采用综合影响力指标评估的 Scilit

Haochen Dong[1], *, 乔孙[1], *, Yanping Mu[1], Lu Liao[1], Diogo Rodrigues [2], Frank Sauerburger[2], Yi Bu[3], Robin Haunschild[3,4]

1 MDPI，中国北京市朝阳区关荫南北街4号院2号楼保利大都会10-14层，邮编101101；2 MDPI，瑞士巴塞尔 Grosspeteranlage 5，邮编4052；**3** 北京大学信息管理系，中国北京市海淀区颐和园路5号，邮编100871；**4** 马克斯·普朗克固体研究所，德国斯图加特Heisenbergstr. 1，邮编70569**通讯作者：**MDPI，中国北京市朝阳区关荫南北街4号院2号楼保利大都会10-14层，邮编101101。dhcydll1004@gmail.com；stevensunss@163.com

摘要

本研究系统阐释了 Scilit数据库的背景与功能，并评估了 Scilit框架内引入的综合影响力指标 I3指标 和 I3/N 的可行性与优势。我们基于包含 17,816 种期刊的匹配数据集，对 2023年和 2024年的 Scilit I3/N、期刊影响因子和 CiteScore 进行了比较分析，从学科和出版商视角考察了描述性统计与分布特征。比较结果表明，在覆盖范围、方法稳健性和学科公平性方面，Scilit I3 和 I3/N 框架显著优于传统的基于均值的指标。该框架为跨学科期刊影响力评估提供了更准确、可诊断且负责任的解决方案。本研究可作为 Scilit 的“入门指南”，为文献计量学或科学计量学领域的学者、图书馆员和学术出版商提供有价值的视角，帮助他们探索包容性数据库中的 I3 和 I3/N。这有助于更准确、全面地理解学科发展与科学进步。我们倡议在更广泛的评价情境中试点并验证这一方法，以促进对科学进步进行更精准、更多元的呈现。

关键词：Scilit；学术数据库；综合影响力指标；期刊影响因子；CiteScore

1. 引言

在当前的研究生态系统中，学术交流、科学研究工作和学术声誉与研究者生产力密切相关。作为学术产出的主要渠道，期刊受益于基于索引数据库所提供的文献计量指标开展的影响力评估，这些索引数据库包括 Web of Science（WoS）（Birkle、Pendlebury、Schnell 和 亚当斯，2020年）和 Scopus（Baas、Schotten、Plume、Côté 和 Karimi，2020年）。作为其中的代表，WoS（科睿唯安）提供的期刊影响因子（JIF）（科睿唯安，2025年b；Zhou & Leydesdorff，2011年）是年份 Y 中对年份 Y–1 至 Y–2期间全部可引用项目的总数所对应的两年平均引文数；而 Scopus（Elsevier）提供的 CiteScore（Feldner，2025年；Meho，2019年）采用四年时间窗口（即从 Y–1 至 Y–4），具有更广泛的文献覆盖范围和每月滚动更新机制，并最终形成年度快照；二者均属于未经学科规范化的引文指标，最适合用于学科内部的期刊基准比较。

对期刊引文影响力的评价，促使基于引文的指标在学术生态系统中得到广泛应用（Wilsdon，2016）。最直接地说，这类指标支持期刊分层和学科排名（Ahmad 等，2020），因而成为研究人员选择投稿目标期刊时的重要参照。对于出版商而言，突出展示引文指标已成为品牌传播和市场定位的核心策略，旨在吸引高质量投稿和读者。对于学术图书馆而言，引文指标同样构成期刊选择与续订决策的核心参照。此外，在大学和科研机构内部，这些指标经常被用作衡量学术能力的代理指标，应用于聘用、晋升和终身教职评估；在国家或机构层面，它们也被用作分配绩效奖金和科研经费的参照标准。然而，当这些做法被用于评价个别研究人员或进行资源分配时，便存在将期刊层面的指标不恰当地外推至个人和项目层面的风险，而这种用法仍存在争议。尽管持续受到批评，基于引文的指标仍因其被认为具有简便性、可扩展性和定量可比性而继续主导评价实践；然而，数据库覆盖偏差、受学科影响的引文密度、高度偏斜的引文分布、分子—分母不对称性，以及利益相关者的策略性行为（例如社论政策和引文操纵）等结构性问题，都凸显了采用更具科学依据的引文指标的必要性（Fong & Wilhite，2017；Juyal、Thawani、Sayana 和 Pal，2019；Wilhite & Fong，2012）。

老牌的 WoS 和 Scopus 如今面临着竞争者。Joshi（2016）对 Scopus 和 WoS 的功能与覆盖范围进行了系统概述，强调 Scopus 涵盖的期刊、书籍、会议和专利更为广泛，而 WoS 在引文关联和分析方面基础深厚。同样，Aghaei Chadegani 等人（2013）综合比较了 Scopus 和 WoS 的相关文献，发现根据研究领域、时间跨度和用途（检索还是评价）等选择标准，并不存在普遍适用的赢家。根据他们的观察，Scopus 通常收录更多期刊，而 WoS 在历史深度和某些分析功能方面往往更胜一筹。具体而言，Google Scholar（https://scholar.google.de/）（Martín-Martín、Orduna-Malea、Thelwall 和 López- Cózar，2018）、Dimensions（https://www.dimensions.ai/）（Singh、Singh、Karmakar、Leta 和 Mayr，2021）以及 OpenAlex（https://openalex.org/）（Priem、Piwowar 和 Orr，2022）已成为免费的替代选择。然而，Google Scholar 与其说是书目引文数据库，不如说是一个搜索引擎，因此不能归类为开放研究信息数据源（ORIDS）。Adriaanse 和 Rensleigh（2013）研究了南非环境科学期刊（2004–2008），发现 WoS 检索到的总引用次数最高，独有条目也最多；Scopus 的元数据一致性最佳，而 Google Scholar 存在的问题最多，且有大量重复副本。他们在 WoS 或 Scopus 中均未发现重复记录，而 Google

Scholar 显示出许多重复项，并且作者顺序、卷/期和页码方面也存在不一致。Dimensions 并非完全免费。部分功能需要付费订阅。下载受到限制。因此，Dimensions 也不算是真正的 ORIDS。不过，Google Scholar 和 Dimensions 的覆盖范围比 WoS 和 Scopus 更广。就科学出版物的收录广度而言，OpenAlex 甚至超过了 Dimensions，并且是真正的 ORIDS。Google Scholar 的覆盖范围可能大于 OpenAlex，但由于 Google Scholar 不提供快照下载，因此只能通过测试案例进行评估。然而，Google Scholar 的元数据质量值得质疑，所收录文献的科学性也同样如此。研究得出结论：Google Scholar 等搜索引擎的元数据质量低于 OpenAlex 等书目引文数据库。迄今为止，OpenAlex 的可持续运营主要依靠捐赠、资助以及高级服务收费。

尽管文献计量数据库日益普及，现有平台仍存在若干局限。学者、机构和出版商等关键学术参与者的档案往往不完整，其成果及关系的详细信息也较为有限。此外，许多数据库仅按一般类别对出版物进行分类，未能区分更具体的子类型。再者，由于消歧能力有限，作者元数据常常存在歧义，妨碍了出版物归属和指标的准确判定。数据访问通常需要付费，进一步限制了其可用性。

事实上，数据库的收录范围和元数据质量会影响指标的计算。另一方面，非排他性数据库包含规模更大、更全面的引文网络，能够更细致地了解期刊层面的引文影响力。由多学科数字出版研究所（MDPI）开发的 Scilit 汇集来自广泛来源的学术记录，以最大限度地扩大覆盖范围并确保时效性。其数据摄取管道主要从 DOI 注册机构（尤其是 Crossref 和 DataCite）获取元数据，整合来自 PubMed/Medline 的生物医学文献记录，并索引出版商通过信息源和 API 提供的内容。此外，Scilit 还纳入开放获取目录（例如 DOAJ）和精选的预印本服务器，同时定期同步出版商更正和更新，以保持数据的时效性。这种多来源、非排他的模式不受专有封闭式索引政策的限制，因此 Scilit 能够收录大型和小型出版商的期刊、学会期刊、地区性及非英语出版物、会议论文集、丛书和预印本。换言之，Scilit 通过汇集多个来源的数据来弥补这些不足，从而提供更完整的学者、机构、来源和出版商档案。它采用高级作者消歧技术来解决歧义，并将出版物划分为 14 个具体子类型，包括研究论文、综述论文、会议论文、临床试验、病例报告、信函、书评摘要、社论、撤稿论文、撤稿、更正、撤回和其他，使用户能够进行更精细的检索并获取详细、可靠的信息。此外，Scilit 完全免费使用，由 MDPI 运营可确保其获得财政支持并实现可持续发展，从而广泛、可靠地提供全面的学术数据。因而，与许多具有选择性或基于订阅的数据库相比，Scilit 有意扩大了覆盖范围，提升了不同学科、语言和出版形式的可发现性，并减少了狭窄筛选标准所造成的偏差。目前，该数据库包含超过 1.81 亿条记录，主要包括期刊论文、会议论文集论文、书籍、书籍章节和预印本；Scilit 和 OpenAlex 的数据快照见表1。总体而言，OpenAlex 收录的作品多于 Scilit。然而，OpenAlex 中并非所有作品都是学术材料。OpenAlex 一个已知的缺陷是参考文献匹配不完整。Scilit 的每篇论文引文数和每篇论文的施引论文数高于 OpenAlex，这表明

Scilit 的参考文献匹配完整度高于 OpenAlex。准确且完整的参考文献匹配对于计算基于引文的指标非常重要。

表1. 2015-2024年期间 Scilit 与 OpenAlex 多个维度数据对比概览。

	OpenAlex（2015-2024）	**Scilit（2015-2024）**
作品总数	100,059,695	72,244,250
包含作者或机构的作品数:	91,405,199	39,348,564
至少有一篇论文	53,186,532	15,702,524
至少有一篇论文	93,606	73,376
至少有一位作者的作品数量作者	90,134,941	31,519,239
至少有一位作者的作品占比作者	0.901	0.436
至少有一个机构	49,430,281	21,620,963
至少有一个机构	0.494	0.299
至少有一个作者 AND 至少一个机构	49,427,038	30,513,489
至少拥有一位作者的作品占比作者 AND 至少一家机构	0.494	0.422
引文关系数量	549,809,112	1,529,927,725
每篇论文的引文关系数	5.495	21.177
被引论文数量	36,332,876	30,765,138
每篇论文的被引论文数	0.363	0.426
施引论文数量	36,702,097	31,167,211
每篇论文的施引论文数	0.367	0.431
文章	66,707,420	46,056,219
数据集	5,763,661	1,542,655
预印本	4,471,388	4,226,705
书籍	1,328,610	1,005,810
书籍章节	10,694,341	11,529,965
社论	305,838	1,034,235
撤稿	19,577	38,370
学位论文	2,710,536	524,211
参考条目	288,908	181,897
撤稿文章	31,925	41,522

另一方面，综合影响指标（I3）（Leydesdorff & Bornmann，2011年）为评估期刊影响力提供了一种基于百分位数、对分布敏感的替代方案，可用于取代期刊影响因子和CiteScore。I3通过百分位排名，汇总引文分布整体范围内文章层面的引文表现，从而减轻偏斜的引文分布造成的扭曲，避免均值固有的分子—分母不对称性，并通过参考集合实现领域标准化比较。I3的计算方式是对各引文百分位类别中的加权出版物计数求和；它既能反映出版物数量，也能体现影响力，同时不会受少数高被引异常值左右，因此可为期刊影响力提供一种统计上更稳健、科学依据更充分的衡量方法，适用于跨领域评估和负责任的科研评价。

尽管现有研究已系统比较了JIF与CiteScore在WoS/Scopus中的表现，但对于对分布敏感、按领域标准化的I3指标在非排他性、广覆盖数据库中的适用性及优势，目前仍缺乏系统证据。尤其是在非英语期刊、区域性期刊、会议论文集和书籍章节占比较高的领域，现有基于均值的指标可能会放大由数据库覆盖范围差异和学科引文密度差异引发的偏差，因此亟需建立新的实证评估框架。基于上述背景，本研究旨在实现两个目标：i）介绍Scilit数据库及其核心功能；ii）在Scilit数据库中实施并计算I3，并将其与期刊影响因子和CiteScore进行基准比较，以评估I3作为期刊影响力指标的适用性。本研究的贡献有两方面。首先，这是首次将Scilit与I3框架相结合。其次，通过计算并比较2023年和2024年的指标（即期刊影响因子和CiteScore），本研究展示了在非排除性、多来源数据库中采用基于百分位数、对分布敏感的指标所具备的优势。这些结果共同凸显了I3与Scilit广泛覆盖范围之间的方法适配性，并为负责任的跨领域期刊影响力评价提供了另一种途径。

2. Scilit数据库简介

*2.1 Scilit*历史

Scilit最初于2012年秋季以“OAlit”之名提出，由MDPI开发，旨在支持开放获取出版生态系统，目标是收集所有开放获取论文（包括全文），并建立可检索的开放数据库。2014年，其范围随后扩展至所有科学论文，服务也更名为Scilit；它通过抓取CrossRef API、PubMed、PMC以及包括arXiv、HAL、PeerJ和RePEc在内的预印本服务器平台的元数据来获取信息。

该服务持续改进并不断扩展：通过抓取更多出版商网站的数据、与向Scilit提交元数据的出版商合作、整合更多数据来源、改进抓取算法以及提升元数据消歧能力。继2014年引入Crossref引文数据并于2022年对其进行增强后，Scilit通过整合引文数据和参考文献数据，逐步扩充其引文数据集。

2.2 数据来源与架构

数据摄取主要来自 Crossref、PubMed、PMC 和 DOAJ 等权威来源（图 1）。元数据通过实时采集机制获取，并经过聚合、去重和消歧（尤其是作者和机构消歧），以提供高质量数据。

