影响图书馆与信息科学研究论文引文影响力的特征

作者：Kalervo Järvelin，坦佩雷大学；
Yu-Wei Chang，国立台湾大学；Pertti
Vakkari，坦佩雷大学

地址：Järvelin&Vakkari：信息技术与传播科学学院|传播科学，坦佩雷大学，
FI-33014坦佩雷大学，芬兰Chang：图书馆与信息科学系，国立台湾大学，罗斯福路
四段1号，台北10617，台湾（中华民国）

电话：+358 50 4034353 (Järvelin)+886-
2-3366-1743 (Chang)+358 50 5288354
(Vakkari)

邮箱：yuweichang2013@ntu.edu.tw
kalervo.jarvelin@tuni.fi
pertti.vakkari@tuni.fi

通讯作者：Vakkari 是负责接收通信和校样的作者。

摘要

本文分析了2015年发表于31种国际顶尖图书情报学期刊的图书情报学（LIS）研究论文的引文影响力。主要研究问题是：作者的学科构成与图书情报学论文的其他内容特征相结合，在多大程度上会影响其引文影响力？本文从获得的引文数量及其权威性方面分析影响力，并采用离群值标准化和子领域标准化。通过定量内容分析考察的文章特征包括主题、方法论、贡献类型和作者团队的学科构成。论文获得的引文追溯自2015年至2021年5月。施引文献的权威性根据其截至2021年5月获得的引文数来衡量。总体发现是，作者的学科构成与引文得分显著相关。不同学科构成之间的引文得分差异通常出现在信息检索和科学传播领域。在这两个主题中，图书情报学与计算机科学的组合所获引文得分显著高于许多其他学科构成，例如在信息检索领域，图书情报学单独作为学科或与人文学科相比；在科学传播领域，自然科学、医学或社会科学单独作为学科。本文的创新之处在于，对内容、作者构成和影响力进行了联合分析。

1. 引言

由于引文影响力分析在研究评价中处于核心地位，因此了解哪些因素会影响学术出版物的引文影响力十分重要。文献计量学文献中已有充分记录关于引用的理由（例如，Weinstock，1971年；Tahamtan& Bornmann，2019年）。这些理由包括向先前的研究致敬、评价当前的研究、为论断提供依据等。然而，这些更属于引用的规范性而非描述性理由，并且在研究实践中往往未被遵循。描述性理由包括自引、裙带关系和语言偏见（Latour，1987年；Phelan，1999年）。它们对引文影响力的作用存在争议。无论是哪种情况，仅基于文献（或其书目信息）并不总能分析这种影响，而是需要结合情境信息。

在本研究中，我们将按可从文献中收集数据的因素分析引文影响力。我们将这些因素分为情境因素和内容因素。前者的例子包括作者名称、期刊名称、机构名称、国家名称、出版年份等。这些因素易于收集，适用于自动化信息计量分析，并且已在文献中得到充分研究（例如，Phelan，1999年）。这些因素并不直接反映研究内容。学术出版物中的内容因素涉及研究主题、研究问题、解决问题的方法、数据收集方法、分析类型以及贡献类型。这些因素存在于文献文本中，但对于影响力分析而言，收集起来并非易事，而且对于其对引文的影响也尚未得到充分研究（Yan，2015年）。鉴于后者在研究设计、实施、报告和评估中发挥着核心作用，我们预计它们与研究影响力相关。否定的回答将表明，研究领域内对所有类型的研究都持宽泛的评价态度。至少，这种解释将超出研究主题和方法论类型的范围。另一方面，肯定的回答将表明，在研究人员中，某些研究设计与方法论类型的组合比其他组合具有更大的影响力。

在出版物层面，引文影响力意味着统计某篇出版物收到的引文数量，并将结果表示为某种指标值。由于引文计数通常呈偏态分布，并且需要进行调整才能获得可比结果，因此在解读引文计数时必须谨慎：必须考虑各领域的规模（例如，Phelan，1999年；Waltman，2016）。目前已有多种引文影响力标准化指标。这些指标会对领域规模、时间因素（出版窗口和引文窗口；引文老化）、出版物类型或引文权重进行调整（例如，Järvelin & Persson，2008；Waltman，2016）。一些指标衡量单篇出版物的影响力，例如领域加权引文影响力；而另一些指标则衡量来自同一来源的一组出版物的总体影响力，例如 H 指数及其衍生指标（例如，Waltman，2016）。在本文中，我们分析单一领域（图书情报学）较窄出版窗口（2015年）内、有限引文窗口（2015年–2021年5月）内论文的引文影响力。因此，所有论文的条件均具有可比性，无需针对出版年份、出版物类型或引文年份进行标准化。

图书情报学出版物的主题可能会影响其影响力，因为图书情报学包含几种截然不同的研究传统，而不同的学术和专业受众对这些传统感兴趣。Vakkari（1994）从理论层面探讨了图书情报学的性质，Jarvelin and Vakkari（1990、1993、2021）和 Tuomaala 等人（2014）则从实证特征层面进行了探讨。这些研究揭示了学科中的知识生产状况，但只能间接反映研究质量。尽管存在衡量出版论坛质量的综合指标

（例如*JIF*），但这些指标将整个论坛作为整体进行评价，并未报告论坛所涵盖主题层面的影响力。总体研究策略和贡献类型也可能影响出版物的引文影响力，因为某些研究策略显然比其他研究策略更受欢迎（例如，Jarvelin & Vakkari，1990；1993；2022；Tuomaala 等人，2014）。

学科构成文章作者团队可能会影响研究主题的选择、研究问题的定义以及方法选择（Vakkari、Chang和Järvelin 2022a、b）。这意味着所产生的贡献借鉴了多样化的文献，也可能对这些文献作出贡献，因此会与引文影响力相互作用。

图1阐释了本研究的主要概念。分析单位是一篇被引的图书情报学期刊文章（图中的主要文章）。它对知识生产的影响力，即引文影响力，以引用它的二级文献数量表示。然而，施引文献的权威性各不相同，这取决于三级文献对它们的引文计数。我们不使用单纯的引文计数来衡量一篇图书情报学文章的影响力，而是根据其按权威性加权的引文来计算其影响力得分。我们的目标是 выяс明，被引图书情报学文章的内容特征——主题、研究策略、贡献类型或作者团队的学科构成——是否与其影响力得分的方差相关。

