛，同比一年前 14% 的市场占比，2014 年 5 月已达到 25%。

专业人群对移动互联网设备的接受度甚至比大众更快。例如，美国的医生超过 80% 的人使用智能手机，62% 拥有平板电脑，超过 90% 的医生利用手机或者平板电脑在患者中交流获取临床信息 (Wolters Kluwer Health, 2013)。在美国的护士中也有类似比例的人群使用移动设备，65% 的护士每天在工作中出于专业目的使用移动设备超过半个小时，20% 的护士甚至使用超过两个小时。2014 年 Wolter Kluwer 的护士药物手册移动设备软件发布，在线下载量已经超过 45 万次 (Wolters Kluwer Health, 2014)。

与大多数学术研究人员相比，繁忙的专业人士对成本/效益比更为重视，但在学者中移动设备的使用量也在快速上升，这种增长主要来自平板电脑的使用。截至 2014 年，一流 STM 平台上虽然只有大概 10% 的流量来自移动设备，但这个数字每年都在快速上升。调查显示，通过台式机、笔记本电脑获取在线文章较平板电脑或手机更受欢迎；但在医学领域移动设备的使用率较学术领域的高，但整体还在较低的水平 (Inger & Gardner, 2012)。

移动设备的广泛使用方兴未艾 (Outsell 2012b)。第一代的应用程序只能简单地用于获取信息 (即呈现内容)，而不是让用户在他们的工作流程中获取什么 (即做一些事情)。所以 STM 出版商最初满足的核心需求是“找到并保存”，即寻找事实和少量的信息，并使用户通过 RSS 或 eToC 的推送了解新的进展。临床计算器具有更多的互动性，但其作用基本类似。

目前大部分兴趣集中在迅速扩大的平板电脑市场，尽管平板电脑的使用还在继续扩大，但其中的许多应用程序同样非常适用于智能手机，比如医疗点药物信息在理想情况下应是通过“随身携带”的一种装置传递的。

平板电脑的其他用途还包括长篇阅读、更深度的自学和其他教育上的应用、对研究内容的积极参与 (尚处于起步阶段，可以添加注释和标注亮点，管理参考文献，添加文献标签，或许还可创建演示文稿或其他新的内容)。将来还会增加与工作流程和企业系统集成的移动设备应用程序 (如病历和电子处方系统以及类

似的其他应用程序)。

使用移动应用程序与使用个人电脑上网获取期刊内容还有一个很大的区别。移动设备属于个人的，很少共享，数据的使用是个人行为，而不与接入网络的机构（通常是通过 IP 地址控制访问）发生联系。应用程序（App）环境允许更多的用户与应用程序/内容的交互数据（在用户允许的前提下）得到收集，并且 App 的生态系统（即设备+云存储+App 商店等）使人们只需要简单的点击（包括应用程序中内嵌的按键或链接）就可以使用个人的信用卡（通过 App 商店）而不是图书馆预算来完成交易。

移动应用的商业模式与其使用情况一样，仍处于发展阶段。学术期刊出版商已经将提供移动设备接入的（必要）附加服务作为“默认”的选项提供。移动用户的不断增加是招揽个人和会员订阅用户的新契机。移动设备订阅量正在增长，同时给个人和会员订阅提供了一个新的机会。报告显示(Edgecliffe-Johnson 2012)，医学期刊平板电脑版本的广告合作机会明显多于网站版本，因此移动设备上的广告费和广告续约价格也更高，这表明平板电脑可能为含有广告内容的期刊转变为全数字化版本提供了一个途径（学会会员可能借此放弃订阅印刷版）。在一般公众移动应用市场，“程序内购买”（in-app purchases）的收入远远超过购买或订购应用程序的收入，这种模式将在 STM 出版中具有很大的潜力（如可以用于销售个人出版物、文章的附加章节或参考文献等）。

有一些重要的技术可以为出版商解决移动接入和应用程序重叠的问题，这些内容本报告将不予阐述。在写作方面，许多大的 STM 期刊和其数字出版平台提供移动优化界面，如使用响应或自适应设计¹¹¹。对于应用程序开发，出版商必须在本地应用程序（为每个设备编写开发语言）、Web 应用程序（使用开放标准特别是 HTML5 编写）或混合应用程序之间进行选择（结合网页内容和本机代码编写）。目前本地应用程序依然提供最佳的用户体验（如最快的访问速度和响应能力以及与设备功能最紧密的集成），而标准化的 HTML5 版本一经开发就可以适用于所有浏览器，其开发和维护成本相对较低，发展前景广阔。

4.6 研究数据

数据密集型科学

计算机、网络、各种自动传感器和研究仪器所产生的海量数据已在科学研究中产生了数据爆炸。这不仅造成了数据管理问题（就如“大科学”项目中需要在实验室内进行的化学基础研究中的问题一样），而且有可能改变科学研究的方式。

¹¹¹ http://en.wikipedia.org/wiki/Responsive_web_design

在传统的科学研究方式中，研究人员首先需要提出假设，再收集必要的的数据验证假设。在数据密集型科学研究中，丰富的数据可以用来创建和检验新的假设。已故的微软研究员 Jim Gray 认为，继早期的理论、实验、计算机模拟三步科学研究范式后，这是一个全然不同的崭新的科学研究方法，可以称作科学研究的“第四范式”（Lynch, 2009）。

对期刊出版的影响

尽管研究数据是多种多样的，但其对 STM 出版潜在的影响是巨大的：

- 研究人员越来越希望拥有访问基于期刊文章结果数据（机器可读）的权限，以对数据进行个人探索并大规模挖掘数据。出版商、数据库以及各种研究团体都需要对期刊（例如补充材料类文件）以及数据库中保存的数据各自的角色达成共识。在绝大多数情况下，这些数据保存在管理得当的数据库中更合适。
- 在期刊文章和研究数据库中需要有一个双向的链接。目前除了下面列出的措施，已经有针对这一领域的一系列研究项目，如 SURFshare 和 OpenAIREplus（见数据存储部分）。
- 数据集将会成为一个带有质量控制和归属性的主流出版形式。当以上情况实现的时候，数据库或许会更像一本期刊（反之亦然），因此需要设置同行评审（编辑和编委会、审稿人等）。至少有两种可能的商业模式：一种是简单地对相关期刊文章的同行评审进行质量控制；另一种是向用户提供服务的会员模式（例如 PANGAEA）。见后面的“数据期刊”部分。

研究数据应该公开给其他的研究人员（在一些特定的情况下受到限制，如为保护病人的隐私）已被广泛认可。STM 出版者已经在 2007 年 Brussels 宣言¹¹²中声明同意这项原则，并在 2012 年的 STM/Datacite 声明中报道¹¹³。

研究数据联盟

于 2013 年创建的研究数据联盟（RDA）目前已经取得了良好的进展（Treloar, 2014）。RDA 拟在基础设施层面（基础设施不仅包括硬件也包括软件、内容、格式标准以及参与者）改善跨边界的数据互通性（如国家、学科、生产者/消费者之间）。它通过一系列的工作组进行操作，每个小组在较短（18 个月）的时间内完成制定政策和技术规范等任务。RDA 兴趣小组有更宽泛的任务和更长（无限

¹¹² <http://www.stm-assoc.org/public-affairs/resources/brussels-declaration/>

¹¹³ <https://www.datacite.org/joint-statement-stm-and-datacite>

期)的时限,并可以建造新的工作组以及协调多个工作组的活动等等。研究数据联盟/世界数据中心系统(RDA/WDS)的出版数据兴趣组¹¹⁴是对STM出版商最重要的工作组之一,它的相关工作组包括数据引用、出版数据文献计量学、出版数据服务、出版数据 workflow 以及 BioSharing 注册小组。

数据存储库

数据储存库的数量已经有了迅猛增加,其中主要有两个目录:一是列有 995 个数据库的 Databib.org;二是超过 1000 个数据库的 re3data.org,据报道每周平均有 10 个数据存储库注册添加到其目录处。

数据存储库已经发展到能够容纳“孤儿数据”,即那些不属于(目前)公认的学科知识库的数据集。例如 DRYAD¹¹⁵,专门用来存储基础和应用生物学领域经同行评议文章的相关数据;另外 Zenodo¹¹⁶是与 OpenAIRE 相关联的一个欧洲项目;还有 figshare,它是数字科学领域的商业服务器。

除了主要的孤儿数据存储库,机构数据库提供数据存储服务(通常由高校图书馆管理)已经越来越普遍了。

数据引用

自本报告的上一版发布之后,数据引用和链接方面取得了实质性的进步(已成为多方合作的一个很好的案例)¹¹⁷,由此促成了“数据引用联合声明”的发布¹¹⁸。Callaghan(2014)与 Murphy(2014)均很好地阐述了上述相关的发展情况。这个声明为数据引用的目的、功能和归属列出了 8 项原则:重要性、信誉和归属性、证据、唯一标识、使用权、持久性、特异性和可验证性、互通性和灵活性。

引用数据的优点是使数据更容易被发现、增加研究的透明度、鼓励研究人员共享数据从而有助于重现性、允许数据引证计量的创建或提供信誉。

在期刊文章中引用数据集是相对直接的,因为它只是简单地延伸已有的引用实例。它对同时建立的反向引用也是很有价值的,比如:将数据库与引用它们的文章相关联。目前这种关联是人工实现的,期刊只是简单地通过发电子邮件提示存储数据库被引用,即使其中一些工作流程可以实现自动化,但是它显然是不可

¹¹⁴ <https://rd-alliance.org/groups/rdawds-publishing-data-ig.html>

¹¹⁵ <http://datadryad.org/>

¹¹⁶ <http://zenodo.org/>

¹¹⁷ 在这一领域较有成就的机构包括: CODATA Task Group 的 Data Citation、FORCE11 数据引用合成小组、DataCite、数字管理中心(Digital Curation Centre)等。

¹¹⁸ <https://www.force11.org/datacitation>

伸缩的。为了解决这个问题，人们提出让注册中心在期刊和数据之间充当中介。

数据引用还面临其他一些挑战：

- 动态数据：有一些数据库并不是静态的，而是持续更新的。由于理想的状态是数据应该是最新的，这为科学记录的固定性带来了问题。RDA 的数据引用工作组正在研究可能解决上述问题的技术，并颁布了一些指南。
- 微小-归属和微小-引用：主要是指投稿者对大型数据库中的一小部分内容负责的状况。

研究人员似乎确实越来越多地引用数据库中的数据。2014 年的一项研究交叉比较了 Scopus 文献列表和 Databib 数据库，发现从 1996 年到 2013 年数据库被引用次数每年增长 19%（复合年增长率），在 2013 年达到了 30000 篇次（Huggett 2014）。

数据共享与期刊政策

随着期刊论文的开放获取，研究的资助者在批准公开分享研究数据中起到了重要作用。资助者出台了一些政策（或者说强化了现有的政策），要求储存与共享研究数据。比如，美国国立卫生研究院（NIH）的数据共享政策，期望数据共享的时间“不迟于从最终数据集中得出的主要结论被接受发表的时间”，同时要求研究人员在所有的基金申请书中必须包含一份数据管理计划¹¹⁹。其他主要强制要求数据共享的指令包括美国白宫科技政策办公室（OSTP）备忘录和欧盟的研究与创新框架计划——展望 2020（Horizon 2020）。数字策展中心（Digital Curation Center）的网站提供了对英国主要研究资助者的数据政策的概述¹²⁰。

虽然研究数据可获取性的益处已经众所周知，但是在实践中仍有许多障碍。研究者们通常不会积极主动地共享他们的数据。2014 年的一个调查发现有一半的研究者进行了数据共享（这些数据主要是提供给期刊的补充材料），但是他们列出了很多种不情愿分享数据的理由。最常见的原因是涉及知识产权或者机密（占 42%），资助人/机构不允许共享（36%），担心研究数据被盗取或者误解/曲解（26%），另外还有其他 8 个方面的原因（Ferguson, 2014）。

在另一方面，进行数据共享的重要原因包括它是研究团体的标准做法，能够增加研究的影响和可见度，同时也是期刊的要求。正如第三点所建议的，越来越多的期刊采取政策鼓励或要求将与文章相关的研究数据共享，同时这也被越来越

¹¹⁹ http://grants.nih.gov/grants/policy/data_sharing/data_sharing_guidance.htm#

¹²⁰ <http://www.dcc.ac.uk/resources/policy-and-legal/overview-funders-data-policies>

多的期刊纳入规定 (Sturges 等, 2014)。出版商可以帮助建立研究团体的规范, 并在他们的政策中反映出来, 但是强硬的数据政策可能需要仔细的介绍与沟通, 这在 2014 年初人们对美国科学公共图书馆 (PLOS) 数据政策介绍的反馈中有相关说明 (Bloom, 2014)。其中一个方法是在提出要求的同时列出奖励政策: 比如, 美国科学公共图书馆 (PLOS)、加利福尼亚数字图书馆 (CDL) 和新墨西哥大学图书馆 (DataONE) 一起协作来制定数据的衡量指标 (与文章的计量指标相似), 从而对数据的使用、浏览和影响力状况提供反馈信息 (Lin, 2014)。

数据期刊

随着人们对研究数据的发现和再利用的兴趣增加, 以及研究人员公开他们的研究数据并因此获得公众认可的机会越来越多, 越来越多的数据期刊开始创建。

传统的期刊与研究数据的链接 (或嵌入) 是在我们熟悉其语言结构的科技论文中进行的, 数据期刊则为“数据文章”或者“数据集文件”提供了一个出版的平台 (通常是可以开放获取的), 这些数据文章主要是一些短的可以提供数据集的技术描述性文章。一些数据期刊也出版它们自己的数据库, 另外一些数据期刊则与专用数据存储库中的数据集链接。那些与外部数据集链接的数据期刊通常对第三方托管有最低要求 (例如 *Geoscience Data Journal* 要求其数据中心必须可以生成 DOI)。

表 12 列出了一些数据期刊。数据期刊能否保持为一种有别于传统期刊的期刊类型并不明确, 因为它们的功能可能会被现存的期刊所取代 (通过扩充其文章类型以包含数据文件), 或者通过其他方法如进行分层描述、扩增数据存储库中的元数据以及归属/奖励机制。

其他进展和计划

在研究数据领域其他的进展和计划包括:

- DataCite¹²¹ 于 2009 年 12 月推出, 用以解决研究数据可视化和可获取性的问题。它为数据集分配永久标识符 (目前是基于 DOI) 并提供注册和搜索服务。
- 目前已建立的新的查询工具: 汤森路透 (Thomson Reuters) 于 2012 年 10 月推出的数据引文索引 (Data Citation Index, DCI)¹²² 支持数据的发现、再利用和解读; OpenAIREplus 跨多个存储库从可开放获取的数据集中获得和索引元数据; DataCite 元数据搜索工具可对其注册了的数据集进行查询。

¹²¹ <http://datacite.org/>

¹²² http://wokinfo.com/products_tools/multidisciplinary/dci/

- “为数据交换创造机会”项目（Opportunities for Data Exchange, 简称 ODE, 是英国永久获取联盟的一个项目）¹²³ 针对期刊提出了一系列建议, 包括对提供基础数据实行更加严格的编辑政策, 推荐数据档案, 为使用持久性标识符的数据提供引用指南, 创办或赞助数据期刊。

表 12 数据期刊（举例）

期刊名称	出版商	网址
生物多样性数据期刊	Pensoft	http://biodiversitydatajournal.com
数据论文系列期刊（11 个主题期刊）	Hindawi	http://www.datasets.com/
地球系统科学数据	Copernicus	http://earth-system-science-data.net/
生态档案——数据论文	Ecological Society of America	http://esapubs.org/archive/instruct_d.htm
F1000 研究	science Navigation Group	http://f1000research.com
基因组数据——短小数据文章	Elsevier	http://www.journals.elsevier.com/genomics-data
地球科学数据期刊	Wiley	http://www.geosciencejournal.com
大数据	BGI/Biomed Central	http://www.gigasciencejournal.com
国际机器人研究期刊	SAGE	http://ijr.sagepub.com
开放数据系列期刊（4 个期刊）	Ubiquity Press	e.g. http://openarchaeologydata.metajnl.com

其他的一些引人注目的计划包括：BioSharing, 旨在期刊、具有数据分享政策的资助者和生物科学的标准化之间建立稳定的联系；BioDBCore, 是一个由团体定义的、非正式的、通用的, 针对生物学数据库核心属性的描述；ISA Commons, 是利用通用 ISA-Tab 文件格式进行数据分享的开放源代码框架。

