

BDM 编辑张月红解译 | 《自然》独家发文：21 世纪被引用最多的文章

原创 BDM 编辑部张月红 [生物设计与制造 BDM](#)

2025 年 04 月 18 日 20:59 浙江公众号：生物设计与制造 BDM

编者按

《自然》杂志的一项分析揭示了本世纪发表的 25 篇被引用次数最高的论文，并探讨了它们打破记录的原委。

原文链接（点击最下方阅读原文可跳转）

Exclusive: the most-cited papers of the twenty-first century——A *Nature* analysis reveals the 25 highest-cited papers published this century and explores why they are breaking records.
<https://www.nature.com/articles/d41586-025-01125-9>

文章开篇介绍：21 世纪的前 25 年出现了一些巨大的科学突破，从第一批 mRNA 疫苗和基于 crispr 的基因编辑技术，到希格斯玻色子的发现和第一次引力波的测量。但你不会在 2000 年以来发表的被引用最多的论文中找到这些进展（即在这张引用表中无踪迹）

这是《自然》新闻团队对 21 世纪发表的 25 篇被引用次数最多的论文进行分析后得出的结论之一。

被引用次数最多的文章是人工智能（AI）的发展；提高研究质量或系统评价的方法；癌症统计数据；研究软件。然而，2004 年一篇关于石墨烯实验的开创性论文也是 21 世纪被引用最多的论文之一，该论文的作者因此获得了 2010 年的诺贝尔物理学奖。

引用是作者在文献中承认先前研究工作来源的方式——也是衡量一篇论文影响力的一个指标。但被引用次数最多的论文通常并非最著名的科学发现。相反，这些作品更倾向于描述科学方法或软件，科学家们所依赖的工具。“科学家们说他们重视方法、理论和实证发现，但在实践中，方法被引用的更多。”

《自然》的分析依据所选的五家数据库的数据（Web of Science, Scopus, OpenAlex, Dimensions, Google Scholar），它们涵盖了 21 世纪发表的数千万篇论文，依据此并选取中位数，最后综合评出了“21 世纪十大被引用次数最多的作品”（表 1）。

见表 1 所示，21 世纪被引用最多的论文是 2016 年科技巨头微软研究人员关于“深度残差学习”网络（ResNets）的一篇报告（Deep residual learning for image recognition）。这是一种基于神经科学启发的算法所构建的人工神经网络，为深度学习及后续人工智能的发展提供了重要支撑。该文描述了一种能训练约 150 层神经网络的方法，其深度是常规模型的 5 倍。ResNets 架构解决了信号在多层传输时耗散的问题。这项研究在 2015 年末首次以预印本形式发表，当时微