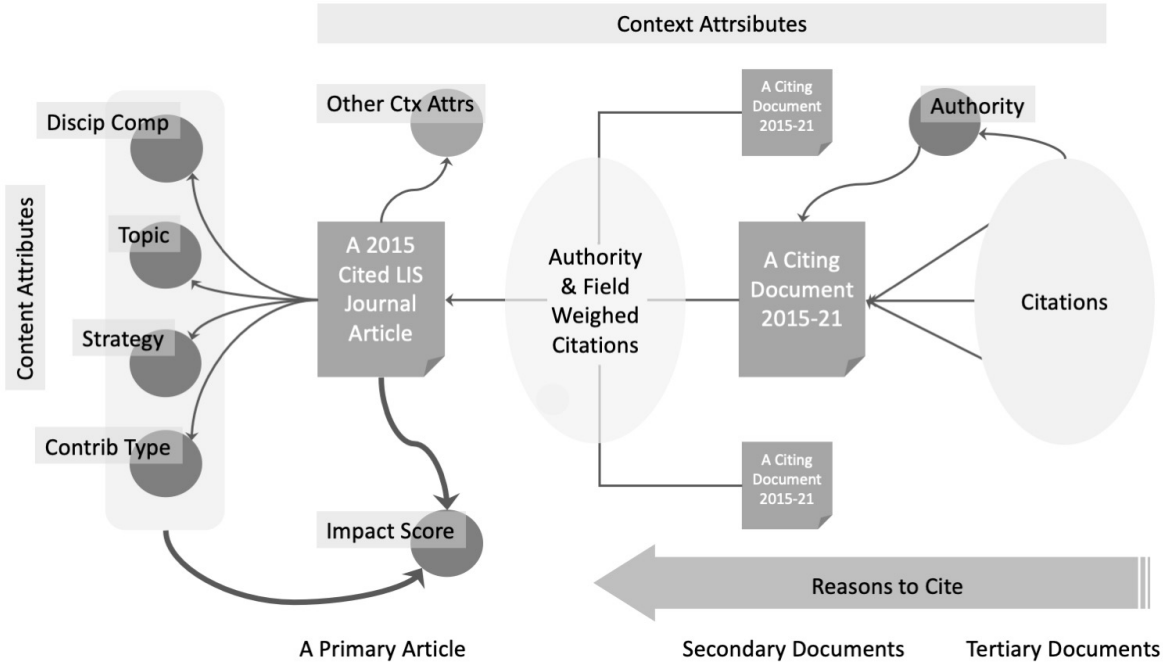


图1。 本研究的主要概念

左侧的内容属性是通过对主要文章进行内容分析得出的。图 1 中间和右侧的情境属性侧重于引文及其

通过领域加权实现权威性和标准化，生成主要文章的影响力得分。

本研究的研究发现有助于：a) 理解图书情报学的本质；b) 设计知识组合恰当的研究项目；c) 规划博士项目并招收博士生；以及 d) 帮助学者个人规划职业生涯。此外，本研究还对信息计量学的 e) 实证研究和 g) 方法论知识体系作出了贡献。

2 既有研究

2.1 图书情报学实证研究的方法

识别主要主题及其变化长期以来一直是图书馆与信息科学研究的关注重点。研究人员采用多种方法探讨图书情报学相关主题，包括内容分析（Järvelin & Vakkari，1990年、1993年、2022年；Tuomaala 等人，2014年）、共被引分析（Åström，2007；Hou 等人，2018年）、书目耦合（Chang 等人，2015年）、关键词分析（Liu & Yang，2019年）、主题索引词分析（Blessinger & Frasier，2007）、共词分析（Mokhtarpour & Khasseh，2021年）、潜在狄利克雷分配建模（Figuerola 等人，2017年；Han，2020年），以及这些方法的各种组合（Chang 等人，2015年）。内容分析是一种用于深入分析的传统方法。通过专家评判，它能够揭示研究论文的多种特征。例如，Järvelin 和 Vakkari（1990年）提出的分类框架包含主题、方法、观点和策略，并经常被其他研究人员采用（Hider & Pymm，2008；Lund & Wang，2021年；Ma & Lund，2021年）。

2.2 期刊文章作为数据源

期刊文章一直是科学计量分析中典型的观察单位。所观察的文章继承了其发表期刊的学科属性。这至少有三项缺点。期刊中的文章既不全都代表该期刊在 Web of Science 等数据库中的学科分类，也不全都是研究论文。例如，Järvelin 和 Vakkari（2022年）指出，2015年，31种主要图书情报学研究期刊中有13%的文章不属于图书情报学，另有7%并非研究论文。2005年，非研究类文章的比例更高，达到30%（Tuomaala 等人，2014年）。第三项局限是，基于期刊的文章特征描述掩盖了文章在子领域（主题）上的差异，

这些因素会影响引文和影响力（Yan，2015年）。因此，基于期刊的分析可能会模糊学科之间的界限，将非研究论文纳入分析，并忽视不同学科子领域之间影响力的差异，从而降低研究发现的效度。在本文中，我们避免了这些陷阱。

2.3 图书情报学出版物中的图书情报学作者与非图书情报学作者

图书馆与信息科学研究中非图书情报学作者的存在日益增加，导致非图书情报学作者主导了图书馆与信息科学研究领域（Chang，2019年；Chang & Huang，2012年；Urbano&Ardanuy，2020年；Vakkari等人 2022a、b）。作者所属机构信息通常用于确定作者的学科，进而用于识别通过跨学科合作产生的论文。研究已证实，来自众多学科的非图书情报学作者对图书情报学出版物的贡献，以及隶属于计算机科学相关院系和研究机构的非图书情报学作者所占的较大比例（Chang，2018年；Vakkari 等人2022a）。Chang（2019年）考察了2015年75种图书情报学期刊的论文，也报告称，约70%的图书情报学期刊由非图书情报学作者主导。Chang（2018年）证实，非图书情报学作者和图书情报学作者具有不同的主题偏好。

Vakkari 等人（2022b）使用本文的数据集发现，图书情报学的大多数论文由非图书情报学领域的作者撰写（57%）。不过，在图书馆与信息服务及机构研究中，图书情报学学者占明显多数（68%）；而在信息检索（73%）和科学传播（科学计量学，69%）领域，外部学者占主导地位。Vakkari 等人（2022a）使用同一数据集发现，图书情报学领域的占比为三分之一，计算机科学领域贡献了五分之一，商业／经济学领域则贡献了六分之一。后两者在信息检索和科学传播领域的贡献中占主导地位，这表明它们对图书情报学有着强大的影响。

2.4 图书情报学中的引文影响力

非 LIS作者因其对主题和方法的偏好（Vakkari 等人 2022a、b）而对LIS 的发展产生了相当大的影响，因此，我们按主题和作者学科评估了研究影响力，并考察了引文计数。科学出版物的引文计数被广泛应用于 Scientometrics 中，作为衡量有影响力出版物的快速指标。然而，一些研究人员认为，使用与引文相关的指标来评估研究影响力并不恰当（MacRoberts & MacRoberts，2018年）。

然而，若干研究（Jirschitzka 等人，2017年；So，1998年）表明，引文计数与专家评判之间存在正相关。

无论哪些因素促成了引文数量的增加，高被引文章都会受到多个领域的关注。然而，尚无研究比较图书情报学相关主题在科学影响力方面的差异。在图书馆与信息科学研究中，研究人员已经对高被引文章及其特征（例如主题）进行了探讨。Blessinger 和 Hrycaj（2010）确定了32篇高被引文章，这些文章是1994年至2004年间发表于10种图书情报学期刊的论文引用次数最多的文献。在将这些文章中的主题索引词划分为七类后，他们发现，高被引文章主要聚焦于图书馆学和用户（68%），其次是技术（22%）。Sahoo等人（2020）分析了四种图书情报学期刊中166篇高被引文章的主题。这些主题通过内容分析确定。在所识别出的14个主题中，研究影响力测量和研究合作是最常见的主题（26%），并且影响力最大（占引文计数总计的29%）。鉴于以往研究主要关注图书情报学出版物中包含的特定特征，本研究旨在确定多个特征之间的联系，例如研究主题、研究策略、贡献类型和作者所属学科。

2.5 引文影响力指标

基于出版物的影响力指标与本研究相关。 解读引文计数时需要谨慎，因为其分布通常呈偏态，而且数据集涵盖的学科范围可能较大或较小。前者会导致离群值扭曲研究发现（Phelan，1999年），后者则会影响出版物预期获得的引文数量（例如，Waltman，2016年）。为缓解这一问题，研究者提出了各种按领域加权或标准化的引文影响力指标（例如，Waltman，2016年）。引文计数也可以按出版物类型和年份进行标准化。此外，有人可能会认为，高影响力出版物的引文比低影响力出版物的引文权重更大（例如，Waltman，2016年）。引文所累积的认可度也可以根据引用时间进行折减。为此，折减累积影响力指数可降低旧引文的价值（Järvelin & Persson，2008年；Ahlgren & Järvelin，2010年）。我们的数据涵盖单一领域——图书情报学，以及该领域顶级期刊在单一年份（2015年）发表的学术期刊文章。因此，我们在影响力分析中采用对数平滑处理离群值，并基于子领域进行领域标准化。

