

基层科技报告资源建设中元数据质量评估研究 ——以中国科学技术信息研究所为例

宋立荣

(中国科学技术信息研究所, 北京 100038)

摘要: 我国各级科技管理部门正积极推进国家科技报告资源建设工作。在这过程中, 基层科研单位作为科技报告资源建设的生产源头, 是保证科技报告质量的关键环节, 尤其是元数据质量建设起到十分重要的作用。为此, 文章就科技报告元数据项要素进行分析, 对其元数据质量评估确定基本内容, 提出从完整性、准确性和一致性等维度进行质量评估。最后, 以中国科学技术信息研究所内部约410份科技报告资源为例, 通过调查统计、评估分析其元数据质量情况。

关键词: 基层科研单位; 科技报告资源; 元数据; 元数据质量; 质量评估

中图分类号: G203

文献标识码: A

DOI: 10.3772/j.issn.1674-1544.2016.01.011

Metadata Quality Assessment of the Scientific and Technical Report Resource in Basic Research Units: Taking ISTIC as an Example

SONG Lirong

(Institute of Scientific and Technical Information of China, Beijing 100038)

Abstract: At present, China is working to promote the construction of national scientific and technical reports. In this process, the basic research units, as the source of scientific and technological report resources, play a very important role to ensure the quality of scientific and technical reports. This paper, analyzes the metadata elements of scientific and technological reports, and the basic content of metadata quality assessment from the aspects of completeness, accuracy and consistency. At last, the paper takes about 410 internal scientific and technical report resources of ISTIC as an example, surveys, statistics and evaluation of the metadata quality of these information resources, and hope that this method can facilitate the quality control of scientific and technical reports, evaluate and improve the quality management of basic research units.

Keywords: basic research units, scientific and technical report resource, metadata, metadata quality, quality assessment

作者简介: 宋立荣(1971—), 男, 中国科学技术信息研究所研究员, 博士, 研究方向: 科技信息共享、信息质量。

基金项目: 国家软科学项目“国家科技报告资源建设中关键质量影响因素及其测评体系研究”(2014GXS4K052); 国家社科基金“网络环境下科技信息资源建设中的质量元数据及评估应用研究”(12BTQ016)。

收稿时间: 2015年12月22日。

1 引言

科技报告作为国家重要的战略信息资源,其制度体系的建设和推进可以为科研人员提供科研基础信息,为社会公众了解和利用国家科研成果提供服务平台,对于提升国家科技实力 and 创新能力具有重要的意义^[1]。我国正在加快建立统一的国家科技报告制度^[2],构建完善的国家科技报告制度体系,各级科研管理部门积极开展科技报告资源建设工作^[3]。作为科技报告资源的生产源头,基层科研单位在科技报告资源积累、科技报告质量控制以及科技报告开放共享服务等知识资产积累和传承方面将承担关键作用^[4]。从目前的情况来看,由于科技报告类型众多,来源广泛,提交加工环节多样等特点,使得提交上来的科技报告不规范、资源质量良莠不齐,大大地降低了科技报告的科技含量和应用价值。因此,提高国家科技报告资源建设质量,研究探讨科技报告资源建设中的质量问题显得十分重要,也是十分必要的。但是,如果仅仅从科技报告产生过程进行质量控制、监督管理,则难以起到有效的质量控制作用,加上科技报告自身还存在保密性、专业性等特点,外部用户并不能直接看到科技报告全文,更多地是通过元数据来间接获取科技报告相关信息。由此可见,元数据描述信息资源的外部和内容特征,能够帮助用户发现、定位、获取资源的关键内容,在科技报告资源建设中发挥着重要作用。科技报告元数据质量成为科技报告资源建设工作的关键问题之一。通过元数据质量评估是其中较为重要的、具有可操作性的有效方法。通过元数据质量的评估、审核和改善将有效促进科技报告元数据结构的补充与完善、元数据内容填写的规范。同时,研究元数据质量评估的方法对于元数据规范标准的制定具有非常重要的意义。

目前,对于科技报告体系研究主要集中于政策法规、制度框架、管理模式、运行机制、标准规范等宏观层面^[5-7]。而在不多的涉及部分科技报告元数据内容的文献中,大多文献主要是针对科研机构如何撰写、保存科技报告等方面的研