4.7 语义网和语义富集

将“语义网”与语义技术的使用和内容的语义富集加以区分并不困难。

语义网是通过结构化的方式标记网上发布的信息（包括文章和数据），对可被计算机自动提取的语义进行编码。最初通用语义网的正式通用概念是由互联网创始人 Tim Berners-Lee 爵士提出的，至今看来仍然是复杂、昂贵且难以实现的愿景，但更务实的域名绑定手段已经使 STM 出版乃至整个网络显著增值。

¹²³ <http://www.alliancepermanentaccess.org/index.php/community/current-projects/ode>

语义富集就是对内容添加标记，通过识别相关的术语和概念并链接到一个特定的按结构分类的域中，以提高内容的附加值。虽然这可以通过人工完成(例如：索引)，但在实践中要实现语义富集在STM中的大规模植入与应用还需要自动化工具的使用。

这是一个包含多个步骤的过程：

- 元数据（具体到每一个域）的自动提取：识别术语并将其对应到已定义的实体（也是具体到域）；（实体（entity）表示数据库中描述的现实世界中的对象或概念。——译者注）
- 定义实体之间的关系（例如，状态 X=疾病 Y 的症状+药物 Z 的副作用）；
- 创建文档内部和跨文档实体之间的链接，以建立一个结构化的知识基础；
- 使用分析来获得新知识。

语义富集的优点见表13，主要表现在以下三大方面：

- **更智能的内容：**语义技术的一个重要价值是改善搜索和发现，为相关资料的发现、新领域的开拓、研究范围的拓宽等提供了更有效的新方法。即使文档中不存在搜索术语，用户应用语义分类依然能够找到相关的内容。将内容与用户的兴趣进行匹配关联可以实现个性化内容的推荐。用户可以自行设置感兴趣的内容，但运用自动化技术可以使工作变得更为高效，能够对已阅的文章内容、行为数据或来自于协同过滤的内容等进行语义或统计分析。
- **提供新产品与新服务：**例如，将广告内容与用户的身份和/或显示的内容进行更密切的匹配，这已被证明可以增加实际点击量并因此提高收益。按照语义界定进行内容分组能够动态地更新内容匹配并生成新的订购产品。
- **内部生产力：**出版商也可以利用语义富集技术提高其内部编辑和生产流程的自动化程度，例如可以（半）自动地进行编辑标记或向审稿人提供审稿建议。

采用这些语义技术也将有助于提升文本和数据挖掘技术。这实际上将已出版的文献转移到结构化的数据库中。除了技术上的挑战和版权问题，可能还需要新的商业模式来支持该技术（见“文本和数据挖掘”部分）。

数据链接

（开放的）数据链接是在互联网上以结构化方式发布数据的一种方法，这种方法有利于数据与其他资源的互联，进而使数据更加有用。使用标准的 Web 技术构建的数据链接可以呈现出事物之间的关系，极大地方便了多种信息源数据之

间的导航和整合。开放数据链接除采用“开放”许可协议外也同样允许共享、扩展和再利用（Miller 2010）。

表13 语义内容富集技术发展为出版商带来的利益（来源：TEMIS/Outsell; Outsell 2012a）

领域	服务案例	获益
更智能的内容	搜索引擎优化 分面搜索 链接 推荐 个性化	提高使用率 降低市场成本 提高更新速度 增加交易收入 作者认知
语义衍生产品	语义（靶向）广告 知识库 主题页 收藏和“切割”插件	新的收入来源 增加收益 现有资产再利用
工作流程生产力	自动内容处理（如：标记、链接） 内容查找 同行评议 审稿人推荐	降低成本 增加相容性 可扩展性 缩短发布时间

对出版商来说，开放数据链接是使出版内容更容易被发现并增加新服务使用的潜在方法。因此，它等同于自动提供元数据给 A&I（摘要和索引系统）和图书馆系统供应商。该技术的更强大之处在于，它能够使第三方（如图书馆人员或应用程序开发人员）整合多种资源以创建新的服务。

目前主要的搜索引擎似乎已经由完全致力于链接开放数据（虽然微软的学术搜索利用其提供一些其他的服 务，如搜索结果的图形可视化）转向另一种更简单的、使用 Schema.Org 规范的“微数据”来使数据结构化。（Schema.org 是 Bing、Google 和 Yahoo 在 2011 年 6 月 2 日发起的一项倡议，以“创造并支持将一组通用模式作为网页上数据结构化的标记”。——译者注）

有人认为数据链接缺乏能够使其得到更快普及的“杀手级应用软件”。尽管如此，它仍然得到主要图书馆的支持，包括美国国会图书馆、大英图书馆和联机计算机图书馆中心（OCLC）。

出版商也开始挖掘开放数据链接的潜力：例如，自然出版集团通过开放数据链接发布了其论文的元数据，Elsevier 也在探索它的使用，Thomson Reuters（Web of Science 的出版商）也在某些领域支持数据链接（包括 OpenCalais 服务）。

Carol Goble等人则认为数据链接在科研中的应用还不充分。他们提议互联数据发布后进一步从更多层次添加相关信息，以支持数据的出处、质量、信誉、归属并提供重现数据的方法，使得数据集成为“一流的研究对象”（Bechhofer et al., 2013）。

4.8 新的文章格式和特征

出版商和其他机构不断地创造和研究一些潜在的新方法来探索、展示、格式化和共享网站上的文章及其相关内容。近期的一些进展包括：

增强型 HTML 格式：比如 Elsevier 的未来文章（Article of the Future）框架、Wiley 的 Anywhere Article 和 eLife 的镜头格式（Lens formats），还有其他的出版商推出类似的文章格式。这些格式都是基于研究科研人员怎样使用在线文章而建立的，其主要目标是改善和简化用户体验，例如，通过划分屏幕区域以便让用户能够同时看到图片或参考文献。使用 HTML5 重新设计网页布局的另一个好处就是这种模式能够自然友好地支持移动设备阅读。

- **增强型 PDF 格式：**由于研究人员通常更倾向于使用 PDF 格式（尤其是利用其进行本地存储和注释等），已经开发了功能更丰富且具有网络链接功能的新的 PDF 版本，其中 Readcube 和 Utopia Docs 是最好的两款 PDF 阅读软件。在这一领域的出版商拥有大量的 3D 信息，如地球科学、地球物理学、地球空间学、工程学、医学扫描等领域都开始采用 3D PDF 格式。
- **文章版本：**数字出版平台的发展使得文章的版本更新或得到扩展，但是必须严格保留文章的原始版本和出版记录。尽管许多出版商平台已经在支持这一技术，PLOS 和 F1000 研究中心仍在积极地探索这一领域。
- **eLife 的 Research Advance 文章类型**是指允许研究人员对基于原始研究性论文的新成果进行发表的一项功能，从而使他们能够快速而有效地报道研究项目的进展，而无需另写一篇完整的新文章。
- **动态（实时的）图：**如果以平面图片的形式发表图数据，使得再利用其中包含的数据变得比较困难或者不太可能；另外也可将文章中的数据以动态生成的图像展示出来。F1000 研究中心就是利用这一想法允许用户获得生成图像的原始编码，举例来说，用户利用不同的参数作图就可以看到不同的结果展示¹²⁴。
- **数据可视化：**目前有大量的文件格式来储存实验数据，可以通过应用可视化

¹²⁴ 使用这种格式的文章见 <http://f1000research.com/articles/3-176/v1>

工具来加强文章的补充数据有效包含数据集的功能。一些出版商(比如 Taylor & Francis、NPG、PLOS)通过 Figshare (图形、表格检索数据库)提供的服务来存储并使数据可视化。

- 文章查看和共享: ReadCube 的内容共享计划使得非订阅用户通过一个链接即可分享需要订阅才能看到的内容。目前这项服务仅被 ReadCube 的姐妹公司自然出版集团所采用。
- 微文章: 这是 Elsevier 开发的一种新型短论文的名称, 它旨在使作者发布有用的数据, 描述实验方法或其他有价值的研究成果(包括一些中间过程以及无效/阴性的结果), 这些结果可能一直不会被发表。
- 地理标记: 在众多的研究领域如考古学、流行病学、环境科学和地球科学等都包含了具体的定位信息。直到现在, 定位一个研究所涉及地域的唯一方法仍是关键词搜索, 而这种方法是不精确的且具有偶然性。基于地理标记的搜索具有搜索精确性、借助地图界面(比如谷歌地图)和其他一些优点。比如 JournalMap, 这是一个科技文献搜索引擎, 它基于地理位置和生物物理学属性, 并结合传统的关键词搜索来搜索研究论文; 又如 Elsevier 的 Geofacets, 它面向地学科学家提供含有来源出版物相关内容的同行评审地图。

4.9 文本和数据挖掘

文本和数据挖掘(TDM)有可能改变科学家利用文献的方式(*Nature*, 2012)。随着数字化资源的增加、计算机运算能力的不断提高、易用软件的增多以及获取内容渠道的扩充, 可以预见 TDM 技术的重要性将会不断提高。在出版研究联盟协会的报告“文本挖掘与学术出版”中对 TDM 有很好的介绍(Clark, Jensen, Campbell, 2014; Smit & van der Graaf, 2011; JISC, 2012)。

TDM 利用自然语言处理技术和信息提取技术, 从大量文本内容中识别不同模式并发现新知识。语义富集和内容标签可能会增强 TDM 的功能。目前 TDM 应用最多的是生命科学研究领域, 尤其是制药公司内部, 在其他领域的使用相对较少。

限制其更广泛应用的主要障碍是法律上还不确定什么是被允许的, 并且缺乏一个有效的许可制度(见“文本和数据挖掘的权利”部分); 还有技术问题, 如标准的内容标准, 包括基本的共同本体; 内容聚合的需求以跨出版商挖掘语料库; 挖掘所需的成本和技术问题; 出版商方面对 TDM 技术缺少了解。国际学术出版协会(ALPSP)的报告中对于上面的最后一点指出: “调查发现大量出版商较少或者没有理解文本挖掘, 从他们的调查反馈中发现从来没有客户要求他们进行文

本挖掘”（Inger & Gardner, 2013）。

以下是针对上述挑战的一系列解决措施：

- STM 出版商们在 2013 年 11 月份发布了一份声明，承认签署了 STM 样品许可条款，另外允许在无额外成本的条件下将非商业化的 TDM 技术应用于订阅内容，开发内容的自我更新技术，并开发一个可供实现 TDM 的跨机构整合平台（STM, 2013b）。
- CrossRef（一个致力于推进在线学术资源相互链接方便性和有效性的出版者协会）的文本和数据挖掘工具（最初的设想）¹²⁵：提供元数据应用程序接口（API）和相关服务，能够为出版商全文本提供 TDM 工具的自动链接，以及存储元数据里许可信息的机制，还有可选择的速率限制机制以防止 TDM 工具压垮出版商的网站。
- 版权结算中心（Copyright Clearance Center, CCC）正在为生命科学公司试行一项新的服务¹²⁶。即从利用标准元数据进行单一服务的多个版权所有者手中聚集文章的内容，并且授权商业性生命科学组织的研究人员可以获取并下载需订阅获得的内容。因此在向版权所有者支付将其内容用于 TDM 的版权费时，该系统可能会降低一次性授权谈判的必要性以及相关的管理费用。
- Infotrieve（一个国际文献检索门户网站）新开发了一种服务，允许研究人员对全文进行搜索并将多个出版商的内容进行集结建立用于文本挖掘的语料库¹²⁷。Infotrieve 后来归 CCC 所有，因此这项服务已经整合到 CCC 的其他服务当中。
- PLSclear（一个出版商信息检索网站）是一个网络服务项目，旨在简化访问出版商的内容以用于 TDM 的流程，它是一个用于科学研究的在线情报交换场所¹²⁸。

4.10 可重复性

在期刊上发表的科学研究缺乏可重复性是一个越来越严重的问题（有时也叫“可重复性危机”）。这一问题在2013年作为封面文章发表在《经济学家》（Anon,

¹²⁵ <http://tdmsupport.crossref.org/>

¹²⁶ http://www.stm-assoc.org/2013_05_20_FACT2_Billington_CCCs_Text_and_Data_Mining_Pilot_Service.pdf

¹²⁷ <http://www.businesswire.com/news/home/20140707005558/en/Infotrieve-Spearheads-Text-Mining-Initiative>

¹²⁸ <http://www.plsclear.com/Pages/ClearWizard.aspx>

2013) 后开始受到广泛的关注, 但过了很久才有人资助这项研究的继续进行。NIH已经率先制定政策来解决这一问题 (Collins & Tabak, 2014), 包括更好地培训调查人员、更加系统地评价基金申请书、让研究数据更加透明, 建立一个数据发现索引 (Data Discovery Index) 项目并更严格地执行数据分享的各项规定, 推出PubMed Commons项目来支持对发表文章的公开讨论 (见 *已发表内容的同行评议* 部分)。

可重复性是一个在研究过程、组织和文化中根深蒂固的复杂、多维的问题, 但也受到出版活动某些方面的影响。这些影响包括争取提早发表及其所带来的压力、选择性发表阳性结果和同行评审过程中统计分析能力薄弱所带来的问题。

出版商和期刊已在许多方面采取相应措施解决上述问题, 包括介绍或强制实施对试验对象进行注册的政策; 介绍数据存储和共享方面的政策 (参见 *数据分享和期刊政策* 部分); 鼓励或要求分享计算机代码和研究数据; 加强同行评审, 比如在评审期间采用评审专家备忘录并尽可能多地聘用统计学专家。

其他出版商采用的方法还包括允许发表阴性结果, 将注册政策由现有的临床试验研究扩展到其他类型研究中¹²⁹, 对试剂、抗体等这样的词语实体进行语义标记以确保对其进行准确的唯一标识。

4.11 大数据和分析

“大数据”指由于数量过大而无法用传统的数据库系统进行处理和分析的数据集¹³⁰。雅虎 (Hadoop)、谷歌 (MapReduce) 和亚马逊都开发了大数据处理工具, 为了满足搜索引擎和大型消费网站实时处理海量用户数据的需求。这种大规模的数据集来自网站本身、客户和用户数据 (如 Walmart、Facebook)、医疗保健数据、地理位置数据、设备数据 (“物联网”), 当然还有科学研究 (如欧洲核子研究中心 (CERN) 的大型强子对撞机每秒处理 4000 万张图片)。

麦肯锡 (McKinsey) 等咨询公司曾预测采用大数据技术将会给企业和社会带来巨大的经济利益。例如, 他们估计大数据每年带给美国医疗保健的年收益为 3 千亿美元, 通过个人位置数据获得的年度消费者盈余 (consumer surplus) 为 6 千亿美元 (McKinsey, 2011)。

大数据带来的挑战似乎还未降临到 STM 出版和信息供应商的头上, 但事实未必如此。Elsevier 出版的 *Research Trends* 专辑讨论了科研领域的例子, 其中包

¹²⁹ 例如Elsevier的期刊*Cortex*开始发表一种新格式的文章, 叫做注册报告, 见 <http://www.journals.elsevier.com/cortex/news/registered-reports-a-new-article-format-fromcortex/>

¹³⁰ 数据量的大小尚不能确定, 目前公认在 PB (前兆字节) ~EP (艾字节)。

括在制定科技政策、科研投入和文献计量分析等方面对大数据的利用（Halevi, 2012）。

多种不同类型的 STM 数据都可以适用于大数据技术，包括研究数据、全文集合数据（即文本挖掘）、元数据（包括引文）、使用和行为数据。以下为 STM 领域大数据的一些具体实例：

- PlantWise 是由国际农业和生物研究中心（Centre Agriculture Bioscience International, CABI）发起的旨在提高食品安全和通过减少作物损失提高农村贫困人口生活品质的项目。CABI 从野外植物保护站（植物医院）整理包括虫害、病害信息以及其他信息的数据，并通过扫描上传到中央数据库。然后将这些数据与来自其出版物和第三方的信息合并。通过利用自身的 CAB 词库，提取信息并存储为语义结构化数据。这些数据和其他数据集的结合可以用于生成病虫害预测地图，并建立虫害危险指数。
- Elsevier 的 SciVal 能够通过分析海量的引文和其他数据来创建各国家相对竞争实力的图谱，这些数据涵盖了来自 220 个国家的 4600 家科研机构的研究成果。
- 期刊/数据库的 GigaScience 是 BGI（原北京基因组研究所，世界上最大的测序中心）与 BioMed Central 之间合作的产物。它将巨大的数据集结合到期刊论文中，并提供数据分析工具和基于云计算的数据资源。
- 数据挖掘已经在前面的“文本和数据挖掘”部分进行详细的讨论。这里举一个例子，Ariadne Genomics 公司（2012 年被 Elsevier 收购）为生命科学的研究人员（特别是在医药领域）提供从文献中挖掘信息的服务。
- Mendeley 使用大数据技术（Hadoop 和 MapReduce）对全文数据库（约 4.7 亿篇篇）与用户（约 350 万）之间互动所产生的大量数据进行处理。这可以生成有价值的统计数据并进行论文推荐，同时还能为机构提供服务，例如它能够帮助图书管理员了解自己的馆藏被读者使用的情况。