3 方法

3.1 研究问题

主要研究问题是：图书情报学研究论文的哪些内容因素与其获得的影响力得分相关？我们通过以下子问题进一步展开分析：

- 研究问题1：作者的学科构成对影响力得分有何影响？
- RQ2：学科构成与其他各项内容因素的综合效应如何？这些因素包括：(a) 研究主题；(b) 研究策略；以及 (c) 贡献类型。它们对影响力得分有何影响？

图2展示了研究设计。左侧的云朵代表研究工作，并详细呈现其四个理论特征。研究人员在某一学科背景下开展工作，将其作为某项研究计划的一部分；他们依据某种策略设定项目层面的目标，并在其中运用一些方法。将项目成果以文章形式报告时，这些特征会体现为文本内容。我们将其中部分特征操作化为内容因素：学科构成、主题、策略和贡献类型。随后，我们分析这些因素对文章影响力得分的主要效应和联合作用。

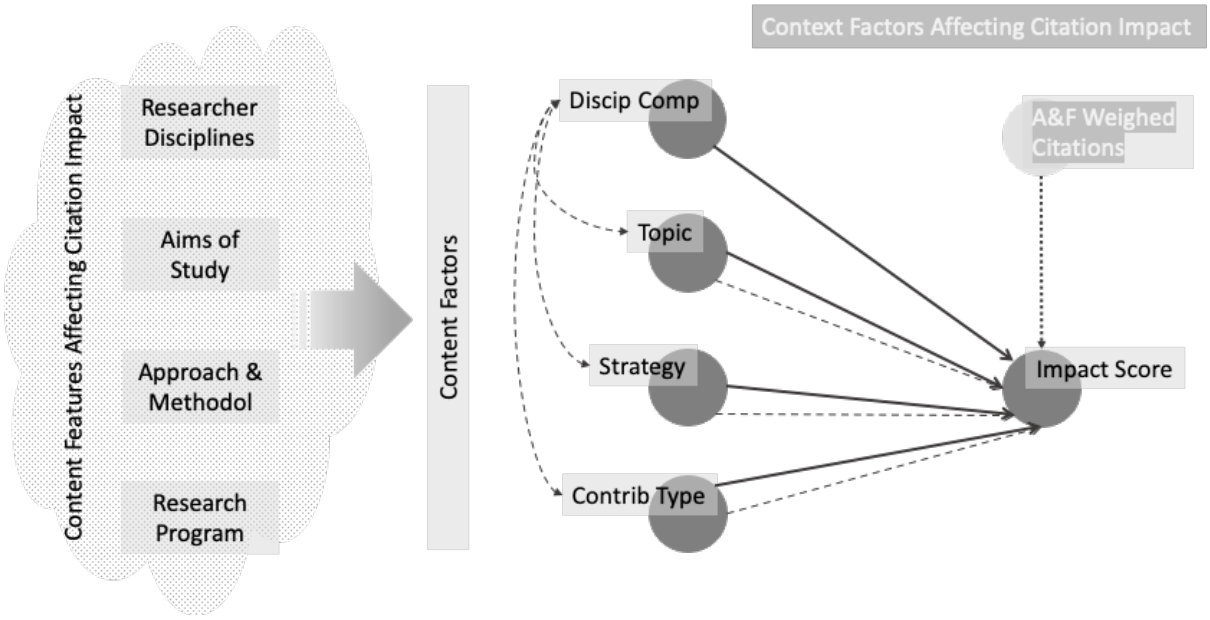


图2。 研究设计。实线箭头表示主效应，虚线箭头表示组合效应，点线箭头表示常数。

3.2 数据收集

研究过程如图3所示，数据集的各个维度概述于表1。该数据集由以下部分组成：

- 第一阶段收集了2015年发表于学术图书情报学期刊的论文的定量智识内容分析数据（Järvelin & Vakkari, 2022），
- 第二阶段收集了作者团队学科构成的数据（Vakkari & Chang & Järvelin, 2022a、b），而
- 第三阶段从 Scopus 获取了2015年至2021年5月期间图书情报学论文集的 ~25 K 条引文数据，本论文报告了这些数据。

对图书情报学文章集的内容分析聚焦于研究主题、方法和贡献类型。根据作者的隶属机构，将文章作者分配到八个学科类别中的一个，从而形成每个作者团队的学科构成。最后，追踪了这些论文从2015年至2021年5月所获得的引文。施引文献权威性通过这些施引文献截至2021年5月所获得的引文数量进行测量。接下来，我们将更详细地描述这些数据。

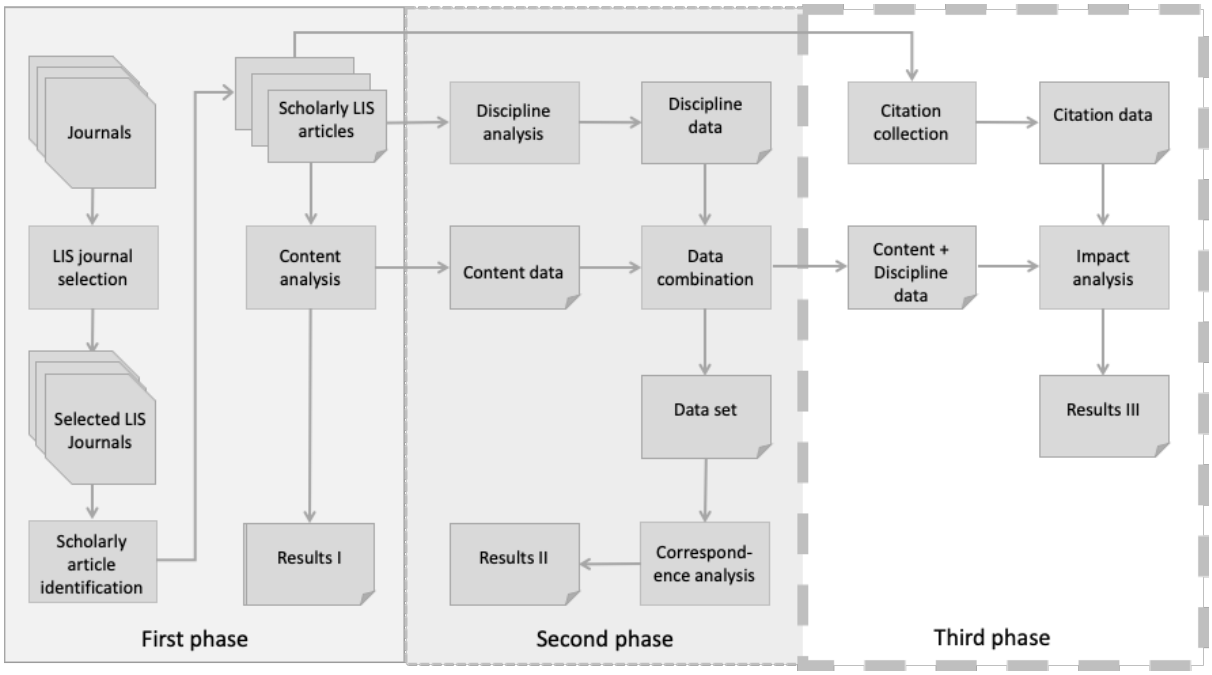


图3。数据收集、准备与分析过程（基于 Vakkari & Chang & Järvelin, 2022a）

对象／属性	**值**

期刊	
卷	2015
观测单位	一本期刊
标题总数	31
论文	
观测单位	一篇文章
总数	1514
- 不包括非图书情报学领域的研究	1322
- 排除非研究类文章	1210
内容维度	3
分类器，等额份额	2
学科构成	
观测单位	（文章，学科）对
总数	1533
内容维度	3
分类器	1
引文	
观测单位	这一对（被引文献，施引文献）
总数	24965
内容维度	1
分类器	-