究^[4]。在实践方面,尽管我国已颁布了《中华人民共和国国家标准:科技报告元数据规范》(GB/T 30535-2014),但也仅仅针对科技报告资源描述元数据制定的,尚缺乏基于科技报告生产全生命周期的元数据建设规范。总体来看,目前理论研究尚缺乏对科技报告资源元数据框架体系的完整的、系统的研究。为此,本文将以中国科学技术信息研究所(以下简称“中信所”)内部约410份科技报告资源为例,通过调查统计、评估分析其元数据质量情况,有效改进元数据填报质量,提升科技报告建设管理水平,促进科技报告资源质量改善。

2 元数据质量评估研究综述

元数据是信息资源建设的基础和关键,而元数据能否发挥重要作用取决于元数据质量的好坏。国内外对“元数据质量”概念目前还没有一个标准的定义。美国学者Bruce和Hillman将“元数据质量”定义为“元数据满足某个特定需求和目标的程度”。换言之,元数据的质量与它能在多少程度上方便用户发现、确认、选择和使用信息资源的功能密切相关^[8]。

就元数据质量评估而言,目前学术界的研究主要集中在以下几个方面。

(1)元数据质量评估方法:随着开放环境中数字资源的高速增长,过去那种由专家、编目专业人员等组成的团队对抽样元数据质量进行手工评估方法^[9]因其评估成本高、效率低、准确性低而逐渐不被普遍使用。目前,多采用人工和自动相结合的综合方法来提高评估的可操作性和有效性,并积极探索使用一些软件工具和方法对元数据的语法和结构质量进行简单统计的自动评估(如相关字段的必备性、语法错误、链接是否可用等),如:国内CALLS针对特色数据库导入元数据的规范性和必备性设计的数据质量检查工具;检查DC元数据记录、协议、规则正确性的DC checker;地理空间数据库元数据标准FGDC的质量检查工具CNS与MP等^[10]。但自动化评估目前主要是对诸如“完整性”这样基本的、易于

实现计量统计的维度进行评估，评估实践并没有包含到元数据内容质量的主要影响因素，而往往只评估了影响元数据质量的某些方面，尚无法实现全面评估^[9]。

(2) 元数据质量评估维度的确定：元数据质量是一个多维的问题，不同评估人员、不同评估对象以及不同的评估、方法对元数据质量维度的选择都不一样。需要根据评估目的、内容、元数据资源情况、元数据获取方式等进行综合考虑。比如：有的学者认为^[11]元数据质量可以从资源集合、记录与元素3个不同层次的正确性、完整性、重复性及一致性等指标进行考察。Bruce和Hillmann提出了7个元数据质量特征：完整性、准确性、来源、与预期的符合程度、一致性和连贯性、时效性和可获得性^[8]。

(3) 元数据质量评估模型：Moen等基于对GILS元数据的评估结果提出一个21个指标的评估模型^[12]。该模型评估从用户、政策、技术、内容和标准等5方面展开，多是评估元数据的完整性、外观、准确性和可服务性方面；Stvilia等借鉴信息质量评估模型评估元数据质量^[12]，指标分别是内在指标（9个）、相关指标（12个）和信誉指标（1个）；黄莺等提出一个以完整性、准确性、一致性和期望满足度等4个维度为核心的两层结构模型^[9]，核心层的维度是对元数据自身质量的评估，其涵义不受元数据外部环境的影响，而外层维度为评估的可选维度，包括但不限于可获取性、易用性、来源、安全性等评估维度，这些维度与元数据的外部属性密切联系。

由此可见，国内外对元数据质量评估的关注尚处于探索阶段，从不同角度出发提出的一些元数据质量评估方法、体系难以在科技报告资源建设领域得到直接推广应用，尤其是对评估体系中各维度的具体量化操作、评估方法的应用更是亟待深入研究和实验论证。

3 科技报告元数据质量评估认识

3.1 科技报告元数据项基本构成

元数据是对科技报告的外表特征和关键内容

的描述，是科技报告资源开发、组织、利用和管理的基础。

科技报告元数据主要用于对科技报告的文献特征信息和项目来源基本信息进行描述、组织和管理，包括提供科技报告的内容、载体、位置与获取方式、制作与利用方法以及项目等方面的信息，它支持科技报告基本信息在计算机信息系统中的存储、管理、定位、调用等功能，帮助用户检索、识别和确认所需要的科技报告。因此，元数据质量直接影响到科技报告资源的共享和利用效率。