4.12 身份识别和消歧

在数字化的学术交流环境中，在不同的系统中准确地识别研究人员和他们所做的工作有以下几个难点：不同人使用相同的姓名；同一个人的姓名有不同的排列方式或译名；研究人员更改姓名（如结婚后更改姓名）。

虽然有不少措施来解决这一问题，但对 STM 出版来说最重要的是 ORCID（开

放的研究人员和贡献者 ID) 系统¹³¹。ORCID (组织) 是一个由科研人员和从事学术交流工作的人员 (包括大学、资助机构、研究机构、数据库、专业学协会以及出版商等大约 150 家机构) 参与建立的非营利性合作组织。它提供两种服务: 为每位研究人员建立并维护 ID 及相关联数据的一个登记系统; 提供一个 API (应用程序编程接口) 平台, 以支持跨系统的交流和认证。ORCID 的注册服务于 2012 年底正式开通, 当时注册了超过 110 万个 ID¹³²。

个人可以得到自己的 ID 并免费管理相关信息, 机构可以将其记录链接到 ORCID 系统, 获取更新信息, 并为其员工和学生注册 ID。

ORCID 的重要性不仅在于它可以解决简单的人名歧义问题, 它还提供了一种独一无二的有效方案, 用以鉴别个体贡献和研究人员间的联系网, 这将促进或改善一系列的服务, 包括研究分析 (见下节)、社交媒体、网络服务, 等等。例如, ORCID 囊括了一些新的服务功能, 如通过标识符、对重复工作更好的管理、一些社交功能, 以及可以输入 Bibtex (一种参考文献管理软件) 目录文件的功能来支持工作的分组 (Cochrane 2014)。

4.13 研究管理和分析

作为一个基于 STM 出版信息服务的新兴市场, 研究分析是链接到分析工具上的科研信息管理系统。设计者的想法是为学术机构及其科研管理人员、资助机构和政府提供一种深度评估科研项目质量和影响力的手段。分析工具使用的文献数据包括引文和基于早期粗略方法计算的数据 (如使用期刊影响因子), 并利用更复杂的数据分析和现今各机构内研究信息系统的整合 (CRIS, 也叫做研究信息管理或 RIM) 来评估科研产出的质量¹³³。CRIS 系统能够整合有关机构科研人员及科研小组的研究活动及成果, 并提取机构内部的信息, 如人力资源信息、财务信息、资助项目跟踪系统以及研究项目进展报告, 另外还可提取一些外部数据, 特别是发表的文献数据及其他一些外部的专有和公共数据集 (如专利或基金数据)。

在这个市场上活跃的 3 个主要公司是 Elsevier、Thomson Reuters 和 Digital Science。Elsevier 拥有 SciVal 分析工具套装 (由 Scopus 网络数据库支持), 并于 2012 年收购了丹麦的 CRIS 供应商 Atira 及其 PURE 服务器; Thomson Reuters 的 CRIS Converis (之前为 AVEDAS, 后来被 Thomson Reuters 收购并且和它的

¹³¹ <http://about.orcid.org/>

¹³² ORCID 的统计数据: <https://orcid.org/statistics>

¹³³ 欧洲国际研究信息组织 (euroCRIS) 每年都会举办会议并管理欧洲共同研究信息格式标准 (CERIF): <http://www.eurocris.org>

Research in View 融合)和 InCites 分析套装由 Web of Knowledge 数据库提供支持。拥有 Symplectic(英国的 CRIS 供应商)的自然出版集团的姊妹公司 Digital Science 公司也在这个新兴市场中有自己的一席之地。另外也有一些非商业性的国家级项目如 METIS (荷兰) 和 CRISin (挪威)。

这些公司提供的主要服务是基于订购的工具和服务(如分析科研项目的相对竞争实力、识别合作者、测度个人/团队的研究表现等); 另外还提供定制研究和分析¹³⁴, 以及为内部分析所需的数据进行授权等服务。

另一个相关的服务是由 Uber Research (另一家数字科技公司) 提供的, 该项服务通过与已受资助的研究项目(不一定是发表出来的)相比较, 帮助资助者决定应该给予哪个项目提供资金支持。

除此之外, 还有一个为公司的研究提供分析服务的独立市场, 比如在医药以及高科技工程领域, 本报告将不予具体阐述。

诸如制度比较这样通常被用来辅助管理与政策制定的计量指标, 被用于其他机构时也应该是可靠且具有可比较性的。Snowball Metrics (书目计量指标) 项目的建立就是为了满足以上要求, 它旨在创建与分享全球公认的研究相关计量指标, 为计算方法和相关的数据来源建立标准化“食谱”(Jump, 2014b)。

4.14 FundRef

FundRef是由学术出版商和基金会联合发起的一项实验项目, 由CrossRef推动, 提供一种标准化的方法来报道发表的学术研究资金的来源。实际上该系统由一整套标准化的代码所组成, 不同的代码代表不同的基金资助机构, 最初覆盖大约4000家基金会, 其后台系统通过DOI将文章与基金会联系起来。该项目旨在帮助基金会报道他们资助的研究及所获得的成果, 并更加系统地在出版物中报道基金资助的科研成果以及进行文本挖掘。

所有的元数据只有被广泛采用后它才会变得有价值。最近, 一些期刊已经要求作者在文章提交的过程中使用FundRef数据来获取其受基金资助的信息。其实, CrossRef统计数据显示已经有38600个DOI与FundRef中的元数据建立了联系。FundRef的使用有望快速增长, 因为科研基金会对它相当重视, 特别是它的使用对支持CHROUS来说十分必要。

¹³⁴ 一个有趣的例子是 Elsevier 为英国商业、创新和技能部所作的有关英国研究基地的国际竞争力的报告 (Elsevier, 2011)

4.15 图书馆出版

在过去大概4年的时间里，越来越多的学术图书馆开始提供出版服务（Jones, 2014b）。2011年学术研究图书馆（ARL）的报道中强调了这一趋势，但只是将其描述为在探索中进化的领域（Ivins & Luther, 2011）。2015年的《图书馆出版目录》则展现了一幅图书馆出版的生动画面，它报道了对大约124个案例的研究，大多是在美国与加拿大（Lippincott, 2014）。大多数图书馆出版商（90%）与其本单位学术部门合作，但是超过半数为第三方组织如学术团体和科研机构提供出版服务，由此出版了大约194种期刊。

图书馆调查后发布了432个由其员工在办的校内期刊（由学生办的有214个），而且几乎所有的都是开放获取期刊，且这些OA期刊很少收取APC（只有10%），而用图书馆预算来支付出版的费用。图书馆出版的期刊总数很有可能大于期刊目录中给出的数字，如开放期刊系统中的数千种期刊，以及bepress数据库中的大约700种期刊，其中有很多可能是图书馆出版的。

加利福尼亚数字图书馆最近创办的开放获取综合性期刊*Collabra*，即是一种与传统出版业中的出版物相当的图书馆出版物的典范。大多数图书馆出版物将轻量级的出版服务和轻量级的技术解决方案进行结合，如Open Journal Systems, bepress, DSpace, Wordpress和Ubiquity Press。这些服务正变得越来越细致，包括元数据的分配（80%的图书馆出版商），同行评审管理（25%）和市场化（41%）。更重要的是，这些服务同样重视数据的可发现性，例如Primo和Summon均向网络发现服务器提供元数据。

4.16 开放注解

开放注解，是一种基于网络的开放注解形式¹³⁵，为对期刊内容（或其他学术著作）进行更丰富的评论和论述提供条件。

开放注解与形式较简单的注解（例如社交书签服务）有一些相似的特征，但开放注解支持更多的注解形式，包括添加书签、突出显示、标签与评论。注解既不需要加注网站的许可也不需要安装任何新软件。然而，出版商们会选择将他们自己的更具特色的开放注释服务提供给订阅者或注册用户。

注释不仅可以链接到网页文本文件，也可以链接到页面内的任何指定位置，对于句子它则允许更有意义以及更具互动性的注释。此外，注释和链接并不仅仅局限于文本内容，还可以支持在图片、地图和视频这样的非文本材料上进行注释，

¹³⁵ <http://openannotation.org/>

并且这些注解可以被引用和作为学术记录来保存。

当前提供注释服务的领先机构是非营利性的 Hypothes.is，它每年都会组织相关的研讨会（iAnnotate）。其他在学术领域提供注释工具和服务的组织包括 Annotator（开放知识基金会提供），Domeo（马萨诸塞州总医院提供）和 Pubpeer。可能也会被用于学术的通用型网页注释工具则包括 Genius 和 Diigo。

目前针对 STM 平台用户的强化版 PDF 阅读器可以提供多种可选择的（非标准的）注释分享方式，比如 Readcube、Utopia Docs、Colwiz 和 Mendeley 等。

4.17 学术团体

学术社团和协会的工作有很多任务和目的，主要是提升该研究领域或学科的水平，或是提高组织成员对学术的兴趣，或者两者兼备。长期以来出版期刊一直是他们的主要任务（还有举办会议、培训等），尽管学术期刊对提升学科水平仍然非常重要，但面对出版业格局的改变，学术团体期刊出版对社团成员不再有优惠，尤其是传统印刷出版的减少，期刊通过联盟许可模式即可获得，以及开放获取期刊（目前影响程度较小），都影响了期刊出版在学术团体中的地位。

学术团体成员仍然继续支持学术社团出版期刊。例如，Ithaka 的调查显示，学术期刊所扮演的两大重要角色包括“组织研讨会和其他面对面会议”以及“出版经过同行评审的学术期刊”（Ithaka S+R 等, 2013）。此外，许多其他的调查显示，期刊本身和社团之间的关系对于作者选择用于发表文章的期刊仅有很小的影响（NPG, 2014）。

社团中新的角色和成员利益是很难去实现的。网络社区和网络服务通常是作为社群的真实会议和关系网络的数字模拟而被推进向前发展，然而在此方面取得成功的社群还凤毛麟角。组织成员对其学术团体提供的这项服务往往不感兴趣，或许是因为实现更广泛的人际网络相较于学术团体某些局限性的服务重要得多，这一客观事实为 Academia、Mendeley 和 ResearchGate 这样的社交服务网站提供了更加开阔的发展机会（Ithaka S+R 等, 2013）。

如上这些问题（以及陈旧的管理结构）使得很多学术团体在出版期刊中秉承低风险、低创新的原则。尤其是在开放获取方面，如果学术团体中没有领导人和勇于创新的人，其成员经常会担心承担经济和信誉上的风险，或对团体品牌造成损害（Outsell, 2014a）。

这样的压力会促进学术团体期刊从自我出版向和出版商合作出版的模式转变。那些专门和学术团体合作的出版商则会开始找寻帮助学术团体改进会员服务的新方法（会员期刊订阅除外，例如 Wiley 和 Knode 合作的案例）。鉴于出版商

之间在与学术团体期刊合作市场中的竞争，我们将能够看到将来在这个领域会出现更多的创新。

4.18 作者服务及工具

出版商理所当然会给作者提供同行评审和编辑校对的服务，但是由于出版商对作者、研究的全球化（带来了大部分英语水平较低的作者）以及新技术的竞争日益激烈，面向作者的服务项目越来越多。包括投稿前服务、发表服务、信息和提醒服务、市场营销和推广服务、折扣和其他服务。具体如下。

投稿前服务 这项服务包括期刊选择工具（如 Research square 的期刊指南、CoFactor、Edanz 期刊选择器；Edanz 期刊选择器的定制版在 Springer 作者学术网站中可以使用），语言润色和翻译服务（大部分出版商都将这项服务外包，仅少数出版商自己承担，如牛津大学出版社），投稿前咨询和筛选服务，期刊信息页（向作者提供最新的同行评审时间和发表时间等信息）¹³⁶。

发表服务 重新绘制和重新标注文章中的图这样的人工服务现在已经很少，取而代之的是规范参考文献格式（去掉提交的文章中不必要的格式）、电子授权这类的自动化服务。

信息和提醒 跟踪并显示稿件处理流程，以及在发表后文章被引用的提醒服务。

市场营销和推广 针对作者对他们成果的可见度和影响力的最大化越来越多的需求，提供以下一些服务项目：文献计量及被使用的统计数据；促进作者提高论文质量的建议和工具/服务，以及类似 Kudos、Publiscize 和 ImpactStory 这样的集成服务工具；订阅内容以及可分享版本的免费链接（如 ReadCube/NPG）。

折扣以及其他服务 图书折扣是长期存在的；另外还有一些新的电子服务包括参考文献数据库的链接，编辑加工和审稿系统工具。

为了进一步提升为作者服务的范围和质量，出版商也会去寻找提升用户体验的方法（如简化流程，删去不必要的步骤，提高易用性等）。因此，在线投稿和追踪系统也是相当重要的一项服务。

4.19 合作创作和分享工具

尽管曾有段时间对为科研者提供专门的合作创作工具所可能产生的利益进

¹³⁶ 例如BMJ的网页（<http://heart.bmj.com/site/about/>）或Elsevier的网页（<http://journalinsights.elsevier.com/journals/0377-0257>）

行了一些讨论,但是微软 Word 软件的统治地位却限制了这一需求的发展(Perkel, 2014)。Google Docs 作为一款可以免费使用的软件,它让人们对在线写作的益处有了一些了解,但其并不能满足科研创作的需求。

新兴的网络公司 Overleaf (原先是 WriteLaTeX 公司,最近由 Macmillan 数字科技公司注资)提供了这样的一项服务,该服务宣称拥有 1000 多个机构大于 15 万的用户,并且已经通过他们的服务完成了 200 万篇文章的创作。

其他在线学术写作工具还包括 Authorea、FidusWriter 和 ShareLaTeX。Authorea 平台不仅可以提供写作服务,还可以进行出版服务¹³⁷; Annotum 则是可以基于 WordPress (可支持学术内容的扩展版本)提供写作、同行评审以及出版服务的平台。此外,像 Plot.ly 这样的网站则允许用户基于云服务器进行图片的合作创作;图片既可以在平台上被免费分享也可以加密后分享,以使用户获取图片相关的数据。

目前,这些服务仅被极少数的科学家使用。这个现状可能随着出版服务被认可和多项服务的融合而改变。例如,WriteLatex 提供给出版商一个网页服务以完成从 Overleaf 到出版系统的“一键投稿”。排版人员 River Valley 曾经开发过一个类似的服务 RVPublisher,主要向出版商推销。一些出版商也在积极探索这方面的应用:Elsevier 表示正在开发一款支持语义标记的作者工具,而 Wiley 则致力于开发可以从作者那里获取更多结构化信息的工具。

4.20 开放笔记本科学

开放笔记本科学(有时也称为开源研究)基于共享和合作,它将是比保密和竞争模式更加成功的理念,其灵感显然来源于计算机软件的开源运动,旨在共享所有研究成果,包括中间过程和详细的实验结果,而不仅是最终归纳在期刊中发表的论文。

开放笔记本科学只被极少量(微乎其微的)的科研人员采纳。在本报告的前两个版本中,我们不看好这种情况会发生迅速改变;果不其然,这方面几乎没有进展,并且一些试验已经暂停(包括之前列举的两个例子)。大多数研究人员都过于担心和保密知识产权;也担心研究结果被他人窃取而影响论文发表;并且,更根本的原因是不确定研究过程中的阶段性结果是否具有共享价值。然而,正如前文讨论,增加研究数据共享和再利用已成为一种主流核心观点。

¹³⁷ Authorea开发了一种有趣的互动型文章,在其网站上发布了一篇名为“未来论文”的文章对其特点进行描述: https://www.authorea.com/users/23/articles/8762/_show_article

5. 结论

我们试图每 3 年左右更新一次这个报告，如果借此机会回顾一下过去的 3~5 年，我们可以看出大量明显的趋势。

互联网已经成为学术交流的主要手段。虽然传播媒介发生着翻天覆地的变化，作者在学术期刊发表文章的动机以及他们对同行评议重要性的看法仍然没有大的改变。如果说有什么改变的话，对同行评议认证的需要更加强烈，这由竞争加剧、全球化、研究评估和资助机构日益强调的“影响力”所驱动。