表 1。 数据维度

3.3 内容分析

表 2 介绍了本研究使用的内容分析变量。附录 II 通过列出各类别对这些变量进行了定义。研究主题 的分类采用四个主要研究主题进行分析： 图书馆与信息背景及服务、信息检索（信息检索）、信息寻求和科学传播。研究主题类别非图书情报学研究也进行了编码，但未纳入分析。

内容分析变量（参见 Järvelin 和 Vakkari，2022年）	
Name	**说明**
图书情报学 主题	文章的焦点，例如信息寻求，可表述为 一个主要主题

研究策略	表示数据收集与分析的整体组合方式 研究方法
贡献类型	表示实证研究、理论、方法论、建构性、 等。研究的贡献类型
学科分析变量	
Name	**说明**
学科	根据文章合著者的所属机构，为每个独特的学科名称提供信息。
学科数量	表示为某篇 文章

表2。 数据集的内容与学科分析变量

为增加分析的自由度，我们合并了部分变量的类别。在研究策略中，历史研究和评估策略与其他实证研究策略合并为其他实证研究策略；引文分析与其他文献计量策略合并为引文分析；言语论证和概念分析合并为概念策略；文献综述和书目策略与其他策略合并为其他策略（附录 II）。

每篇文章在每个内容变量下均被归入一个内容类别。分类可靠性采用 Fleiss' 卡帕系数进行测量（表3）。主题，和主题的分类一致性良好，而 研究策略 和 贡献类型的一致性中等。

内容分析变量（*N* = 32）				
Name	**卡帕系数**	***p*-值**	**评审者数量级**	
图书情报学主要主题	0.73	0.000	2	good
图书情报学主题	0.62	0.000	2	good
研究策略	0.53	0.000	2	中等程度的
贡献类型	0.60	0.000	2	中等的
学科分析变量 (*N* = 40)				
Name	**卡帕系数**	***p*-值**	**评审者数量级**	
学科	0.71	0.000	3	good
学科数量	0.64	0.000	3	good

表3。 分类可靠性（Fleiss 卡帕系数）

3.4 作者学科的编码

采用 Chang（2018年）的方法确定每位作者的学科归属。该方法在本数据集中的应用详见 Vakkari 等人（2022a）。主要内容如下：

（1）与图书情报学相关机构的附属关系被归类为图书情报学附属关系。（2）其他作者则按以下学科归属进行分类：商业与经济学、计算机科学、工程学、人文学科、医学、自然科学和社会科学（见附录三）。（3）当文章提供的信息不完整时，本研究采用了各种参考来源和互联网来确定部分作者的学科归属。

此后，每篇文章均按其所属的各个不同学科归属进行分类。作者中代表同一学科的人数及其排序均不予考虑。因此，在剔除非学术文章和非图书情报学文章后，1513篇文章中共有1210篇来源文章。其中一些主要来源文章缺少引文数据，最终得到1181篇文章。为避免合并学科时损失自由度，分析仅限于涉及少于三个学科的文章（n=1145）。这些文章占全部文章的97%。此外，因引文计数属于离群值，另有三篇文章被剔除。

3.5 收集施引文献权威性的引文数据

为评估一篇图书情报学文章的影响力，我们计算了其影响力得分。该得分依据文章获得的加权引文数量计算。每条引文都是影响力的基本表征；影响力越大，引文越多，反之亦然（见图1）。此外，施引二级文献的权威性越高，主要图书情报学文章的影响力就越大。我们根据Scopus数据库记录的二级文献自身在2015年至2021年5月期间获得的引文数量，计算其施引权威性。引文数量可近似反映影响力，但不考虑三级施引出版物的质量、二级或三级文献的出版年份，亦不考虑引用原因的差异。不过，所有施引文献的条件都相同，其中最重要的是主要文章的出版时间窗口（2015年）以及二级和三级文献的引文时间窗口（2015年至2021年5月）。因此，可以可靠地分析图2所示因素对图书情报学文章引文影响力的贡献。

然而，以原始引文计数来判定文献的权威性存在问题，见图4。二级文献获得的引文数量范围为[0, 1397]，且分布极度偏斜。近25,000篇二级文献中，近7,000篇没有引文，三分之二的文献引文数不超过6。我们通过三种方式对引文计数进行标准化：首先取对数以平滑数据；其次按主要文章主题进行分析，以控制领域规模效应；

第三，将特定主题的得分与平均得分进行比较，计算相对得分，并将其作为主要文章的引文影响力。

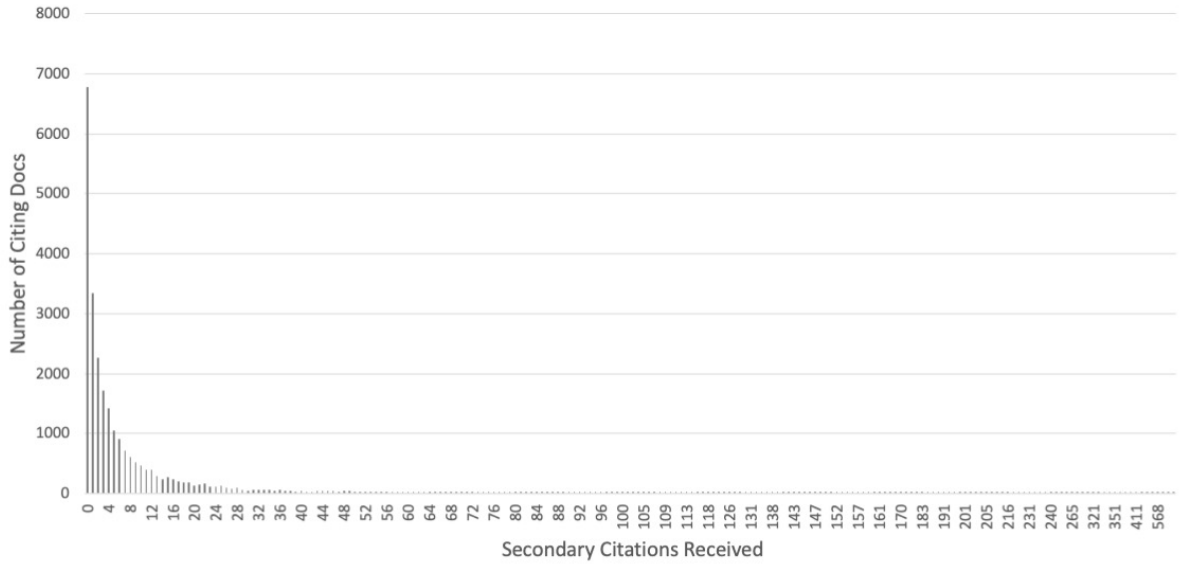


图4. 权威性判定：按所获引文数量统计的引用二级文献分布。x 轴截断于 568 次引文。

通过取对数进行平滑处理，旨在减弱离群值的影响。我们尝试了几个不同的对数底数 b（1.5、2 和 10），发现$b = 2$ 是兼顾保留差异与有效平滑离群值的折中选择。设cc_{sd}表示二级文献sd.的引文计数。二级文献sd的平滑引文计数scc计算如下：

