



图片来源:《自然》

迎接开放获取大时代——科学文献引用现状及趋势扫描

去年5月,当Heather Piwowar开始调查公开研究数据是否能增加论文引用率时,她从未想到会遭遇困难。作为加拿大开放指标服务ImpactStory的共同创始人,那时的Piwowar是美国北卡罗来纳州杜克大学的一名博士后。由于缺乏从机构层面进入爱思唯尔SCOPUS数据库的途径,Piwowar最终通过和加拿大国家科学图书馆的一份研究协议获得了使用权。

但这需要提取Piwowar的指纹以获得一份证明。“我浪费很多时间在获取研究所需的引文数据上。这很可笑。”Piwowar需要分析1万篇文章的引用次数,但是其他主要的引文来源,诸如汤森路透科学网,在当时并不支持使用PubMed文献服务系统独特数字标识符的查询。Piwowar解释:“要是能公开的引文数据,我就可以写我自己的文章了!”

在美国马萨诸塞州波士顿哈佛大学医学院神经病学家Steven Greenberg准备研究假说是如何被转变为“事实”(仅通过重复引用)时,他有着和Piwowar相似的经历。Greenberg构建和分析了一个引文网络,包括242篇论文、675份引文、553个截然不同的引文路径(和一个特定的假说相关)。如果这些引文数据能够轻易在网上获取,那将能使Greenberg少费许多周折。

《自然》杂志在线报道称,在这个开放获取的时代,学者不能免费使用期刊论文上的参考文献,这着实是一桩不可思议的事情。为了改变这一现状,引文数据需要被视为公有物的一部分,并安放在一个公开的资源库中。为了达到这个目的,自2010年起,David Shotton领导了一个JISC(一家英国信息技术研究与开发资金组织)资助的价值13.2万英镑(约合21.2万

美元)的项目,建立并发展了开放引文语料库(OCC)。在开放的学术引用数据方面,OCC还是一个“羽翼未丰”的资源库,目前正在寻求可持续发展的资金。

开放式获取

直接引用一直是衡量产出重要性的一个主要指标。学术交流包括引用网络信息和观点的流动,以及随着时间的推移分析网络的变换,从而揭示学者间沟通模式的变换和学科的发展与消亡。这种信息对学术研究至关重要,对于制定正确的研究投资和战略、促进创新、增长和繁荣,尤其是在日益国际化的研究合作中也非常重要。

最权威的学术引用数据来源要属汤森路透科学网,其从美国科学家Eugene Garfield于1964年创建的科学文献索引发展而来,最初由科学信息研究所(ISI)出版;它的主要商界对手——爱思唯尔数据库,诞生于2004年。这两大巨头都囊括主要的学术文献,但由于都不完整,它们也互为补充。

为了使用这两大资源,英国每所研究型大学每年需花费数万英镑。这一高昂的花费严重损害了那些不在这些机构工作的群体的利益,包括大多数企业和公众。其他引文信息的重要来源也由商业公司运营,但是无需订阅即可获取,例如谷歌学术搜索和微软学术搜索。谷歌学术搜索的资源库比同类型的资源库大,因为其包括书籍、论文、预印本、技术报告和其他非同议的“灰色”文献。

所有这些资源都有许可限制,阻止对其引文数据的重新发布。更糟的是,可用的引用数

据是不准确的。在汤森路透科学网、斯高帕斯数据库、谷歌学术搜索和微软学术搜索上,Shotton的引用记录竟截然不同。例如,Shotton于2009年发表的一篇语义学论文,在这4家平台的引用次数分别是22、37、88和16。究竟该相信哪一个?更令人担心的是,一份更早的关于蛋白质结晶学的论文在汤森路透科学网上有3个单独的条目,每个条目的引用次数分别为59、19和0。Shotton据此怀疑汤森路透期刊影响因素的可信度。

渐成主流

理想情况下,出版商会把自己的出版书目和引用数据拿出来分享,比如自然出版集团把自己的数据发布在data.nature.com上。它是第一家也是唯一一家将数据共享的公司。

但是,如果能够将这些数据整合在一起,则会获得更大的收益。OCC将为用户浏览引用数据提供完整的权限,这些数据的来源非常广阔,既有传统学术出版物的数据,也有其他数据,并且所有数据都会标明来源。OCC将把文章与文章、文章与数据库、数据库与文章引用的关系清晰展现出来。此外,还将提供附加信息,例如作者、机构成员、共同的资助人、文章间语义关系,以及数据来源。

一旦引用数据得以开放,有益的分析服务也将开展,包括面搜索与浏览工具、建议与趋势识别服务,以及时间表可视化服务。对于这些服务,目前已经开发出了其中几种的原型。OCC对统计引用指标的巨大价值也将随着其覆盖面的延伸而增加。

