

国外科学数据引用研究进展

张静蓓 吕俊生 田野

[摘要]: 文章总结目前数据引用所面临的引用意识、引用对象以及长期保存方面的三个关键问题。从数据引用规范的发展、文献管理软件的调研、数据引用利益相关者的责任与义务三个方面概括目前国外数据引用的研究进展,发现数据引用规范格式不兼容,文献管理软件缺乏对数据引用的支持,并且目前数据引用利益相关者责任与义务不明确。最后,针对关键问题提出对策建议以及今后研究的方向。

[关键词]: 数据引用 引用规范 文献管理工具 利益相关者

[分类号]: G254

1.引言

随着计算机辅助科研的大量应用,许多科研过程会产出海量的科学数据,科学数据是信息时代最基本、最活跃、影响面最宽的一种战略性资源,对于科技创新具有显著的支撑作用,从一定角度来说,未来的科研活动将是科学数据驱动的科研活动^[1]。同时,数据密集型科学研究范式的兴起,使得数据引用问题日益受到关注。科学数据引用(data citation)^[2]是指类似于研究人员通常为印本资源提供书目参考的方式提供数据参考的做法,通过一定的标识技术和机制,对所使用的科学数据资源进行描述,标识数据的来源,从而在一定程度上促进数据的知识产权保护,便于进行数据引用情况的统计和分析,数据引用对象可以是数据仓储、数据集或数据记录等。目前,国际社会对数据共享的需求日益强烈,一些地区性组织已经开展的关联数据(linked data)项目、政府数据开放(data.gov)运动等,也增加了人们对数据共享的意识。因此,数据引用的规范化制定显得更加紧迫。

数据引用日益增长的需求及重要性与现实引用实践的反差引起了众多国际组织、科研项目、数据仓储或数据中心以及学术期刊出版商等的关注。2011年起,众多国际组织如Data cite^[3]、IQSS^[4]、ANDS^[5]等纷纷开展以“Data Citation”为主题的研讨会与相关活动。2012年10月16日,全球领先的智能信息供应商汤森路透旗下的知识产权与科技事业部日前宣布推出Data Citation Index^[6](DCI,数据引文索引)。国外“数据引用”的理论与实践研究早于国内,并已经取得了一些重要进展,文章从数据引用关键问题、数据引用规范、文献管理工具、数据引用利益相关者、四个方面对国外目前数据引用研究现状进行梳理与总结,以便对数据引用领域国际上的最新研究前沿动态进行实时跟踪,及时了解国内研究的不足之处并加以补充与追赶。

2.数据引用关键问题

2.1 缺乏数据共享和引用意识

数据引用目前最大的挑战就是数据的生产者和使用者没有数据共享和引用的意识。没有数据的共享,就没有数据的引用。研究人员不愿意共享自己的数据的原因有很多,在《国外数据共享行为影响因素研究综述》^[7]中,作者总结了影响数据共享的三大因素,分别是①制度因素(数据引用政策与规范的缺乏);②技术因素(没有合适的可提交数据的数据仓储);③个人因素(得不到预期利益与承担共享风险)等。此外,即使数据被广泛的共享,比如政府部门数据,那么数据使用者也不一定用正确合理的方式引用它们。Parsons^[8]在2010年的一项研究中表明,在2002到2009年期间,引用了美国航天局(NASA)的遥感数据的作者中,只有很少一部分使用者提供了正式的引用方式。总体来说,数据使用者通常缺少一些数据引用的意识:①不知道是否应该引用该数据集;②不确定应该怎么去引用数据集;③缺乏

学术报偿体系 (Academic Reward System), 即使引用了数据, 但是没有相应的学术激励体系, 因此缺乏引用的动机。

2.2 引用对象不明确

目前, 还没有关于“数据集”统一的定义。数据集一般来说是“复合”型的, 每个数据集下面可能会包含若干个子数据集。数据的使用者在面对引用时, 不知道该引用完整的数据集, 抑或是该引用这个数据集的子集。所以, 正如 P.J.Hayes 等人^[9]所说, 数据引用本身是一个比较“模糊”的过程。此外, 数据集具有动态性、不确定性以及边界的模糊性, 这些都从一定程度上抑止了数据引用规范的向前发展。首先, 数据集往往是高度动态化的, 因此很难将数据集“语义化”。R.E.Duerr^[10]等人将这种问题总结成“科学唯一性”, 并指出现有的一些技术, 比如数字唯一标识符等, 并不能从根本上有效地解决此类问题。其次, 为了保证学术优先权, 将数据集嵌入作者身份也是同样比较困难的。以往的在期刊论文、会议报告中, 可以将作者信息加入到指定位置, 但是在科学数据中, 很难进行此类做法, 此外, 在数据集中, 署名本身也是比较有争议, 数据集通常是合作的产物, 同一个项目中的不同生产者可能对数据集的署名以及责任归属有着不同的认识。