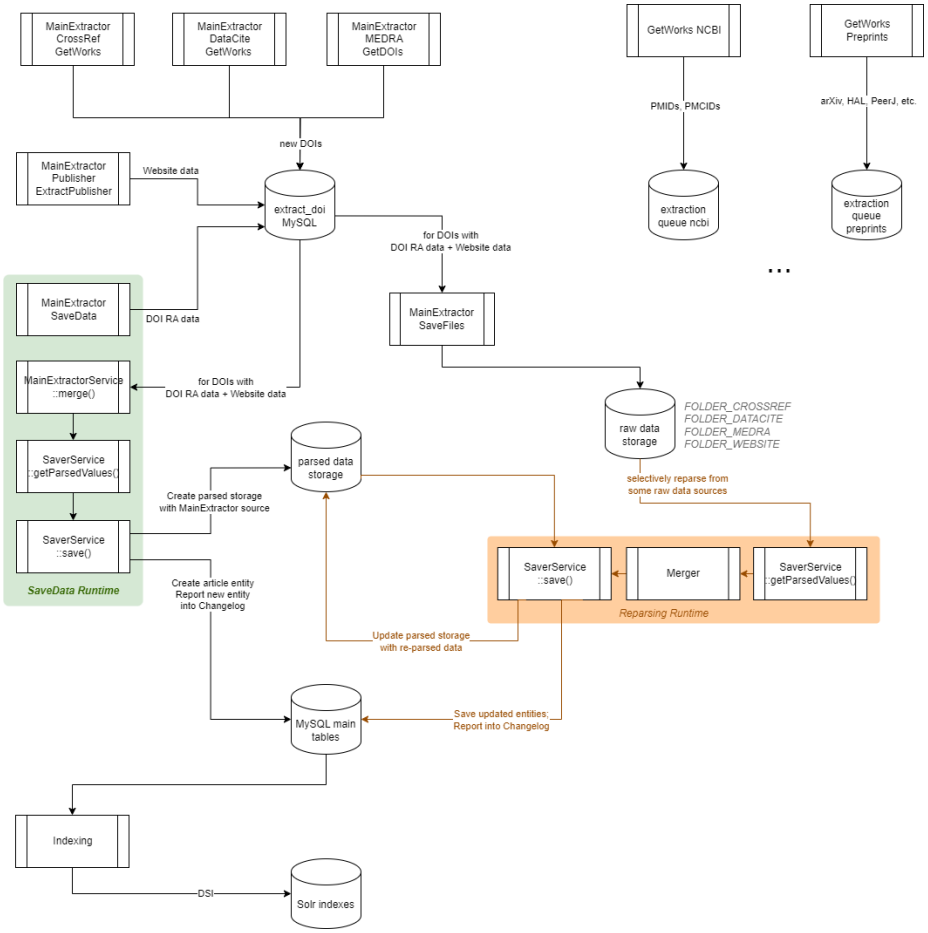


图 1. Scilit 数据处理示意图

2.3 搜索功能

Scilit 提供基础和高级搜索功能。Scilit 的基础搜索可帮助用户快速发现学术研究内容。用户只需输入关键词，系统便会自动检索所有已索引出版物的标题、关键词、摘要、DOI 和 ROR ID 字段，并返回相关研究。用户还可以根据多种筛选条件细化结果，例如出版物类型、学科分类、可持续发展目标分类、出版年份、国家/地区、出版商、来源和医学主题词，从而获取更精准的结果。Scilit 的高级搜索功能支持对多种元数据字段进行精细查询，包括标题、摘要、关键词、主题、DOI、ROR ID、作者姓名、来源标题、出版商和出版年份。通过灵活选择字段，该功能支持复杂且结构化的文献检索。此外，高级搜索面板支持使用布尔运算符 AND、OR 和 NOT，帮助用户构建具体且精确的查询。

2.4 Scilit 导出选项

Scilit 支持用户以 CSV、RIS、EndNote、BibTeX 和出版物列表文件格式导出搜索结果。导出的结果可包含以下字段：出版物标题、作者、来源标题、出版商、出版物关键词、卷、页码、开放获取状态、被引次数、文章 DOI、发布日期和 Scilit 文章页面链接。每次最多可导出 1000 条记录，但可以分页导出。通过应用程序编程接口访问和提供可下载快照的功能正在规划中。

2.5 Scilit 排名

Scilit 提供全面的排名系统，展示学术出版领域各类实体的表现。排名帮助用户：1）追踪表现领先实体的出版量和影响力；2）比较各实体的输出和影响力；3）发现新兴期刊、活跃组织或具有影响力的预印本平台。

此外，Scilit 提供比较工具，用户可以选择多个实体，并并排查看指标、趋势和排名随时间的变化。此功能对于科研评估、机构基准评估和编辑策略尤其有用。其主要亮点包括：

1. Scilit 排名涵盖五类维度，包括出版商、期刊、预印本服务器、组织和国家。每个子排名部分主要根据以下出版产出和引文指标计算：i) 出版产出——某一版本年份中的出版物/OA 出版物总数（Scilit 排名的版本年份，等于我们存档数据中的最近一个完整年份，即截至2021年7月，版本年份为2020年）。在 Scilit 中；ii) 引文指标——包括五年被引文章数、五年引文数、H5 指数和月度引文指标。2. Scilit 排名提供比较功能：在此部分，用户可以直接根据出版表现，对最多三个出版商、期刊和组织实体进行比较。3. 出版市场：用户还可以全面了解学术出版市场的发展情况，包括文章、期刊和预印本的数量，以及深入了解各出版机构的贡献比例。

2.6 Scilit 学科分类

由 MDPI 开发的 Scilit学科采用全面、稳健的两级分类体系，涵盖该组织发表的所有学科。12 个上级类别分别为：农业；艺术与人文；生物学与生命科学；化学与材料科学；临床科学；计算机与信息科学；地球与环境科学；电气工程与电子学；工程学；数学；物理学；社会科学。每个上级类别又细分为子类别，共计 290 个，用于进一步明确学科范围，例如化学与材料科学下的催化剂、纳米颗粒或纺织化学。这种层级结构既能对跨学科研究进行精确分类，又能清晰呈现各学术领域的整体概况。

为分配这些学科，我们采用一种基于内部微调的 BERT 模型 ScilitBERT 的自然语言处理（NLP）分类模型（de la Broise、Bernard、Dubuc、Perlato 和 Latard，2021年）。ScilitBERT 在涵盖多个学科的多样化、大规模学术文本语料库上进行预训练，其架构还在基础网络上增加了一个分类层，用于分配主题。该模型利用文章的标题和摘要，推断每篇文章内容所属的学科。

该人工智能模型可用作多标签分类器，允许将稿件分配到一个或多个类别。期刊的学科类别通过汇总其过去两年文章的类别构建而成，确保能够反映期刊的实际内容。这种方法以客观、数据驱动的方式呈现期刊的范围。

2.7 Scilit 生成的实体和实体关系图Scili

其采用关系数据库结构组织，以出版物为核心实体

y

与学者、Affiliations、来源、出版商和参考文献相关联。实体–关系（ER）图见图2至图7，展示了关键实体及其关系。

2.7.1 学者实体

作者系统旨在为学者建立学术档案，包括其基本信息、学术成就、学术关系、研究领域、已发表论文及相关指标，以展示其研究成果和学术贡献。

主要实体有三个：UniqueAuthor、UniqueAuthorExtraInfo 和 LogUniqueAuthorDisambiguation。UniqueAuthor 是作者系统的核心实体。它由 Authors 实体生成，并汇总作者姓名、电子邮件、ORCID、Affiliations 和机构等核心详细信息。为优化存储和复用，姓名、电子邮件、Affiliations 和机构等部分属性会拆分为独立实体。例如，AuthorEmails 管理作者的电子邮件地址，Affiliations 记录特定的作者与机构之间的关系，Institutions 存储研究机构的详细信息，包括名称、国家和网站。
UniqueAuthorExtraInfo 跟踪学术指标，例如出版物、引文、自引、H指数、i10 指数、学术年龄、学科及其排名（Subjects 和 UniqueAuthorSubjects 定义研究领域，UniqueAuthorSubjectsRanking 记录作者在其研究领域内的排名）。
LogUniqueAuthorDisambiguation 监控并记录存在身份冲突的作者及消歧进度，包括电子邮件、状态（消歧进度）和备注（补充背景信息）。

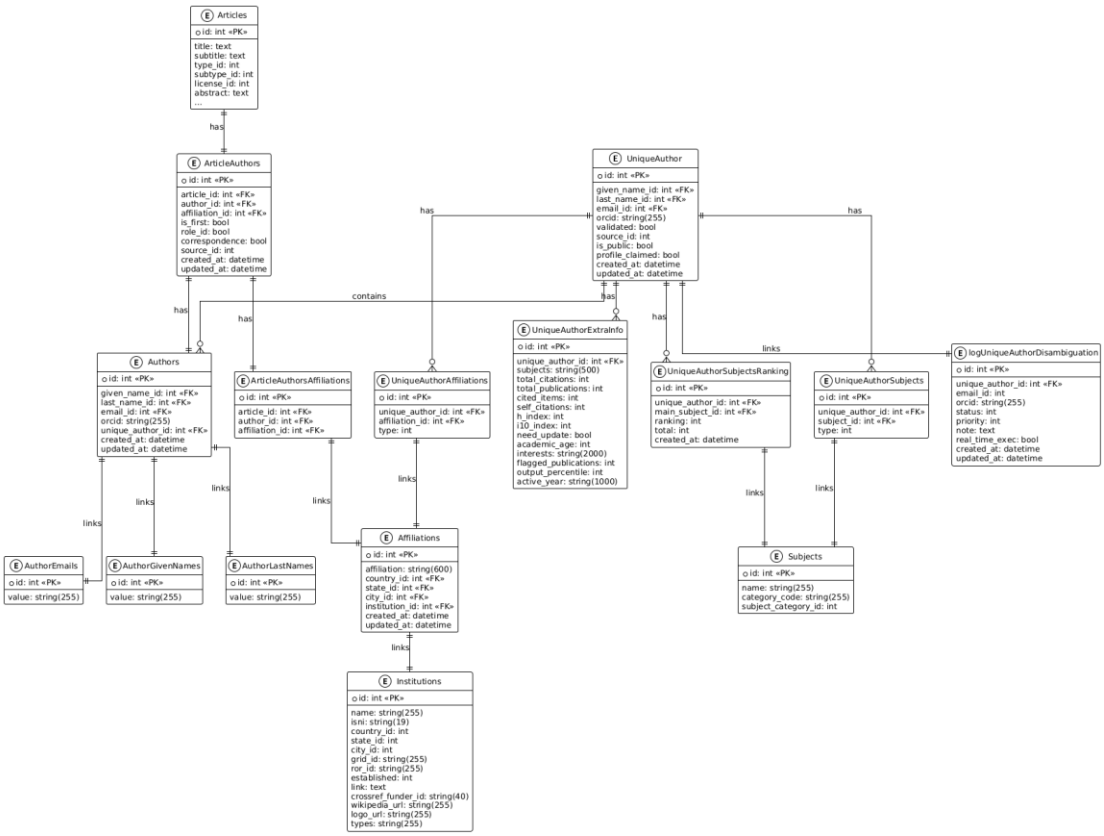


图2. Google Scholar学者实体的实体关系图

为确保学者元数据的高质量，Scilit 对作者信息实施两阶段验证：作者消歧模型和歧义检测模型。作者消歧模型侧重于合并属于同一人的记录，而歧义检测模型则会识别因元数据不足而可能需要进一步

由人工或算法进行澄清的记录。二者共同构成可靠作者档案、准确隶属关系链接和稳定学者指标的基础。

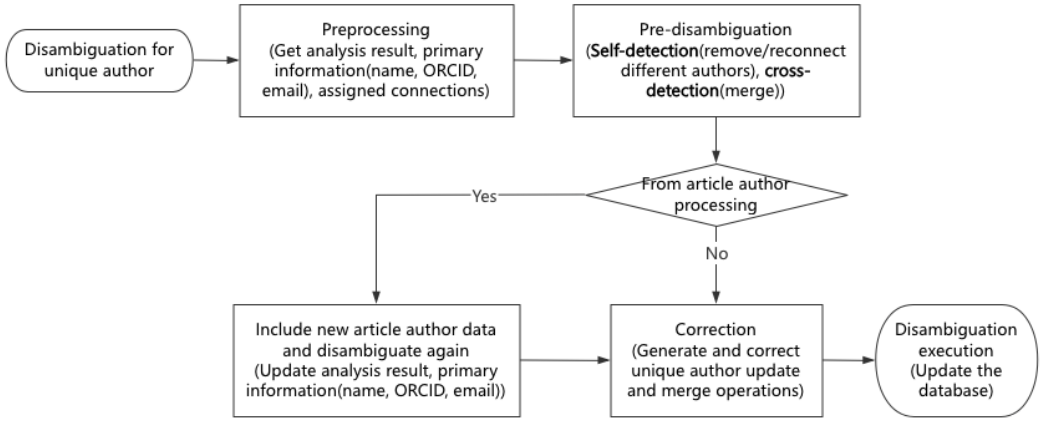


图 3. 作者消歧模型。

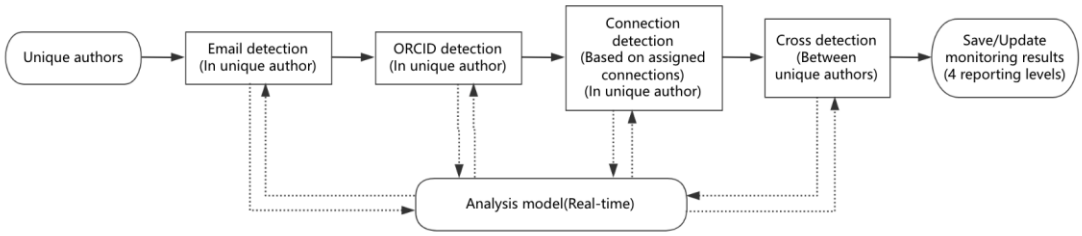


图 4. 歧义检测模型

在现有消歧模型和歧义检测模型的基础上，人工智能团队目前正在开发并应用更先进的方法，以进一步完善作者元数据。这些方法以现有数据集为基础，预计将显著提升作者信息的准确性、完整性和可靠性，从而为下游分析和指标提供更坚实的依据。