社交媒介和网络的用户及使用在大众中快速增长，但仍需假以时日才能对研究人员的职业活动产生巨大的影响。新的科学社交网络（Academia、Mendeley、ResearchGate）近年来发展迅速，但目前很多用户创建帐户的动机似乎是为了增加显示度和扩大影响力，而不是为了活跃在线讨论或合作，或者将其作为更方便的参考文献管理形式。

广泛采用的智能手机已经产生了一定的影响，甚至平板电脑设备已经开始改变某些医生和其他医疗专业人员的工作方式，但他们也似乎对大多数研究人员的工作方式影响不大。

目前研究人员对学术内容的获取处于历史高位。将内容与相关的联盟许可模式进行捆绑销售继续提升着史无前例的获取水平，估计年全文下载量为 25 亿次，并且每次下载的成本处于历史低位（很多大客户支付的每篇文章的下载费用不足 1 美元）。

学术交流系统的全球一体化持续加剧，同时伴随着研究产出的加速增长。最令人瞩目的是东亚尤其是中国的论文产出的增加，中国已成为全世界第二大研究性论文生产国（并且某些学科领域的论文数量已经超过了美国）。印度和巴西研究基础的扩张也引人注目（与俄罗斯的紧缩形成了鲜明的对比）。

Research4Life 项目（HINARI、AGORA、OARE 和 ARDI）再一次持续扩大，内容资源的规模、广度及注册机构和用户的数量等进一步增加。第三世界的数字基础设施（和更广义的科研能力）仍然落后于西方，但该项目的成功意味着最贫穷国家的研究人员不会因负担不起订阅费而无法获取学术文献。

虽然获取和通过大单交易（Big Deal）传输的有价值信息的确在增加，但也面临着日益增加的经济压力，同时图书馆正致力于谋求更大的灵活性和控制力，更合理的定价模式和真正更低的价格。尽管经常有批判的声音，该模式似乎更可能进化（如：捆绑开放获取费用；与历史的印版收藏不关联的新定价机制），而

不是被替代。

尽管如此，通过 OA 提升获取已清楚地主导出版产业和政策的发展。从聚焦绿色 OA 的美国的 OSTP 备忘录到英国研究理事会的金色 OA 导向的政策，资助者和政府的作用仍然是中心。在上一版报告中我们看到一个明确的转向金色 OA 的政策，现在很清楚这一观点过于受到当时英国发生的事件的影响：在英国以外，奥地利和荷兰，占主导地位的政策立场更倾向于绿色 OA。一方面，几乎所有的投资者愿意补偿金色 OA 的论文发表费用（APC），但另一方面，他们的 OA 政策更倾向于绿色。

无论如何，OA 期刊在数量和产出方面在持续增加。估数有变化，但 DOAJ 的 OA 期刊清单超过 1000 种，并且目前金色 OA（包括无需 APC）的文章占 11%~12%，另外绿色 OA 至少有 12%，延迟 OA 占 5%。巨型期刊（megajournals）仍然是增长最快的部分，最近有各种出版商宣布创建举行期刊，其中有 Elsevier，也有加利福尼亚数字图书馆。

至于绿色开放获取，机构知识库的数量持续增长，ROAR 列出的知识库由 2012 年的高达 3000 个增加至目前的 3900 余个。尽管知识库在不断扩张，并且资助方和机构广泛强制推行，但学术界的很多人对于在机构知识库中进行自存储这样的个人行为仍然是兴趣有限（除了几个研究领域之外，共享论文预印本或工作文档已经是常态）。事实上，在一项研究中，诸如作者或部门主页的零散的、非结构化的自存储文档目前可能比机构知识库更重要。学科知识库（如 PubMed Central、arXiv（最近已有超过 100 万篇文章）、SSRN、RePEc 等）对于研究人员仍然更具吸引力；然而，无论是作为作者还是（或许更多地）作为读者，都一直担心出版商关注对订阅的影响。

财政预算的影响无疑使决策者更容易选择绿色 OA，但我们在本报告的前两个版本中所提到的对出版商信任度的降低以及对其盈利动机的怀疑也是一个因素。正如图书馆员 Rick Anderson（Anderson, 2014C）所说的，虽然对于学术交流（尤其是 OA）未来发展形式的辩论一直是缺乏确凿证据和夸大奇谈的，对于出版商的不信任依然成为导致选择绿色 OA 的越来越重要的因素。过于看重少数过激分子的态度是一种危险的行为，但目前（有全球性读者的）全国性报纸上高调地批评、作者的抵制、以及政府-出版商许可协议磋商的难度日益增加这些因素，都营造了一个（例如）学会回避与最大的商业性出版商合作的氛围。远离开放获取行动的一线（小部分活跃的科研人员所感兴趣的一项活动），出版商和研究人员继续卓有成效地在一起工作，即使这一点很容易在激烈的争论中被忽视。

研究数据的管理和共享，以及与研究文献的联系和整合，已经到了决定性的

中心阶段。研究数据联盟的建立是对重要性的一个可见的反映，研究群体正探讨这一话题，同时研究资助机构优先确保研究数据可以获取。潜在的利益是巨大的，正如有影响的皇家学会报告《科学作为一个开放的企业》中所论述。在研究和出版界都有具体的发展，包括在政策方面，数据仓储的扩展，以及包括数据期刊的出版创新。然而，这场革命仍然刚刚开始。在相关领域，也有开发解决方案的进展以促进文本和数据的挖掘，虽然这方面的需求（远离医药/生物技术领域）仍然刚刚开始。