$$scc_{sd} = \log_2(cc_{sd} + 1)$$

由于 $\log_b 0$ 未定义，因此需要“+1”。平滑后的权威性范围变为 [0,10.45]。给定一篇主要文献 pd 和一组引用它的二级文献 $SD_{pd} = \{sd_{1,pd}, sd_{2,pd}, ..., sd_{n,pd}\}$，$pd$ 的权威性加权引文影响力为：

$$AWCI_{pd} = \sum_{sd \in SD} scc_{sd} .$$

$AWCI_{pd}$ 统计所有引用文章sd 的二级文献，并将其计入pd平滑权威值pd之和；如果sd没有引文（$cc_{sd} = 0$），该值可能为零。然而，$AWCI_{pd}$ 不提供领域归一化。给定一组代表某个领域的主要文章$PD = \{pd_1, ..., pd_n\}$，我们使用平均平滑权威值：

$$avg\text{-}AWCI_{PD} = \sum_{pd \in PD} AWCI_{pd} / |PD|$$

用于赋予 pd 权威性与 领域归一化影响力得分（$AFNIS_{pd}$），其计算函数如下：

$$impactscore(pd, PD) = AWCI_{pd} / avg\text{-}AWCI_{PD}.$$

归一化影响分数的取值范围为 [0, z]，其中 z 为实数。分数的平均值为 1，而 z 通常 >> 1。表4列出了图书情报学及其子领域中各类引文计数的统计数据。

	图书馆与信息科学 子领域				
Name	**图书馆与信息科学语境**	**IR**	**ISeek**	**Scim**	**总计**
平均值 引文 cc_{pd}	9.78	13.34	16.77	20.37	15.63
最大值 引文 cc_{pd}	320.00	106.00	184.00	463.00	463.00
$avg\text{-}AWCI_{PD}$	15.98	26.89	30.23	43.34	30.95
最大 $AWCI_{PD}$	728.44	230.40	296.08	873.00	873.00
最大 $AFNIS_{pd}$	45.57	8.57	9.79	20.14	28.20
N	291	273	166	451	1181

表4. 标准化对图书馆与信息科学子领域影响力得分的影响。

3.6 数据分析

通过将作者的学科编码与引文数据和内容分析数据结合起来，构建了用于分析的最终数据矩阵（图3）。

4 研究发现

4.1 学科构成与影响力得分

学科构成与影响力得分显著相关（df=10，F=3.1，p<.001）（表5）。图书情报学作者与计算机科学家合作（1.44），以及计算机科学领域的作者与社会科学家合作（1.46），其论文获得了最高的影响力得分。影响力得分最低的是图书情报学学者与其他学科的学者共同撰写的论文（0.31）、自然科学与医学领域学者撰写的论文（0.56），以及人文学科领域学者撰写的论文（0.54）。事后分析（Tukey HSD检验）表明，自然科学与医学领域作者撰写的论文，其影响力得分显著低于与图书情报学和计算机科学领域合作，以及与计算机科学和社会科学领域合作所产出的论文。

学科构成	**Mean**	**标准差**	**N**
LIS	0.88	1.24	397
图书情报学与社会科学	0.94	0.98	54
图书情报学与计算机科学	1.44	1.65	47
图书情报学与其他学科	0.31	0.34	15
计算机科学	1.07	1.36	232
计算机科学与社会科学	1.46	1.90	43
计算机科学及其他学科	0.98	1.56	24
社会科学	0.90	1.07	204
社会科学与其他学科	0.98	1.35	25
自然科学与医学	0.56	1.20	70
人文学科	0.54	0.87	31
总计	**0.94**	**1.28**	**1142**

[图例：图书情报学 = 图书情报学；社会科学 = 社会科学、商业与经济学；计算机科学 = 计算机科学、工程学；其他学科 = 其他学科] **表 5**。 不同学科构成的影响力得分（研究问题1）

图书情报学与计算机科学的合作，与图书情报学的其他合作组合相比，产生了更高的影响力得分。此外，计算机科学与社会科学的联合研究，与许多其他学科组合相比，取得了影响力较高的研究成果。因此，计算机科学是图书情报学领域中一个具有影响力的参与者。不过，不同学科的影响力可能因图书情报学的子领域而异。接下来，

我们分析不同学科构成之间以及图书情报学的主要主题之间的影响力得分差异程度。

4.2 学科、主题和影响力得分

学科构成和主题的双因素方差分析具有显著性（df=42，F=1.9，p<.001）。主题（df=3，F=1.6，p=.18）和学科构成（df=10，F=1.1，p=.38）对影响力得分的主效应均不显著，而交互效应（df=29，F=1.7，p=.012）显著。因此，在不同的主要主题中，学科构成对影响力得分的影响不同。接下来，我们将对此进行更详细的分析。

在一些主要主题中，学科构成与影响力得分之间存在显著关联。在信息检索领域，不同学科群体之间的影响力得分没有显著差异（df=10，F=1.6，p=.10）。然而，详细的独立样本t检验显示，仅由图书情报学学者撰写的关于信息检索的文章，其影响力得分显著低于由计算机科学（计算机科学）代表撰写的文章（df=142.1，t=2.9，p=.004），以及由图书情报学与计算机科学学者共同创作的文章（df=63，t=2.4，p=.021）。与图书情报学和计算机科学学者合作撰写的文章（df=21，t=3.0，p=.008）以及计算机科学与社会科学（社会科学）学者合作撰写的文章相比，人文学科（人文学科）的影响力得分显著更低（df=25.4，t=2.5，p=.021）。

在信息寻求领域，不同学科组合的影响力得分存在显著差异（df=9，F=4.8，p=.017）。事后分析（Tukey HSD检验）表明，仅由图书情报学学者撰写的论文，其影响力得分显著低于计算机科学与社会科学学者合作撰写的文章。

在科学传播领域，不同学科组合的影响力得分存在显著差异（df=10，F=2.8，p=.002）。独立样本t检验显示，自然科学与医学学者（自然科学与医学）撰写的论文，其影响力得分显著低于图书情报学（df=99.3，t=3.5，p<.001）、图书情报学与社会科学合作（df=21.9，t=2.6，p=.018）、图书情报学与计算机科学合作（df=21.7，t=2.8，p=.01）、计算机科学（df=73.4，t=3.1，p=.003）以及社会科学（df=178.1，t=4.3，p<.001）。图书情报学与计算机科学学者合作撰写的文章，其得分显著高于社会科学家撰写的论文（df=25.9，t=2.2，p=.035），也显著高于计算机科学及其他学科（其他学科）学者撰写的论文（df=23.5，t=2.3，p=.028）。

主题内显著的学科差异汇总于表6中。

主题	显著差异
信息检索	• 图书情报学 < 计算机科学、图书情报学与计算机科学 • 人文学科 < 图书情报学与计算机科学、计算机科学与社会科学
信息寻求	• 图书情报学 < 计算机科学与社会科学
科学的 沟通	• 图书情报学与计算机科学 > 社会科学、计算机科学与其他学科 • 自然科学与医学 < 图书情报学、图书情报学与社会科学、图书情报学与计算机科学、计算机科学，社会科学