2.7.2 来源实体

源实体通常指期刊、预印本、会议论文集、书籍或书籍章节。该系统是一个先进的内容聚合平台，旨在将学术出版物管理并组织到逻辑容器中。该系统围绕两个主要实体构建：作品容器和作品容器组，其中各个容器（如特定期刊期次或预印本合集）归入更广泛的容器组（如期刊系列或会议系列）。该架构遵循工厂方法设计模式，以处理不同的容器类型（期刊、预印本、会议论文集和书籍），并采用专门的

各类型的处理逻辑。每个容器实体都包含元数据，例如期刊、缩写、国际标准期刊号、出版物类型，以及与出版商组和容器组的关系；服务层（WorkContainerService）负责协调容器数据的处理、验证和持久化，并在多个来源对同一容器提供相互冲突的信息时进行冲突解决。系统还采用基于 Solr 的搜索功能、冲突检测机制和后处理工作流，以维护数据质量并确保学术内容得到正确索引。此外，系统还实现了观察者模式，用于事件处理和引文指标追踪，并提供全面的存储库方法，以便在整个容器生态系统中高效地访问和管理数据。

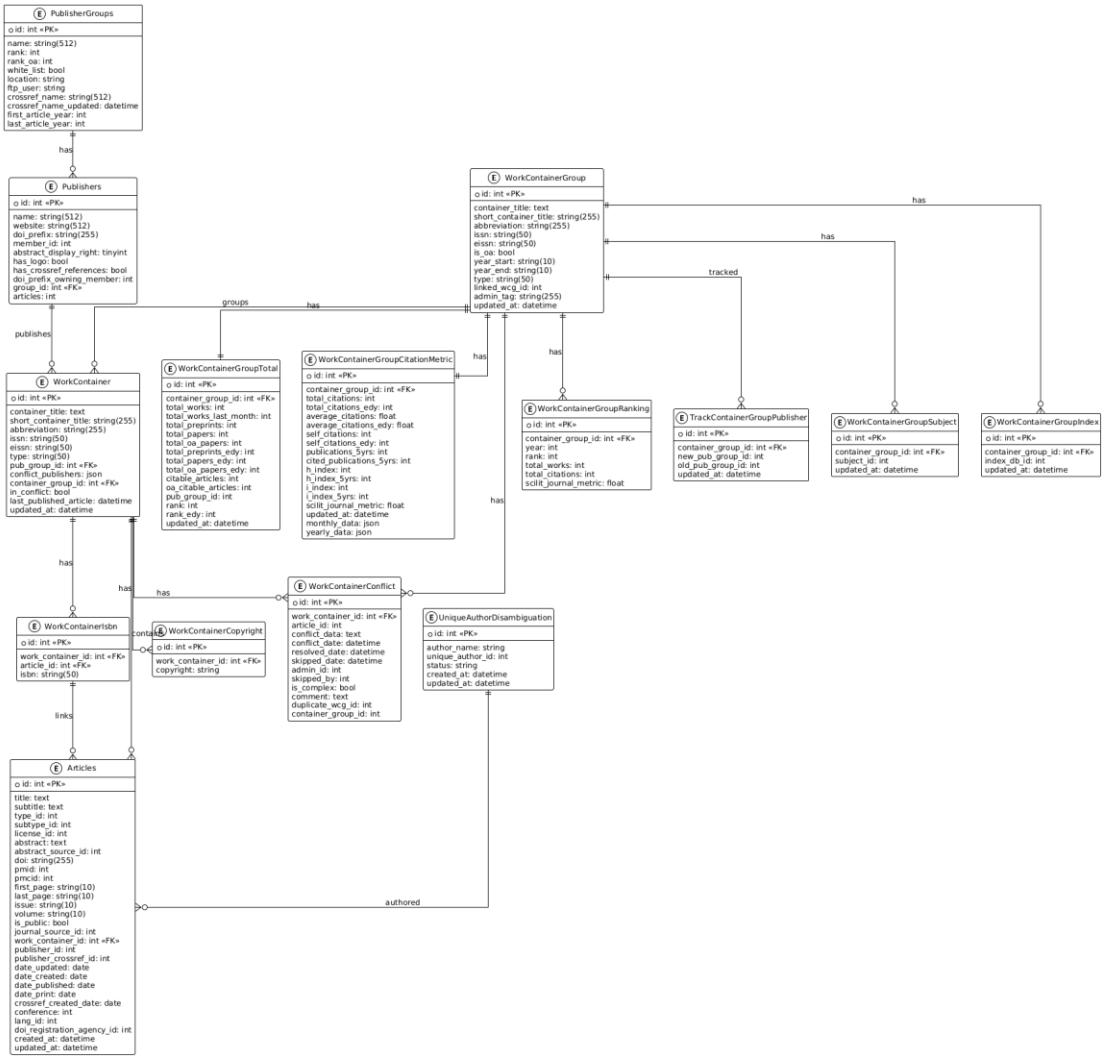


图5. 源实体的实体关系图

2.7.3 出版商实体

Scilit 中的出版商系统是集中收集和管理所有出版商及出版商组信息的核心场所。每条出版商记录都会存储基本信息，例如出版商姓名、网站、DOI 前缀，以及若干用于描述平台应如何处理出版商内容的标志。出版商组充当父组织，可以包含多个出版商，并规定谁有权限管理这些出版商。

系统还会将出版商与文章和工作容器关联起来，从而更轻松地追踪每份出版物的来源及其索引方式。系统还提供工具，用于解决不同来源针对同一出版商提供不同数据时产生的冲突，并提供加入白名单、追踪变更以及更新出版商组之间关系的机制。

用户权限是系统设置中的重要组成部分。系统中的每位用户始终关联一个出版商。如果某个出版商尚未注册用户，用户会在注册时选择自己所代表的出版商。随后，Scilit管理员必须手动审核并批准此关联。获批后，该用户便会与该出版商关联。

共有两种用户角色：所有者和用户。所有者角色拥有对出版商的全面管理权限，包括邀请和移除用户、管理访问权限、导入文章XML

文件，以及添加或管理属于同一出版商组的其他出版商。所有者还可以更新出版商在 Scilit 上显示的名称。普通用户的权限有限，只能在所分配出版商的范围内执行所有者授予的操作。

用户可以拥有出版商组、上传文件，并根据其访问权限请求更改。服务层负责协调所有这些流程，从验证数据、更新记录到在发生变更时触发事件。总的来说，出版商系统提供了清晰的结构，确保出版商信息一致、可靠，并与 Scilit 索引工作流的其余部分紧密集成。

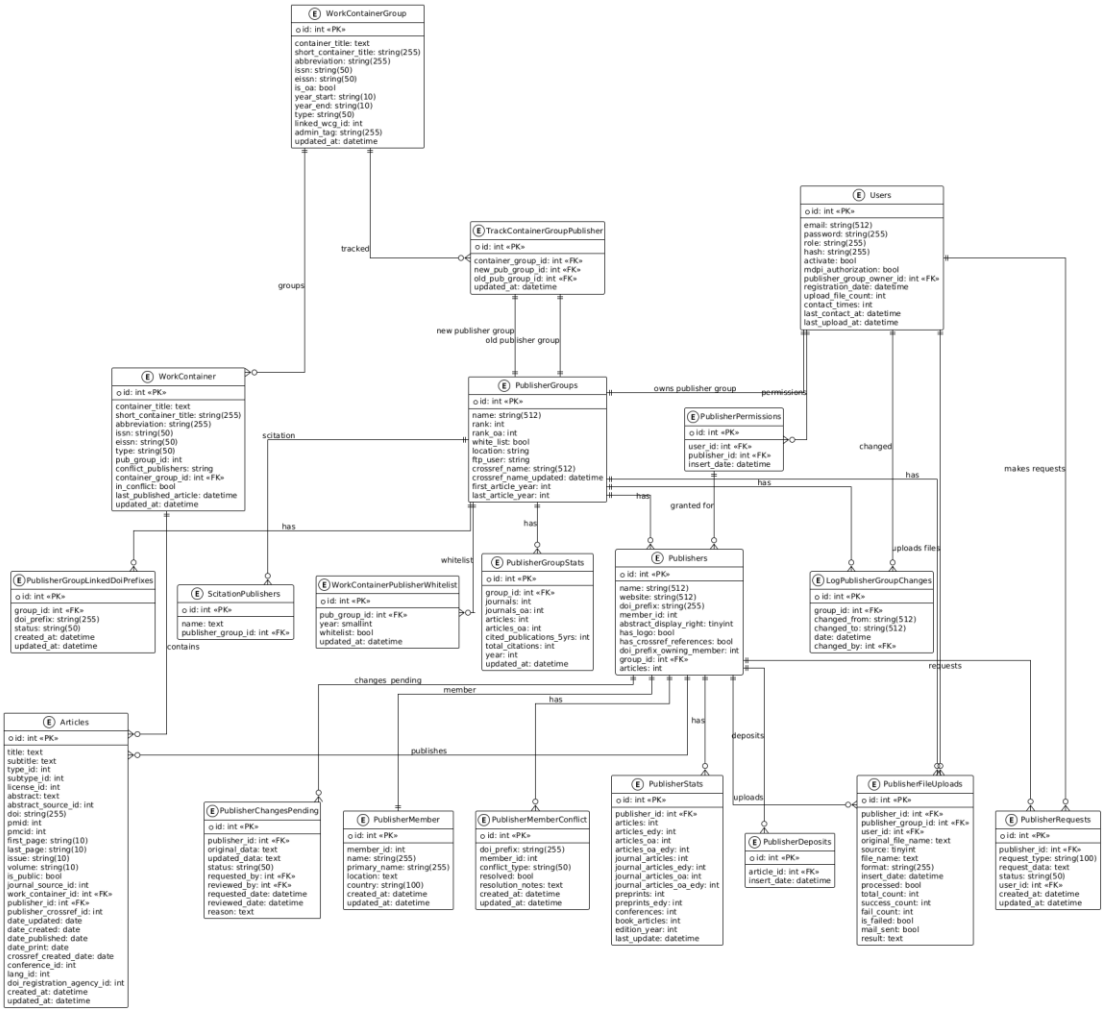


图6. 出版商实体的实体关系图

2.7.4 机构实体

机构系统用于管理和分析与学术与研究机构相关的数据。它围绕能够捕捉机构管理各个维度的实体构建。

核心实体是机构（Institutions），用于存储姓名、国家、Grid ID（研究机构的全球标识符）、ROR ID（研究组织注册库标识）、成立年份（创立年份）和标志图片网址等关键信息。它是所有机构数据的枢纽。同时，它还使用三个地理层级实体：国家（Countries，存储国家名称和代码）、州（States，记录州／省信息，并与国家关联）和城市（Cities，存储城市名称，并与州关联）。此外，Affiliations 是另一个主要实体，它

定义作者与机构之间的具体关系。关联的实体包括文章-作者-隶属关系（连接文章、作者和隶属关系）以及唯一作者-隶属关系（将唯一作者与隶属关系关联起来）。关于机构的统计数据和指标，涉及 InstitutionStats 和 InstitutionMetrics 两个实体。InstitutionStats 提供按年份统计的数据，例如高_被引_学者、高_被引_论文，以及国际_合作_率。InstitutionMetrics 则记录学术指标，例如总_引文_5yrs（过去 5 年的引文数）和 oa_文章（开放获取文章的数量）。

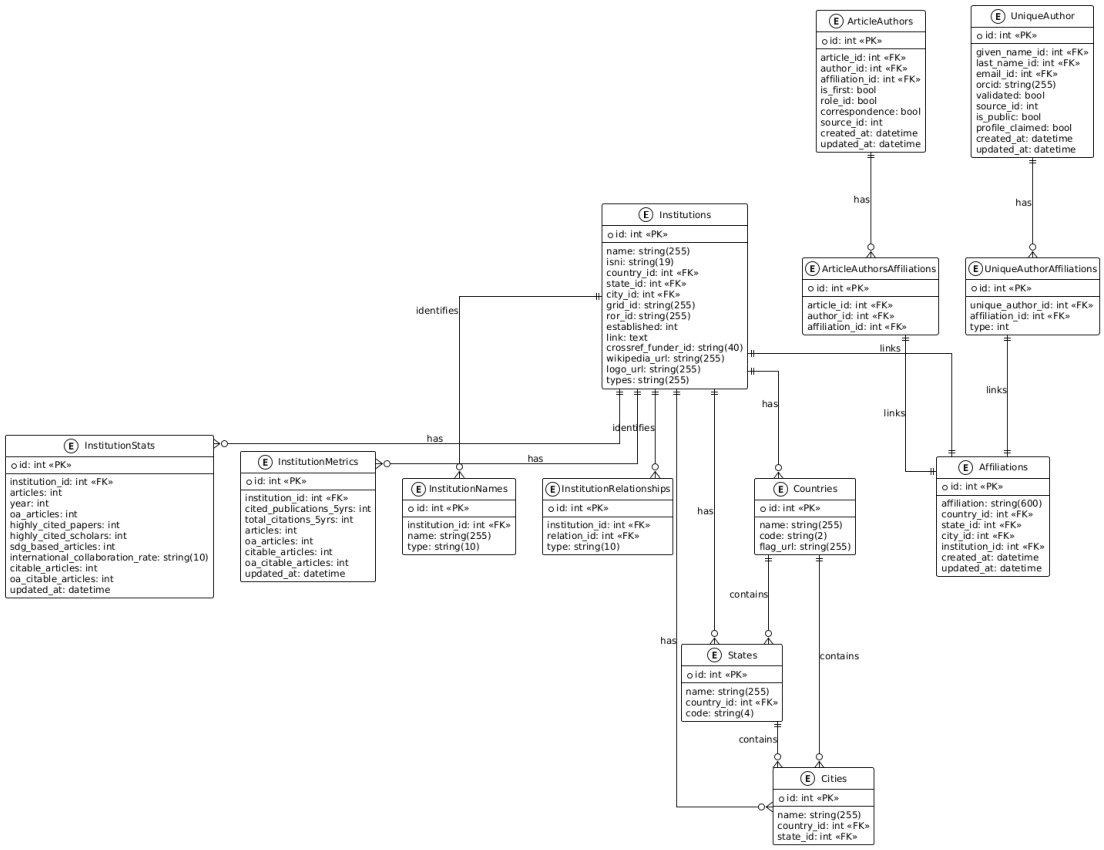


图7. 研究所实体的实体关系图

*2.8 Scilit*产生的指标

Scilit 提供用于定量和定性方面的二维评估指标，如表2所总结。总体而言，我们采用先进的自动化技术创建了四类实体：来源、出版商、学者和组织。在每个实体的独立档案中，用户可以查看基本信息、近几十年的出版产出、最新出版物以及引文指标，从而清晰了解该实体的出版表现。本研究以绝对视角的 I3 指标和相对视角的 I3/N 为例进行说明和演示。