科技报告元数据大部分来自基层科研单位撰写的科技报告。基层科研单位是科技报告生成、提交、审查和管理的基础单位，处于科技报告产生的前端。因此，在科技报告建设中不仅要求基层科研单位必须按照国家制定的统一模式撰写、编号和划分密级，以及按照国家标准管理和元数据集等标准规范为基础进行管理^[3]，确保科技报告有规范、统一的元数据集，而且要求科技报告建设的過程质量控制达到后端用户服务的质量要求。这需要在科技报告资源建设中抽取相应的元数据项来反映上述质量规范要求，显示过程质量在受控状态。同时，还要扩展元数据项以满足后期必要的信息检索、数据挖掘、知识关联等服务功能的需要。

上述这些工作的前提就是要保证所填写的元数据质量是真实、准确的。因此，除了要确定必要的元数据质量优劣的判断标准外，还需要有一套切实可行的质量审核规范流程进行元数据质量评估。有专家建议充分利用管理系统的自动化流程，比如：在科技报告工作系统添加相应的自动审核模块，设置字段和属性，根据《科技报告元数据规范》国家标准及相关科技报告质量审核指标对科技报告全文、元数据项进行审核^[3]。

在实际工作中，我国已制定了国家标准《中国科技报告元数据规范》，以及通用的《电子文件的管理元数据规范》，其中部分内容涉及一些描述元数据的制定规范问题，比如：在国家标准“中国科技报告元数据规范”中科技报告元数据

集由13个元素、27个元素修饰词、8个编码体系修饰词构成，元数据集的元素按功能分类，一般可分为描述科技报告文献特征的元数据、描述科技报告来源项目的元数据和用于科技报告管理保存的元数据。描述文献特征的元数据包括题名、作者、报告类型、日期、摘要、分类号、关键词等；描述项目信息的元数据包括计划名称、项目/课题名称及编号、承担单位等；管理元数据包括科技报告编号、密级、特别声明、馆藏号、收藏日期等^[13]，见表1所示。但这些元数据规范尚未细化到具体各个环节的元数据项要求。加上科技报告资源信息元数据的加工单位、保藏单位和应用单位众多，采用的元数据规范、资源分类、资源标识等标准规范的不统一，且科技报告本身具有多样性、异构性和复杂性的特点，使得元数据在质量方面存在参差不齐的现象，从而加大了实现入库的科技报告资源共享的难度，给科技报告资源整合、共享和服务带来了一定困难。

目前，尚未有针对科技报告资源建立较系统、全面的元数据质量评估体系，有些单位内部建立的科技报告质量评估体系多是基于本系统内部管理考核用途，在核心元数据遴选、指标设

置、权重分配、考核方式等方面尚缺少有力的理论支持，有些仅仅通过专家打分进行评估，缺少可量化的指标体系。

总的来看，在科技报告资源建设过程中，元数据质量除满足一般元数据的质量共性要求（如真实性、完整性、可靠性与可用性），至少满足以下几点质量要求：一是元数据对数据对象描述的准确性与完整性。要求元数据能够对科技报告的来源、数据内容、数据结构、利用、管理及其处理过程等做出准确、全面和详尽的说明。它关系到科技报告是否可以被准确确定的程度。二是元数据对数字对象描述的准确度，其描述的行为与真实值之间的差异。三是元数据描述数据对象的时间精度，可以通过元数据更新的时间和频度来体现科技报告的时效性。

3.2 科技报告资源建设中元数据质量评估的基本内容

作为一种特殊的灰色文献，科技报告要求能够如实、完整、及时地描述科研的基本原理、方法、技术、工艺和过程等，以便科技工程人员之间、政府部门之间快速交流和共享最新的前沿技术和核心研究结果^[5, 7]。

表1 科技报告元数据项标准对比

序号	国家标准 (科技报告元数据规范GB/T 30535-2014)	中信所内部标准 (中信所科技报告管理系统)
1	题名	报告名称
2	作者	报告作者
3	作者单位	作者单位
4	关键词	关键词
5	目次	目录
6	责任者顺序	责任者顺序
7	摘要	摘要
8	参考文献	参考文献
9	起止日期	起止时间
10	完成日期	编制时间
11	提交日期	提交日期
12	范围	公开范围
13	报告密级	报告密级
14	计划名称	计划类型
15	项目/课题名称及编号	项目(课题)名称及编号
16	项目/课题承担单位	课题承担单位
17	---	合作单位
18	---	主管部门
19	---	负责人
20	---	联系人
21	---	支持渠道
22	---	备注