在出版行业内部，并购活动显得很有活力。这部分反映了诸如规模经济和内容范围的长期结构性趋势（尽管这些已经被网络强化），如 Springer 和 Macmillan Science & Education 的合并计划，但也反映了技术创新的重要性日益增加以及从内容到内容加服务的转变（如 Elsevier 收购了 Knovel 和 Mendeley，汤森路透社收购了 AVEDAS）。

~~~

展望未来，我们希望在这不断变化的世界中保持一些永恒不变的东西，尤其是期刊的核心功能（注册、传播、认证和提供档案记录）。作者发表论文的核心动机确实始终没有改变，即：需要归属和认同，需要包括同行评议的质量控制，需要显示度和最广泛地传播学术思想。从“不加区分”这一品牌和轻视同行评议看，开放获取巨型期刊的成功似乎有悖于这些基本要素。但是，这些期刊还处在创办的初期，我们或许还可看到推动其发展的其他因素，如：作为更负盛名期刊退稿的论文的另一选择（级联审稿模式）、新作者入行的途径、或淡出为了品牌延伸（如 *Nature Communications*, *Royal Society Open Science*）。

其他一些趋势已经存在了足够长的时间，使人感觉如同整个发展图景的一部分：研究产出的数量和复杂程度都在不断地增长，并且越来越多地具有以数据为中心的属性。来自新兴市场尤其是中国、印度和巴西的研究产出将继续增长。由企业资助的研发（虽然不是基础研究）经费所占比例将继续长期上升，尽管它比公共资助更具周期性。实际上，传统的学术购置预算将增长缓慢（如果有的话），尽管从 2008/09 年的衰退中有一些迟来的恢复，特别是在欧洲，复苏似乎出现停滞（并且，国际货币基金组织最近警告说，全球经济增长可能再也不会回到危机前的水平了）。各类买家将寻求他们所购买的价值的展示（通过使用，或许以更复杂的计量驱使的方式）。新兴市场将继续提供最佳的增长机会，但与近期相比未来增长的速度可能会放缓。面对内容资源的过度丰富，读者将注重相关性、可用性、见解和答案，而非基本的获取。

尽管如此，OA 仍将是下一阶段 STM 出版的显著特征之一。这将是一个复

杂的过渡，肯定不会在未来几年内完成，并且也许仍然保持其复杂的市场地位。但作者、出版商和资助者采纳 OA 的势头明显，并至少将有一个短时期的加速发展。

研究界和出版界仍在探索金色 OA 的稳定的可持续发展的经济发展模式，具体来说：如何将对作者的资助精确地分配到研究经费、机构补助金、图书馆和部门预算以及其他来源的经费中？出版费的市场价为多少，目前的收费为 250~5000 美元（甚至实际为 0）？市场的力量会将促使 APC 总体降低吗，如果是这样，后果如何；或者如同其他竞争激烈的市场一样，价格将分层吗？

目前要求赋予开放获取内容再利用权的（如对文章的 CC-BY 授权，对数据的 CC0 授权）的做法今后还将继续，主要原因是同样的：研究资助机构的要求（正确或错误地视 CC-BY 及同类授权作为再使用的重要方面，尤其是对诸如文本和数据挖掘的新方法而言）；并且由于同样的原因赋予开放获取内容再利用权似乎可能会被承认，尽管在相当一部分研究人员中还不受欢迎。

绿色 OA 和知识库的作用将仍然与出版商存在争议。对机构知识库（尽管数量和存储的内容都在增加，但作者仍然利用较少，并且对读者发现和使用而言是一个相对不重要的渠道）的争论比对学科知识库（尤其是 PubMed/PMC）少。学科知识库的作用更像是一个出版商，投资建设一个出版平台并具备相关的功能（如数据集成、先进的文章格式、评论平台）。保密时间的长度也将继续为争论的主题，虽然我们不期望投入太多的行动。

开放获取也将不只是针对期刊文章：OA 专著模式正显露端倪并预期会发展。开放图书和开放教育资源等其他类型也将愈来愈重要。

并非所有的出版商都是平等的：ALPSP 的调查（Inger & Gardner, 2013）表明，小型出版商比大中型出版商的出版政策和措施明显缺少前瞻和创新性。学会出版社（最大者除外）仍将发现很难适应行业和研究团体的变化（包括开放获取），而其母体学会一直努力地与其会员进行联系。

图书馆的作用不断发展。虽然提供和支持获取学术内容的核心功能仍然存在，图书馆将在扩大机构知识库的使用（包括支持研究数据的本地管理）、管理 OA 资助和支付方案、以及图书馆出版的新兴领域发挥潜在的重要作用。

在消费领域，成功的新媒体品牌正在不断超越内容的“容器”（即类型的分类，如杂志、新闻、电视、书籍、音频等）。STM 的用户和购买者会有一个类似的转变（虽然不是新的转变），远离“容器”（期刊、专著）而转向研究内容本身及围绕内容建立的服务。这可能会导致综合内容许可的进一步扩大，并且也许会导致瞄准小企业或个人的“通吃”订阅服务的增长。

由印刷出版向数字出版的过渡当然仍将继续。虽然基本上所有的期刊都有电子版，但传统的印刷版仍在继续，尤其是图书。对于研究类期刊，几乎所有有意义的使用都已经迁移到在线（甚至是下载 PDF 后打印出来），以后的几年会看到这类期刊越来越多的放弃印刷版。对于学会期刊，尤其是临床和专业领域，以及一般性的期刊，至少有一些残存的印刷版的使用似乎可能转向新的移动终端设备。

虽然移动获取可能还没有对 STM 产生某些预期的影响，鉴于其在消费媒体的部分变革性影响，拓展移动设备的应用将是以后三年的一个重要趋势，特别是临床和医疗职业领域。它仍然处在早期阶段：通过移动设备对 STM 平台的访问占总访问量低于 10%（虽然在一些临床平台较高），并且增长十分迅速。

有关移动设备使用的案例仍在不断出现和发展，仍以“查找和保持了解新进展”为主，不过长篇阅读和教育方面的应用正在快速增长，然而与研究内容的互动（例如注释，参考文献管理或写作）仍然处于起步阶段。目前有限的证据显示，平板电脑的使用比其他电子设备更有可能取代印刷版，有可能与台式电脑和智能手机并肩占领一块新领地。

商业策略也在不断发展。最常见的商业案例是为现有订阅增加价值或便利，也许还希望增加使用和参与。目前已有一些这方面的实验，如定价、程序内置购买、免费增值模式等，不过成功的案例有限；临床医学期刊在这方面也有令人鼓舞的迹象，即期刊的平板电脑版本最终有可能成为发布数字广告的地方。总体而言，出版商在移动媒体的投资回报方面尚缺乏连贯的短期策略；对于一些出版商，提供一个移动访问的网站（例如通过平台响应的设计）可能会简单地成为一个标准平台的成本。

新的科学社交网络有如此大的注册用户基础，似乎必定会在不断演进的景观中发挥一定的作用。至少他们很可能会成为内容新发现和分享的另一个渠道，并成为发现潜在合作者的一个渠道。

考虑到技术与商业模式的密切关联，我们可期望出版商有足够的相关能力寻求为核心期刊的内容增值，包括动态的内容、可视化和分析、以及向 workflow 工具和系统的发展。这些新技术的发展将有利于资源整合（认识到单一出版商的产出往往是不够的）和不同类型内容的融合（或至少更大的聚合——图书、期刊论文、会议论文将不会失去它们的独立角色和特性）；结盟与合作的重要性将因此增加。开发这样的平台将引导出版商越来越多地思考服务而不是（现有的）产品，并且也将有从“内容”到“软件”进一步转变的趋势，同时（更大）的出版商越来越像科技公司。增值的能力也将受益于更好、更详细和更细致地了解用户的需求和行为。

STM 技术策略的明确特征将是一个开放、具有开放 API 的可互操作平台、

与广泛采用语义技术的组合。语义富集使得内容更加智能，提高可见性和使用率，也将是内容更加互动的一种方式；这将衍生出新的产品和服务，并有助于提高内部生产力。

出版平台将在内容的种类（即：容纳的期刊、书籍和数据）上不断收敛，并且在商业模式上越来越中立。诸如混合 OA 的论文被挡在付费墙之外、或需要付费才能再使用 CC-BY 材料的平台等悲剧的发生是将 OA 的特性后加在按照征订模式建立的平台中的结果：可以理解，但需要一个新的方法。

STM 出版平台现在开始追赶消费者网络中常见的用户界面/用户体验的发展，虽然这方面一直相对地被忽视。它反映了对研究人员越来越多的关注（相对于图书馆），部分地由 OA 模式消费者的驱动，而且也由关注研究评估和计量的驱动。出于类似的原因，引文和信誉将会变得更细化，对于原创作者（或投稿者）更细化的定义以及什么是可以引用的（例如较大的、动态的、数据集内的个体元素）都是如此。

更多的平台将使用诸如在 eLife Lens, Wiley Anywhere Article 等类似项目中探索互动内容的特点。然而，找到真正提高读者体验或提高增加研究人员产出与那些更复杂或不熟悉界面之间的平衡是不容易的。它也需要出版商在平台方面继续投资，以保持跟上趋势。

数据将在 STM 出版商的生命中发挥更大的作用，其原因有两个：首先，数据将成为研究产出中日益重要的核心部分；期刊不仅需要引用并提供基础数据的获取（通常是托管在一个数据仓储中），而且还直接整合一些类型的数据。数据出版将变得越来越普遍，例如初步已有“数据论文”模型出现。

其次，STM 出版商将比以往任何时候都更多地获得有关其用户和内容使用的数据。将这些数据能够用于分析和发展的能力将提供一个优势。

目前尚不清楚所有这一切是否为某些政策制定者和其他倡导者所设想的向“科学 2.0”或“开放科学”的转型。不可否认的驱动因素有：研究产出的增长、研究数量的增加和研究的全球化、以及向数据密集型科学的迁移。但是，即使认定这些与论文和数据开放获取的扩张有关系，也不会带来决策者一直寻求的经济创新的增长（由更快、更具社会响应性的科学所驱动），对于出版部门而言这仍然是一个明确的挑战。

总而言之，未来几年的显著特征是即使核心的价值保持不变，市场和技术创新的步伐也将不断加快。STM 出版业正自豪地处于其创新的历史阶段，但游戏正在发生变化，未来收入的增长将更多地由创新引导；并且，潜在的颠覆性创新会更为常见。在数字世界，用户的期望越来越多地被领先的消费者品牌引导。出

版商将必须面对更快的变化速度、更频繁的开发和发布周期、以及更多的外部创新。创新成功的关键因素、来自传统 STM 出版部门以外的人才招募、合作伙伴、技术初创企业的收购等都将变得更为常见。

## 6. 信息资源

除了参考文献，以下的信息资源或许也有助于理解本报告。

### 6.1 出版商组织

- 国际 STM 协会：覆盖面广，包括版权、公共事务、以及标准和技术  
[www.stm-assoc.org](http://www.stm-assoc.org)
- ALPSP（学术与专业出版者协会）：可从其网站获取广泛的信息资源，包括定期的“学术期刊出版实践调查”（目前是第四版）[www.alpsp.org](http://www.alpsp.org)
- OASPA（开放获取学术出版商协会）[oaspa.org/](http://oaspa.org/)

### 6.2 全球统计与趋势

- Battelle 全球研发资助基金预测。每年发布  
[www.battelle.org/media/publications/global-r-d-funding-forecast](http://www.battelle.org/media/publications/global-r-d-funding-forecast)
- NSF 科学与工程指标 2014。每两年更新一次（上一版本与当前的版本都很值得阅读）<http://www.nsf.gov/statistics/seind14/>
- OECD 统计摘要，主要科技指标 [http://stats.oecd.org/index.aspx?datasetcode=msti\\_pub#](http://stats.oecd.org/index.aspx?datasetcode=msti_pub#)
- 联合国教科文组织的科学报告，最近的版本（2010）有点过时；2015 版将于 2015 年 11 月发布  
<http://www.unesco.org/new/en/natural-sciences/science-technology/prospective-studies/unesco-science-report/>

### 6.3 开放获取

- OA 新闻与资源 (CCC/ALPSP)  
[http://www.copyright.com/content/cc3/en/open\\_access.html](http://www.copyright.com/content/cc3/en/open_access.html)
- CREATE 2014 工作报告《OA 出版：文献回顾》涵盖了 750 份深度报道的文献 (Frosio 2014) <http://www.create.ac.uk/publications/000011>
- OpenDOAR (OA 仓储名录)：包括仓储的增长、内容的类型等统计数据 and 图表等等 [www.opendoar.org/](http://www.opendoar.org/)
- ROAR (OA 仓储登记)：类似 OpenDOAR 但包括 ROARMAP，一个 OA 规定和政策的登记 [roar.eprints.org/](http://roar.eprints.org/)



- SHERPA/RoMEO: 出版商有关自存储和再使用的版权政策数据库  
<http://www.sherpa.ac.uk/romeo/>
- SHERPA/JULIET: 研究资助机构的 OA 政策  
<http://www.sherpa.ac.uk/juliet/index.php?la=en>

## 6.4 出版产业研究和分析

- Outsell: 专注于信息内容策略和使用的研究和咨询服务, 服务范围广泛, 有供应商、买家、信息用户等 [www.outsellinc.com/](http://www.outsellinc.com/)
- Simba Information: 媒体与出版业的市场情报与预测  
[www.simbainformation.com/](http://www.simbainformation.com/)
- Publishing Research Consortium (出版研究协会): 免费提供大量有用的报告, 包括同行评议、文本和数据挖掘, 以及 OA 方面的信息  
[www.publishingresearch.org.uk](http://www.publishingresearch.org.uk)
- 研究信息网络: 学术交流的发展研究与报告。当前的项目  
<http://www.researchinfonet.org>; 2006 至 2011 年间更早版本的 RIN 有影响的报告列于此(可能需要一些搜索)<http://www.researchinfonet.org/links/>

## 7. 参考文献

Acharya, A., Verstak, A., Suzuki, H., Henderson, S., Iakhiaev, M., Lin, C. C. Y., & Shetty, N. (2014). Rise of the rest: The growing impact of non-elite journals. Eprint arxiv: 1410.2217. Retrieved from arXiv.

Adams, J., & Gurney, K. (2014). Evidence for excellence: Has the signal overtaken the substance?. Retrieved from <http://www.uberresearch.com/first-digital-research-report-launched-bydigital-science-looking-wether-there-is-evidence-for-excellence/>

Albanese, A. (2009). Institutional repositories: Thinking beyond the box. *Library Journal*, 23(2), 107-116. Retrieved from Ingenta: <http://www.libraryjournal.com/article/CA6639327.html>

Allen, L., Scott, J., Brand, A., Hlava, M., & Altman, M. (2014). Publishing: Credit where credit is due. *Nature News*, 508(7496), 312. doi:10.1038/508312a

American Society for Cell Biology, & et al. (2012). San Francisco declaration on research assessment. Retrieved from <http://am.ascb.org/dora/>

Amin, M., & Mabe, M. (2007). Impact factors: Use and abuse. *Perspectives in publishing*. Elsevier Science. (Original work published October, 2000) Retrieved from <http://is.gd/Tlp9f6>

Anderson, K. (2012). The race to the bottom — data, pertussis, roads, fires, and scholarly publishing. *The Scholarly Kitchen Blog*. Retrieved from <http://scholarlykitchen.sspnet.org/2012/05/16/the-race-to-the-bottom-data-pertussis-roads-firesand-scholarly-publishing/>

Anderson, K. (2014a). 82 things publishers do (2014 edition). *Scholarly Kitchen Blog*. Retrieved from <http://scholarlykitchen.sspnet.org/2014/10/21/updated-80-things-publishersdo-2014-edition/>

Anderson, R. (2014b). Housecleaning at the directory of open access journals. Retrieved from <http://scholarlykitchen.sspnet.org/2014/08/14/housecleaning-at-the-directory-of-openaccess-journals/>

Anderson, R. (2014c). Is rational discussion of open access possible? *Insights: The UKSG Journal*, 27(2), 171-180. doi:10.1629/2048-7754.162

Anon. (2013, October 19). How science goes wrong. *Economist*. Retrieved from <http://www.economist.com/news/leaders/21588069-scientific-research-has-changed-world-nowit-needs-change-itself-how-science-goes-wrong>

Archambault, E., Amyot, D., Deschamps, P., Nicol, A., Provencher, F., Rebut, R., & Roberge, G. (2014). Proportion of open access papers published in peer-reviewed journals at the european and world levels 1996–2013. Retrieved from [http://science-metrix.com/files/science-metrix/publications/d\\_1.8\\_sm\\_ec\\_dg-rtd\\_proportion\\_oa\\_1996-2013\\_v11p.pdf](http://science-metrix.com/files/science-metrix/publications/d_1.8_sm_ec_dg-rtd_proportion_oa_1996-2013_v11p.pdf)

- ARL. (2011). ARL annual statistics 2010-11. Association of Research Libraries. Retrieved from <http://www.arl.org/stats/annualsurveys/arlstats/arlstats11.shtml>
- Aspesi, C. (2012). Reed Elsevier: Transitioning to open access - are the cost savings sufficient to protect margins?. Retrieved from <http://www.richardpoynder.co.uk/OAcosts.pdf>
- Aspesi, C. (2014). Reed Elsevier - goodbye to Berlin - the fading threat of open access. Retrieved from <http://www.richardpoynder.co.uk/Aspesi.pdf>
- Bailey, C. (2006). SPEC kit 292: Institutional repositories. Association of Research Libraries. Retrieved from <http://www.arl.org/bm~doc/spec292web.pdf>
- Baker & McKenzie. (2014). Internal draft of European Commission's copyright policy paper revealed. Retrieved from <http://www.lexology.com/library/detail.aspx?g=ad633f55-0f79-48f1-8ede-93b5606aa818>
- Baldwin, C. (2004). What do societies do with their publishing surpluses?. LPSP/Blackwell.
- Battelle. (2011). Global R&D funding forecast. Retrieved from [http://battelle.org/docs/default-document-library/2012\\_global\\_forecast.pdf?sfvrsn=2](http://battelle.org/docs/default-document-library/2012_global_forecast.pdf?sfvrsn=2)
- Battelle. (2013). Global R&D funding forecast. Retrieved from <http://www.rdmag.com/articles/2013/12/2014-r-d-magazine-global-funding-forecast>
- Beall, J. (2012). Predatory publishers are corrupting open access. *Nature News*, 489(7415), 179. doi:10.1038/489179a
- Bechhofer, S., Buchan, I., De Roure, D., Missier, P., Ainsworth, J., Bhagat, J.,... Goble, C. (2013). Why linked data is not enough for scientists. *Future Generation Computer Systems*, 29(2), 599-611. doi:10.1016/j.future.2011.08.004
- Beckett, C., & Inger, S. (2006). Self-Archiving and journal subscriptions: Co-existence or competition?. Publishing Research Consortium. Retrieved from [http://www.publishingresearch.net/self\\_archiving2.htm](http://www.publishingresearch.net/self_archiving2.htm)
- Bennett, L. (2012). ALPSP PA potential results of six month embargo. Retrieved from <http://www.publishingresearch.org.uk/documents/ALPSPPApotentialresultsofsixmonthembargofv.pdf>
- Biosciences Federation. (2008). Surveys of the financial contribution of bioscience societies to UK universities and of researchers' attitudes to open access and self-archiving. Retrieved from <http://www.publishingresearch.net/links.htm>
- Bird, C. (2008). Oxford journals' adventures in open access. *Learned Publishing*, 21(3), 200-208. doi:10.1087/095315108X288910
- Bird, C. (2010). Continued adventures in open access: 2009 perspective. *Learned Publishing*, 23(2), 107-116. doi:10.1087/20100205

BIS. (2011). Innovation and research: Strategy for growth. UK Department for Business, Innovation and Skills. Retrieved from <http://www.bis.gov.uk/assets/biscore/innovation/docs/i/11-1387-innovation-and-research-strategy-for-growth.pdf>

Björk, B. -C. (2011). A study of innovative features in scholarly open access journals. *J Med Internet Res*, 13(4), e115. doi:10.2196/jmir.1802

Björk, B. -C. (2014). Open access subject repositories: An overview. *Journal of the Association for Information Science and Technology*, 65(4), 698-706. doi:10.1002/asi.23021

Björk, B. -C., & Solomon, D. (2012a). Open access versus subscription journals: A comparison of scientific impact. *BMC Med*, 10, 73. doi:10.1186/1741-7015-10-73

Björk, B. -C., & Solomon, D. (2012b). Pricing principles used by scholarly open access publishers. *Learned Publishing*, 25(2), 132-137. doi:10.1087/20120207

Björk, B. -C., & Solomon, D. (2012c). A study of open access journals using article processing charges. *Journal of the American Society for Information Science and Technology*, 63(8), 1485-1495. doi:10.1002/asi.22673

Björk, B. -C., & Solomon, D. (2014). Developing an effective market for open access article processing charges. Wellcome Trust. Retrieved from [http://www.wellcome.ac.uk/stellent/groups/corporatesite/@policy\\_communications/documents/web\\_document/wtp055910.pdf](http://www.wellcome.ac.uk/stellent/groups/corporatesite/@policy_communications/documents/web_document/wtp055910.pdf)

Björk, B. C., Laakso, M., Welling, P., & Paetau, P. (2014). Anatomy of green open access. *Journal of the Association for Information Science and Technology*, 62(2), 237-250. doi:10.1002/asi.22963

Björk, B. -C., Roos, A., & Lauri, M. (2009). Scientific journal publishing: Yearly volume and open access availability. *Information Research*, 14(1). Retrieved from <http://informationr.net/ir/14-1/paper391.html>

Björk, B. -C., Welling, P., Laakso, M., Majlender, P., Hedlund, T., & Gudnason, G. (2010). Open access to the scientific journal literature: Situation 2009. *PLOS ONE*, 5(6), e11273. doi:10.1371/journal.pone.0011273

Blodgett, H., & Cocotas, A. (2012). The future of mobile. Retrieved from [www.businessinsider.com: http://www.businessinsider.com/the-future-of-mobiledeck-2012-3?op=1](http://www.businessinsider.com/the-future-of-mobiledeck-2012-3?op=1)

Bloom, T. (2014). PLOS' new data policy: Part two. *PLOS EveryONE Blog*. Retrieved from <http://blogs.plos.org/everyone/2014/03/08/plos-new-data-policy-public-access-data/>

Bollen, J., de Sompel, H. V., Hagberg, A., & Chute, R. (2009). A principal component analysis of 39 scientific impact measures. *PLoS ONE*, 4(6), e6022. doi:10.1371/journal.pone.0006022

Bornmann, L., & Mutz, R. (2014). Growth rates of modern science: A bibliometric analysis

based on the number of publications and cited references. Eprint arxiv:1402.4578. Retrieved from arXiv.

Callaghan, S. (2014). Preserving the integrity of the scientific record: Data citation and linking. *Learned Publishing*, 27(5), 15-24. doi:10.1087/20140504

Campbell, P. (2006). Report of nature's peer review trial. *Peer-to-Peer*. Retrieved from [http://blogs.nature.com/peer-to-peer/2006/12/report\\_of\\_natures\\_peer\\_review\\_trial.html](http://blogs.nature.com/peer-to-peer/2006/12/report_of_natures_peer_review_trial.html)

Campbell, R., & Wates, E. (2009). Journal publishing: Implications for a digital library policy. In D. Baker & W. Evans (Eds.), *Digital library economics*. Chandos.

Carpenter, T. (2013). IAnnotate — whatever happened to the web as an annotation system. *Scholarly Kitchen Blog*. Retrieved from <http://scholarlykitchen.sspnet.org/2013/04/30/iannotate-whatever-happened-to-the-web-as-an-annotation-system/>

CIBER. (2008). Information behaviour of the researcher of the future. Retrieved from [http://www.jisc.ac.uk/media/documents/programmes/reppres/gg\\_final\\_keynote\\_11012008.pdf](http://www.jisc.ac.uk/media/documents/programmes/reppres/gg_final_keynote_11012008.pdf)

CIBER Research. (2012a). PEER usage study: Descriptive statistics for the period march to august 2011. *PEER*. Retrieved from <http://www.peerproject.eu/reports/>

CIBER Research. (2012b). PEER usage study: Randomised controlled trial results. *PEER*. Retrieved from <http://www.peerproject.eu/reports/>

CIBER Research Ltd. (2011). The Journal Usage Factor: Exploratory data analysis. Retrieved from [http://ciber-research.eu/CIBER\\_news-201103.html](http://ciber-research.eu/CIBER_news-201103.html)

Citron, D. T., & Ginsparg, P. (2014). Patterns of text reuse in a scientific corpus. *Proceedings of the National Academy of Sciences*, 201415135. doi:10.1073/pnas.1415135111

Clark, A., Jensen, L., & Campbell, D. (2014). Member briefing: Text and data mining. *ALPSP*. Retrieved from <http://www.alpsp.org/Ebusiness/ResearchPublications/AdviceNoteGuidelines.aspx>