表6。 各主题内不同学科构成之间影响力得分的显著差异（RQ2a）

4.3 学科、研究策略与影响力得分

学科构成和研究策略的双因素方差分析具有显著性（df=91，F=1.6，p<.001）。学科构成的主效应（df=10，F=1.6，p=.10）和研究策略的主效应（df=9，F=1.8，p=.07）均不显著，而二者的交互效应显著（df=72，F=1.4，p=.012）。

在某些研究策略中，学科构成与影响力得分之间存在显著关联。在调查策略中，学科组合对影响力得分的差异具有显著影响（df=10，F=2.5，p=.008）。事后分析（Tukey HSD检验）表明，计算机科学与社会科学的合作文章所产生的影响力显著高于图书情报学、图书情报学与社会科学、计算机科学、社会科学以及自然科学与医学。

引文分析策略（df=10，F=3.0，p=.09）并未显著影响不同学科群体的影响力得分。然而，事后分析（Tukey HSD检验）表明，在引文分析策略中，由图书情报学与计算机科学合作撰写的论文，其影响力得分显著高于自然科学与医学。

在概念研究策略中，学科构成显著影响了影响力得分（df=6，F=6.2，p<.001）。事后分析（TukeyHSD检验）显示，自然科学与医学所产生的影响力得分显著高于图书情报学、图书情报学与社会科学、计算机科学、社会科学以及人文学科。除此之外，由计算机科学和社会科学合作署名的论文，其影响力得分也显著高于仅由图书情报学学者单独署名的论文。

Sig表格总结了研究策略内显著的学科差异。 7.

研究策略	显著差异
调查研究	• 计算机科学与社会科学 > 图书情报学，图书情报学与社会科学，计算机科学，社会科学，自然科学与医学
引文分析	• 图书情报学与计算机科学 > 自然科学与医学
概念性的	• 自然科学与医学 > 图书情报学、图书情报学与社会科学、计算机科学、社会科学> 人文学科 • 图书情报学 < 计算机科学 & 社会科学

表7。 研究策略内不同学科构成之间影响力得分的显著差异（RQ2b）

我们已经表明，在主题内部以及研究策略内部，不同学科群体之间的影响力得分存在一些差异。可能是因为这些学科群体更擅长运用这些策略来解决特定主题中的研究问题，从而产生更高的影响力得分。我们的数据量不足以直接阐述这一假设。然而，如果不同研究策略下的主题影响力得分不存在差异，这表明不存在所谓能够解决这些主题中的研究问题的优越策略。结合我们此前的结果，这将间接表明，对于某些学科而言，某些策略能够产生影响力得分更高的结果。为检验这一假设，我们对主题和研究策略的影响力得分进行了双因素方差分析，结果不显著（df=39、F=1.1、p=.28）。这佐证了我们的推测：一些学科群体在某些主题中采用特定的研究策略，其方式能够产生较高的影响力得分。

4.4 学科、贡献类型和影响力得分

学科构成与贡献类型的双因素方差分析结果显著（df=69，F=2.6，p<.001）。学科构成（df=10，F=2.5，p=.006）和贡献类型（df=7，F=4.9，p<.001）的主效应均显著，二者的交互效应也显著（df=52，F=2.3，p<.001）。事后分析（Tukey HSD检验）显示，与描述性、理论性和方法论贡献相比，解释性贡献获得的影响力得分显著更高。

在某些贡献类型中，学科构成与影响力得分之间存在显著关联。在描述性研究类型中，学科组合显著影响影响力得分（df=10，F=2.6，p=.002）。事后分析（Tukey HSD检验）表明，与图书情报学、自然科学与医学以及人文学科的文章相比，图书情报学与计算机科学合作的文章获得的影响力得分显著更高。

学科构成显著影响了方法论贡献的影响力得分（df=9，F=2.1，p=.042）。与自然科学与医学相比，图书情报学在产出方法论贡献方面获得的影响力得分显著更高（df=17.2，t=2.2，p=.044）。

表8总结了各贡献类型中影响力得分的显著差异。

贡献类型	显著差异
描述性	• 图书情报学与计算机科学 > 图书情报学、自然科学与医学、人文学科
方法论方面的	• 图书情报学 > 自然科学与医学

表8。不同贡献类型内学科构成之间影响力得分的显著差异（RQ2c)

5 讨论

5.1 主要研究发现

我们通过2015年一批领先的图书情报学期刊论文的内容因素，分析了研究论文的引文影响力。我们重点关注论文作者的学科、主题、方法论以及所声称的贡献类型。这些因素在研究设计、执行、报告和评估中发挥着核心作用。我们考察了作者的学科构成与主要主题、研究策略及贡献类型之间的关联在多大程度上与研究的影响力相关。我们将影响力操作化为每篇文章的影响力得分，并以该文章获得的引文数量及其权威性为依据。

计算机科学分别与图书情报学或社会科学合作时，其影响力得分显著高于自然科学与医学。这些组合产生了最高的影响力得分，而人文学科、自然科学与医学以及图书情报学在

与其他学科的合作产生了最低的影响力得分。这些研究发现揭示了计算机科学在图书情报学中对于创造具有影响力的研究成果所发挥的关键作用。这也表明，图书情报学学者与计算机科学家合作开展具有影响力的研究是有益的。

不同学科构成之间的影响力得分差异通常出现在信息检索和科学传播主题中。在这两个主题中，图书情报学与计算机科学的联合学科构成所获得的影响力得分均显著高于许多其他构成，例如，在信息检索中，图书情报学单独构成或人文学科构成；或者在科学传播中，自然科学与医学构成或社会科学单独构成。在信息检索中，计算机科学与社会科学合作产生了较高的影响力得分；在信息寻求中，其影响力得分也显著高于图书情报学单独构成。

计算机科学与社会科学的结合，在采用调查研究研究策略方面产生的影响力得分显著高于诸如图书情报学、单独的计算机科学或社会科学等许多学科。这可能与这一学科组合在信息寻求领域产生的较高影响力得分有关。Järvelin和Vakkari（2022年）表明，调查研究是信息寻求中最受欢迎的研究策略。因此，计算机科学与社会科学通过熟练运用调查研究来共同解决信息寻求问题，能够取得富有成效的成果。

反映出科学传播中更高的影响力，图书情报学与计算机科学相结合，在使用引文分析研究策略方面产生了显著更高的影响力得分。

在描述性贡献类型中，图书情报学与计算机科学结合的研究获得了最高的影响力，而在方法论类型中，单独的图书情报学最为有效。

5.2 解读关联关系

主题、研究策略和贡献类型之间存在一些学科差异模式。在信息检索领域，与其他一些学科构成相比，图书情报学与计算机科学的组合以及单独的计算机科学获得了显著较高的影响力得分，而人文学科的得分则显著较低。2015年，数学策略、实验策略和系统分析策略是信息检索领域最常见的三种策略（Järvelin & Vakkari 2022年）。我们的结果表明，在数学策略中，计算机科学以及与计算机科学合作的图书情报学获得的影响力得分略高于其他学科构成。在系统分析策略中，

与单独的图书情报学相比，计算机科学以及与计算机科学合作的图书情报学获得了明显更高的影响力得分。

在信息检索领域，通常会通过实验比较系统性能。因此，比较贡献类型和系统分析贡献类型较为常见。在比较类型中，图书情报学、计算机科学以及图书情报学与社会科学共同构成的学科构成，相较于其他学科构成得分明显更高；而在系统分析类型中，图书情报学与计算机科学合作的得分明显高于仅包含图书情报学的情况。