一般情况下，由于呈交科技报告的基层单位的规模、行业、领域不同，数据类型也较繁杂，因此，评估其科技报告的评估体系不尽相同。这需要每个基层科研单位结合自身情况建立适宜的元数据质量评估体系和评估内容。本文主要是评估体系基于科技报告资源建设中的几个主要环节提出几个评估要点。在实际应用中，考虑到某些基层单位的具体情况，有必要在评估流程上进一步增减一些环节或步骤，以使评估更有效率。但一些具有共性的基本评估内容必不可少，包括以下几个方面。

(1) 评估对象及范围。根据评估需求及评估目的，确定评估对象及其范围，评估对象既可以是某一份科技报告也可以是基层单位整体科技报告资源。

(2) 关键元数据项的遴选。可从目标科技报告资源数据库的主要质量特征进行分析。目前，一般是通过科技报告元数据规范和标准来初步确认一些关键元数据项，也可以根据该科技报告资源数据库常出现的质量问题建立相应的元数据项作为质量控制的关键要素。本文确定的几个关键元数据项见表1所示。

(3) 评估对象的关键质量维度的确定。首先依据相关国家或国际标准来确定质量维度。对无标准参考的数据资源的质量维度的把握通常可从质量维度来统一定义、分类（包括形式、内容、关联环境）、从属性（指标冲突或上下关系等）和筛选（包括实际感知程度和理论重要程度两方面）等方面进行确定。

针对元数据质量评估有很多质量维度，基于各自的视角和方法不同所选取的质量维度也不一样。而且也注意到，过多的质量维度的评价设置常常会因为指标间存在相互冲突、评价成本等因素而使得评价结果差强人意。本文基于一些文献内容的梳理，多数研究集中在“准确性”、“一致性”和“完整性”等几个最基本维度。J.R.Park等人分析了各种元数据质量评估中经常重合的一些标准和指标，发现准确性、一致性和完整性是衡量元数据质量时最常用的标准^[14]。也是影响元数

据质量最重要的3个因素。为此，本文拟从完整性、准确性、一致性等几个质量维度入手对目标科技报告资源的元数据质量进行评估。

①完整性：它是衡量元数据全面、详尽地描述其目标资源的程度。它包括元数据项数目的完整性和填写内容的完整性两个方面。元数据项的完整性要求科技报告元数据项中必选、可选以及条件选择等必须完整，并且符合相应标准的要求，从而保证全面地描述整个科技报告。元数据项的完整性可以用元数据项中非空字段的多少来衡量。

②准确性：准确性衡量的是元数据提供的内容正确、客观地反映目标资源的程度。准确性衡量多采用定性方法，通过专业人员和专家审核判断。内容的准确性要求元数据项内容（如摘要、关键词等说明）必须准确表达资源本身内容，确保记录真实可靠或者内容有无实质意义，必须完整记录科技报告资源相关方面的信息。内容的完整性多采用定性方法，通过负责科技报告质量审核的专业人员和专家审核判断。

③一致性：一致性主要考察的是元数据遵循元数据规范、应用指南以及按照规范、应用指南从受控词表取值的情况。对一致性的量化是统计元数据记录违背元数据规范和应用指南的频率。

(4) 元数据质量评估指标函数的确定。元数据质量评估指标计算方法通常有加权平均法、最大最小运算法、简单比率法等评估函数。由于本文评估指标涉及定量和定性元素，故本次调查对每一评估指标项统计采用的是百分制加分评估标准方法，对每一项指标根据其内容完善程度进行加分评估。主要采用分项检查指标内容有（或无）情况以及元数据内容与科技报告资源匹配度进行数据调查收集。权重采用直接赋权法。

项目总分值采用加权平均计算得出，即：对各个指标评分结果×权重的加总计算。

$W=A_1 \times \lambda_1 + A_2 \times \lambda_2 + \dots + A_j \times \lambda_j + \dots + A_n \times \lambda_n$ ，其中， A_j 为各个评估指标分值， λ_j 为对应的指标权重。