表 2. Scilit产生的指标。

维度	指标	描述	备注 / 示例
输出	出版物总数	近几十年所有索引记录的数量	按时间划分的总数趋势
输出	开放获取出版物总数	近几十年开放获取记录的数量	按时间变化的开放获取覆盖率
输出	按出版物数量排名前十的国家	近几十年各国的出版物数量	带数量的排名列表
输出	按出版物数量排名前 10 的学科领域	近几十年各学科领域的出版物数量	附数量的排名列表
影响力	总引用次数	累计获得的引用次数	基本引用指标
影响力	平均引用次数	每篇出版物的平均引用次数	容易受偏态分布影响；解读时请谨慎
影响力	自引	各实体的自引次数	可按实体类型报告
输出	国际合作率	跨国共同署名的比例/水平	量化跨境合作模式
输出	国内机构间合作	同一国家内不同组织之间的合作	组织层面网络强度
输出	国内组织内合作	同一组织内的合作	内部组队强度
影响力	引文深度	引用该核心对象的出版物中，同时也引用该对象的出版物所占的绝对和相对比例 论文的参考文献	反映深入的领域内影响力
影响力	引文广度	跨研究领域的概念广度	体现跨领域覆盖范围
影响力	引文依赖性	绝对值：也引用焦点论文参考文献的施引论文数量； 相对值：上述数量所占的比例	表示对先前研究的依赖；针对高被引论文计算（≥100 条引文）
影响力	引文独立性	绝对值：未引用焦点论文参考文献的施引论文数量； 相对值：上述数量所占的比例	表示独立贡献；针对高被引论文计算（≥100 条引文）
影响力	H指数	h 篇出版物的引文数至少为 $\geq h$	职业生涯影响力
影响力	i10 指数	含有 ≥ 10 条引文的出版物	更简单的基于阈值的影响力
影响力	H5 指数	过去 5 年的 h 指数	侧重近期表现的期刊影响力
影响力	月度引文指标	月度引文累积指标	跟踪短期势头
影响力	I3	各引文百分位类别中加权出版物计数的总和	对分布敏感且领域可比；兼顾规模与影响力
影响力	I3/N	按出版物数量标准化的 I3指标	按规模标准化且领域可比的影响力度量

3. 数据获取与处理

我们按期刊层面计算了2023年和2024年的期刊影响因子、CiteScore、I3指标和I3/N，具体如下：

3.1 期刊影响因子

期刊影响因子的定义和计算方法（科睿唯安，2025a）如公式（1）所示：

$$JIF(Y) = \frac{Citations\ received\ in\ year\ Y\ to\ items\ publis\ \ ed\ in\ years\ Y-1\ and\ Y-2}{Number\ of\ citable\ items\ publis\ \ ed\ in\ years\ Y-1\ and\ Y-2} \quad (1)$$

h h其中，Y表示期刊影响因子年份。

我们从官方期刊引证报告（JCR）网站（https://jcr.clarivate.com）下载并获取了21,591种期刊的2023年期刊影响因子和22,224种期刊的2024年期刊影响因子。下载的数据还包括期刊ISSN/eISSN、出版商和学科类别等信息，用于后续的数据匹配与分析。

3.2 CiteScore

CiteScore 的定义和计算方法如公式（2）所示：$CiteScore(Y) = \frac{Citations\ received\ in\ year\ Y\ to\ documents\ publis_{h}}{ed\ in\ years\ Y-1\ Y-2\ Y-3\ and\ Y-4\ \ ,,,h,„}$ (2)

其中，Y 表示 CiteScore 年份。

$Number\ of\ documents\ publis\ \ ed\ in\ years\ Y-1\ Y-2\ Y-3\ and\ Y-4$

我们通过 Scopus API（https://dev.elsevier.com/documentation/ScopusSearchAPI.wadl）获得了 29,555 种期刊的 2023年 CiteScore 和 30,509 种期刊的 2024年 CiteScore。下载的数据还包括期刊ISSN/eISSN、出版商和学科类别等信息，这些信息用于进一步的数据匹配和分析。

3.3 Scilit I3 和 *I3/N*

我们按照公式（3）所示定义了 Scilit 中的 I3指标：

$$I3 = \sum_i (x_i \times n(x_i)) \quad (3)$$

其中，x_i 是根据处于前 1%、前 10%、前 50% 和后 50% 的论文，为百分位等级类别i分配的权重；$n(x_i)$ 是百分位等级类别i中的出版物数量。

I3 指标的核心概念是为处于不同引文区间的出版物赋予不同权重，以体现引文分布高度偏斜的特点（Leydesdorff、博恩曼和亚当斯，2019）。我们在计算 Scilit I3 时采用的加权方案与上述方案略有不同。根据 Scilit 顾问委员会成员（ABM）的建议，我们为引文影响力方面赋予比产出方面更高的权重，将排名后 50% 的论文得分设为 0，以提高该指标对引文影响力的敏感度：

$$Scilit\ I3 = \sum_i (x_i \times n(x_i)) = \sum_{i=1}^{4} w_i \quad (4)$$

其中，w_i 表示$x_i \times n(x_i)$。我们定义了四组权重：$x_1 = 100$、$x_2 = 10$、$x_3 = 2$和$x_4 = 0$，分别对应引文数排名前 1%、前 10%、前 50% 和后 50% 的论文。$n(x_i)$表示与各权重对应组别中的论文数量。同时，在确定 Scilit I3 定义中的时间窗口时，我们采纳了 ABM 的建议，选择两年的出版窗口和四年的引文窗口。这样既能确保被引出版物有充足的时间积累引文，又能维持该指标所评估对象的时效性和信息价值。此外，Scilit I3 指标中出版窗口和引文窗口的选择，也借鉴了期刊影响因子（五年）和 CiteScore 的做法。

在确定权重和时间窗口后，为计算 Scilit I3 中的参考文献，我们采用施引侧标准化方法进行领域标准化（Zhou & Leydesdorff，2011年）。对于每次引文，我们按照公式（5）计算分数引文计数：

$$\text{分数引文计数} = \sum 1/\text{米}_i \qquad _i \qquad (5)$$

其中，m_i 是施引文章i中的参考文献数量。

我们采用分数计数法，原因如下：i）这种方法会针对领域依赖的引文“潜力”（即不同的参考文献列表长度和引文密度）进行标准化，从而无需依赖并不完善的基于期刊的领域分类，也能减少跨领域偏差。ii）通过在单篇施引论文层面进行标准化，分数计数法能够更公平地比较不同学科和专业领域，并将规模效应（总引用次数）与相对影响力（每篇论文的引文数）区分开来；在跨单位汇总 I3指标 时，这一点尤为重要。iii）该方法能够提供分布数据（每项内容的分数引文权重），以便对不同单位之间的差异进行统计检验；研究也表明，它能显著降低影响力指标中的领域间方差，并以合理的方式改变排名（例如，提升人文学科或工程学等学科的排名，因为这些学科的引文密度与其他领域差异较大）（Zhou & Leydesdorff，2011年）。因此，我们认为，对于 Scilit I3 这种汇总不同百分位等级或论文的加权贡献的指标，将分数计数作为引文输入，或可提高综合影响力的跨学科可比性、可加性和可解释性。

获得每篇论文的分数引文计数总和后，我们在相同的出版年份和文献类型内对论文进行排名，并赋予百分位排名，以计算 Scilit I3。Scilit I3 的计算纳入了 2021—2022年 的以下文献类型：研究论文（6,569,561）、综述文章（559,475）、会议论文（59,274）、病例报告（49,272）和临床试验（353）。

为与期刊影响因子和 CiteScore 进行比较，我们定义了期刊层面的 Scilit I3/N，即用 Scilit I3 除以期刊论文数量（N），以体现考虑期刊特征后平均每篇论文的贡献：

$$\text{j}$$

$$Scilit\ I3/N = \frac{1}{n_k}\sum_{i=1}^{n_k k} w_i \qquad (6)$$

其中，n_k 表示期刊k在 Scilit I3时间窗口内、属于相应文献类型的期刊论文数量。我们分别计算了 2023年 55,842 种期刊和 2024年 61,739 种期刊的 Scilit I3 和 I3/N。Scilit 数据还包含期刊 ISSN/eISSN、出版商和学科类别等信息，这些信息用于进一步的数据匹配和分析。

3.4 数据处理

根据期刊标题、ISSN 或 eISSN，共匹配了 17,816 种期刊，用于基于期刊影响因子、CiteScore、Scilit I3 和 Scilit I3/N 的比较分析。为了在不同维度中以相同尺度比较不同指标，我们采用最大值-最小值标准化方法，对期刊影响因子、CiteSore、Scilit I3、Scilit I3/N 以及时间窗口内的期刊论文数量同时进行标准化，以供后续讨论和分析。分析：

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \qquad (7)$$

其中，x 是期刊指标之一。我们使用 R（4.5.1版）（R CoreTeam，2025年）进行数据匹配和数据可视化。

3.5 主题分类检索与期刊级匹配

为确保三个数据库中的期刊归属能够在共同的主题维度内进行比较，我们首先从每个来源获取期刊的原生主题分类。对于 WoS，我们从 JCR数据库中获取期刊层面的 Web of Science类别；该数据库为每种被索引期刊提供一个或多个主题类别。对于

Scopus，我们从 Scopus 来源列表（Scopus 来源标题列表）中获取期刊学科分类；该列表为每种期刊分配一个或多个全科学期刊分类（ASJC）代码。随后，我们将这些 ASJC 代码作为 Scopus 学科分类体系。对于 Scilit，我们按照第 2.6 节所述，使用相应数据库文档或来源列表中以期刊（来源）层级提供的官方主题类别。我们使用 ISSN（以及可用时的电子 ISSN）将这三组主题分类与我们的期刊总体进行匹配。当某种期刊拥有多个 ISSN 时，匹配过程中会考虑所有有效的 ISSN，并删除重复项。在给定数据库中被分配至多个主题类别的期刊将保留其全部学科归属；因此，我们的跨数据库映射是在期刊–主题配对层面进行的，而非针对单一且相互排斥的类别。基于这些期刊–主题归属，我们随后通过识别出现在多个数据库中的期刊并比较其分配的类别，在三个分类体系之间构建了对照表。因此，本研究所呈现的比较分析建立在同一期刊在 WoS、Scopus 和 Scilit 中实际观测到的主题归属之上。

4. 结果

4.1 统计描述

表3报告了2023年和2024年的描述性统计数据，并按指标类型整理。所有主要数据库中，2024年覆盖的期刊数量均较2023年有所增加。值得注意的是，拥有 Scilit I3指标（来自 Scilit）的期刊数量约为拥有 CiteScore（来自 Scopus）期刊数量的两倍，凸显了非排他性数据库在覆盖范围方面的优势。

对于期刊影响因子，2024年的集中趋势略有上升，同时极端值明显减少，表明分布略微趋于集中。相比之下，CiteScore 在这两年间的均值和中位数总体保持稳定，但离散程度有所增加，这表明高端 CiteScore 值进一步上升。

与2023年相比，2024年 Scilit I3/N 的均值和中位数均有所下降，而离散程度有所增加，表明期刊群体内部的差异正在扩大。对于 Scilit I3，整体均值基本不变，但中位数下降、上尾延伸，说明尽管均值稳定，极端值却更加突出。

表3。 指标的描述性统计。

指标	Year	N	Min	Mean	中位数	Max	95%置信区间	SD	SE
JIF	2023	21,951	0	2.27	1.4	521.60	[2.20, 2.34]	5.19	0.04
JIF	2024	22,224	0	2.37	1.5	232.40	[2.31, 2.42]	4.26	0.03
CiteScore	2023	29,555	0	3.64	2	873.20	[3.55, 3.72]	7.57	0.04
CiteScore	2024	30,509	0	3.64	2	1,154.20	[3.54, 3.73]	8.68	0.05
Scilit I3/N	2023	55,842	0	1.39	0.76	100.00	[1.36，1.41]	2.91	0.01
Scilit I3/N	2024	61,739	0	1.34	0.67	100.00	[1.32, 1.37]	3.17	0.01
Scilit I3	2023	55,842	0	246.94	12	161,566.00	[230.55，263.32]	1,975.16	8.36
Scilit I3	2024	61,739	0	244.89	10	183,360.00	[228.67，261.11]	2,056.77	8.28

4.2 JIF、CiteScore 与 Scilit I *3(/N)*

4.2.1 指标分布

图8展示了三组指标的散点图，并标注了拟合曲线的决定系数（R^2）、斯皮尔曼秩相关系数和林氏一致性相关系数（CCC）。与斯皮尔曼秩相关系数不同，林氏一致性相关系数衡量的不是相关性，而是一致性。三组指标在两年的分布总体上相似，整体表现始终较强。在相关性和一致性方面，排序大致为 CiteScore ≳ I3/N >期刊影响因子，这可能反映了 CiteScore 采用了更宽的引文窗口，从而降低了时间波动性。