因此，在信息检索领域，图书情报学与计算机科学携手合作，有利于提升研究影响力。这种影响力可能与图书情报学和计算机科学共同署名的论文中采用数学和系统分析策略有关，因为与包括单独的图书情报学在内的其他一些学科构成相比，这些论文获得的影响力得分略高。这表明，与计算机科学家的合作为图书情报学带来了信息检索领域所需的有益方法知识。

在信息寻求领域，与社会科学合作的计算机科学，其影响力得分显著高于图书情报学。根据 Järvelin和Vakkari（2022年）的研究，调查研究和定性研究策略是最受欢迎的策略。在采用调查研究时，与社会科学合作的计算机科学所产生的影响力得分显著高于图书情报学、图书情报学与社会科学的合作，以及其他几种学科构成。这表明，在这一子领域，对于计算机科学家和社会科学家开展合作而言，调查研究是一种有效的方法论选择。

在科学传播中，图书情报学与计算机科学的学科构成所获得的影响力得分明显高于其他几种学科构成。另一方面，自然科学与医学所获得的得分低于包括图书情报学在内的其他几种学科构成。Järvelin 和 Vakkari（2022年）表明，在科学传播中，引文分析、调查研究和案例研究策略是最常见的策略。我们的研究发现表明，在采用引文分析策略的情况下，图书情报学与计算机科学所获得的影响力得分显著高于自然科学与医学。因此，科学传播同样受益于与计算机科学开展合作，从而产生具有较高影响力的成果。这似乎与图书情报学和计算机科学合作文章所采用的引文分析策略有关。与许多学科构成所对应的其他文章相比，这些文章获得了更高的影响力得分。

总之，我们想 выяснить文章内容因素——学科构成、研究主题、研究策略和贡献类型——是否与文章的引文影响力相关。我们发现了上文所讨论的显著相关性，但这种相关性较少且较弱。

比我们预期的更……这表明，只要执行或产出得当，所有策略和贡献类型都值得认可。策略和贡献的类型并不是解释引文影响力差异的万能钥匙。其他内容因素也有发挥作用的空间，例如研究问题和研究设计的新颖性、数据集的质量以及方法论的质量，都可以用来解释引文影响力的方差。此外，其他背景因素也有发挥作用的空间，例如作者、作者所在机构以及期刊在学术共同体中的地位，都可以用来解释这种方差。其中一些数据相当容易收集，而另一些则非常费力——尤其是评估与新颖性相关的因素，向来尤其棘手。

5.3 局限性

用于观察引用2015年文章的论文所获得引文数量的引文窗口有限。因此，影响力数据反映的是研究界对这些文章的早期反响。此外，引文数据缺乏细节：“文献 x 引用文章 y”并未说明 x 中的哪些内容类别引用了 y 中的哪些内容，也未说明该引文是否合理或重要，或其立场如何。因此，研究发现不如预期的那样精确，其解释力也有所降低。第三，我们不知道作者所属的哪个学科应对某一引文负责或受到认可。第四，尽管数据集提供了主要文章的学科构成，却没有施引文献的学科数据，因此无法探讨一些有趣的研究问题。最后，尽管数据集中的主要文章发表于期刊的2015年卷期，但它们可能更早在线发表，也可能直到2015年之后才上线。这种情况可能因期刊而异，并影响实际的引文窗口大小。

5.4 后续研究

基于当前数据集或对其进行扩展，可以开展一些有趣的后续研究。一种是对选定的图书情报学主题中少数低被引文章与高被引文章的研究设计创新性进行定性研究。另一种是对引用文献的二级和三级文献所属学科进行分类，以便分析图书情报学研究中的思想流动。当然，也可以将分析视角转向参考文献以及涌入图书情报学领域的思想。最后，可以为1995年和2005年的相应数据集补充参考文献／引文数据，以推动对思想流动的纵向分析。还可以考察以下背景因素：作者名称、期刊名称、机构名称，以及

国家名称等与引文影响力的关系比本文分析的内容因素更强。关于方法论发展，可将现有数据集用作真实基准，用于开发数据挖掘/机器学习方法，以自动分类内容维度和识别贡献（Goh 等，2020年）。

6 结论

我们分析了2015年发表在31种国际领先图书情报学期刊上的学术论文的引文影响力。

总体而言，我们发现作者的学科构成与引文影响力得分显著相关。不同学科构成

本研究的主要局限——引文数据未经限定、无法确定哪些学科带来了哪些引文，

参考文献

Ahlgren, P.&Järvelin, K.(2010)。使用 DCI 指数衡量 12位信息科学家的影响力。美国信息科学与技术学会期刊（*JASIST*） 61(7), 1424-1439。DOI: 10.1002/asi.21333

Åström, F.（2007年）。图书馆与信息科学研究前沿的变化：对图书馆与信息科学期刊文章进行分时段共被引分析，1990年–2004年。*JASIST*，58(7)，947–957。

Blessinger, K., & Frasier, M. (2007). 图书馆文献十年分析：1994年-2004年。《大学与研究图书馆》，68，55–169。Blessinger, K., & Hrycaj, P. (2010). 图书情报学领域的高被引文章：内容与作者署名趋势分析。《图书馆与信息科学研究》，322)，156–162。https://doi.org/10.1016/j.lisr.2009.12.007Chang, Y-W. (2018). 基于非图书情报学作者撰写的图书情报学文章，考察图书情报学（LIS）的跨学科性。*Scientometrics*，116 (3)，1589– 1613。Chang, Y-W., Huang, M-H & Lin, C-W. (2015). 基于关键词、书目耦合和共被引分析的图书情报学研究主题演变。*Scientometrics*，105(3)，2071–2087。Chang, Y. W. (2019). 图书情报学（LIS）期刊中的文章是否主要由图书情报学作者撰写？*Scientometrics*，121(1)，81–104。Figuerola, C., Marco, F. & Pinto. M. (2017). 运用 LISA 主题建模绘制图书情报学的发展脉络。*Scientometrics*，112(12)，1507–1535。Goh, Y.C., Cai, X.Q., Theseira, W. 等 (2020). 评估人类与机器学习在研究摘要分类方面的表现。*Scientometrics* 125(2)，1197–1212。DOI:10.1007/s11192-020-03614-2Han, X. (2020). 1996年至2019年间图书情报学研究主题的演变：基于狄利克雷分配主题模型的分析。*Scientometrics*，125，2561–2595。Hider, P. and Pymm, B. (2008). 高影响力图书情报学期刊文献中报告的实证研究方法。*LISR*，30，108–114。Hou, J., Yang, X. and Chen, C. (2018). 信息科学的新兴趋势与新发展：文献共被引分析（2009年-2016年）。*Scientometrics*，115(1)，869–892。Huang, M-H., Shaw, W-C., & Lin, C-S. (2019). 一个类别，两个社群：期刊引证报告中的信息科学与图书馆科学子领域差异。*Scientometrics*，119(2)，1059–1079。Järvelin, K. & Persson, O. (2008). DCI指数：基于折扣累积影响力的研究评估。《美国信息科学与技术学会会刊》（*JASIST*）59(9)，1433-1440。勘误：JASIST 59(14)：2350-2352。Järvelin, K. & Vakkari, P. (1990). 图书馆与信息科学研究论文的内容分析。《图书馆与信息科学研究》，12，395–421。