Clark, J. (2013). Text mining and scholarly publishing. Retrieved from <http://www.publishingresearch.net/documents/PRCTextMiningandScholarlyPublishinFeb2013.pdf>

Clarke, M. (2013). Game of papers: ELife, BMC, PLOS and EMBO announce new peer review consortium. *Scholarly Kitchen*. Retrieved from <http://scholarlykitchen.sspnet.org/2013/07/15/game-of-papers-elife-bmc-plos-and-embo-announce-new-peer-reviewconsortium/>

Cochran, A. (2014). Implementing CHORUS: Big decisions loom for publishers. Retrieved from <http://scholarlykitchen.sspnet.org/2014/07/09/implementing-chorus-big-decisions-loomfor-publishers/>

Cochrane, A. (2014). Interview with Laurel Haak of ORCID: Supporting the efforts with

membership and integration. Scholarly Kitchen Blog. Retrieved from <http://scholarlykitchen.sspnet.org/2014/09/18/interview-with-laurel-haak-of-orcid-supportingthe-efforts-with-membership-and-integration/>

Coles, B. R. (1993). The STM information system in the UK [R&D Report No. 6123] (R&D Report No. 6123). London: British Library.

Collins, F. S., & Tabak, L. A. (2014). Policy: NIH plans to enhance reproducibility. *Nature News*, 505(7485), 612. doi:10.1038/505612a

COPE. (2011). Code of conduct and best practice guidelines for journal editors. Retrieved from [http://publicationethics.org/files/Code\\_of\\_conduct\\_for\\_journal\\_editors\\_Mar11.pdf](http://publicationethics.org/files/Code_of_conduct_for_journal_editors_Mar11.pdf)

COUNTER. (2014). Release 1 of the COUNTER code of practice for usage factors. Retrieved from [http://www.projectcounter.org/usage\\_factor.html](http://www.projectcounter.org/usage_factor.html)

Cox, J., & Cox, L. (2008). Scholarly publishing practice: Academic journal publishers' policies and practices in online publishing. Third survey. ALPSP. Retrieved from [http://www.alpssp.org/ngen\\_public/article.asp?id=0&did=0&aid=2446&st=scholarly%20publishing%20practice&oid=0](http://www.alpssp.org/ngen_public/article.asp?id=0&did=0&aid=2446&st=scholarly%20publishing%20practice&oid=0)

Cox, J., & Cox, L. (2010). Scholarly book publishing practice. ALPSP. Retrieved from <http://www.alpssp.org/Ebusiness/ProductCatalog/Product.aspx?ID=41>

Craig, I. D., Plume, A. M., McVeigh, M. E., Pringle, J., & Amin, M. (2007). Do open access articles have greater citation impact? A critical review of the literature. *Journal of Informetrics*, 1(3), 239-248. doi:10.1016/j.joi.2007.04.001

Creaser, C., Maynard, S., & White, S. (2006). LISU annual library statistics 2005. Retrieved from <http://www.lboro.ac.uk/microsites/infosci/lisu/pages/publications/als06.html>

Cressey, D. (2014). Journals weigh up double-blind peer review. *Nature News*. doi:10.1038/nature.2014.15564

Crotty, D. (2008). Web 2.0 for biologists—are any of the current tools worth using? BenchMarks blog. Retrieved from <http://cshbenchmarks.wordpress.com/2008/04/03/web-20-forbiologists-are-any-of-the-current-tools-worth-using/>

Crotty, D. (2010). Rules of thumb for social media in science. The Scholarly Kitchen Blog. Retrieved from <http://scholarlykitchen.sspnet.org/2010/04/05/rules-of-thumb-for-socialmedia-in-science/>

Crotty, D. (2012). Are university block grants the right way to fund open access mandates? The Scholarly Kitchen Blog. Retrieved from <http://scholarlykitchen.sspnet.org/2012/09/13/are-university-block-grants-the-right-way-to-fund-open-access-mandates/>

Czujko, R., & Chu, R. (2015). Scholarly and scientific publishing workforce in the U.S.

AAP/AIP. Retrieved from <http://publishers.org/press/158/>

Dallmeier-Tiessen, S et al. (2010). Open access publishing – models and attributes (SOAP). Retrieved from <http://edoc.mpg.de/478647>

Davis, P. (2014). Peak PLOS: Planning for a future of declining revenue. Scholarly Kitchen. Retrieved from <http://scholarlykitchen.sspnet.org/2014/12/02/peak-plos-planning-for-a-future-of-declining-revenue/>

Davis, P. M. (2011). Open access, readership, citations: A randomized controlled trial of scientific journal publishing. *FASEB J*, 25(7), 2129-34. doi:10.1096/fj.11-183988

Davis, P. M., & Walters, W. H. (2011). The impact of free access to the scientific literature: A review of recent research. *J Med Libr Assoc*, 99(3), 208-17. doi:10.3163/1536-5050.99.3.008

Department for Business, Innovation & Skills. (2014). Science and innovation: Rates of return to investment. Retrieved from <https://www.gov.uk/government/publications/science-and-innovation-rates-of-return-to-investment>

Dickson, D. (2012). Developing world gains open access to science research, but hurdles remain. *The Guardian*. Retrieved from [http://www.guardian.co.uk/global-development/2012/sep/03/developing-world-open-access-research-hurdles?CMP=tw\\_t\\_gu](http://www.guardian.co.uk/global-development/2012/sep/03/developing-world-open-access-research-hurdles?CMP=tw_t_gu)

Economist. (2014). No peeking.... *The Economist*. Retrieved from <http://www.economist.com/news/science-and-technology/21593408-publishing-giant-goes-after-authors-its-journals-papers-no-peeking>

Edgecliffe-Johnson, A. (2012, July 24). Tablets lift journals' ad revenues. *Financial Times*. Retrieved from <http://on.ft.com/MFMldR>

Edwards, D. C. (2014). How can existing open access models work for humanities and social science research? *Insights: The UKSG Journal*, 27(1), 17-24. doi:10.1629/2048-7754.135

Eisen, M. (2013). A CHORUS of boos: Publishers offer their “solution” to public access. Retrieved from [www.michaeleisen.org](http://www.michaeleisen.org): <http://www.michaeleisen.org/blog/?p=1382>

eLife. (2014). A foundation for change: The eLife Sciences 2013 annual report. eLife. Retrieved from <http://2013.elifesciences.org/#toc/all>

Elliot, L. (2014, October 7). IMF says economic growth may never return to pre-crisis levels. *The Guardian*. Retrieved from [www.theguardian.com](http://www.theguardian.com/business/2014/oct/07/imf-economic-growth-forecasts-downgraded-crisis): <http://www.theguardian.com/business/2014/oct/07/imf-economic-growth-forecasts-downgraded-crisis>

Elsevier. (2011). International comparative performance of the UK research base, report for UK department of business, innovation and skills. Retrieved from <http://www.bis.gov.uk/assets/biscore/science/docs/i/11-p123-international-comparative-performance-uk-researchbase-2011.pdf>

Elsevier. (2013). International comparative performance of the UK research base – 2013. Retrieved from <https://www.gov.uk/government/publications/performance-of-the-uk-researchbase-international-comparison-2013>

EPS. (2006). UK scholarly journals: 2006 baseline report. Research Information Network. Retrieved from <http://www.rin.ac.uk/our-work/communicating-and-disseminatingresearch/uk-scholarly-journals-2006-baseline-report>

European Commission. (2012a). Areas of untapped potential for the development of the European research area: Preliminary summary and analysis of the response to the ERA framework public consultation. Retrieved from [http://ec.europa.eu/research/era/pdf/era-summaryreport-2012\\_en.pdf](http://ec.europa.eu/research/era/pdf/era-summaryreport-2012_en.pdf)

European Commission. (2012b). Online survey on scientific information in the digital age. Retrieved from [http://ec.europa.eu/research/science-society/document\\_library/pdf\\_06/survey-on-scientific-information-digital-age\\_en.pdf](http://ec.europa.eu/research/science-society/document_library/pdf_06/survey-on-scientific-information-digital-age_en.pdf)

European Commission. (2014). Public consultation – “Science 2.0”: Science in transition. Retrieved from <http://ec.europa.eu/research/consultations/science-2.0/background.pdf>

Eysenbach, G. (2011). Can tweets predict citations? Metrics of social impact based on twitter and correlation with traditional metrics of scientific impact. *J Med Internet Res*, 13(4), e123. doi:10.2196/jmir.2012

Ferguson, C., Marcus, A., & Oransky, I. (2014). Publishing: The peer-review scam. *Nature News*, 515(7528), 480. doi:10.1038/515480a

Ferguson, L. (2014). How and why researchers share data (and why they don't). Wiley Exchanges Blog. Retrieved from <http://exchanges.wiley.com/blog/2014/11/03/how-andwhy-researchers-share-data-and-why-they-dont/>

Ferwerda, E. (2014). Open access monograph business models. *Insights: The UKSG Journal*, 27, 35-38. doi:10.1629/2048-7754.46

Finch Working Group. (2012). Accessibility, sustainability, excellence: How to expand access to research publication. Retrieved from <http://www.researchinfonet.org/wp-content/uploads/2012/06/Finch-Group-report-FINAL-VERSION.pdf>

FORCE 11. (2011). FORCE 11 manifesto: Improving future research communication and e-scholarship.

Frass, W., Cross, J., & Gardner, V. (2014). Taylor & Francis open access survey June 2014. Retrieved from <http://www.tandf.co.uk/journals/explore/open-access-surveyjune2014.pdf>

Frosio, G. (2014). Open access publishing: A literature review. CREATE. Retrieved from <http://www.create.ac.uk/publications/000011>



Fry et al. (2009). PEER behavioural research: Baseline report. Retrieved from [http://www.peerproject.eu/fileadmin/media/reports/Final\\_revision\\_-\\_behavioural\\_baseline\\_report\\_-\\_20\\_01\\_10.pdf](http://www.peerproject.eu/fileadmin/media/reports/Final_revision_-_behavioural_baseline_report_-_20_01_10.pdf)

Gantz, P. (2012). Digital licenses replace print prices as accurate reflection of real journal costs. Retrieved from [http://publishers.org/\\_attachments/docs/library/2012%20bulletin.pdf](http://publishers.org/_attachments/docs/library/2012%20bulletin.pdf)

Gantz, P. (2013). Journal print subscription price increases no longer reflect actual costs. *Learned Publishing*, 26(3), 206-210. doi:10.1087/20130309

Gargouri, Y., Larivière, V., Gingras, Y., Carr, L., & Harnad, S. (2012). Green and gold open access percentages and growth, by discipline. ArXiv Preprint <http://arxiv.org/abs/1206.3664>

Gargouri, Y. A., Hajjem, C., Larivière, V., Gingras, Y., Carr, L., Brody, T., & Harnad, S. (2010). Self-selected or mandated, open access increases citation impact for higher quality research. *PLoS ONE*, 5(10), e13636. doi:10.1371/journal.pone.0013636

Global Coalition of Access to Research, Science and Education Organizations Calls on STM to Withdraw New Model Licenses. (2014). [Web page]. Retrieved from <http://www.plos.org/global-coalition-of-access-to-research-science-and-education-organizations-calls-on-stm-to-withdraw-new-model-licenses/>

Goble, C. (2008, September 23). The future of research (science & technology); presentation to British Library board. Retrieved from <http://www.slideshare.net/dullhunk/the-future-of-research-science-and-technology-presentation>

Graham, D. (2013). Academic publishing: Survey of funders supports the benign open access outcome priced into shares. HSBC Global Research.

Group of Eight, & Outsell. (2010). Building value frameworks for academic libraries – cost–benefit study. Retrieved from [http://go8.edu.au/\\_documents/go8-committees/go8-librarians/anu\\_outsell\\_8\\_april\\_2010.pps](http://go8.edu.au/_documents/go8-committees/go8-librarians/anu_outsell_8_april_2010.pps)

Halevi, G. (2012). Research trends: Special issue on Big Data. Elsevier. Retrieved from [http://www.researchtrends.com/wp-content/uploads/2012/09/Research\\_Trends\\_Issue30.pdf](http://www.researchtrends.com/wp-content/uploads/2012/09/Research_Trends_Issue30.pdf)

Hargreaves, I. (2011). Digital opportunity: A review of intellectual property and growth. Retrieved from <http://www.ipo.gov.uk/ipreview-finalreport.pdf>

Harley, D., & et al. (2010). Assessing the future landscape of scholarly communication: An exploration of faculty values and needs in seven disciplines. Retrieved from [escholarship.org: http://escholarship.org/uc/item/15x7385g](http://escholarship.org/uc/item/15x7385g)

Houghton, J. (2009). Open access: What are the economic benefits? A comparison of the United Kingdom, Netherlands and Denmark. doi:10.2139/ssrn.1492578

Houghton, J. (2010). Economic and social returns on investment in open archiving publicly

funded research outputs. SPARC. Retrieved from <http://sparc.arl.org/sites/default/files/vufrpaa.pdf>

Houghton, J., Rasmussen, B., Sheehan, P., Oppenheimer, C., Morris, A., Creaser, C.,...  
Gourlay, A. (2009). Economic implications of alternative scholarly publishing models. JISC.  
Retrieved from [http://www.jisc.ac.uk/publications/reports/2009/  
economicpublishingmodelsfinalreport.aspx](http://www.jisc.ac.uk/publications/reports/2009/economicpublishingmodelsfinalreport.aspx)

Houghton, J., Swan, A., & Brown, S. (2011). Access to research and technical information in Denmark. Retrieved from <http://eprints.soton.ac.uk/272603/>

Huggett, S. (2014). A quick look at references to research data repositories. Research Trends (Elsevier). Retrieved from <http://www.researchtrends.com/issue-38-september-2014/aquick-look-at-references-to-research-data-repositories/>

ICMJE. (2013). Recommendations for the conduct, reporting, editing and publication of scholarly work in medical journals. Retrieved from <http://www.icmje.org/recommendations/>

Inger, S., & Gardner, T. (2012). How readers discover content in scholarly journals. Renew Training. Retrieved from <http://www.renewtraining.com/How-Readers-Discover-Contentin-Scholarly-Journals-summary-edition.pdf>

Inger, S., & Gardner, T. (2013). Scholarly journals publishing practice: Academic journal publishers' policies and practices in online publishing. ALPSP. Retrieved from <http://www.alpsp.org/Ebusiness/ProductCatalog/Product.aspx?ID=359>

Intellectual Property Office. (2014). Intellectual property – guidance: Exceptions to copyright. Retrieved from <https://www.gov.uk/exceptions-to-copyright>

Ioannidis, J. P. A., Boyack, K. W., & Klavans, R. (2014). Estimates of the continuously publishing core in the scientific workforce. PLOS ONE, 9(7), e101698. doi:10.1371/journal.pone.0101698

Ithaka S+R. (2010). Faculty survey 2009: Key strategic insights for libraries, publishers, and societies. Retrieved from <http://www.sr.ithaka.org/research-publications/facultysurvey-2009>

Ithaka S+R. (2011). Hindawi publishing corporation: Growing an open-access contributor-pays business model. Retrieved from [www.ithaka.org](http://www.ithaka.org): [http://www.ithaka.org/ithaka-s-r/research/case-studies-in-sustainability-2011/SCA\\_IthakaSR\\_CaseStudies\\_Hindawi\\_2011.pdf](http://www.ithaka.org/ithaka-s-r/research/case-studies-in-sustainability-2011/SCA_IthakaSR_CaseStudies_Hindawi_2011.pdf)

Ithaka S+R. (2013). US faculty survey 2012. Ithaka S+R. Retrieved from [www.sr.ithaka.org](http://www.sr.ithaka.org): <http://www.sr.ithaka.org/research-publications/us-faculty-survey-2012>

Ithaka S+R, JISC, & RLUK. (2013). UK survey of academics 2012. Retrieved from <http://www.sr.ithaka.org/research-publications/ithaka-sr-jisc-rluk-uk-survey-academics-2012>

Ivins, O., & Luther, J. (2011). Publishing support for small print-based publishers: Options for ARL libraries. ARL. Retrieved from <http://www.arl.org/component/content/article/6/1170>

Jalalian, M., & Mahboobi, H. (2014). Hijacked journals and predatory publishers: Is there a need to re-think how to assess the quality of academic research? *Walailak Journal of Science and Technology*, 11(5), 389-394. doi:10.14456/WJST.2014.16

Jefferson, T., Rudin, M., Brodney Folse, S., & Davidoff, F. (2007). Editorial peer review for improving the quality of reports of biomedical studies. *Cochrane Database Syst Rev*, (2), MR000016. doi:10.1002/14651858.MR000016.pub3

Jha, A. (2012). Open access to research is inevitable, says Nature editor-in-chief. *The Guardian*. Retrieved from <http://www.theguardian.com/science/2012/jun/08/open-access-researchinevitable-nature-editor>

JISC. (2012). The value and benefits of text mining. Retrieved from <http://www.jisc.ac.uk/publications/reports/2012/value-and-benefits-of-text-mining.aspx>

Jisc. (2014). Jisc's evidence to the review of the RCUK OA policy. Retrieved from <http://scholarlycommunications.jiscinvolve.org/wp/2014/09/18/jiscs-evidence-to-the-review-of-the-rcuk-oa-policy/>

Jisc, & British Library. (2012). Researchers of tomorrow: The research behaviour of generation Y doctoral students. Retrieved from [www.jisc.ac.uk: http://www.jisc.ac.uk/publications/reports/2012/researchers-of-tomorrow.aspx](http://www.jisc.ac.uk/publications/reports/2012/researchers-of-tomorrow.aspx)