2.3 长期保存

目前全球范围内都在进行数字资源的保存研究, 与传统学术出版不同, 数据集有其特殊性, 如何更好的进行数据出版, 这也是需要仔细考量的。并且目前的数据引用技术中, 很少涉及到数据的保存, 因此这些都是需要进一步协商和研究的。N.M.Weber^[11]等人 2010 年的一项研究发现, 大多数期刊没有明确的数据归档或引用政策, 他们认为期刊的数据归档或引用政策在细节上可能是各不相同的, 但保障正式出版和数据的持久性却是它们共同的要求。数据的长期保存不仅仅需要解决“去哪儿”的问题, 还需要解决“如何检索”的问题, 这就涉及到了为科学数据分配可持续性标识符问题, 目前唯一标识符分配最常见的唯一标识符是 DOI (Digital Object Unique Identifier)^[12], 此外, 诸如 Handle system^[13]、Uniform Resource Name^[14]等也常被使用。

3. 数据引用规范研究

在专门针对科学数据引用规范出现之前, 国际通用的学术论文写作规范中已经有部分针对数据引用的要求。例如人文社科领域三大格式 (The Big3): APA Manual(第六版)^[15]、MLA Manual(第三版)^[16]、Chicago Manual(第 16 版)^[17]中都有给出与“数据引用”有关的说明。但是这些格式指南文件都存在着一些普遍问题:(1) 由于并不是专门针对科学数据的引用规范, 这三大格式针对科学数据的引用对象分别为数据集、网络资源以及科学数据库。对引用对象的定义与划分缺少一致性, 不能相互兼容。(2) 由于科学数据的类型多样, 数据的引用缺乏系统的、统一的标准。(3) 三大格式都没有包含唯一标识符。

国外专门针对数据引用规范的研究自 2007 年开始, M.Altman^[18]等发表论文, 对数据引用包含的最小元素集、可选元素等进行了具体分析, 该文是最早探索数据引用标准的论文之一。此后, 众多致力于数据引用的国际组织和数据中心等在此基础上提出了自己的数据引用标准。自 2011 年起, 一些致力于数据引用实践的国际组织如 DataCite^[19]、OECD^[20]、DCC^[21]等纷纷发布了专门针对科学数据的引用规范指南文件。Data Cite^[22]作为一个专门为推进数据引用规范化而设立的机构, 是目前数据引用领域的研究主力。它的国际数据引用元数据工作组于 2011 年 6 月推出“DataCite Metadata Schema for the Publication and Citation of Research Data” Version 2.2, 已于 2013 年 7 月更新到 3.0 版本。该规范给出了一些核心的元数据属性以及使用规则建议, 以实现数据的准确及持久的识别, 从而方便数据引用及查询。它的引用对象是在线的科学研究数据 (Scientific Data on the Internet), 被引用资源可以是任何类型, 典型的是数据集。推荐引用格式为: Creator (PublicationYear): Title. Version. Publisher. ResourceType. Identifier。对这些国际组织的引用规范/指南文件的比较中, 发现这些文件都包

含以下三个主要组成部分:制定规范的目的及意义、科学数据/数据集的定义、具体的引用格式介绍。对比这些文件,会发现其制定逻辑具有高度的一致性,并且也与先前出现的一些通用的文献参考格式标准,如 APA、Chicago Mannul 等所定义的规则保持一致,并就数据这一日益重要的科学资源进一步细化与丰富。