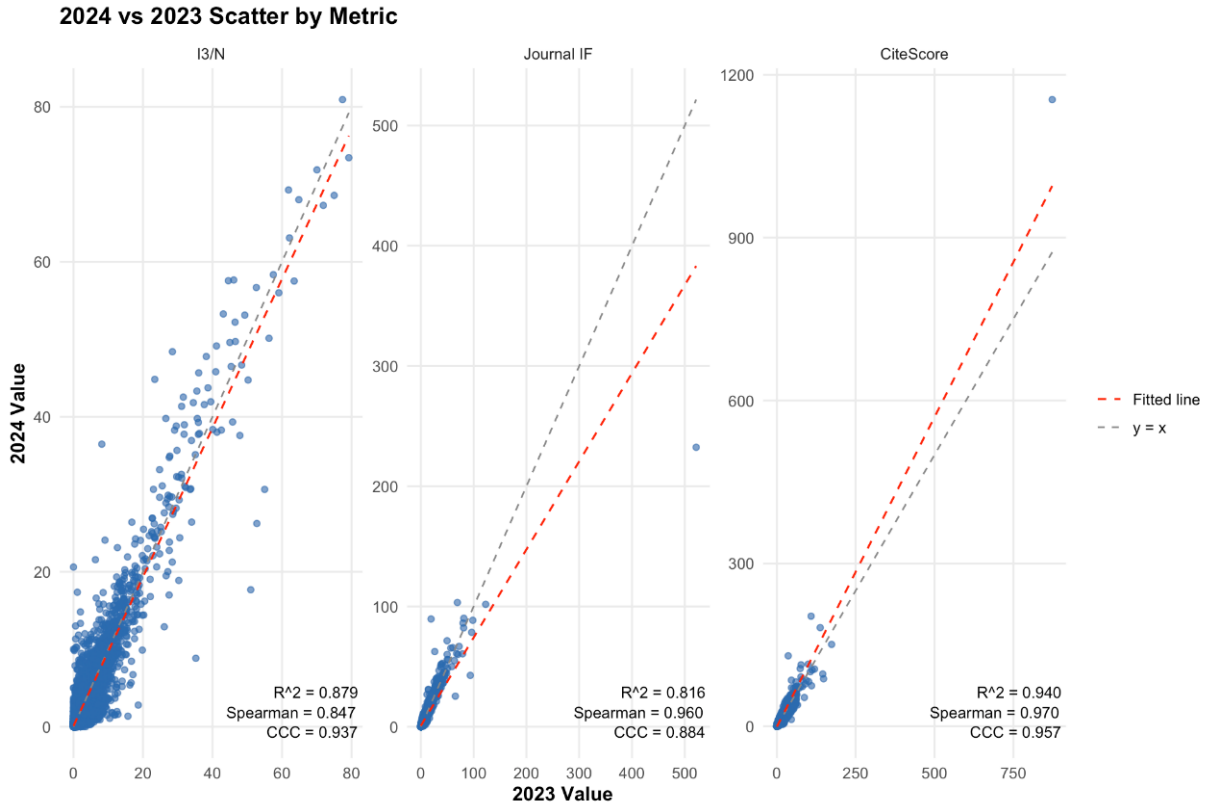


图8. I3/N、期刊影响因子和 CiteScore 的 2023–2024年散点比较。

表4评估了这三个指标之间的相关性和一致性。尽管2024年整体一致性显著下降（CCC从0.81降至0.56），期刊影响因子与CiteScore仍保持很高的秩相关。相比之下，Scilit I3/N与期刊影响因子/CiteScore仅呈中等相关性，表明I3/N反映的结构性影响力有别于基于平均引文的指标。值得注意的是，2024年I3/N与期刊影响因子的尺度趋于一致（CCC升至0.72），这可能是因为两者的趋势更加相似；然而，它们与CiteScore之间的差距扩大了（CCC降至0.30），这与图8中的拟合曲线模式相符。

表4. I3/N、期刊影响因子和 CiteScore (CiteScore) 之间的相关性和一致性

Year	Pair	斯皮尔曼	CCC
2023	I3/N 与 CiteScore	0.65	0.36
2023	I3/N 与期刊影响因子	0.64	0.54
2023	期刊影响因子与 CiteScore	0.94	0.81
2024	I3/N 与 CiteScore	0.62	0.30

2024	I3/N 和期刊影响因子	0.62	0.72
2024	期刊影响因子和 CiteScore	0.94	0.56

图9展示了这三个指标的经验累积分布函数（ECDF），并对比了它们两年间的动态变化。图9A中，I3/N的ECDF曲线更平滑，严格单调递增，且拐点较少，表明其分布更连续，没有明显的不连续性。I3/N的分布看起来也更集中，横轴跨度较短，说明期刊间异质性较低。相比之下，期刊影响因子的ECDF曲线分段更明显，局部增幅较陡，意味着期刊聚集在特定数值范围内；其上尾较长，但所占期刊总数比例相对较小。CiteScore的ECDF曲线比期刊影响因子的更平滑，这可能得益于其更长的聚合窗口和更大的分母。图9B进一步表明，I3/N的两年分布在低值至中高值范围内最为显著，这可能反映了该指标对引文结构变化或其归一化方案的敏感度较高。期刊影响因子在中位数附近的负差异与其在2024年略微向右偏移相符，表明整体期刊影响力略有增加，但增幅并不显著。相比之下，CiteScore基本保持稳定，意味着其计算设计或有效样本量的年际变异有限。

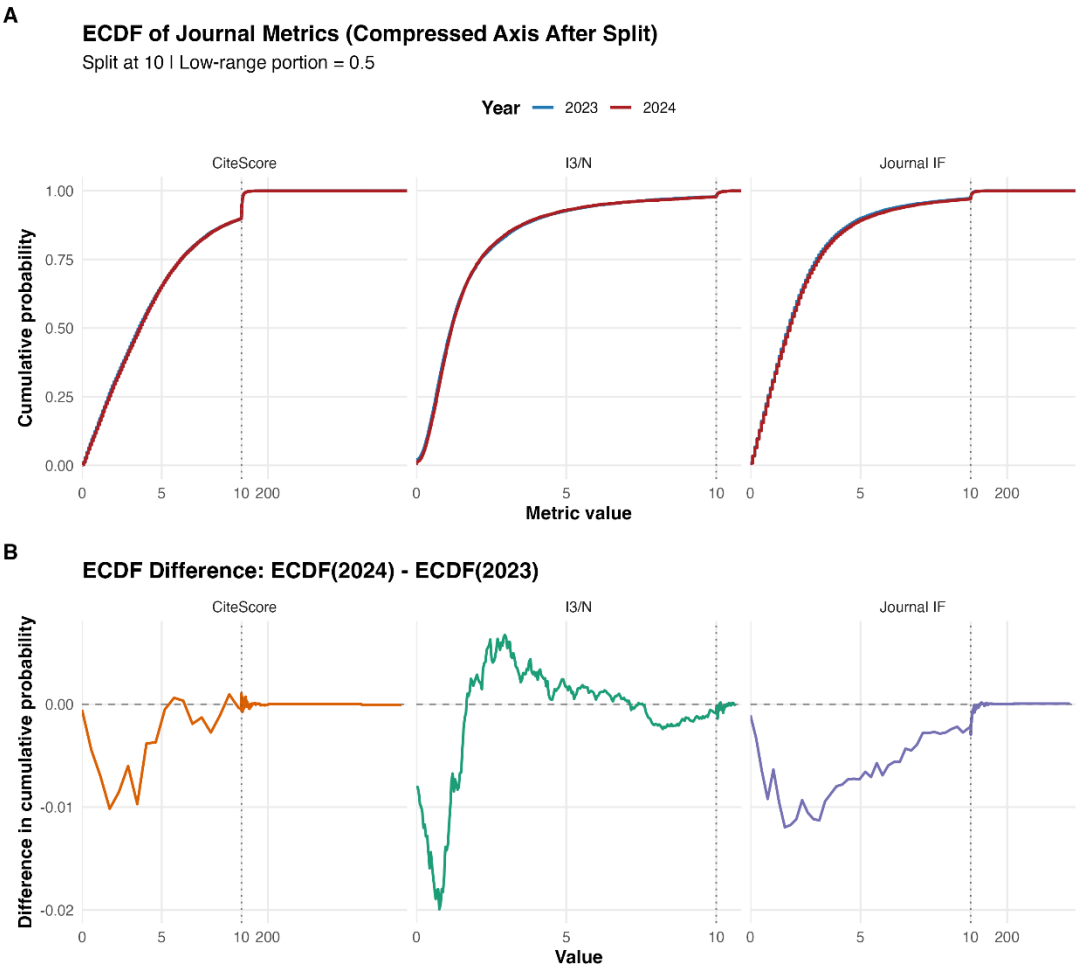


图9. 用于年度差异比较的期刊指标经验累积分布函数（ECDF）分布：A）2023年与2024年的CiteScore、Scilit I3/N 和期刊影响因子 ECDF，横轴在数值 10 之后进行压缩，以突出低中值区间的细节；B）2024年相对于2023年的 ECDF 差值，即 ECDF(2024) - ECDF(2023)，用于判断分布向左或向右偏移。

4.2.2. 基于象限的评估

从概念上看，Scilit I3是一项对规模敏感的总影响力指标，反映出版量和产出规模；而期刊影响因子、CiteScore和I3/N则属于能够减弱规模效应的密度指标。不过，这些密度指标在不同年份呈现出各具特点的分布动态。综合来看，Scilit数据库中的Scilit I3和I3/N能够更全面地评估期刊，并有助于揭示规模效应与单位质量表现之间的差异。

图10使用Scilit I3和Scilit I3/N评估期刊的规模效应和单位质量，并将2023年和2024年的匹配期刊可视化呈现（图10A–B）。在图10A中，第一象限（右上方）的期刊（例如《自然-通讯》、*Nature*和《科学》）同时表现出较高的I3（总体影响规模）和较高的I3/N（每篇出版物的影响力），表明这些出版渠道既“规模大”又“实力强”。第二象限的期刊（例如*PLOS ONE*和《科学报告》）在I3方面表现合理（即具有明显的规模效应），但其Scilit I3/N相对有限，说明与顶级期刊相比，其单位质量并不突出。这类期刊通常具有多学科特征，或在特定研究领域内保持较大的出版量（例如环境科学领域的*Sustainability*和《整体环境科学》）。第四象限聚集了一批知名医学期刊，体现出较高的单位质量，同时主要伴随着较大的出版量。图10B展示了2024年Scilit I3和Scilit I3/N的分布，整体上与2023年的模式相似；不过，第一象限中高影响力期刊的相对位置似乎更加分散，这可能意味着头部期刊之间的进一步集中趋势。

为进一步阐明 Scilit I3 和 Scilit I3/N 的特性，我们对时间窗口内的 Scilit I3、Scilit I3/N、CiteScore 和出版物数量进行了标准化，并从规模效应和单位质量两个维度，将基于 Scilit 的指标与标准化后的 CiteScore 和出版物数量进行比较，以评估它们的独特特征。两年比较的结果见图10C–D。如图10C所示，第一象限（单位质量 > 0 和规模效应 > 0）中的期刊在这两个维度上均优于参照指标（CiteScore 和出版量）。代表性期刊包括《自然-通讯》、*Nature*、《科学》和*IEEE Access*，这表明 I3/N 能够体现相较于 CiteScore 的单位质量优势，而 I3 则表明相同数量的出版物可以产生更大的总体影响。

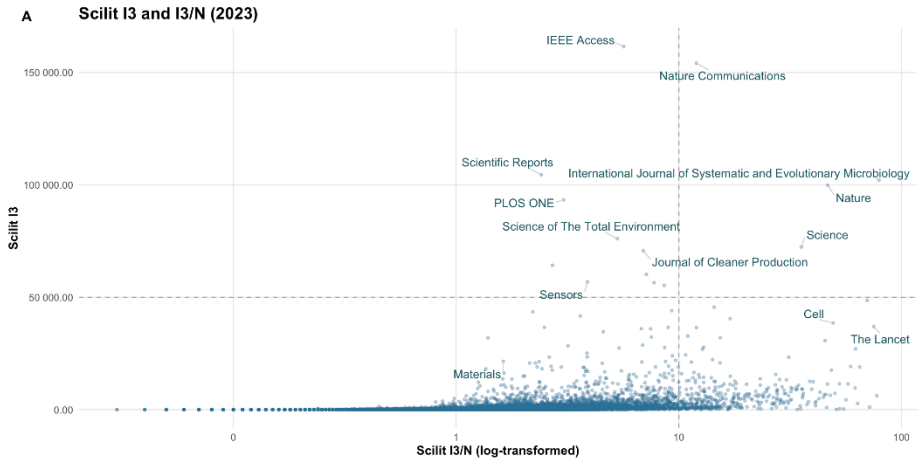
相比之下，第二象限（单位质量 < 0 但规模效应 > 0）反映了规模效应显著，但相较于 CiteScore 不具备单位质量优势。这类期刊通常出版量大、汇总引文多，但单篇文章表现相对有限。最突出的例子是《*CA*：临床医师癌症杂志》。值得注意的是，尽管该期刊在2023年CiteScore榜单中排名第一，且其 Scilit I3 高于出版量，这表明规模效应使 Scilit I3 成为相较于出版量更客观的影响力指标。另一方面，其 Scilit I3/N 大幅削弱了基于相应 CiteScore 得出的单位质量印象。换言之，I3/N 减弱了长尾效应，从而对单篇文章质量作出更均衡的评估。