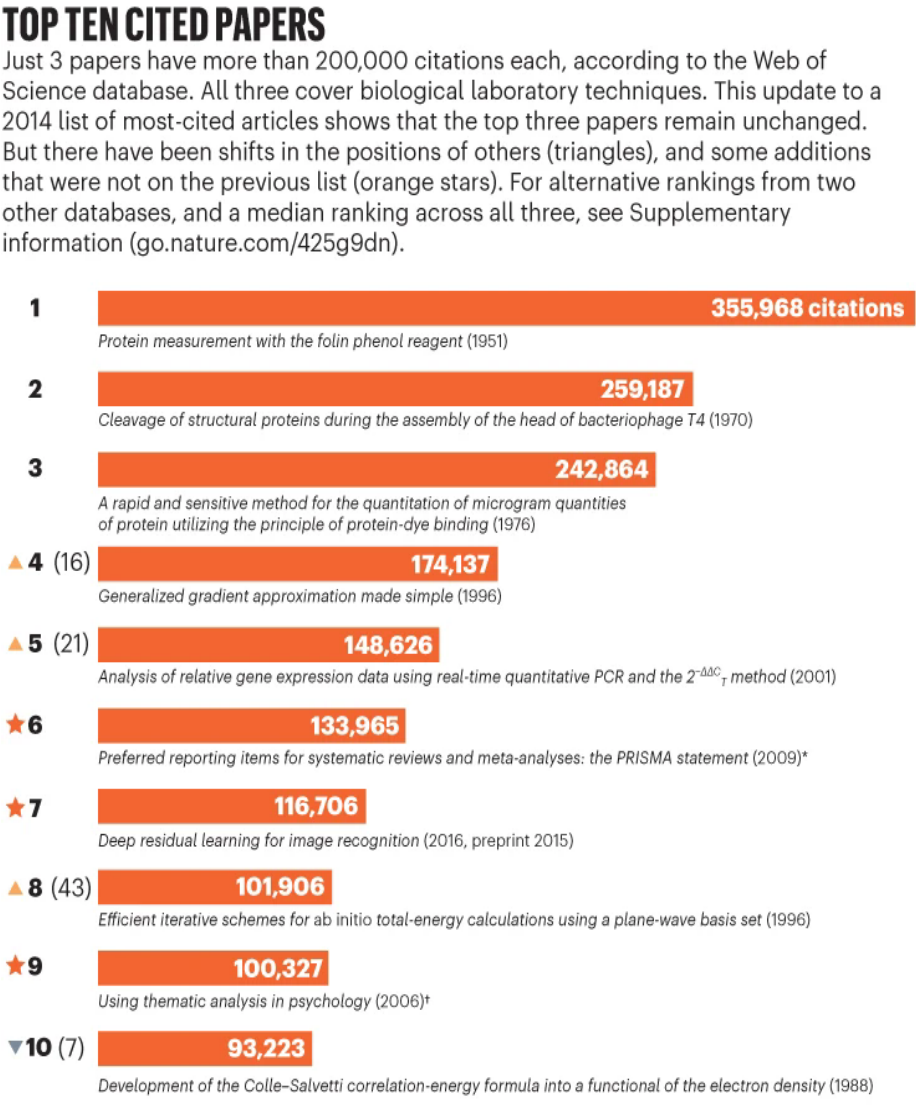
软的研究人员宣布，他们利用这种方法在一场图像识别竞赛中获胜。ResNets 背后的概念是导致人工智能工具能够玩棋盘游戏（AlphaGo），预测蛋白质结构（AlphaFold）并最终建模语言（ChatGPT）的因素之一。论文作者之一、现供职于麻省理工学院（Massachusetts Institute of Technology）的 Kaiming He 表示，在他们之前，“深度学习并非这么深奥”。

表 1. 21 世纪被引用次数最多的 10 项工作

Top ten most-cited works of the twenty-first century		
Rank (median)	Citation	Times cited (range across databases)
1	Deep residual learning for image recognition (2016, preprint 2015)	103,756–254,074
2	Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_T}$ method (2001)	149,953–185,480
3	Using thematic analysis in psychology (2006)	100,327–230,391
4	<i>Diagnostic and Statistical Manual of Mental Disorders, DSM-5</i> (2013)	98,312–367,800
5	A short history of SHELX (2007)	76,523–99,470
6	Random forests (2001)	31,809–146,508
7	Attention is all you need (2017)	56,201–150,832
8	ImageNet classification with deep convolutional neural networks (2017)	46,860–137,997
9	Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries (2020)	75,634–99,390
10	Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries (2016)	66,844–93,433

微软的论文不仅是本世纪发表论文中的热门之作，据本期《自然》发表的另一篇分析文章“有史以来被引用次数最多的 10 篇研究论文”（These are the most-cited research papers of all time, <https://www.nature.com/articles/d41586-025-01124-w>）中这篇也是名列前茅。当然，比较引用次数或有不公平和不一致之嫌。正如微软的研究发表发生在十年前，它比年轻的论文有更多的时间积累而被更多引用，同时它也与当下热门领域——计算机科学研究的大量产出有关。为此《自然》委托文献计量学家进行了一些因素调控分析，但这篇“深度残差学习”同样出现在史上被引用次数最多的列表中（见表 2）。

表 2. 有史以来被引用次数最多的 10 篇研究论文



Data show citations from Web of Science ‘Core Collection’ journals as of March 2025, to permit comparison with 2014 list (*Nature* **514**, 550–553; 2014). Orders would change if citation metrics from other databases were included (see Supplementary information).
*Paper was published in multiple journals simultaneously. This total aggregates citations to all journal versions.
†Corrected for data error in Web of Science, which lists a different paper by the same authors.