此外,一些致力于数据引用规范制定的科研项目也发布了相应的数据引用最佳实践与规范。由 JISC 资助的 SageCite Project^[23],它旨在开发和测试一个将数据、方法与出版物关联的引用框架,并且提出数据引用的最佳做法:(1)数据引用所引用的对象必须具有唯一性;(2)数据集引用的元数据字段必须完整;(3)所引用的数据必须可以访问或获取;(4)数据引用除了能够引用某一数据集,还必须能够提供对其子集的引用;(5)数据引用能够被计算机处理,可以支持传统式的计量学分析。还有一些科研项目如由 DFG(德国研究基金会)资助的 STD-DOI^[24](Publication and Citation of Scientific Primary Data)项目等也提出了数据引用格式: creator(s)(publication year):data set name, publisher.persistent identifier。

一些地球科学(如 PANGAEA^[25]、ESIP^[26]等)、生命科学(如 GBIF^[27]、Dryad^[28]等)以及社会科学(如 Dataverse^[29]等)等呈现数据密集型态势的学科的数据仓储或中心也发布了自己的数据引用规范,这些引用规范与前述国际组织所发布的引用规范的核心元素组成基本相同。

目前很多国外的期刊出版商也在逐步引入数据存档政策,这种存档一般是以两种形式存在^[30]:(1)鼓励研究人员在发表论 文的同时,将相关的科学数据也一并提交,这类期刊大致有 PLOS One、BMC Evolutionary Biology、F1000 Research 等;(2)强制要求研究人员发表论 文的同时,将相关的科学数据提交到相应的机构仓储中,如 Nature、BioMed Central 等

综上所述,不论是国际通用学术论文写作规范最新版本中涉及到科学数据,还是专门针对科学数据的引用标准的相继出现,都预示着科学数据引用问题正受到各方关注。将已有具体数据引用规范的文件进行对比,可以发现其中 Author/Creator(创建者/责任者)、Title(标题)、Publication Year(发布年份)、Publisher/Archive/Distributor(发布机构/存储机构/传播机构)、URL/Electronic Retrieval Location/External Links(解析网址/获取地址)这四个属性,在所有引用规范、写作指南等文件中都被提及并作为“强制要求性”元素。Persistent Identifier(唯一标识符)作为比较新的事物,在较早发布的三大格式中都并未提及,但被所有专门针对科学数据引用规范指的文件作为强制性元素。另外,Resource Type/Material Designator(资源类型/载体代号)、Version(版本)也被多数规范所要求。随着数据引用规范的元数据组成进一步丰富与细化,数据引用过程变得更加清晰,但正如前述关键问题所述,科学数据这一引用对象具有高度的动态性、不确定性以及边界模糊性,因此数据引用规范目前还面临着数据集粒度和动态数据集处理等诸多问题,另外,由于科学数据本身的复杂性,数据引用规范的制定目前呈现出各自为阵的状态,各种引用规范之间的兼容也是亟待解决的问题。

4.文献管理工具

目前有许多文献管理软件(Reference Management Software)可以对各种资源进行引用,帮助用户高效管理和快速生成参考文献。绝大部分的 RMS 都是根据特定的资源类型,比如“期刊”“专利”“报告”“电子资源”等,从而生成由不同的元数据构成的参考引文实体。随着目前对“科学数据”引用的日益增加,在文献管理软件中专门集成“科学数据”等非文本资源类型显得更为急迫。

笔者根据下载使用、相关论文综述以及软件介绍说明等,对目前常用的 18 种文献管理软件的引用规范进行调研,结果显示目前很少有专门集成“科学数据”引文类型的软件。如表 1 所示,

表 1 常用的文献管理软件引用规范调研

文献管理软件	是否有专门科学数据或数据集的引用功能
Biblioscope	没有
Bibus	没有, 但用户可以自定义“数据集”选项
Bookends	没有, 但用户可以自定义“数据集”选项
Citavi	没有
Docear	没有
EndNote	有
JabRef	没有, 但用户可以自定义“数据集”选项
Mendeley	没有
Papers	有, 作为“数据库”或“表格”形式
Pybliographer	没有
Qiqqa	没有
Rebase	没有
Reference Manager	没有, 但用户可以自定义“数据集”选项
RefWorks	没有
Scholar's Aid	没有, 但用户可以自定义“数据集”选项
Sente	有, 作为'Data file' 选项
WizFolio	没有
Zotero	没有