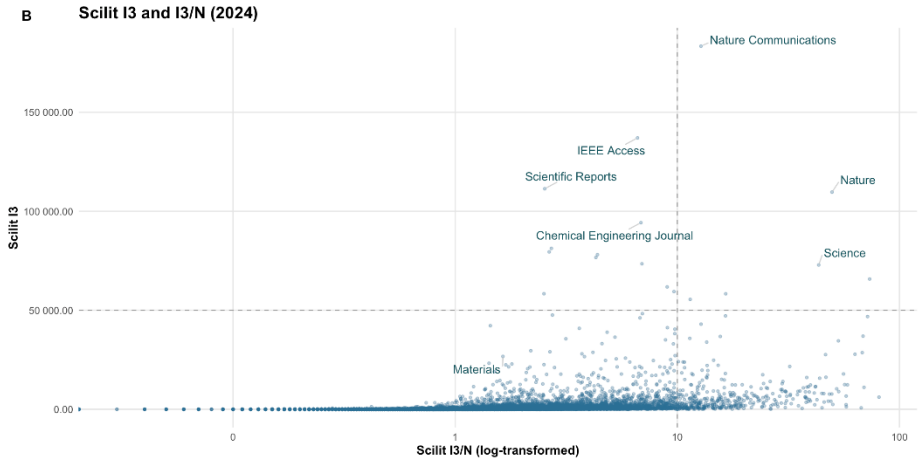
最后，第四象限的期刊（单位质量> 0 但规模效应< 0），例如*PLOSONE*、*Energies*和*Sustainability*，在单位质量方面具有优势，但规模效应不足，即相对于其出版物数量而言，总体影响较为有限。

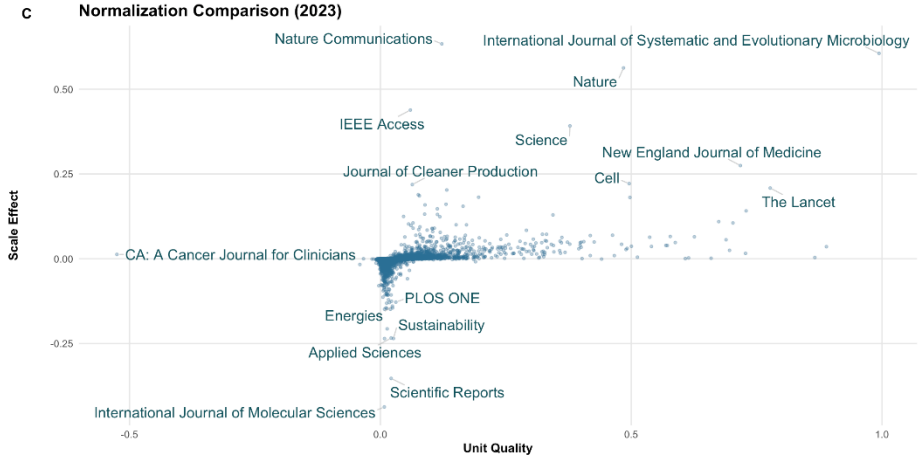
与2023年相比，工程与环境领域的代表性期刊*IEEE Access*和*Journal of Cleaner Production*在规模效应方面的提升更为显著。令人惊讶的是，*International Journal of Systematic and Evolutionary Microbiology*

期刊在单位质量维度上表现出色。这可能是分数引文计数法造成的假象：由于期刊引文的元数据不完整，大多数引文的分数计数都被计算为 1。因此，该期刊的影响力在 Scilit I3 和 I3/N 中被高估了。

基于 Scilit数据库的象限评估从单位质量（I3/N）和规模效应（I3）两个方面提供了互补视角。一些期刊凭借较高的 I3/N，体现出平均质量优势；另一些则凭借较高的 I3，展现出规模化聚合影响力的特点；而第一象限中的顶尖期刊两者兼具。总之，Scilit I3-I3/N 四象限框架能够更细致、便于诊断地呈现期刊影响力，并与出版物数量和 CiteScore 相结合所具有的解释力形成互补。


A
Scilit I3 and I3/N (2023)
IEEE Access
Nature Communications
Scientific Reports
International Journal of Systematic and Evolutionary Microbiology
Nature
PLOS ONE
Science of The Total Environment
Science
Journal of Cleaner Production
Sensors
Cell
The Lancet
Materials
150 000.00
100 000.00
50 000.00
0.00
Scilit I3
0
1
10
100
Scilit I3/N (log-transformed)


B
Scilit I3 and I3/N (2024)
Nature Communications
IEEE Access
Scientific Reports
Nature
Chemical Engineering Journal
Science
Materials
150 000.00
100 000.00
50 000.00
0.00
Scilit I3
0
1
10
100
Scilit I3/N (log-transformed)


C
Normalization Comparison (2023)
Nature Communications
International Journal of Systematic and Evolutionary Microbiology
Nature
IEEE Access
Science
New England Journal of Medicine
Journal of Cleaner Production
Cell
The Lancet
CA: A Cancer Journal for Clinicians
PLOS ONE
Energies
Sustainability
Applied Sciences
Scientific Reports
International Journal of Molecular Sciences
0.50
0.25
0.00
-0.25
Scale Effect
-0.5
0.0
0.5
1.0
Unit Quality

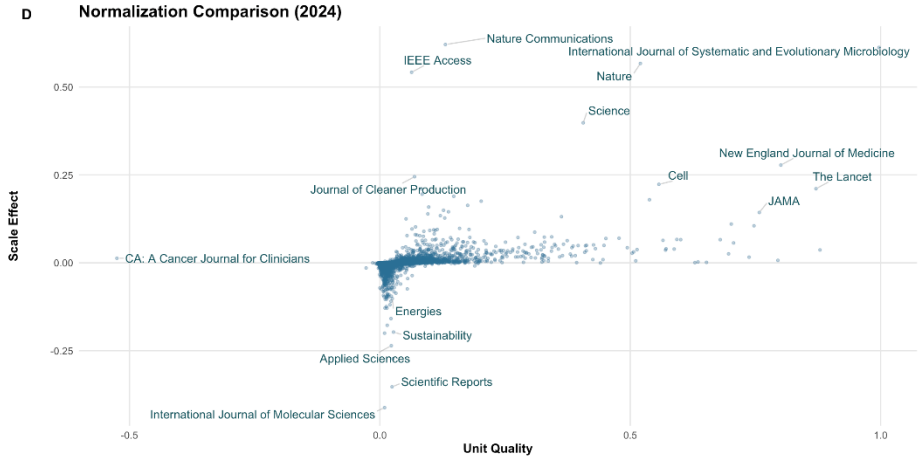

D
Normalization Comparison (2024)
Nature Communications
IEEE Access
International Journal of Systematic and Evolutionary Microbiology
Nature
Science
New England Journal of Medicine
Cell
The Lancet
Journal of Cleaner Production
JAMA
CA: A Cancer Journal for Clinicians
Energies
Sustainability
Applied Sciences
Scientific Reports
International Journal of Molecular Sciences
0.50
0.25
0.00
-0.25
Scale Effect
-0.5
0.0
0.5
1.0
Unit Quality

图10. 基于象限的评估：A）根据2023年数据，对匹配期刊的 Scilit I3 和 I3/N 进行可视化；B）根据2024年数据，对匹配期刊的 Scilit I3 和 I3/N 进行可视化；C）比较 Scilit I3 和 I3/N 与 CiteScore 及期刊出版物（2023年）的标准化情况；D）比较 Scilit I3 和 I3/N 与 CiteScore 及期刊出版物（2024年）的标准化情况。

4.2.3. 学科内部数据库间的异质性

为使三个数据库中的期刊归属能够在共同的主题维度上进行比较，我们依据同一期刊在各数据库中的学科归属，对三种学科分类系统进行了映射，并据此开展跨数据库比较分析。图11展示了重叠期刊数量超过200时的学科（学科编号）跨数据库共现网络。节点代表学科编号。当两条或更多不同颜色的边汇聚于同一节点时，表明三个数据库的学科定义之间存在较强的对应关系。例如，76（Scilit：经济学）、386（Scopus：经济学与计量经济学）和648（WoS：经济学）之间的三角形连接表明，经济学相关期刊存在大量重叠，说明尽管名称和粒度有所不同，跨数据库对齐仍然一致。此外，相对较粗的蓝色边表明，在某些领域，WoS 与 Scopus 之间的重叠更多；绿色边的分布则凸显了 Scilit 与其他两个数据库相互对应的领域。总体而言，该图显示多个关键学科之间存在较强的共现关系和大量重叠，为后续在学科层面开展标准化映射、差异分析和指标对齐奠定了基础。

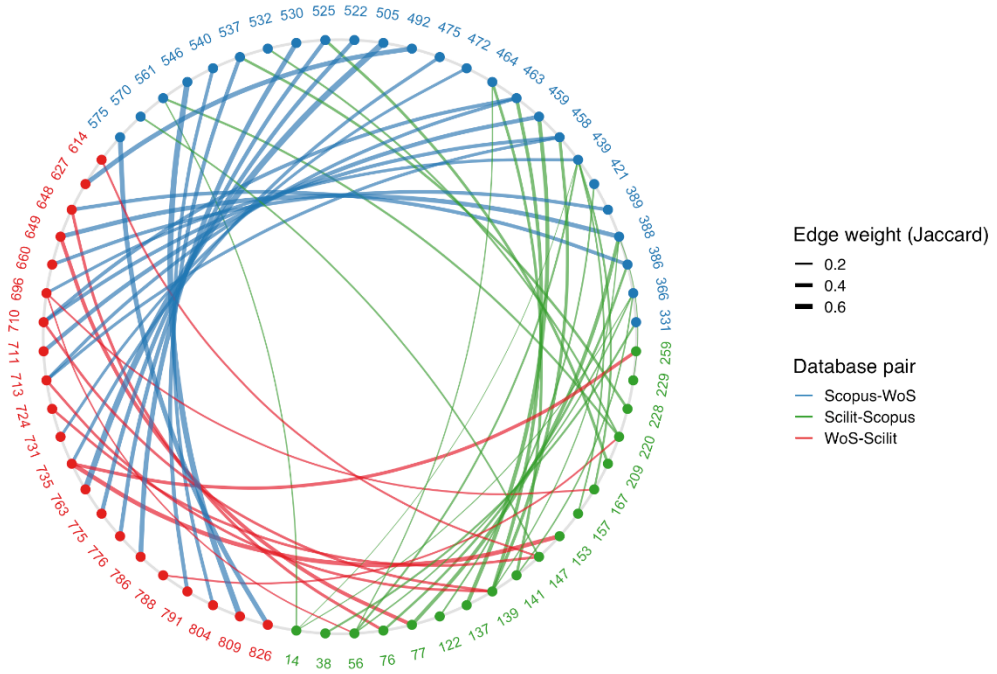


图11. 期刊重叠超过200时的跨数据库学科共现。每个点代表一个学科ID：1–290 = Scilit学科；291–587 = Scopus学科；588–841 = WoS类别。

在建立跨数据库学科定义映射后，我们考察了2024年与2023年相比，各组指标在配对期刊中的表现。表5列出了一些具有代表性的学科领域和指标特征，反映了不同数据库和不同指标之间的一些共性与差异。在人文社会科学（如语言学、法学）以及一些工程学和材料领域（如能源、电气与电子、冶金学）中，三项指标在2024年均显示出显著

改善。这表明，当某一学科处于稳步增长期时，期刊集群的引文生态、平台覆盖度和传统影响力都会提升。生命医学的某些子领域（如传染病）则反映出，当主题出现波动、稿源结构发生调整时，三组指标均呈现显著下降趋势；这表明，后新冠阶段稿源结构的变化和期刊策略的调整，使该领域的影响力呈现减弱态势。在人工智能的前沿方向上，Scilit的I3/N和CiteScore往往先反映热度和引文活跃度，而期刊影响因子的反馈则有所滞后。在教育学的跨数据库映射中，WoS的期刊影响因子和Scopus的CiteScore呈上升趋势，而Scilit的I3/N呈下降趋势，这意味着头部期刊和均值质量有所提升，但平台侧的扩散有所减弱；另一方面，Scilit数据库中该领域的期刊数量十分突出，2024年比2023年增加了488种期刊，这在一定程度上影响了原有期刊数据的分布。总之，期刊影响因子锚定质量和声望，CiteScore衡量广谱均值，而I3/N则捕捉扩散率。在映射后的学科分类中解读这三项指标，可以有效避免对单一指标的误判，从而全面把握学科增长情况。

表5 代表性跨数据库主题映射，包括重叠期刊和趋势信号。↑*/↓* 表示2024年相较于2023年呈现显著上升（下降）趋势；—表示无显著变化。

数据学科映射	重叠期刊	Scilit I3/N	JIF	CiteScore
Scopus：教育学 WoS：教育学与教育研究 Scilit：教育学与教育学	405	↓*	↑***	↑*
Scopus：经济学与计量经济学 WoS：经济学 Scilit：经济学	283	—	—	—
Scopus：语言学与语言 WoS：语言与语言学 Scilit：语言学与语言研究	214	↑*	↑***	↑***
Scopus：法学 WoS：法学 Scilit：政治学与国际关系	108	↑**	↑***	↑***
Scopus：传染病 WoS：传染病 Scilit：病毒学	64	↓***	↓***	↓***
Scopus：材料力学 WoS：材料科学（综合） Scilit：冶金工程	62	↑*	↑***	↑**
Scopus：电气与电子工程 WoS：电气与电子工程 Scilit：网络安全	46	↑*	↑*	↑*
Scopus：能源工程与电力技术 WoS：能源与燃料 Scilit：电力系统与电动汽车	33	↑*	↑*	↑**
Scopus：人工智能 WoS：计算机科学，人工智能 Scilit：人工智能与机器学习	61	↑***	—	↑**

4.2.4. I3/N 与 CiteScore 的差异

为考察所提出的 I3/n 指标在多大程度上会改变期刊评价，我们研究了期刊 I3/n 指标与 CiteScore 指标的百分位排名差异。对于每本同时具有 I3/n 指标和 CiteScore 指标数据的期刊，我们分别计算这两个指标的百分位排名，并求出两者之差。计算公式如下：

$$rank_difference_i = rank_i3n_i - rank_citescore_i \quad (8)$$

其中，$rank_difference_i$ 表示期刊 i 的百分位排名差异，$rank_i3n_i$ 和 $rank_citescore_i$ 分别表示期刊 i 的 I3/n 百分位排名和 CiteScore 百分位排名。

我们按 $rank_difference_i$ 升序排列期刊，并将其划分为四分位数，因此，第一四分位至第四四分位的期刊，其平均 $rank_difference_i$ 由低到高。表 6 列出了各期刊四分位数的一些汇总统计数据。