《自然》的这篇独家文章也探索了这些业已成为 21 世纪引用大师的作品，概括了几个特点：

特点 1：人工智能的崛起 (The rise of AI)

因为深度学习革命而获得 2024 年诺贝尔物理学奖的计算机科学家杰夫·辛顿 (Geoff Hinton) 指出“人工智能论文在被引用方面具有天然优势，缘由 AI 在 21 世纪取得了极其迅速的进展，论文数量也非常多。”从表 1 的 21 世纪引用大师的作品中不难看出，首篇加第 7 和第 8 篇，人工智能的文章几乎占据了列表的三分之一。

特点 2：研究软件 (Research software)

根据 Web of Science 的数据，21 世纪被引次数排行表排位第 2，发表于 2001 年的这篇利用实时定量方法分析相关基因表达数据的文章并非偶然上榜，它的写作显然是为研究人员提供了一些可引用的东西。大约 25 年前，制药科学家托马斯·施米特根 (Thomas Schmittgen) 提交了一篇论文，其中包括一种名为定量 PCR 的技术的数据，该技术允许研究人员量化样本中的 DNA 含量。为了分析数据，施米特根使用了一本技术手册中的方程式。回忆“其中一位审稿人返回意见说，‘你不能在论文中引用用户手册’”。于是，施米特根联系了这些方程式的创造者，共同发表了这篇为生物学家提供了计算基因活性在不同条件下的变化的这种简单的公式，至今引用次数超过 16.2 万次。

的确，正如研究人员发现，人工智能 (AI)、科学软件及提高研究质量的方法等主题确实主导了今天 21 世纪研究论文的参考文献列表。

如本期《自然》的第 3 篇引文分析文章，题为“科学的黄金老歌：几十年前的研究论文今天仍然被大量引用” (Science’s golden oldies: the decades-old research papers still heavily cited today, <https://www.nature.com/articles/d41586-025-01126-8>) 中所示，即使到现在，2000 年之前发表的一些研究仍然得到广泛认可。施米特根这篇很受欢迎的文章也在 2023 年发文中引用排名第 7 (参见表 3)

特点 3：癌症与健康研究 (Cancer and health research)

三篇被引最多的论文几乎是癌症研究工作中引言的常客。其中两篇 (第 9 和第 10) 是全球癌症数据库 (GLOBOCAN) 2020 年和 2018 年的报告，这也是世界卫生组织的一个项目，GLOBOCAN 每 1-2 年发布跟踪全球癌症的统计数据，给特定癌症类型发病率和死亡率的研究人员无疑提供了医疗倡导和政策制定的参考数据。

表 3. 2023 年发文中哪些文章被引用最多？

Which articles are referenced the most?		
The top-ten cited references across three databases* for all papers published in 2023.		
Rank	Title	Number of citations (range)
1	Deep residual learning for image recognition (2016)	19,826–33,339
2	Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries (2021)	10,952–24,830
3	Attention is all you need (2017)	9,395–19,348
4	Generalized gradient approximation made simple (1996)	14,338–17,540
5	Using thematic analysis in psychology (2006)	10,660–17,347
6	The PRISMA 2020 statement: an updated guideline for reporting systematic reviews (2021)	13,115–13,443
7	Analysis of relative gene expression data using real-time quantitative PCR and the 2 ^{-ΔΔC_T} method (2001)	11,851–13,082
8	Random forests (2001)	6,266–12,294
9	Efficient iterative schemes for <i>ab initio</i> total-energy calculations using a plane-wave basis set (1996)	10,063–10,789
10	Long short-term memory (1997)	2,394–12,355

*Web of Science, Cross-Ref and Scopus. Ordered according to median of rank in the three databases.