Jones, J. (2014a). To flip or not to flip. In ALPSP international conference. Retrieved from <http://blog.alpssp.org/2014/09/open-access-daily-challenge-new.html>

Jones, P. (2014b). What's going on in the library? Part 1: Librarian publishers may be more important than you think. *Scholarly Kitchen Blog*. Retrieved from <http://scholarlykitchen.sspnet.org/2014/12/01/whats-going-on-in-the-library-part-1-librarianpublishers-may-be-more-important-than-you-think>

Jump, P. (2014a). IOP launches 'offsetting' scheme to cut cost of open access. *Times Higher Education*. Retrieved from <http://www.timeshighereducation.co.uk/news/iop-launchesoffsetting-scheme-to-cut-cost-of-open-access/2013546.article>

Jump, P. (2014b). Snowball metrics: No pinch of salt needed. *Times Higher Education*. Retrieved from <http://www.timeshighereducation.co.uk/news/snowball-metrics-no-pinchof-salt-needed/2014496.article>

Khabsa, M., & Giles, C. L. (2014). The number of scholarly documents on the public web. *PLoS ONE*, 9(5), e93949. doi:10.1371/journal.pone.0093949

King, D. W., & Alvarado-Albertorio, F. M. (2008). Pricing and other means of charging for scholarly journals: A literature review and commentary. *Learned Publishing*, 21, 248-272. doi:10.1087/095315108X356680

- King, D. W., & Tenopir, C. (2004). An evidence based assessment of author pays. *Nature Web Focus on access to the literature*. Retrieved from <http://www.nature.com/nature/focus/accessdebate/26.html>
- King, D. W., Tenopir, C., & Clarke, M. (2006). Measuring total reading of journal articles. *DLib Magazine*, 12(10). doi:10.1045/october2006-king
- Kurtz, M. J., Eichhorn, G., Accomazzi, A., Grant, C., Demleitner, M., Henneken, E., & Murray, S. S. (2005). The effect of use and access on citations. *Information Processing & Management*, 41(6), 1395 - 1402. doi:10.1016/j.ipm.2005.03.010
- K|N Consultants. (2014). A scalable and sustainable approach to open access publishing and archiving for humanities and social sciences. Retrieved from <http://knconsultants.org/towards-a-sustainable-approach-to-open-access-publishing-and-archiving/>
- Laakso, M., & Björk, B. -C. (2012). Anatomy of open access publishing: A study of longitudinal development and internal structure. *BMC Medicine*, 10(1), 124. doi: 10.1186/1741-7015-10-124
- Laakso, M., & Björk, B. -C. (2013). Delayed open access: An overlooked high-impact category of openly available scientific literature. *Journal of the American Society for Information Science and Technology*, 64(7), 1323-1329. doi:10.1002/asi.22856
- Lawrence, S. (2001). Free online availability substantially increases a paper's impact. *Nature*, 411(6837), 521. doi:10.1038/35079151
- Lawson, S. (2015). 'Total cost of ownership' of scholarly communication: Managing subscription and APC payments together. *Learned Publishing*, 28(1), 9-13. doi: 10.1087/20140103
- Levine-Clark, M., McDonald, J., & Price, J. (2014). Examining the impact of discovery systems on usage. *Wiley Exchanges Blog*. Retrieved from <http://exchanges.wiley.com/blog/2014/07/21/examining-the-impact-of-discovery-systems-on-usage/>
- Lin, J. (2014). Making data count: PLOS, CDL, and DataONE join forces to build incentives for data sharing. *PLOS Tech Blog*. Retrieved from <http://blogs.plos.org/tech/making-datacount-plos-cdl-and-dataone-join-forces-to-build-incentives-for-data-sharing/>
- Lippincott, S. (2014). *Library publishing directory 2015*. Library Publishing Coalition. Retrieved from <http://www.librarypublishing.org/resources/directory/lpd2015>
- Lynch, C. (2009). Jim Gray's fourth paradigm and the construction of the scientific record. In T. Hey, S. Tansley, & K. Tolle (Eds.), *The fourth paradigm: Data-Intensive scientific discovery*. Richmond: Microsoft.
- Lynch, C. (2014). The need for research data inventories and the vision for SHARE. *Information Standards Quarterly*, 26(2). Retrieved from [www.niso.org](http://www.niso.org): <http://www.niso.org/>

publications/isq/2014/v26no2/lynch/

Mabe, M. (2003). The growth and number of journals. *Serials*, 16(2), 191-197. Retrieved from <http://uksg.metapress.com/link.asp?id=f195g8ak0eu21muh>

Mabe, M. (2004). Presentation at 6th Nordic Interlending Conference, Trondheim.

Mabe, M. (2006). Journal futures: How will researchers communicate as the internet matures? Presentation at Council of Science Editors Annual conference, Tampa 2006

Mabe, M. (2008). STM publishing: The known knowns, the known unknowns,... And all points in between. Presentation at Academic Publishing in Europe conference, Berlin

Mabe, M. (2010). Scholarly communication: A long view. *New Review of Academic Librarianship*, 16, 132-144. doi:10.1080/13614533.2010.512242

Mabe, M. (2012). Does journal publishing have a future? In R. Campbell, E. Pentz, & I. Borthwick (Eds.), *Academic and professional publishing*. Oxford: Chandos.

Mabe, M., & Amin, M. (2002). Dr Jekyll and Dr Hyde: Author-reader asymmetries in scholarly publishing. *Aslib Proceedings: New Information Perspectives*, 54(3), 149-157. doi: 10.1108/00012530210441692

Mabe, M., & Mulligan, A. (2011). What journal authors want: Ten years of results from Elsevier's author feedback programme. *New Review of Information Networking*, 16, 71-89. doi: 10.1080/13614576.2011.574495

Mandler, P. (2014). Open access: A perspective from the humanities. *Insights: The UKSG Journal*, 27(2), 166-170. doi:10.1629/2048-7754.89

Marcus, A., & Oransky, I. (2012). Bring on the transparency index. *The Scientist*, (12 August 2012). Retrieved from the-scientist.com: <http://the-scientist.com/2012/08/01/bring-on-thetransparency-index/>

McCabe, M., & Snyder, C. (2013). Identifying the effect of open access on citations using a panel of science journals. [preprint]. Retrieved from [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2269040](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2269040)

McCabe, M. J., & Snyder, C. M. (2014). Does online availability increase citations? Theory and evidence from a panel of economics and business journals. *Review of Economics and Statistics*. doi:10.1162/REST\_a\_00437

McEntyre, J. R., Ananiadou, S., Andrews, S., Black, W. J., Boulderstone, R., Buttery, P.,... Vaughan, P. (2011). UKPMC: A full text article resource for the life sciences. *Nucleic Acids Research*, 39(Database issue), D58-65. doi:10.1093/nar/gkq1063

McKinsey. (2011). *Big Data: The next frontier for innovation, competition, and productivity*.

Retrieved from [http://www.mckinsey.com/Insights/MGI/Research/Technology\\_and\\_Innovation/Big\\_data\\_The\\_next\\_frontier\\_for\\_innovation](http://www.mckinsey.com/Insights/MGI/Research/Technology_and_Innovation/Big_data_The_next_frontier_for_innovation)

Meadows, A. (2014). An interview with Amy Brand on a proposed new contributor taxonomy initiative. Scholarly Kitchen. Retrieved from <http://scholarlykitchen.sspnet.org/2014/08/20/an-interview-with-amy-brand-on-a-proposed-new-contributor-taxonomyinitiative/>

Meadows, A., Campbell, R., & Webster, K. (2012). The access question. *Learned Publishing*, 25(3), 189-193. doi:10.1087/20120305

Meeker, M. (2014). Internet trends 2014. Retrieved from [http://kpcbweb2.s3.amazonaws.com/files/85/Internet\\_Trends\\_2014\\_vFINAL\\_-\\_05\\_28\\_14-\\_PDF.pdf?1401286773](http://kpcbweb2.s3.amazonaws.com/files/85/Internet_Trends_2014_vFINAL_-_05_28_14-_PDF.pdf?1401286773)

Meyer, C. A. (2011). Distinguishing published scholarly content with CrossMark. *Learned Publishing*, 24(2), 87-93. doi:10.1087/20110202

Miller, P. (2010). Linked data horizon scan. JISC. Retrieved from <http://cloudofdata.s3.amazonaws.com/FINAL-201001-LinkedDataHorizonScan.pdf>

Milloy, C. (2013). Innovative approaches to publishing open access monographs – it's not business as usual. *Jisc Inform*, (37). Retrieved from <http://www.jisc.ac.uk/inform/inform37/OpenAccessMonographs.html>

MIT. (2003). MIT's DSpace experience: A case study. Retrieved from <http://dspace.org/implement/case-study.pdf>

Moed, H. F. (2007). The effect of “open access” on citation impact: An analysis of arXiv's condensed matter section. *Journal of the American Society for Information Science and Technology*, 58(13), 2047-2054. doi:10.1002/asi.20663

Morgan, C., Campbell, B., & Teelen, T. (2012). The role of the academic journal publisher and open access publishing models. *International Studies Perspectives*, 13(3), 228-234. doi:10.1111/j.1528-3585.2012.00495.x

Morris, S. (2006). Data about publishing. *ALPSP Alert*, (112), 8. Retrieved from [www.alpsp.org: http://www.alpsp.org/Ebusiness/ResearchPublications/ALPSPAlert/ALERTARCHIVE.aspx](http://www.alpsp.org/Ebusiness/ResearchPublications/ALPSPAlert/ALERTARCHIVE.aspx)

Morris, S. (2007). Mapping the journal publishing landscape: How much do we know? *Learned Publishing*, 20, 299-310. doi:10.1087/095315107X239654

Morris, S. (2009). Journal authors' rights: Perception and reality. *Publishing Research Consortium*. Retrieved from [http://www.publishingresearch.net/author\\_rights.htm](http://www.publishingresearch.net/author_rights.htm)

Morrison, H., & et al. (2014). Open access article processing charges: DOAJ survey may 2014. [preprint]. Retrieved from <http://sustainingknowledgecommons.files.wordpress.com/2014/10/oa-apcs-article-2014-october-171.pdf>

- Mulligan, A., & Mabe, M. (2011). The effect of the internet on researcher motivations, behaviour and attitudes. *Journal of Documentation*, 67, 290-311. doi: 10.1108/00220411111109485
- Murphy, F. (2014). Data and scholarly publishing: The transforming landscape. *Learned Publishing*, 27(5), 3-7. doi:10.1087/20140502
- Nature. (2012). Gold in the text? (Editorial). *Nature*, 483(7388), 124. doi:10.1038/483124a
- Nature Editors. (2014). Nature Publishing Index 2013 Asia-Pacific. Retrieved from <http://www.natureasia.com/en/publishing-index/asia-pacific/supplement2013>
- Nature Publishing Group. (2014). 2014 Author Insights survey. figshare. doi: 10.6084/m9.figshare.1204999
- Newman, M. L., & Sack, J. (2013). Information workflow of academic researchers in the evolving information environment: An interview study. *Learned Publishing*, 26(2), 123-131. doi:10.1087/20130208
- Nicholas, D., & Clark, D. (2012). 'Reading' in the digital environment. *Learned Publishing*, 25(2), 93-98. doi:10.1087/20120203
- Nicholas, D., Watkinson, A., Jamali, H. R., Herman, E., Tenopir, C., Volentine, R.,... Levine, K. (2015). Peer review: Still king in the digital age. *Learned Publishing*, 28(1), 15-21. doi: 10.1087/20150104
- NISO. (2008). Journal article versions (JAV): Recommendations of the NISO/ALPSP JAV technical working group (April 2008), NISO-RP-8-2008. Retrieved from <http://www.niso.org/publications/rp/RP-8-2008.pdf>
- NISO SERU Standing Committee. (2012). SERU recommended practice (RP-7-2012). NISO. Retrieved from <http://www.niso.org/workrooms/seru/NPG>. (2014). Author Insights 2014. figshare. Retrieved from <http://dx.doi.org/10.6084/m9.figshare.1204999>
- NSF. (2012). Science & Engineering Indicators 2012. Retrieved from <http://www.nsf.gov/statistics/seind12/>
- NSF. (2014). Science and Engineering Indicators 2014. Arlington, VA: NSB 14-01. Retrieved from <http://www.nsf.gov/statistics/seind14/>
- OECD. (n.d.). OECD.StatExtracts: Main Science and Technology Indicators. Retrieved from [stats.oecd.org: http://stats.oecd.org/Index.aspx?DataSetCode=MSTI\\_PUB#](http://stats.oecd.org/Index.aspx?DataSetCode=MSTI_PUB#)
- OECD. (2011). Main Science and Technology Indicators. OECD. Retrieved from [www.oecd.org/sti/msti](http://www.oecd.org/sti/msti)
- OECD. (2014). OECD science, technology and industry outlook 2014. doi:10.1787/sti\_outlook-2014-en

OpCit. (n.d.). The effect of open access and downloads ('hits') on citation impact: A bibliography of studies. Retrieved from <http://opcit.eprints.org/oacitation-biblio.html#most-recent>

Orduña-Malea, E., Ayllón, J. M., Martín-Martín, A., & López-Cózar, E. D. (2014). About the size of Google Scholar: Playing the numbers. arXiv eprint <http://arxiv.org/abs/1407.6239>

OSTP. (2012). US office of science and technology policy requests for information on public access to scholarly journals and data. Retrieved from <http://www.whitehouse.gov/administration/eop/ostp/library/publicaccess>

OSTP. (2013). Increasing access to the results of federally funded scientific research. Retrieved from [http://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp\\_public\\_access\\_memo\\_2013.pdf](http://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf)

Outsell. (2009). Open access primer (public version). Retrieved from [http://www.outsellinc.com/images/user\\_assets/Outsell\\_Mkt\\_Report\\_14dec09\\_Open\\_Access\\_Primer\\_Public.pdf](http://www.outsellinc.com/images/user_assets/Outsell_Mkt_Report_14dec09_Open_Access_Primer_Public.pdf)

Outsell. (2011). Scientific, technical & medical information: 2011 market forecast and trends report. Retrieved from [www.outsellinc.com: http://www.outsellinc.com/store/products/1040](http://www.outsellinc.com/store/products/1040)

Outsell. (2012a). Evolution of the STM publishing platform – an industry overview and roadmap. Retrieved from <http://www.outsellinc.com/store/products/1083>

Outsell. (2012b). Mobile in STM: Case studies of accelerating change. Retrieved from <http://www.outsellinc.com/store/products/1096>

Outsell. (2012c). STM 2012 market size, share, forecast & trend report. Retrieved from [www.outsellinc.com: http://www.outsellinc.com/products/index/1107-scientific-technicalmedical-information-2012-market-size-share-forecast-and-trend-report](http://www.outsellinc.com/products/index/1107-scientific-technicalmedical-information-2012-market-size-share-forecast-and-trend-report)

Outsell. (2012d). STM e-books: 2012 market size, share, and forecast. Retrieved from [www.outsellinc.com: http://www.outsellinc.com/store/products/1100](http://www.outsellinc.com/store/products/1100)

Outsell. (2012e). STM in China: 2012 market size, share & forecast. Retrieved from [www.outsellinc.com: http://www.outsellinc.com/store/products/1091](http://www.outsellinc.com/store/products/1091)

Outsell. (2013). Open access: Market size, share, forecast, and trends. Retrieved from <http://www.outsellinc.com/store/products/1135>

Outsell. (2014a). Professional associations and learned societies: Current issues and outlook. Retrieved from <http://www.outsellinc.com/store/products/1276>

Outsell. (2014b). Sensored world, sensible choices: Information industry outlook 2015. Retrieved from <http://www.outsellinc.com/store/products/1272>

Outsell. (2014c). STM 2014 market size, share, forecast, and trend report. Retrieved from <http://www.outsellinc.com/1280/products/1268-scientific-technical-medicalinformation-2014-mar>



ket-size-share-forecast-and-trend-report

Outsell. (2014d). STM platform providers: Market update. Retrieved from <http://www.outsellinc.com/store/products/1248>

Outsell. (2014e). Web-Scale discovery services: The evolution continues. Outsell. Retrieved from <http://www.outsellinc.com/1178/products/1264-im-report-web-scale-discovery-services-the-evolution-continues>

Parsons, D., Willis, D., & Holland, J. (2011). Benefits to the private sector of open access to higher education and scholarly research. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.368.1820&rep=rep1&type=pdf>

Patel, J. (2014). Why training and specialization is needed for peer review: A case study of peer review for randomized controlled trials. *BMC Medicine*, 12(1), 128. doi: 10.1186/s12916-014-0128-z

Pearson. (2012). Pearson launches higher education college. Retrieved from [www.pearson.com: http://www.pearson.com/news/2012/august/pearson-launcheshigher-education-college.html?article=true](http://www.pearson.com/news/2012/august/pearson-launcheshigher-education-college.html?article=true)

Perkel, J. M. (2014). Scientific writing: The online cooperative. *Nature*, 514(7520), 127-8. doi: 10.1038/514127a

Phillpotts, J., Devenport, T., & Mitchell, A. (2015). Evolution of the Transfer Code of Practice. *Learned Publishing*, 28(1), 75-79. doi:10.1087/20150112