表 6. 期刊组的汇总统计数据。

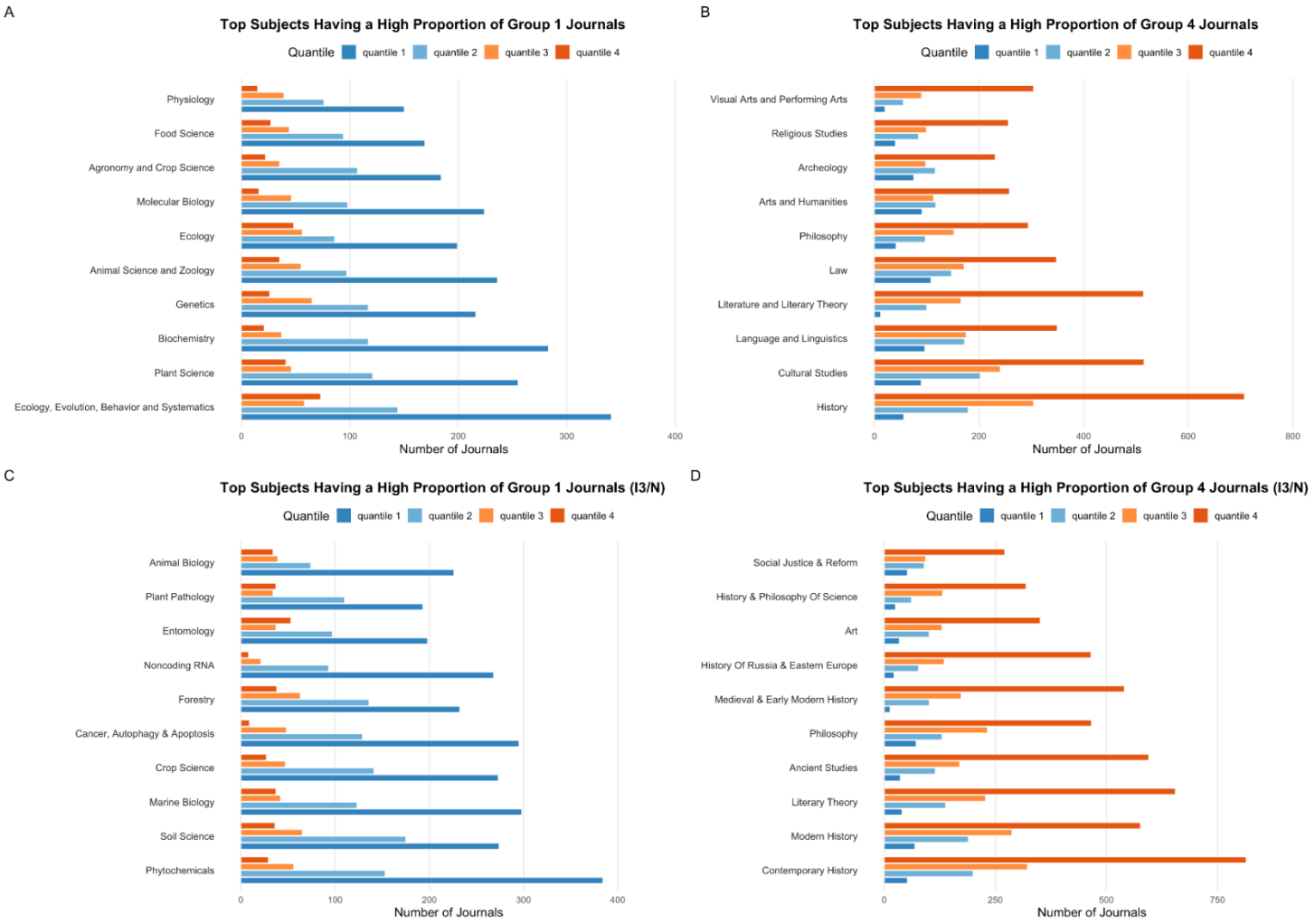
组别	(a) 学术产出的平均数量	(b) 引文计数的平均值	(c) I3指标/n 的平均值	(d) CiteScore 的平均值	(e) 平均百分位排名（I3/n）	(f) 平均百分位排名（CiteScore）
1	445.98	2697.4	0.7	4.43	0.31	0.61
2	542.19	4405.8	1.29	4.81	0.45	0.54
3	571.92	5402.89	3.32	6.18	0.59	0.55
4	260.05	581.04	2.09	1.54	0.64	0.28

在此表中，我们根据$rank_difference_i$将24,687种期刊划分为四分位数，该指标根据方程（8）计算得出：$rank_difference_i = rank_i3n_i - rank_citescore_i$ 其中，$rank_difference_i$ 是期刊i的百分位排名差异，$rank_i3n_i$ 和$rank_citescore_i$ 分别是期刊i的I3/n 百分位排名和CiteScore 百分位排名。

表 6 显示，第一四分位的 I3/n指标平均百分位排名明显低于 CiteScore的平均百分位排名，而在第四四分位中，这一顺序则明显颠倒。这一结果并不令人意外，因为根据期刊分组方法，第一四分位应由那些基于 I3/n指标的期刊排名普遍显著下降的期刊组成，而第 4 组则应包含那些在引入 I3/n指标后，相较于其 CiteScore排名显著上升的期刊。

就学术产出数量和引文计数而言，尽管第四四分位的期刊受益于从基于 CiteScore的评价转向基于 I3/n指标的评价，但其表现仍远低于其他组的期刊。从证据中可以得出的一个重要结论是，第四四分位的学术产出数量并不高；而学术产出数量已被证明是造成 I3指标变化的关键因素，因此在很大程度上已由 I3/n指标加以控制。我们认为，这一现象可能是因为这些受益期刊由于获得的引文数量较少，在 CiteScore指标中的评级较低，但这些期刊拥有较高比例的高被引出版物，而 I3/n指标会强调这一点。列（c）和列（d）也证实了这一推断。这两列表明，与其他组的期刊相比，受益期刊的 CiteScore值平均较低；CiteScore值是一种会根据学术产出数量进行调整的基于引文数量的测度。然而，在考虑学术产出数量后，受益期刊的 I3/n指标优于第 1 组和第 2 组的期刊；I3/n指标主要是一种基于引文排名的测度。

此外，我们还考察了引入 I3/n 是否会对不同学科的期刊产生不均衡的影响。为减轻某些学科学术产出数量较少可能造成的偏差，我们仅考虑学术产出数量在整个学科样本集中排名处于第一四分位的学科。对于每个 CiteScore 学科，我们计算上述各期刊四分位中的期刊数量。随后，我们计算每个学科在各四分位中期刊数量所占的比例。我们首先按照第一四分位中期刊数量所占比例的降序对学科进行排序，然后对第四四分位采用相同的方法。为便于展示，我们仅呈现各排序中排名前十的学科。结果如图12所示。



Fig图 12. CiteScore 和 Scili 中不同四分位数内期刊占比最高的主题 t.

我们统计每个主题下的期刊数量，并根据期刊数量按降序排列主题，选取期刊数量排名位于第25百分位的主题，以确保数据具有足够的代表性。接着，我们根据主题在第一四分位（子图A）和第四四分位（子图B）中的期刊比例对主题进行排序。我们选取在所研究四分位数中占比最高的前十个主题进行展示。横轴表示每个四分位数所涵盖的期刊数量，纵轴表示主题。我们还在多个柱形中展示了每个四分位数中的期刊数量。子图C和子图D与Scilit学科相对应。

子图12A显示，期刊占比较大的主题集中于生物学、生态学、农业等领域，这些期刊将因I3/n而被显著降级。相比之下，子图12B表明，人文学科相关主题将从由CiteScore转向I3/n的评价变化中大幅受益。子图12A和子图12B所涉及的主题在被引次数和学术产出方面存在差异。子图12C和子图12D展示了关于Scilit主题分类的类似分析。引入I3/n后期刊排名将大幅下降的主题涉及生物学、生态学、农业和动物科学，而在I3/n评价框架下表现优于其他主题的期刊，其所属主题与历史学、读写能力、艺术和哲学相关。这两个子图所呈现的模式总体上与子图12A和子图12B一致。因此，引入I3/n将对具有不同引用文化和出版频率的主题产生不均衡的影响；与社会科学相关的研究在CiteScore指标中可能曾被低估，但通过I3/n指标可以得到更公平的评价。

4.3 基于出版商的分析

图13展示了主要出版商旗下期刊过去两年的 I3/N 分布，并采用该指标两年中位数差异的相对变化量（Δ）以及 Wilcoxon 检验标注显著性（例如“稳定”“增加”或“减少”）。不难看出，主流综合出版商（如 Elsevier、施普林格·自然、威立、SAGE）的分布呈现出典型特征：低端集中，上尾延长，年度总体中位数变化接近于零。这表明，在大样本中，期刊组合质量具有显著的惯性和稳健性。相比之下，Frontiers Media SA 的中位数下降最为明显（~11.5%），同时期刊数量有所增加；Wolters Kluwer Health 和牛津大学出版社的中位数则适度下降，而 MDPI 和美国化学会基本保持稳定。这一证据表明，一方面，期刊数量与稳定性呈正相关，规模越大，中位数越不容易受到年度扰动的影响；另一方面，快速扩张可能带来“稀释效应”：新增期刊的影响力有限，无法抵消现有期刊所贡献的影响力，从而拉低出版商层面的分布。

从出版商的角度综合分析 I3/N 分布后，我们发现，覆盖范围广且发展成熟的期刊组合往往能逐年保持稳定的集中趋势，即使尾部特征有所变化。当期刊组合快速扩张，且新增期刊主要是较新的或影响力较低的期刊时，尽管期刊总数增加，I3/N 中位数仍可能下降，这凸显了规模与短期影响力整合之间的权衡。因此，学术出版商可以结合 Scilit数据库的年度 I3/N指标和期刊数量，为结构优化和期刊发行政策提供依据。

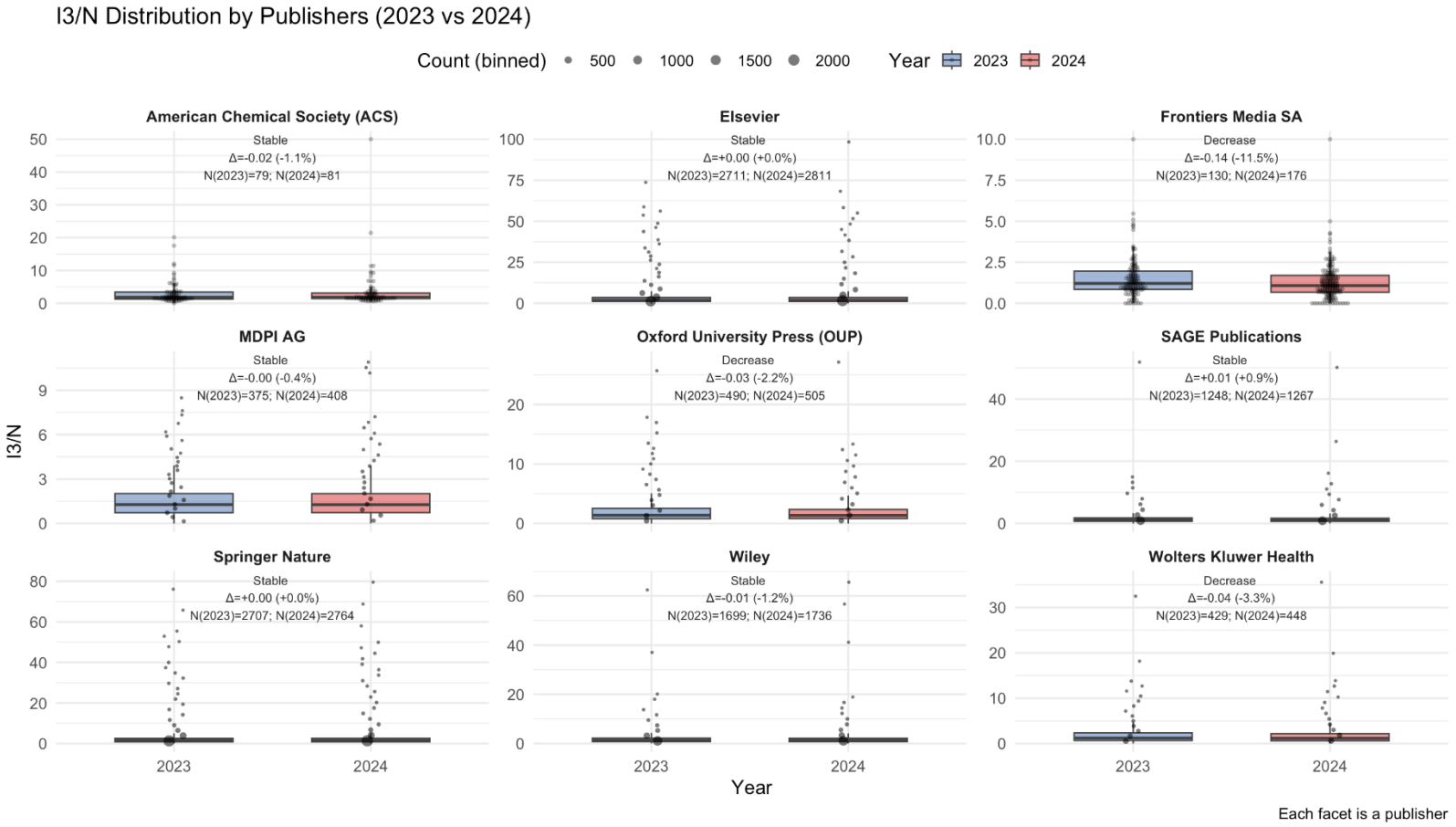


图13. 按顶级出版商划分的 I3/N 分布。Δ 表示通过 Wilcoxon 检验得出的中位数差异的相对百分比（“稳定”、“增加”或“减少”）， N 表示该年份每个出版商的期刊总数。