各个评估指标值的计算公式为：

$$A_j = \frac{\sum_{i=1}^n P(i)}{N} \times 100$$

即: (符合标准数量/调查总数) × 100, 结果越大表明该指标的质量越好。其中, j 为评估维度指标, 如完整性、准确性和一致性。 i 为元数据项

比如: 在“完整性”公式计算中, 若第 i 个字段为空, 则 $P(i)$ 为 0, 非空则为 1。 N 为元数据大纲规定的字段项个数。

对“准确性”进行量化的最直接方法就是逐一对比判断关键元数据项记录内容与科技报告资源本身内容的准确度, 主要采取人工判断方法, 打分采取分级标度法, 可以是三级、五级或更细化内容。本文主要采取“1—(1/2)—0”三级标度法。其中: “1”表示两者完全匹配; “1/2”表示二者基本匹配; “0”表示二者没关联性”。公式中若第 i 个字段描述错误, 则 $P(i)$ 为 0; 若第 i 个字段描述不完全正确, 则 $P(i)$ 为 1/2; 若第 i 个字段描述正确, 则 $P(i)$ 为 1。

“一致性”进行量化计算的方法主要是判断各个元数据项是否遵循各项元数据质量约束规范。公式中若第 i 个字段遵循第 i 条规则的 N 种情况, 取值为 1, 则 $P(i)$ 为 1, 否则为 0。 N 为元数据大纲规定的字段规则数量。其中对第 i 个字段如果出现以下 4 种情况之一, 则可判断没有遵循第 i 条规则: 一是元数据中包含了指定元数据规范并没有定义的字段; 二是记录里没有元数据规范规定的必备字段; 三是某些字段没有从元数据规范规定的受控词表中取值; 四是不遵循元数据规范的应用指南, 将受控词表中的多个值整合赋给某个或某些字段。分项检查四类指标有(或无)情况进行统计。

(5) 进行评估分析, 根据前面 5 步确定的质量对象、质量范围、测度指标及其评估方法进行评估分析。

(6) 评估结果分析及报告, 基于评估结果对目标数据资源进行评估、分级、质量标示, 并将评估结果进行分析, 找出差距, 进行故障分析, 并分析质量较低的元数据项或关键环节、过程, 进行针对性的质量改进。

4 中信所案例分析

为了更清晰说明科技报告资源元数据质量评估的过程和结果, 本文以中信所内部约 410 份科技报告资源进行调查评估, 调查工作主要有两个部分: 一是对中信所内部科技报告资源的元数据进行归类整理, 确定关键元数据信息, 其中, 因每一份科技报告名称、作者、作者单位都是每份必填, 基本都能够比较完整准确, 故本文次选择了“关键词”等 19 个元数据项进行调查统计分析, 见图 1 所示; 二是对关键元数据项的字段信息进行调查、对 410 份中信所科技报告元数据按统计打分和测试打分, 采用 EXCEL、ACCESS 和 SPSS 软件进行处理以及按均分法确定各个元数据项的权重, 进行总体质量评估。

4.1 综合评估结果分析

从统计分析结果(图 1)总体看, 410 份中信所科技报告在元数据规范方面做得较好, 在一些必填字段项填写中都能很好完整填写内容。存在的主要问题是: 首先对一些可选项(如“备注”“合作单位”等)没有填写内容的没有添加“无”, 出现字段项空值情况。其次对一些元数据项的填写格式规范不熟悉, 出现不规范填写情况, 造成准确性、一致性较差。如“起止时间”要求填写“年-月-日”, 而很多报告中只是出现“年-月”, 缺少具体日期; 再次对一些元数据项出现缺失情况, 如有些科技报告类型正文中缺少“目录”“参考文献”内容, 或没有标注“无”。

从“完整性”来看, 大部分科技报告都完成对必选元素的填写, 只是对一些可选元素的填写因没有注明“无”而造成该字段项为空值显示。

从“准确性”来看, 多数字段项基本做到了对目标资源本身内容的正确描述。仅有个别项(如“主管部门”“起止时间”“支持渠道”等)因填写者对该字段项的理解不同, 出现填写内容不准确或错误描述的情况。