由以上调研可以看出,在此次调研的常用的 18 个文献管理软件中,只有 3 个专门具备“科学数据”引用的功能,分别作为“数据库”“表格”和“数据文件”的选项进行引用。在其余 15 种软件当中,有 5 种允许用户自定义引用类型如“数据集”选项,但并没有专门针对“科学数据”的引用元数据,另外 10 种工具则根本没有涉及“科学数据”选项。综上所述,在目前常用的文献管理软件中,缺乏对数据引用的支持,对于“科学数据”等非文本资源类型的引用的关注度还远远不够。

到目前为止,虽然还没有出现可以在学术论文中自动索引数据引用的工具,但是科学数据的引用越来越受到各大研究机构和企业关注。比如汤森路透在 2012 年推出的数据引文索引工具 (Data Citation Index) [6],作为汤森路透搭建的 Web of Knowledge 信息平台上的一个新的资源发现工具,数据引文索引 (DCI) 将使数据集的发现和利用方式发生革命性变化。DCI 支持研究人员在文献的上下文中发现、引用和查看相关数据,从而推动对数据集和数据研究的发现、使用及归属,促进对全球研究数据的再利用。此外还有一些提供数据发现服务的工具,比如 ChemSpider^[31]、CrystalEye^[32]等。这些工具主要是面向特定的学科领域,对该领域的数据进行下载和聚合,并以增值资源的方式重新发布出来。

5. 数据引用利益相关者探讨

从很多实际情况来看,数据引用实践的滞后并不仅仅是引用规范制定的问题,它还涉及到多方利益相关者。许多致力于数据引用的国际组织都开展了有关数据引用利益相关者责任与能力的讨论,如 ODE (Opportunities for Data Exchange) 在 2012 年有关数据引用最佳实践的报告^[33]中从不同利益相关者的角度探讨了数据引用实践给它们带来的机遇与挑战,CODATA-ICSTI 数据引用标准和规范任务组也在 2012 年有关数据引用实践、政策以及技术现状分析报告^[34]中探讨了现有与数据引用相关的机构性质的基础设施,但总的来说,目前对于数据引用利益相关者责任与义务的探讨比较散乱,没有形成统一的体系。经过大量的调研,笔者大致归纳出以下 6 个与数据引用活动最相关的机构与它们的责任与义务,如表 2

所示，只有多方携手，才能从根本上推动数据引用实践活动的不断向前发展。

表 2 数据引用利益相关者责任与义务

利益相关者	责任与义务
研究资助机构（推动数据引用的外部力量）	<ul style="list-style-type: none"> ● 大力促进和推动科学数据的开放获取 ● 鼓励规范的数据引用行为 ● 对数据管理、归档和引用进行学术激励。
研究机构（推动数据引用的内在力量）	<ul style="list-style-type: none"> ● 一方面与研究资助机构一样，促进和推动科学数据的不断开放获取 ● 另一方面也需要制定促进科学数据的永久标识和引用的长期政策，作为数据引用规范的根本保证。
数据生产者（科学数据的来源，是与数据引用有着最直接关系的利益相关者）	<ul style="list-style-type: none"> ● 合理的管理数据以便共享和复用 ● 了解并且同意数据引用的条款 ● 在研究中使用科学数据时进行合理规范的引用
数据中心/数据仓储（科学数据的存储中介）	<ul style="list-style-type: none"> ● 提供强有力的数据策展服务 ● 提供科学数据的推荐引用格式 ● 参与建立统一的数据引用标准 ● 与数据发布者、收集方、使用方一起协作，增加数据引用的意识
图书馆（数据使用者和数据中心的联络人）	<ul style="list-style-type: none"> ● 围绕数据引用存在的问题，与相关机构进行协调，提供专业的标识技术 ● 在一些情况下作为数据仓储库 ● 发挥传统的计量分析作用，从而更好的激励数据使用者
学术出版商（数据出版的媒介）	<ul style="list-style-type: none"> ● 以经济杠杆的作用来促进数据引用规范的执行 ● 参与制定和实施数据引用的政策和标准、促进数据集的永久标识、建立数据集和出版者的第三方链接