Pinfield, S., Salter, J., Bath, P. A., Hubbard, B., Millington, P., Anders, J. H., & Hussain, A. (2014). Open-access repositories worldwide, 2005-2012: Past growth, current characteristics, and future possibilities. *Journal of the Association for Information Science and Technology*, 65(12), 2404-2421. doi:10.1002/asi.23131

PLOS. (2012). PLOS progress update 2011-2012. Retrieved from [http://www.plos.org/wpcontent/uploads/2012/09/2011\\_PLOS\\_Progress\\_Update\\_hi.pdf](http://www.plos.org/wpcontent/uploads/2012/09/2011_PLOS_Progress_Update_hi.pdf)

PLOS. (2014). PLOS progress update 2013/2014. Retrieved from <http://www.plos.org/about/plos/progress-update/>

Plume, A., & van Weijen, D. (2014, September). Publish or perish? The rise of the fractional author.... *Research Trends*, (38). Retrieved from [www.researchtrends.com: http://www.researchtrends.com/issue-38-september-2014/publish-or-perish-the-rise-of-the-fractional-author/](http://www.researchtrends.com/issue-38-september-2014/publish-or-perish-the-rise-of-the-fractional-author/)

Priem, J. (2010). Altmetrics: A manifesto. Retrieved from [altmetrics.org: http://altmetrics.org/manifesto/](http://altmetrics.org/manifesto/)

Procter, R., Williams, R., Stewart, J., Poschen, M., Snee, H., Voss, A., & Asgari-Targhi, M. (2010). Adoption and use of Web 2.0 in scholarly communications. *Philosophical Transactions of*

the Royal Society A: Mathematical, Physical and Engineering Sciences, 368(1926), 4039-4056.  
doi:10.1098/rsta.2010.0155

Pulverer, B. (2010). Transparency showcases strength of peer review. *Nature*, 468(7320), 29-31. doi:10.1038/468029a

Pulverer, B. (2014). The cost of open access. Wiley Exchanges. Retrieved from <http://exchanges.wiley.com/blog/2014/10/23/the-cost-of-open-access/>

RCUK. (2012). Research councils UK policy on access to research outputs. Retrieved from [http://www.rcuk.ac.uk/documents/documents/RCUK%20\\_Policy\\_on\\_Access\\_to\\_Research\\_Outputs.pdf](http://www.rcuk.ac.uk/documents/documents/RCUK%20_Policy_on_Access_to_Research_Outputs.pdf)

Renear, A. H., & Palmer, C. L. (2009). Strategic reading, ontologies, and the future of scientific publishing. *Science*, 325(5942), 828. doi:10.1126/science.1157784

Rieh, S., Markey, K., St Jean, B., Yakel, E., & Kim, J. (2007). Census of institutional repositories in the U.S. *D-Lib Magazine*, 13(11/12). Retrieved from <http://www.dlib.org/dlib/november07/rieh/11rieh.html>

RIN. (2008). Activities, costs and funding flows in the scholarly communications system in the UK. Research Information Network. Retrieved from <http://rinarchive.jisc-collections.ac.uk/ourwork/communicating-and-disseminating-research/activities-costs-and-funding-flowsscholarly-commu>

RIN. (2009a). Communicating knowledge: How and why UK researchers publish and disseminate their findings. Retrieved from <http://rinarchive.jisc-collections.ac.uk/our-work/communicating-and-disseminating-research/communicating-knowledge-how-and-whyresearchers-pu>

RIN. (2009b). E-journals: Their use, value and impact. Research Information Network. Retrieved from <http://rinarchive.jisc-collections.ac.uk/our-work/communicating-anddisseminating-research/e-journals-their-use-value-and-impact>

RIN. (2009c). Patterns of information use and exchange: Case studies of researchers in the life sciences. Retrieved from <http://rinarchive.jisc-collections.ac.uk/our-work/using-andaccessing-information-resources/patterns-information-use-and-exchange-case-studie>

RIN. (2010). If you build it, will they come? How researchers perceive and use Web 2.0. London: Research Information Network. Retrieved from <http://rinarchive.jisc-collections.ac.uk/ourwork/communicating-and-disseminating-research/use-and-relevance-web-20-researchers>

RIN. (2011a). Access to scholarly content: Gaps and barriers. Retrieved from <http://rinarchive.jisc-collections.ac.uk/node/1172>

RIN. (2011b). E-journals: Their use, value and impact – final report. Retrieved from <http://rinarchive.jisc-collections.ac.uk/our-work/communicating-and-disseminating-research/ejour>

nals-their-use-value-and-impact

RIN. (2011c). Heading for the open road: Costs and benefits of transitions in scholarly communications. Retrieved from [http://www.rin.ac.uk/system/files/attachments/Dynamics\\_of\\_transition\\_report\\_for\\_screen.pdf](http://www.rin.ac.uk/system/files/attachments/Dynamics_of_transition_report_for_screen.pdf)

RIN. (2011d). Reinventing research? Information practices in the humanities. Retrieved from <http://rinarchive.jisc-collections.ac.uk/our-work/using-and-accessing-informationresources/information-use-case-studies-humanities>

RIN. (2012). Physical sciences case studies: Information use and discovery. Retrieved from <http://rinarchive.jisc-collections.ac.uk/our-work/using-and-accessing-information-resources/physical-sciences-case-studies-use-and-discovery->

RIN. (2014). Nature Communications: Citation analysis. Nature Publishing Group. Retrieved from [http://www.nature.com/press\\_releases/ncomms-report2014.pdf](http://www.nature.com/press_releases/ncomms-report2014.pdf)

Romary, L., & Armbruster, C. (2009). Beyond institutional repositories. Retrieved from [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1425692](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1425692)

Rowlands, I., & Nicholas, D. (2005). New journal publishing models. CIBER. Retrieved from [http://www.ucl.ac.uk/ciber/ciber\\_2005\\_survey\\_final.pdf](http://www.ucl.ac.uk/ciber/ciber_2005_survey_final.pdf)

Rowlands, I., Nicholas, D., & Huntingdon, P. (2004). Scholarly communication in the digital environment: What do authors want,. CIBER. Retrieved from <http://www.ucl.ac.uk/ciber/ciber-pa-report.pdf>

Royal Society. (2011). Knowledge, networks and nations: Global scientific collaboration in the 21st century. The Royal Society. <http://sro.sussex.ac.uk/id/eprint/45410>

Salo, D. (2008). Innkeeper at the roach motel. *Library Trends*, 57(2), 98-123. doi:10.1353/lib.0.0031

Samuelson, P. (2012). Reforming copyright is possible. *The Chronicle of Higher Education*, 9 July 2012.

ScienceWatch. (2012). Multiauthor papers: Onward and upward. Thomson Reuters. Retrieved from <http://sciencewatch.com/multiauthor-papers-onward-and-upward>

Seglen, P. O. (1992). The skewness of science. *Journal of the American Society for Information Science*, 43(9), 628-638. doi:10.1002/(SICI)1097-4571(199210)43:9%3C628::AID-ASIS5%3E3.0.CO;2-0

Sense About Science. (2009). Peer review survey 2009: Preliminary findings. Sense About Science. Retrieved from <http://www.senseaboutscience.org/pages/peer-review.html>

Shotton, D. (2009). Semantic publishing: The coming revolution in scientific journal

publishing. *Learned Publishing*, 22, 85-94. doi:10.1087/2009202

Shotton, D. (2010). CiTO, the citation typing ontology. *Journal of Biomedical Semantics*, 1(Suppl 1), S6. doi:10.1186/2041-1480-1-S1-S6

Simba. (2011). Global STM publishing 2010-2011. Retrieved from <https://www.simbainformation.com/Global-STM-Publishing-6059305/>

Simba. (2013a). Global medical publishing 2013-2014. Retrieved from <http://www.simbainformation.com/Global-Medical-Publishing-7681197/>

Simba. (2013b). Global scientific and technical publishing 2013-2014. Retrieved from <http://www.simbainformation.com/Global-Scientific-Technical-7681199/>

Simba. (2014). New Simba report sizes the market for open access journal publishing. Retrieved from <http://www.simbainformation.com/about/release.asp?id=3618>

Smit, E., & van der Graaf, M. (2011). Journal article mining: A research study into practices, policies, plans... and promises. Publishing Research Consortium. Retrieved from <http://www.publishingresearch.org.uk/documents/PRCSmitJAMreport2.30June13.pdf>

Somerville, M., & Conrad, L. (2014). Collaborative improvements in the discoverability of scholarly content. A SAGE white paper. SAGE. Retrieved from <http://www.sagepub.com/repository/binaries/pdf/improvementsindiscoverability.pdf>

SPARC, & PLOS. (2013). How open is it? Retrieved from <http://www.plos.org/wp-content/uploads/2013/12/hoii-guide.pdf>

Sparks, S. (2005). JISC disciplinary differences report. JISC/Rightscom. Retrieved from <http://www.jisc.ac.uk/media/documents/themes/infoenvironment/disciplinarydifferencesneeds.pdf>

STM. (2012). Text and data mining: STM statement & sample licence. Retrieved from <http://www.stm-assoc.org/text-and-data-mining-stm-statement-sample-licence/>

STM. (2013a). A statement of commitment by STM publishers to a roadmap to enable text and data mining (TDM) for non-commercial scientific research in the European Union. Retrieved from [http://www.stm-assoc.org/2013\\_11\\_11\\_Text\\_and\\_Data\\_Mining\\_Declaration.pdf](http://www.stm-assoc.org/2013_11_11_Text_and_Data_Mining_Declaration.pdf)

STM. (2013b). Text and data mining for non-commercial scientific research. Retrieved from <http://www.stm-assoc.org/text-and-data-mining-stm-statement-sample-licence/>

STM. (2014). STM response to 'Global coalition of access to research, science and education organisations calls on STM to withdraw new model licenses'. Retrieved from [http://www.stm-assoc.org/2014\\_08\\_07\\_STM\\_response\\_to\\_call\\_for\\_license\\_withdrawal.pdf](http://www.stm-assoc.org/2014_08_07_STM_response_to_call_for_license_withdrawal.pdf)

STM Future Lab Committee. (2014). STM Future Lab trend watch 2014. Retrieved from <http://www.stm-assoc.org/future-lab-trend-watch/>

Strieb, K., & Blixrud, J. (2013). The state of large-publisher bundles in 2012. ARL. Retrieved from <http://publications.arl.org/rli282/13>

Sturges, P., Bamkin, M., Anders, J. H., Hubbard, B., Hussain, A., & Heeley, M. (2014). Research data sharing: Developing a stakeholder-driven model for journal policies. *Journal of the Association for Information Science and Technology*. <http://eprints.nottingham.ac.uk/id/eprint/3185>

Swan, A., & Houghton, J. (2012). Going for gold? The costs and benefits of gold open access for UK research institutions: Further economic modelling. UK Open Access Implementation Group. Retrieved from <http://repository.jisc.ac.uk/610/>

Taylor & Francis. (2014). Open access survey June 2014. Retrieved from <http://www.tandfonline.com/page/openaccess/opensurvey/2014>

Tenopir, C. (2000). *Towards electronic journals: Realities for scientists, librarians, and publishers*. Washington, DC: Special Libraries Association. Retrieved from [http://works.bepress.com/carol\\_tenopir/7/](http://works.bepress.com/carol_tenopir/7/)

Tenopir, C. (2007). What does usage data tell us about our users? Presentation at Online Information, London

Tenopir, C. (2010). Research publication characteristics and their relative values. Publishing Research Consortium. Retrieved from [www.publishingresearch.net/projects.htm](http://www.publishingresearch.net/projects.htm)

Tenopir, C., King, D. W., Edwards, S., & Wu, L. (2009). Electronic journals and changes in scholarly article seeking and reading patterns. *Aslib Proceedings*, 61(1), 5-32. doi: 10.1108/00012530910932267

Tenopir, C., Mays, R., & Wu, L. (2011). Journal article growth and reading patterns. *New Review of Information Networking*, 16(1), 4-22. doi:10.1080/13614576.2011.566796

Tenopir, C., Volentine, R., & King, D. W. (2012). Scholarly reading and the value of academic library collections: Results of a study in six UK universities. *Insights: The UKSG Journal*, 25(2), 130-149. doi:10.1629/2048-7754.25.2.130

Terry, R. (2005). Funding the way to open access. *PLoS Biol*, 3(3), e97. doi:10.1371/journal.pbio.0030097

Thomson Reuters. (2012). Global publishing – changes in submission trends and the impact on scholarly publishers. Retrieved from [http://scholarone.com/media/pdf/GlobalPublishing\\_WP.pdf](http://scholarone.com/media/pdf/GlobalPublishing_WP.pdf)

Thomson Reuters. (2014). The research & innovation performance of the G20. Retrieved from <http://sciencewatch.com/grr/the-g20-nations>

Thorn, S., Morris, S., & Fraser, R. (2009). *Learned societies and open access: Key results*

from surveys of bioscience societies and researchers. *Serials: The Journal for the Serials Community*, 22(1), 39-48. doi:<http://dx.doi.org/10.1629/2239>

Treloar, A. (2014). The Research Data Alliance: Globally co-ordinated action against barriers to data publishing and sharing. *Learned Publishing*, 27(5), 9-13. doi:10.1087/20140503

UNESCO. (2010). UNESCO Science Report 2010. Retrieved from <http://www.unesco.org/new/en/natural-sciences/science-technology/prospective-studies/unesco-science-report/unesco-science-report-2010/>

University of California. (2007). Faculty attitudes and behaviors regarding scholarly communication: Survey findings from the University of California. Retrieved from <http://osc.universityofcalifornia.edu/responses/materials/OSC-survey-full-20070828.pdf>

Van Noorden, R. (2011). Science publishing: The trouble with retractions. *Nature*, 478, 26-28. doi:10.1038/478026a

Van Noorden, R. (2012a). Britain aims for broad open access. *Nature*, 486(7403), 302. doi:10.1038/486302a

Van Noorden, R. (2012b). Journal offers flat fee for 'all you can publish'. *Nature News*, 486(7402), 166. doi:10.1038/486166a

Van Noorden, R. (2013). Open access: The true cost of science publishing. *Nature News*, 495(7442), 426. doi:10.1038/495426a

Van Noorden, R. (2014a). The arXiv preprint server hits 1 million articles. *Nature News*. doi:10.1038/nature.2014.16643

Van Noorden, R. (2014b). Google Scholar pioneer on search engine's future. *Nature News*. doi:10.1038/nature.2014.16269

Van Noorden, R. (2014c). Online collaboration: Scientists and the social network. *Nature News*, 512(7513), 126. doi:10.1038/512126a

Van Noorden, R. (2014d). The scientists who get credit for peer review. *Nature News*. doi:10.1038/nature.2014.16102

Van Orsdel, L. C. (2008). The state of scholarly communications. *The Serials Librarian*, 52(1-2), 191-209. doi:10.1300/J123v52n01\_16

van Rossum, J. (2014). How we can better support and recognize reviewers. Elsevier Editor's Update. Retrieved from <http://editorsupdate.elsevier.com/issue-42-march-2014/can-bettersupport-recognize-reviewers/>

Verstak, A., Acharya, A., Suzuki, H., Henderson, S., Iakhiaev, M., Lin, C. C. Y., & Shetty, N. (2014). On the shoulders of giants: The growing impact of older articles. *ArXiv Preprint*

ArXiv:1411.0275. doi:arXiv:1411.0275

de Vrieze, J. (2012). Horizon 2020: A €80 billion battlefield for open access. Science Insider. Retrieved from <http://news.sciencemag.org/scienceinsider/2012/05/horizon-2020-a-80-billion-battle.html>

Waldrop, M. M. (2008). Science 2.0. Scientific American, 298(5), 68-73. doi:10.1038/scientificamerican0508-68

Wallace, J. (2012). PEER project: Final report. PEER. Retrieved from [www.peerproject.eu](http://www.peerproject.eu): <http://www.peerproject.eu/reports/>

WAME. (2009). Conflict of interest in peer-reviewed medical journals. World Association of Medical Editors. Retrieved from <http://www.wame.org/conflict-of-interest-in-peerreviewed-medical-journals>

Ware, M. (2008). Peer review: Benefits, perceptions and alternatives. Publishing Research Consortium. Retrieved from <http://www.publishingresearch.net/documents/PRCsummary4Warefinal.pdf>

Ware, M. (2009). Access by UK small and medium-sized enterprises to professional and academic information. Publishing Research Consortium. Retrieved from <http://www.publishingresearch.net/SMEaccess.htm>

Ware, M. (2011). Peer review: Recent experience and future directions. New Review of Information Networking, 16(1), 23–53. doi:10.1080/13614576.2011.566812

Ware, M., & Monkman, M. (2008). Peer review in scholarly journals: Perspective of the scholarly community--an international study. Publishing Research Consortium. Retrieved from <http://www.publishingresearch.net/PeerReview.htm>

Weckowska, D. (2014). Open access publishing and innovation

Willinsky, J. (2003). The nine flavours of open access scholarly publishing. Journal of Postgraduate Medicine, 49(3), 263-267.

Wolters Kluwer Health. (2013). 2013 physician outlook survey. Retrieved from <http://www.wolterskluwerhealth.com/News/Documents/White%20Papers/Wolters%20Kluwer%20Health%20Physician%20Study%20Executive%20Summary.pdf>

Wolters Kluwer Health. (2014). Mobile device internet and social media use and habits survey report. Retrieved from <http://www.wolterskluwerhealth.com/News/Pages/Wolters-Kluwer-Health-Survey-Finds-Nurses-and-Healthcare-Institutions-Accepting-Professional-Use-of-Online-Reference--Mobi.aspx>