从“一致性”来看, 主要是个别项(如“主管部门”“起止时间”“支持渠道”等)出现和规范填写不一致的情况。

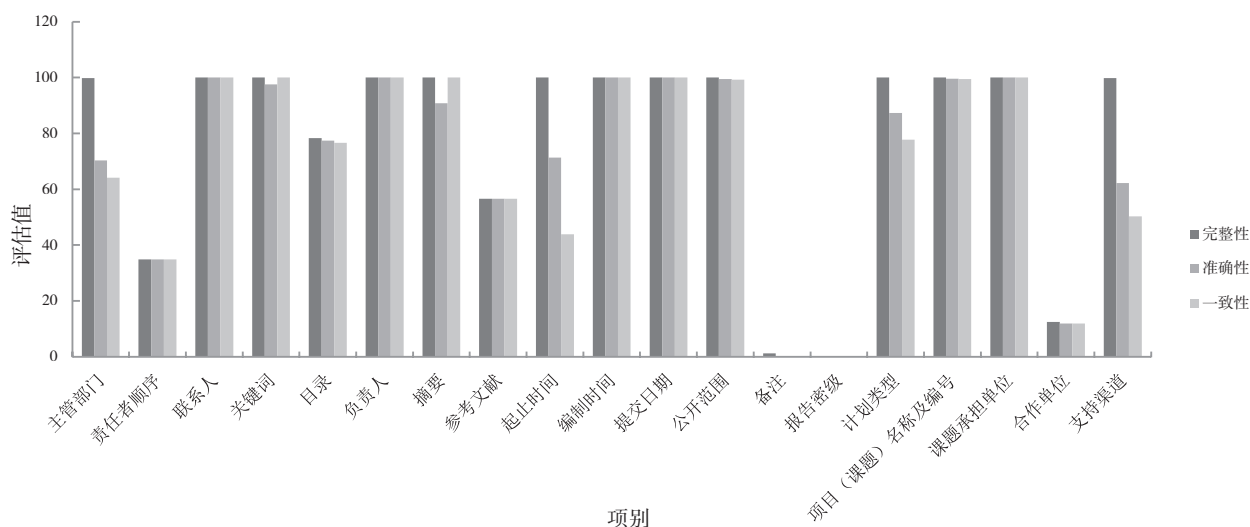


图1 410份科技报告元数据质量总体评估

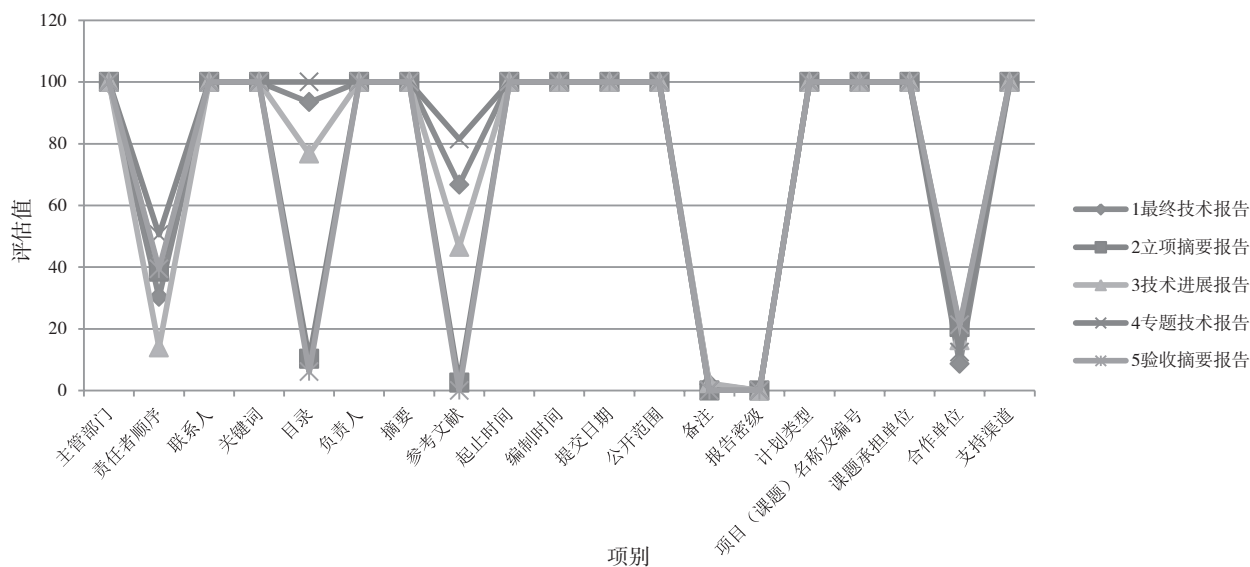


图2 中信所科技报告类型统计的“完整性”评估

4.2 按科技报告类型统计的评估结果

科技报告类型主要有立项摘要报告、技术进展报告、专题技术报告、最终技术报告和验收摘要报告等。按科技报告类型对完整性、准确性、一致性的评估，结果分别如图2、图3、图4所示。