Järvelin, K. & Vakkari, P.（1993）。《图书馆与信息科学的发展，1965–1985：期刊文章的内容分析》。Information Processing & Management，29(1)，129–144。Järvelin, K. & Vakkari（2022）。《图书情报学研究纵览50年：期刊文章内容分析》。Journal of Documentation，78(7)，65-88。https://doi.org/10.1108/JD-03-2021-0062。Jirschitzka, J., Oeberst, A., Göllner, R., & Cress, U.（2017），《跨学科领域中同行评审的评审者间信度与效度》。Scientometrics，113(2)，1059–1092。Latour, B.（1987）。《科学在行动：如何在社会中追踪科学家与工程师》，Open University Press，米尔顿凯恩斯。Liu, G., & Yang, L.（2019）。《图书情报学近期期刊出版物中的热门研究主题》。The

（2021）。《1980年至2010年五种面向从业者的顶级图书情报学期刊所采用的研究方法分析》。Journal of Documentation，77(5)，1196–1208。Luukkonen, T.（1997）。《为什么文献计量学界忽视了拉图尔的引文理论？对引文分析的社会学解释的探讨》。Scientometrics，38(1)，27–37。https://doi-org.libproxy.tuni.fi/10.1007/BF02461121 Ma, J., & Lund, B.（2021）。《图书情报学研究主题与方法的演变和转变》。JASIST，72(8)，1059–1074。MacRoberts, M.H. & MacRoberts, B.R.（2018）。《科学的错误测量：引文分析》。

Khasseh, A. A.（2021）。《图书情报学研究焦点与热点的26年（1990-2016年）：共词分析》。Journal of Information Science，47(6)，794–808。Phelan, T.J.《引文分析的问题化编》。Scientometrics，45(1)，117–136（1999）。https://doi.org/10.1007/BF02458472。Sahoo, J., Mohanty, B., Biswal, O., Dash, N. K., & Sahu, J.K.（2020）。《作者署名趋势与内容分析：图书情报学期刊高被引文章案例研究》。Performance Measurement and Metrics，21(1)，33–51。https://doi.org/10.1108/PMM-06-2019-0021 So, C. Y. K.（1998）。《引文排名与专家评判在评价传播学学者时的比较：研究专长规模与个人知名度的影响》。Scientometrics，41(3)，325–333。

Tahamtan, I. & Bornmann, L.（2019）。引文计数测量的是什么？对2006年至2018年发表的科学文献引文研究的最新综述。*Scientometrics*，121(1)，1635-–1684。Tuomaala, O., Järvelin, K., & Vakkari, P.（2014）。图书馆与信息科学的发展，1965-2005：期刊文章的内容分析。*JASIST*，65(7)，1446–1462。Urbano, C., & Ardanuy, J.（2020）。图书馆与信息科学连续出版物中的跨学科合作与共存：对四个欧洲国家作者署名机构的分析。*Scientometrics*，124(1)，575–602。Vakkari, P.（1994）。图书情报学：内容与范围。Godden, I.P.（编），《图书馆学进展》，18。Academic Press，圣迭戈，加利福尼亚州，1–55。Vakkari, P. & Chang, Y-W. & Järvelin, K.（2022a）。图书情报学研究主题与方法论的学科贡献——是否导致碎片化？《信息科学与技术协会学报》 73(12)：4499-4522。https://doi.org/10.1002/asi.24690Vakkari, P. & Chang, Y-W. & Järvelin, K.（2022b）。以研究论文特征衡量外部学科对图书情报学的最大贡献。 *Scientometrics* 127(8)：4499-4522。 https://doi.org/10.1007/s11192-022-04452-0Waltman, L.（2016）。引文影响力指标文献综述，《信息计量学报》 10(2)：365-391，ISSN 1751-1577，https://doi.org/10.1016/j.joi.2016.02.007。Weinstock, M. 引文索引。载于：Kent, A. 等编。，《图书馆与信息科学百科全书》，第5卷。纽约，纽约州：Dekker，1971年，16-40页。Yan, E.（2015）。研究动态、影响力与传播：主题层面分析。*JASIST* 66(11)，2357-2372。

附录 I

期刊名称 2015年	Vols	文章数量
ACM 信息系统汇刊	33(1) -34(1)	27
Aslib 信息管理期刊（原名：Aslib Proc.）	67	36
《大学与研究图书馆》	76	57
《信息与文化》（原名：《图书馆与文化》）	50	24
《信息处理与管理》	51	65
《信息研究》	20	46
信息检索	18	21
《信息服务与利用》	35	27
信息技术与图书馆	34	19
国际信息与图书馆评论	47	10
国际信息管理期刊	35	71
Journal of Documentation	71	71
图书馆与信息科学教育期刊	56	27
Journal of Information Science	41	57
图书馆学与信息科学期刊	47	32
图书馆管理期刊	55	22
信息科学与技术协会期刊	66	196
图书馆与信息史（原名：图书馆史）	31	11
图书馆与信息科学研究	37	40
图书馆馆藏、采购与技术服务	39	11
图书馆季刊	85	24
图书馆资源与技术服务	59	15
《图书馆趋势》	63	47
Libri	65	24
信息网络新评论	20	27
《在线信息评论》	39	52
项目	49	24
《参考与用户服务季刊》（前身：《参考季刊》）	54(3)-55(2)	12
Scientometrics	102-105	345
《电子图书馆》	33	70
索引编制者	33	30
总计		**1540**

附录 II – 内容分析分类

按主要主题划分的图书情报学主题

图书馆与信息背景及服务研究

010 专业职业 020 图书馆史、图书馆与信息机构史 030 出版

100 图书情报学教育 200 方法论 300 图书情报学学科分析

410 文献传递 420 馆藏 430 信息或参考咨询服务 440 用户教育或信息素养教育 450 图书馆与信息服务建筑 460 管理或规划 470 自动化或数字图书馆 480 其他图书馆与信息服务 490 若干相互关联的活动

图书情报学的其他800个方面

信息存储与检索研究510 元数据 / 编目 520 分类与索引 531 文本检索 532 其他媒体中的检索方法 533 网络检索方法 534 社交媒体检索 540 数字信息资源 550 交互式（以用户为导向的信息检索） 560 信息检索的其他方面)

关于信息寻求的研究610 信息传播 620 渠道或信息来源的使用或用户 630 图书馆与信息服务的使用 641 基于任务的信息寻求 642 其他类型的信息寻求 650 信息利用 660 信息管理

科学传播研究 710 科学／专业出版 720 引文模式和结构 730 网络计量学 740 科学／专业传播的其他方面

研究策略

实证研究11 历史研究
12 调查研究 13 定性研究
14 评估 15 个案研究或行动研究

16 内容或协议分析 17 引文分析
18 其他文献计量分析 21 二次分
析 22 实验 29 其他实证研究策
略

概念性 31 言语论证 32 概
念分析

其他非实证研究 40 数学或逻辑
50 系统分析与设计 60 文献综述
80 书目 90 其他策略 00 不适用

研究类型

实证研究 11 描述
性 12 比较性 13
解释性

非实证研究 20 概念
性 30 理论性 40 方法
论 50 系统设计

其他贡献90 其他
类型 00 不适用

附录三 – 基于隶属关系的学科类别

主要类别与示例子类别（Chang 2018年）

商业与经济学商业 经
济学 管理学

计算机科学 计算机科学与工程学 信息
系统与人机交互

工程学 工程学
建筑学 能源

人文学科人文学
科 文学 艺术 人
类学

语言学 哲学与宗教 历
史

图书情报学（图书情报学） 文献学
信息科学 图书馆学

医学医学 护理
健康科学

自然科学综合科
学 物理学 数学
生物学 农业 化学
动物学 植物学

社会科学教育 一般社
会科学 传播学 法学 心
理学 社会学 政治学 旅
游学

其他任何不适用或未知的学科