此外,还有另一种尤其对作者和编辑有益

“在这个开放获取的时代,学者不能免费使用期刊论文上的参考文献,这着实是一桩不可思议的事情。

的服务:错误引用更正服务。目前,出版论文的引用中有约1%存在着不同程度的错误,小到在引用标题中将“β-淀粉酶”写成“β-淀粉酶”,或者对作者名字中符号的疏忽;大到年份、期刊号、页码或者数字对象标识符的错误。OCC已经在内部使用引用更正方法,以处理多次引用问题,或者通过外部资源获取权威书目记录。

未来展望

接下来会有怎样的发展呢?就在10年前,也有一个具有相似目标,旨在开放引文数据的项目存在,名为开放引文计划。这是一个由英国南安普顿大学、美国纽约州伊萨卡市康奈尔大学以及arXiv网站共同构建的项目,该项目从1999年开始到2002年共维持了3年时间。

该项目开发了一款名为引用库的软件,能将引用信息记录下来,该软件的设计者将它描述为“开放引文计划皇冠上的珠宝”。但是,如今再点击该项目的网址citebase.eprints.org,得到的提示是:该网址目前没有网站链接。

要想使纸面上雄心勃勃的计划真的成为现实,并将它长期维持下去,是一件极为困难的事情。为了避免OCC重蹈引用库的覆辙,使它得到全面的发展并最终向整个学术界各个学科提供令人信任的引用数据开放服务,需要拥护者、管理者、开发者以及监督者共同的努力。此外,还需要怀有相同目标的同道中人精诚合作、来自资助者充足持续的资金、支持者和投资者以提升社会福利而不是经济回报为导向的决心、来自出版界的鼎力支持、重要机构或者国际组织的认同。你能贡献一份力量吗? (段歆滢)

朱莉:好榜样? 坏榜样? 科学家认为检测乳腺癌基因可能带来激烈预防手段



安吉丽娜·朱莉在得知其遗传了BRCA1突变后,决定进行双乳房切除术。图片来源:Wikimedia

但是,与会的其他研究人员并不能很好地接受相关理念。英国萨顿癌症研究所的Nazneen Rahman指出,对于为那些非BRCA的基因进行临床检测而言,“它是不可靠和不能接受的”,且风险并没有很好地被确认。这样做是“危害严重的”,她在接受《科学》杂志采访时表示,因为它可能会让女性采取切除乳房等激烈的措施。King研究小组将“研究检测和临床检测混为一谈”,她说。

另一方面,King研究小组为自己的工作进行了辩护。Walsh表示,内科医生收到的结果仅包括12个基因,这些证据显示了可以加倍或四倍罹患癌症的风险,换句话说,一位女性的乳腺癌发病风险为25%~50%。当该研究小组发现这些基因后,他们没有推荐外科手术,而是建议定期进行影像学检查。

King指出,在1997年学界达成共识之前,对BRCA1基因进行检测也备受争议,“我有一种似曾相识的感觉”。她说她的批评者误认为临

床BRCA测试的返回结果是所有40个基因,不仅仅是12个。基于自己的研究和文献,“对于那些我们已经拥有数据的基因,我们会回馈结果”。她说,“但是我们也非常保守,不会回报不确定特征的变体。”

King研究小组的外科医生Greta Bernier补充道,尽管只检测BRCA基因的测试得到了美国国家综合癌症网络(该网络发行了癌症治疗指导方针)的推荐,但是她的研究小组希望该网络能够迅速将更多的基因加入到其列表中。Bernier还指出,有医生已经从基因测试公司订购了检验其研究小组列表上基因的测试。

澳大利亚墨尔本大学的Melissa Southey相信,在研究中检测所有癌症风险基因是十分重要的,仅对几个基因甚至只是几个特殊变体而言,临床检测得出的证据是足够坚实的。但Southey还提到,问题是有许多其他变体非常稀少,可能人们永远无法足够了解它们,以确定其风险。相反,研究人员需要依靠其他证据证明突

变是有害的,例如有关基因蛋白质编码的实验室研究。

Southey还补充道,遗传学家在探索其他稀有疾病风险遗传变体时,也面临同样的困难。“我们正处在边缘地带,我们有许多信息,但是不足以提出建议。”她说。(张章)

延伸阅读

2002年6月,美国哈佛大学附属组织黛娜-法伯-吉勒特癌症研究所和儿童医院宣布了一项重大的科学突破:最新发现的6种基因可能与乳腺癌的发生有关联。该项目研究员Alan D'Andrea说:“这对于乳腺癌和卵巢癌的诊断来说具有非常重要的意义。”这项发现将会帮助科学家更精确地判断易患癌症的人群。