(1) “完整性”

不同类型科技报告侧重在不同元数据项，例如：“立项摘要报告”“验收摘要报告”普遍在正文中缺少“目录”“参考文献”项内容，多只是文字说明；又如在“技术进展报告”和“最终技术报

告”对参与者贡献（“责任者顺序”）表述也不一致。因此，从完整性来看，主要在“目录”、“参考文献”和“责任者顺序”等方面差别较大。这和报告类型的要求要素不同有关。

(2) “准确性”

从“准确性”来看，差别主要是在“目录”“参考文献”和“责任者顺序”这3项。一般，“最终技术报告”“专题技术报告”在这方面的得分较高，也是最全面、最规范的。而“技术进展报告”则相对简化，多是对项目工作进展的汇报说明。

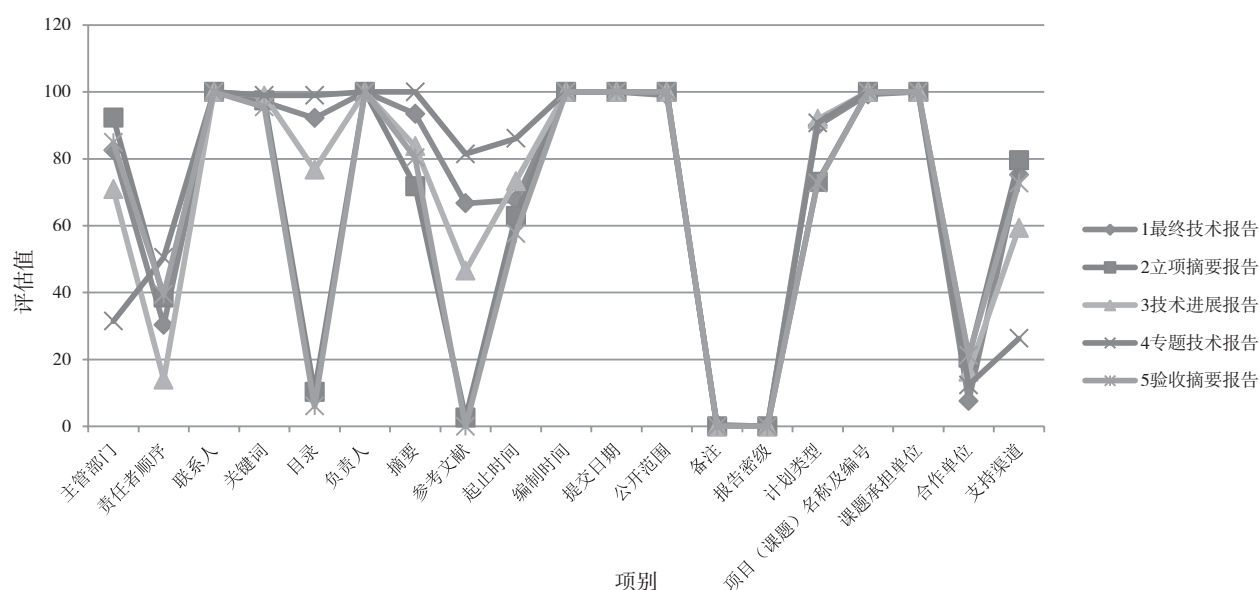


图3 中信所科技报告类型统计的“准确性”评估

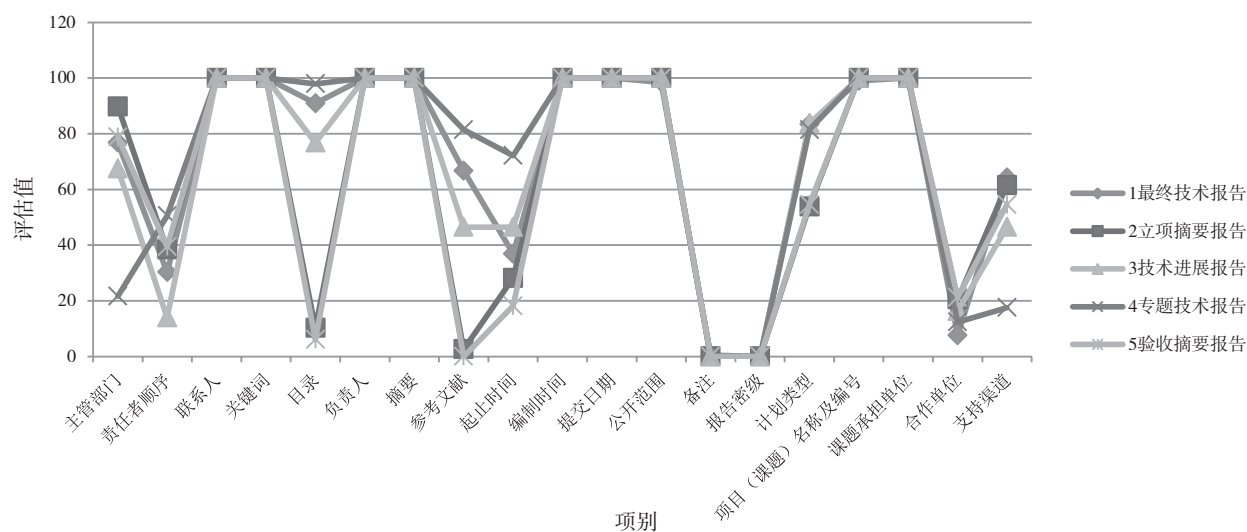


图4 中信所科技报告类型统计的“一致性”评估

(3) “一致性”

从“一致性”来看，除了以上3项差别大外，在“起止时间”的表述规范性不够。一般，在“立项摘要报告”中只是大概说明了“年-月”时间段，而在“最终技术报告”“验收摘要报告”中则能填写较全的“年-月-日”。

(4) 按科技报告类型统计的评估结果

从评估统计结果来看，“最终技术报告”的科技报告资源元数据质量总体较高。大部分字段项指标得分较高，在3个质量维度的得分较均衡。