6. 结语

随着数字化科研时代的到来，科学数据成为越来越重要的科学资源，科学数据的引用也受到了更多的关注。目前，缺乏数据共享和引用意识、引用对象不明确以及数据的长期保存问题是当前数据引用面临的三大难题。纵观目前数据引用的国外现状，众多国际组织、科研项目、数据仓储、数据出版商等都致力于建立各自的数据引用规范，但由于科学数据本身的动态性以及边界的模糊性，这些规范的元数据组成以及兼容性问题还有待解决；其次，目前的文献管理工具中几乎没有涉及对“科学数据”的引用管理，科学数据的引用必须进行手动添加，增加了科研人员引用科学数据的障碍，文献管理工具需要对科学数据这一重要资源做出必要的支持；再次，数据引用的发展与实践还有待研究资助机构、研究机构、数据生产者和研究人员、数据仓储、图书馆以及数据出版商等各个利益相关者的共同努力。

对于科学数据引用关键问题以及未来研究方向，笔者有以下几点思考：首先，针对缺乏数据共享意识和引用意识问题，完善针对科学数据的学术报偿体系，将科学数据引用纳入科研评价体系中是解决该问题的核心对策也是非常有价值的研究方向；其次，针对数据引用规范的兼容性问题，可以运用“本体”的思想，将常用规范之间进行本体映射或者根据已有规范

制定国际通用的引用标准；再次，针对科学数据的长期保存问题，需要促进科学数据唯一标识符以及数据仓储的注册，目前许多国际组织提供科学数据唯一标识符的注册，如 DataCite 等；另外，诸如 OAD^[35]、re3data.org^[36]、Databib^[37]等，能帮助人们识别和查找在线的数据仓储。我国目前的数据引用实践还处于起步阶段，需要借鉴国外的相关理论经验以及成功的实践案例，结合我国的实际情况，重点关注数据引用关键问题的同时，完善数据引用规范的制定、促进数据管理工具的开发以及鼓励各个相关机构共同努力，从而推动我国数据引用实践的不断向前发展。

参考文献

- [1]王卫华, 胡良霖, 沈志宏. 科学数据引用规范的研制[J]. 中国科技资源导刊, 2013, 45(1): 36-40.
- [2]李丹丹,吴振新. 研究数据引用研究[J]. 图书馆杂志, 2013, 5: 65-71
- [3]Data Cite[EB/OL]. [2014.2.15]. <https://www.datacite.org/>
- [4]Institute for Quantitative Social Science 2013. Data citation principles workshop[EB/OL]. [2014.2.14]. <http://thedata.org/presentations/data-citation-principles-and-practices>
- [5]Australian National Data Service 2011. "Building a culture of research data citation" workshop[EB/OL]. [2014.2.15]. <http://www.iassistdata.org/blog/workshop-building-culture-research-data-citation>
- [6]Thomson Reuters. 汤森路透发布 Data Citation Index 促进对全球研究数据的发现[EB/OL]. [2014-2-15]. <http://ip-science.thomsonreuters.com.cn/press/press20121102/>
- [7] 张静蓓, 吕俊生, 田野. 国外数据共享行为影响因素研究综述[J]. 图书情报工作, 2014, 58(4): 136-142.
- [8]Parsons M A, Duerr R, Minster J B. Data citation and peer review[J]. Eos, Transactions American Geophysical Union, 2010, 91(34): 297-298.
- [9]Hayes P J, Halpin H. In defense of ambiguity[J]. International Journal on Semantic Web and Information Systems, 2008, 4(2): 1-18.
- [10]Duerr R E, Downs R R, Tilmes C, et al. On the utility of identification schemes for digital earth science data: an assessment and recommendations[J]. Earth Science Informatics, 2011, 4(3): 139-160.
- [11]Weber N M, Piwowar H A, Vision T J. Evaluating data citation and sharing policies in the environmental sciences[J]. Proceedings of the American Society for Information Science and Technology, 2010, 47(1): 1-2.
- [12]吴立宗, 王亮绪, 南卓铜等. DOI 在数据引用中的应用: 问题与建议[J]. 遥感技术与应用, 2013, 28(3): 377-382.
- [13]Handle system[EB/OL]. [2014-2-14]. <http://www.handle.net/>
- [14]Uniform Resource Name[EB/OL]. [2014-2-14]. <http://www.iana.org/assignments/urn-namespaces/urn-namespaces.xhtml>
- [15]American Psychological Association(6th ed). APA style manual[EB/OL]. [2014-2-15]. <http://www2.bgsu.edu/downloads/lib/file40389.pdf>
- [16]Cornell University. MLA citation style (3th ed) [EB/OL]. [2014-2-15]. <http://www.library.cornell.edu/resrch/citmanage/mla>
- [17]The Chicago Manual of Style(16th ed). Chicago-Style Citation Quick Guide. http://www.chicagomanualofstyle.org/tools_citationguide.html
- [18] Altman M, King G. A proposed standard for the scholarly citation of quantitative data[J]. D-lib Magazine, 2007, 13(3): 5.
- [19]DataCite International Data Citation Metadata Working Group. DataCite metadata schema for the publication and citation of research data version 3.0[EB/OL]. [2014-2-15]. <http://schema.datacite.org/meta/kernel-2.1/doc/meta/kernel-3/meta/kernel-3/meta/kernel-2.2/index.html>