研究人员相信,一些女性携带着1~2种变异的基因,这些基因会在很大程度上增大患乳腺癌的危险。科学家相信,如果BRCA1和BRCA2基因或其中的一种出现功能异常,乳腺和卵巢细胞中就会出现DNA的大量受损,从而导致癌症。美国梅奥诊所乳房健康专家Sandya Pruthi介绍,所有BRCA1基因发生变异的女性中,20%会在40岁之前患乳腺癌,51%在51岁之前,87%在60岁之前。

一旦知道基因发生了变异,人们的医疗选择就会受到影响。美国国立癌症研究院表示,“对于有着较多乳腺癌家族病例的女性来说,预防性乳房切除手术是一种选择”。但是“尽管预防性的乳房切除术能够减小(患乳腺癌的)危险,但谁也不能保证这就一定能让女性免于患乳腺癌”。澳大利亚昆士兰科技大学公共卫生学教授Beth Newman就担心基因测试会引起更多的预防性乳房切除手术。她认为在诊断乳腺癌的发生方面,基因的信息所占的比重应当是有限的。

科学线人

全球科技政策新闻与解析

政事

欧洲食品监管部门 独立性遭质疑



报告称欧洲食品安全局的利益冲突仍“令人沮丧”。图片来源:跨国公司欧洲观察

2012年,欧洲食品安全局(EFSA)批准了一项新独立政策,以设法应对利益冲突问题。但是跨国公司欧洲观察(CEO)日前发布的新报告称,产业界对EFSA的影响依然“令人沮丧”。

这份报告指出,存在经济利益冲突的专家占据了EFSA几乎所有的专家小组,仅有一个除外。调查结果显示,在最糟糕的情况下,饮食产品专家小组20位成员中有17位牵扯到利益关系,而在营养和过敏小组中,全部的113位成员都涉及利益冲突。

EFSA发言人表示,该机构无法立刻对个别申诉作出反应,但是它将会对这份报告进行审查,并对其提出的建议进行考虑。他还指出,EFSA“采取了一系列健全的内部机制和工作程序,来捍卫其科学工作的独立性”。EFSA管理委员会主席Sue Davies则重申,“管理委员会相信,EFSA的政策在确保其科学工作独立性方面是健全的”。

这份报告在EFSA的一个敏感时刻到来。自2006年起担任该局局长的法国食品安全专家Catherine Geslain-Laneelle出乎意料地于7月辞职,EFSA正在寻找其继任者。

不过,CEO的发言人Martin Pigeon承认,EFSA已经改进了其管理潜在利益冲突的方式。“现在,他们确实在认真处理这个问题。”Pigeon说。作为EFSA新政策的一部分,该机构咨询的专家必须提交详细的利益声明。但是CEO指出,这些声明的形式仍计划不周,例如,它们不包括进行会议发言等一些普通活动,即便这些会议是由企业赞助的。

该报告还指责,在10种情况里,EFSA并没有严格按照规定行事。不过,有人认为,该报告的大部分指责都源于CEO定义的“利益冲突”比EFSA的定义宽泛得多。不过CEO坚持,一个与巧克力工厂有联系的科学家,也不能出现在动物食品专家小组中。(张章)

人事

法国数学家将担任 欧洲研究委员会新主席



数学家Jean-Pierre Bourguignon 图片来源:Wikimedia Commons

一位有着国际声望以及丰富国际关系资源的法国数学家被邀请成为欧洲研究委员会(ERC)新主席。ERC是欧洲前沿科学的基金资助机构。在法国伊维特河畔比尔的高级科学研究所(IHES)长期担任所长的Jean-Pierre Bourguignon将接任奥地利社会学家Helga Nowotny的ERC主席职位,后者从2010年开始主持该机构。

Bourguignon毕业于巴黎综合理工大学,专业为微分几何,并且获得了巴黎第七大学的博士学位。

在领导IHES的19年里,他帮助该机构获得新资金,并吸引了世界各地的顶尖人才。他还促使这个传统上主要关注数学领域的机构,将研究范围扩展到理论物理学和理论生物学等学科上。

总部在法国斯特拉斯堡的欧洲科学家组织——欧洲科学在近日发表的声明中表示欢迎这一即将到来的任职。“Bourguignon对维持和加强IHES高水平原创和创造性研究氛围、为IHES吸引精英等至关重要。”

在成立的最初7年里,ERC在欧洲科学家中非常受欢迎,因为和其他欧盟的基金资助机构相比,它将“杰出性”作为唯一的衡量标准,其审核申请的过程相对较严格。但事实上,不成比例的拨款更偏向于欧盟的富裕国家,这偶尔会引起东部和南部成员国的不满。

Nowotny拒绝对新任命进行评论,并称,她的继任者的主要挑战是维持ERC资助的科学研究的杰出性和在审查上的声望。另一个挑战将是欧盟以外吸引更多的捐资者。(张冬冬)