排名第 4 的是有时被称为“精神病学圣经”的《精神疾病诊断与统计手册》（DSM-5）的第五版，它于 2013 年出版，比上一版晚了近 20 年。这本书描述了包括成瘾和抑郁在内的精神障碍的分类和诊断标准，被世界各地的研究人员和卫生专业人员广泛参考使用。

特点 4：提高研究质量（Improving research quality）

心理学家维吉尼亚·布劳恩和维多利亚·克拉克在他们关于性别和性的研究论文中，好像已经习惯被引用不多的事实。因此，当他们得知 2006 年发表的这篇论文竟然成为本世纪被引用次数第 3 的文章时，惊讶地说“它有了自己的生命。”

这篇文章是两人在 2005 年写成的，当时布劳恩正在英国的西英格兰大学休假，克拉克也在那里工作。多年来，他们一直试图教学生“主题分析”，这是一种定性的研究方法，用于探索从数据集合（如访谈）中解决问题的模式。布劳恩和克拉克注意到，这种方法的定义在当时的描述都很差。有时作者会说，从他们的研究中“出现了主题”，但似乎相当神秘，然后给人一种什么都有可能发生的

感觉。那一年，布劳恩和克拉克就针对学生的问题写了这篇可读性的论文，解释了主题分析，并列出了做好主题分析的标准清单。论文发表后，研究人员开始引用这篇主题分析的论文作为他们的使用方法，故而论文的引用数一直飙升。克拉克兴奋地说，这篇论文“完全改变了我的生活。”

特点 5: 综述的兴起 (The rise of reviews)

被引用最多的论文反映了 21 世纪的一个趋势，即系统综述和元分析的兴起。在系统评价中，研究人员综合所有能回答一个问题的严谨研究，比如综合临床试验是否表明一种药物有效，它们通常包括荟萃分析，这是一种用于组合数值结果的统计技术。

特点 6: 统计软件 (Statistical software)

最后一类引用最多的论文与统计或编码软件有关。例如，一篇描述 scikit-learn 的论文（它是一个免费的开源库，为 Python 编程语言的用户提供预编码的函数和机器学习技术）已经获得了超过 5 万次引用（根据 Google Scholar 的数据达 10 万次），排名第 15 位（见原文的补充材料）。

总之，《自然》这 3 篇独家报道反映了 21 世纪、2023 年及引用有史以来的三份前 10 的列表颇有意义，值得我们学术研究者与期刊出版者开拓思维及逆向思考。也如文章结尾时德国斯图加特马克斯普朗克固体研究所的科学计量学家 Robin Haunschild 所说，这些引用例子显示了引文文化和记录是多么反复无常。但他为《自然》杂志提供了一个分析建议。“第一课：如果你写了一个有影响力的项目——那就写一篇关于它的论文吧”（“Lesson one: if you write an influential program — write a paper about it”）。

关于本刊

Bio-Design and Manufacturing (中文名《生物设计与制造》)，简称 BDM，是浙江大学主办的专业英文双月刊，主编杨华勇院士、崔占峰院士，2018 年新创，2019 年被 SCI-E 等库检索，2023 年起改为双月刊，年末升入《2023 年中国科学院文献情报中心期刊分区表》医学一区；2025 再次入选医学大类一区 Top 期刊，同时在工程、生物医学小类也升至一区。2024 年公布的最新影响因子为 8.1，位列 JCR 的 Q1 区，13/122。

初审迅速：初审快速退稿，不影响作者投其它期刊。

审稿速度快：学科编辑 24 小时初审决定投稿是否进入同行评议阶段；平均评审录用周期约 40 天；文章录用后及时在线 SpringerLink，一般两周左右即被 SCI-E 检索。

收稿方向：先进制造（3D 打印及生物处理工程等）、生物墨水与配方、组织与器官工程、医学与诊断装置、生物产品设计、仿生设计与制造等。

文章类型：Research Article, Review, Short Paper (包括 Editorial, Perspective, Letter, Technical Note, Case Report, Lab Report, Negative Result 等)。

期刊主页：<http://www.springer.com/journal/42242>

<http://www.jzus.zju.edu.cn/> (国内可下载全文)

在线投稿地址：<http://www.editorialmanager.com/bdmj/default.aspx>