- [20]Green T. We need publishing standards for datasets and data tables[J]. Learned Publishing, 2009, 22(4): 325-327.
- [21]Duke, M. and Ball, A.How to Cite Datasets and Link to Publications:A Report of the Digital Curation Centre[C]. //In:23rd International CODATA Conference: Taipei,2012
- [22]DataCite.What is DataCite?[EB/OL]. [2014-2-18].
<https://www.datacite.org/whatisdatacite>
- [23]SageCite [EB/OL]. [2014-2-15].<http://blogs.ukoln.ac.uk/sagecite/>
- [24]The project Publication and Citation of Scientific Primary Data. Use of persistent identifiers in the publication and citation of scientific data[EB/OL]. [2014-2-15].
<http://wiki.pangaea.de/wiki/STD-DOI>
- [25]PANGAEA[EB/OL].[2014.3.1].<http://www.pangaea.de/>
- [26]ESIP[EB/OL].[2014.2.12].<http://www.esipfed.org/>
- [27]GBIF[EB/OL].[2014.2.12].<http://www.gbif.org/>
- [28]Dryad[EB/OL].[2014.2.12].<http://datadryad.org/>
- [29]Dataverse[EB/OL].[2014.2.12].<http://thedata.org/>
- [30]黄永文, 张建勇, 黄金霞, 等. 国外开放科学数据研究综述[J]. 现代图书情报技术, 2013, 5: 21-27.
- [31]ChemSpider[EB/OL]. [2014-2-13].<http://www.chemspider.com/>
- [32]CrystalEye[EB/OL].[2014-2-13].
http://download.cnet.com/Acer-Crystal-Eye-webcam/3000-2108_4-188685.html
- [33]Task Group on Data Citation Standards and Practices C I. Out of Cite, Out of Mind: The Current State of Practice, Policy, and Technology for the Citation of Data[J]. Data Science Journal, 2013, 12(0): CIDCR1-CIDCR75.
- [34]Kotarski R, Reilly S, Smit E, et al. Reports on best practices for citability of data and on evolving roles in scholarly communication[J]. 2012,4(13):1-60.
- [35]OAD[EB/OL].[2014-4-5].http://oad.simmons.edu/oadwiki/Data_repositories
- [36]re3data[EB/OL].[2014-4-5].<http://www.re3data.org/>
- [37]Databib[EB/OL].[2014-4-5].<http://databib.org/>

[作者简介] 张静蓓,中国科学院兰州文献情报中心、中国科学院大学硕士研究生
E-mail:zhangjb@mail.las.ac.cn

吕俊生, 中国科学院兰州文献情报中心研究馆员。

田野,中国科学院中国科学院兰州文献情报中心、中国科学院大学硕士研究生

作者贡献声明:

吕俊生: 提出研究思路与研究框架, 参与论文最终版本修订;

张静蓓: 负责论文起草和论文修订;

田野: 负责文献管理软件引用规范调研。

Research Progress of Data Citation in Forging Countries

Zhang JingBei^{1, 2} Lv JunSheng¹ Tian ye^{1, 2}

¹ Lanzhou Library of Chinese Academy of Sciences, Lanzhou 730000, China

²University of Chinese Academy of Sciences, Beijing 100049, China

[Abstract]: This article summarizes three key issues of data citation : citation awareness, citable object and long-term preservation. Discussing the research progress of data citation in foreign countries from the development of data citation standards,the survey of reference management software,the responsibilities and obligations of stakeholders.It finds that data citation standards are incompatible, the reference management software is lack of support for scientific data, responsibilities and obligations of stakeholders are not clear.Finally,the article puts forward suggestions about key issues and then gives a forecast.

[Keyword]: data citation citation standard reference management software stakeholders