5. 讨论

本研究旨在介绍 Scilit 数据库，将其打造为一个全面且不排斥任何群体的学术平台，并在其框架内实施综合影响指标（I3），作为期刊影响力评估的有力替代方案。我们将 Scilit、I3 及其标准化版本（I3/N）与成熟指标期刊影响因子和 CiteScore 进行基准比较，得出了若干重要见解。本文的核心发现是，Scilit 广泛的覆盖范围与 I3 对分布敏感的方法学相结合，能够提供一种更细致、更公平且诊断能力更强的框架来评估期刊影响力，尤其适用于跨学科语境。

我们的结果显示，最直接的优势是期刊覆盖范围大幅增加。Scilit 为超过 6 万种期刊计算 I3 指标，数量约为 Scopus CiteScore 覆盖期刊数的两倍；这不仅是数量上的成就，也是质量上的提升。Scilit 直接回应了对传统数据库的一项根本性批评：其覆盖范围天然偏向大型商业出版商发行的英语期刊，往往以牺牲区域性期刊、学会所有的期刊或规模较小的开放获取出版物为代价（Larivière, Haustein, & Mongeon, 2015）。借助多源数据摄取模型，Scilit 缓解了这些偏差，从而让更广泛的学术成果得以展现。这种广泛的覆盖范围，是开展更具包容性、能够代表全球学术成果的文献计量分析的先决条件。

从方法论上看，I3框架通过从根本上重新思考影响力的聚合方式，彰显了自身价值。与JIF和CiteScore不同，后两者采用简单平均值，因此容易受到少数高被引文章以及分子与分母不对称这一问题的扭曲；I3则考虑了完整的引文分布。我们的实现方法为位于前1%、前10%、前50%和后50%百分位的出版物分别赋予100、10、2和0的权重，有意将重点转向高影响力研究，同时有效忽略未被引用或仅获少量引用的论文。这种方法在统计学上更加稳健，因为它较不易受到偏态分布和通过引文堆叠进行策略性操纵的影响。象限分析（图10）有力地展示了这一双指标系统的诊断能力。它能够清晰地区分“规模大且实力强”的期刊（I3和I3/N均高，例如*Nature*）、“规模大但分布分散”的期刊（I3高、I3/N低，例如PLOS ONE），以及“专注且质量高”的期刊（I3低、I3/N高）。将规模效应与单位质量分开考察，所呈现的图景比任何单一的平均值指标都更加真实，也更具指导价值。

我们的分析所揭示的学科异质性，进一步凸显了采用 I3/N 这类指标的必要性。I3/N 与 CiteScore 的百分位排名存在显著差异，其中呈现的主题特定模式也极具启示性。研究发现，生态学和生物学等领域的期刊在 I3/N 排名中往往排名下降，而人文社会科学领域（如法学、语言学）的期刊则常常排名上升，这直接体现了领域标准化这一问题。期刊影响因子和 CiteScore 等传统指标，往往对引文密度较低、引文窗口较长的领域不利。正如 Zhou 和 Leydesdorff（2011）所倡导的，我们在计算 I3 时采用引文分数计数法，从根本上对不同学科之间参考文献列表长度和引文潜力的差异进行了标准化。因此，I3/N 似乎创造了一个更加公平的

竞争环境，从而纠正了引文评价体系长期以来对社会科学与人文学科研究价值的低估。

对于学术生态系统中的利益相关者而言，这些发现具有深远意义。对于研究人员和图书馆员来说，Scilit-I3组合提供了一种更全面、偏见更少的期刊选择工具。例如，社会科学某个领域的学者如今可以在更广泛的 Scilit语料库中找到这样的期刊：它们的 CiteScore 可能并不高，但单篇文章影响力（I3/N）却很强。对于出版商而言，按年追踪其期刊组合中 I3/N 分布的变化（如图6所示），可以提供战略情报。观察到的“稀释效应”出现在快速扩张的出版商身上：随着新增影响力较低的期刊，I3/N 中位数随之下降。这凸显了增长与整体影响力之间切实存在的权衡。这些数据有助于制定更明智的创刊与收购策略，而不是一味追求提高期刊影响因子。对于科研管理者和政策制定者而言，I3框架符合《旧金山科研评估宣言》（DORA，https://sfdora.org）的原则；该宣言倡导采用避免滥用基于期刊的指标的衡量方式。I3既可以在文章层面计算，也可以汇总至期刊、院系或机构层面，因此为负责任的评价提供了途径，能够关注影响力的分布，而不仅仅是其均值。这也契合了近期关于负责任地使用指标的呼吁，例如《科研评估改革协议》（CoARA，https://coara.eu）。

然而，这种方法并非没有局限性。首先，Scilit 的覆盖范围极广，本身就带来了数据质量方面的挑战。依赖从多种来源进行自动化采集，可能导致元数据不一致；*International Journal of Systematic and Evolutionary Microbiology*出现的异常结果便暗示了这一点：不完整的引文数据可能抬高了其 I3分数。尽管 Scilit 采用了去重和消歧算法，但面对如此庞大的数据规模，确保数据完全准确仍是一项持续的挑战。其次，I3指标的加权方案虽参考了专家意见，但仍包含价值判断。将排名后 50% 的论文权重设为零，等于明确表示完全不考虑那些可能通过其他方式产生影响的出版物（例如借助替代计量指标或实践导向的应用）。针对不同的评价情境，可以探索其他加权方案。第三，跨学科映射虽具创新性，但并非完美的替代方案。对齐 WoS、Scopus 和 Scilit 的学科类别是一项复杂工作，残留的错位可能影响学科规范化比较的精确度。

未来的研究应解决这些局限，并在此基础上进一步发展。下一步的关键工作是开展严格、独立的审计，将 Scilit 的元数据质量与经过整理的数据库进行比较。此外，探索在中观层面（机构）和微观层面（个体研究人员）应用 Scilit I3框架，将检验其用于更广泛科研评价的效用。通过对编辑和研究人员开展访谈的定性研究，可以考察相较于传统指标，I3在公平性和实用性方面的感知。最后，I3对新兴趋势的动态响应——例如我们的结果中所指出的人工智能研究的快速增长——值得进行纵向研究，以了解其作为科学范式变化早期指标的敏感性。

In 总之，本研究表明，将 I3 指标与 Scili t

数据库相结合，是文献计量学和科学计量学领域向前迈出的重要一步；这一转变也可能对科学学、科学社会学以及科学技术研究（STS）产生重大影响。它将讨论从对基于选择性数据库计算出的简单平均值的执着，推进到更加复杂、公平且负责任的影响力评估模式。通过在包容性的数据基础设施中采用基于百分位数的、对分布敏感的指标，学术界可以促进对科学进步更准确、更多元的呈现。

6. 结论

通过将 Scilit、I3 指标和 I3/N 与期刊影响因子和 CiteScore 进行基准比较，我们发现，Scilit 广泛的覆盖范围结合 I3 指标对分布敏感的方法学与标准化，能够为期刊影响力评估提供更细致、更公平且诊断稳健性更高的框架，尤其适用于跨学科情境。Scilit 通过多源聚合显著扩大了期刊覆盖范围，缓解了传统数据库偏向英文出版物和商业出版期刊的结构性偏差。通过百分位数加权整合完整的引文分布，I3 指标和 I3/N 进一步降低了基于均值的指标所固有的偏态性和可操纵性，并能够将规模效应与单位质量评估相分离，包括通过期刊象限诊断实现这一点。通过映射各数据集中的期刊学科归属，我们识别出 Scilit 体系下的 I3/N 与期刊影响因子/CiteScore 在不同学科分类中的绩效异同，为评估学科发展提供了新的视角。此外，I3/N 与分数计数法结合后，可以纠正对引文密度较低领域（如人文学科和人文社会科学）的系统性低估。最后，我们证明了 I3/N 能够捕捉期刊绩效在出版商层面的差异，为组合部署策略提供决策参考。基于这些发现，我们认为，Scilit 和 I3/N 代表了从“基于选择性数据库的简单平均值”转向“植根于包容性数据的对分布敏感的评估”的范式转变，从而为负责任的科研评估提供了可操作的路径。与此同时，未来的研究应包括对 Scilit 元数据质量开展独立审计，将 I3 指标的验证扩展至机构和个人层面，并追踪 I3 指标对新兴领域（例如人工智能）的动态敏感性，以评估其面向未来的潜力。

7. 材料和数据的可用性

用于复现所呈现结果的所有数据和代码均已公开存放于 10.5281/zenodo.18013019。

8. 竞争利益声明

RH 是《Scientometrics》的副编辑。RH 和 YB 是 Scilit数据库的顾问委员会主席。其他作者声明，他们受雇于 MDPI；该组织负责开发和维护本研究所述的 Scilit数据库。作者声明，这一从属关系未影响研究设计、数据分析、结果解释或论文的发表决定。

9. CRediT 作者贡献声明

Haochen Dong：概念化、数据整理、方法学、软件、撰写——初稿。Qiao Sun：概念化、数据整理、方法学、监督、撰写——初稿。Yanping Mu：监督、撰写——初稿。Lu Liao：监督、撰写——初稿。Diogo Rodrigues：撰写——初稿。Frank Sauerburger：撰写——初稿。Yi Bu：监督、撰写——初稿。Robin Haunschild：监督、撰写——初稿。

10. 致谢

作者感谢 Shu-Kun Lin 博士和 MDPI 提供技术支持，感谢 MDPI 作者服务提供专业英语编辑服务，并感谢 Scilit 开发者（Haijiang Li、Mladen Milutinovic、Darko Todoric 和 Fuyong Zhang）绘制实体关系图。作者使用 ChatGPT（OpenAI）作为辅助工具，对部分分析代码进行整理和重构。所有研究设计、数据分析、结果解释以及任何遗留错误，均由作者独自负责。

参考文献

Adriaanse, L. S., & Rensleigh, C. (2013). Web of Science、Scopus 和 Google Scholar：内容覆盖全面性的比较。《电子图书馆》，31(6), 727-744. Aghaei Chadegani, A., Salehi, H., Yunus, M. M., Farhadi, H., Fooladi, M., Farhadi, M., & Ale Ebrahim, N. (2013). 两个主要学术文献集的比较：Web of Science 和 Scopus 数据库。《亚洲社会科学》，9(5), 18-26. Ahmad, S. J., Ahmed, A. R., Kowalewski, K. F., Nickel, F., Rostami, K., Stocker, C. J., Hakky, S. M., Archid, R., McWhinnie, D., & Mohajer-Bastami, A. (2020). 综合医学期刊中的引文经典：证据质量评估；系统性综述。《从基础到临床的胃肠病学与肝病学》，13(2), 101. Baas, J., Schotten, M., Plume, A., Côté, G., & Karimi, R. (2020). Scopus：一个经过整理的高质量文献计量数据来源，可用于定量科学研究。《定量科学研究》，1(1), 377-386. Birkle, C., Pendlebury, D. A., Schnell, J., & Adams, J. (2020). Web of Science：科学与学术活动研究的数据来源。《定量科学研究》，1(1), 363-376. 科睿唯安。(2025a). 科睿唯安影响因子。载于（第2025卷）。https://clarivate.com/academia- government/essays/impact-factor/. 科睿唯安。(2025b). 期刊影响因子（JIF）简介。载于（第2025卷）。https://clarivate.com/academia-government/scientific-and-academic-research/research-funding-analytics/journal-citation-reports/publishers/first-time-publishers/. de la Broise, J.-B., Bernard, N., Dubuc, J.-P., Perlato, A., & Latard, B. (2021). 如何预训练高效的跨学科语言模型：ScilitBERT案例。载于2021年第六届国际信息技术研究会议（ICITR）（第1-6页）：电气电子工程师学会。Delgado-Quirós, L., & Ortega, J. L. (2024). 八个免费学术数据库中出版元数据的完整度。《定量科学研究》，5(1), 31-49. Feldner, D. (2025). 介绍 CiteScore 2024：期刊影响力的全面透明指标。载于（第2025卷）。https://blog.scopus.com/introducing-citescore-2024/. Fong, E. A., & Wilhite, A. W. (2017). 学术研究中的作者身份与引文操纵。《PLOS ONE》，12(12), e0187394. Joshi, A. (2016). Scopus 与 ISI科学网的比较。《全球价值期刊》，7(1), 1-11. Juyal, D., Thawani, V., Sayana, A., & Pal, S. (2019). 影响因子：变异、操纵与扭曲。《家庭医学与初级保健杂志》，8(11), 3475-3479. Leydesdorff, L., & 博恩曼, L. (2011). 综合影响力指标与影响因子的比较：一种具有政策意义的替代性研究设计。《美国信息科学与技术学会会刊》，62(11), 2133-2146. Leydesdorff, L., 博恩曼, L., & 亚当斯, J. (2019). 综合影响力指标再探（I3*）：期刊影响因子的非参数替代指标。《科学计量学》，119(3), 1669-1694. Martín-Martín, A., Orduna-Malea, E., Thelwall, M., & López-Cózar, E. D. (2018). Google Scholar、Web of Science 和 Scopus：对252个学科类别中引文的系统比较。《信息计量学杂志》，12(4), 1160-1177. Meho, L. I. (2019). 使用 Scopus 的 CiteScore 评估计算机科学会议的质量。《信息计量学杂志》，13(1), 419-433. Priem, J., Piwowar, H., & Orr, R. (2022). OpenAlex：一个完全开放的学术作品、作者、出版渠道、机构和概念索引。arXiv 预印本 arXiv:2205.01833. Singh, V. K., Singh, P., Karmakar, M., Leta, J., & Mayr, P. (2021). Web of Science、Scopus 和 Dimensions 的期刊覆盖范围：比较分析。《科学计量学》，126(6), 5113-5142. Team, R. C. (2025). R：统计计算语言与环境。载于。奥地利维也纳：R统计计算基金会。Wilhite, A. W., & Fong, E. A. (2012). 学术出版中的强制引用。《科学》，335(6068), 542-543. Wilsdon, J. (2016).《指标浪潮》：关于指标在科研评估与管理中作用的独立评估。Zhou, P., & Leydesdorff, L. (2011). 科研评价中的引文分数计数：对北京清华大学进行跨学科与学科间评估。《信息计量学杂志》，5(3), 360-368.