出现的一些低分值多一是因为未填“无”造成评估值按照零分统计；二是因为在“立项摘要报告”正文中缺少“目录”“参考文献”项内容，致使这两项为空值。由于“立项摘要报告”是在项目刚开始时提交的科技报告，很多研究项目的具体研究内容尚未展开，故在一些元数据项填写内容描述的准确性较低，比如在“摘要”内容表述和正文研究上有一定的差别；“技术进展报告”正文内容在很大程度上仅是研究工作进展汇报，因此，在“目录”“参考文献”项内容较简单，使得

这两项的“准确性”“一致性”得分较低。另外，大部分“专题技术报告”普遍在正文内容上较完整、齐全，在表述正文内容的字段项（如“关键词”“摘要”“目录”“参考文献”等）得分较高；大部分“验收摘要报告”的正文内容较简单，没有“目录”“参考文献”内容。另外，在“主管部门”“支持渠道”等项填写不规范，“一致性”得分较低。

4.3 案例评估结论分析

经过对410份中信所科技报告的统计评估分析，可以看到，这些科技报告的质量总体较高。究其原因主要是严格执行了中信所内部制定的科技报告管理流程和规范，从而保证了资源质量。从结果分析来看，元数据质量评估方法具有一定的可操作性，在一定程度上可以帮助发现在科技报告资源建设中存在的一些规范性、专业性的质量问题。通过各项统计分析，可以看到尚有以下几点质量问题。

一是对一些非填写内容要求的元数据项不重视，没填入“无”等，而造成字段项内容空项，得分为零，从而影响整体评估结果。

二是对一些诸如“支持渠道”“主管部门”等项填写缺乏统一、规范的约束规则，造成对这些字段项的认识、填写不一致，出现用各种简称代替填写等情况。

三是对一些报告类型的正文内容缺少具体、明确的格式要求，使得一些报告的内容过于简单，缺乏必要的字段项内容，达不到必要的技术内容质量要求。

基于以上分析，提出如下建议。

（1）进一步完善和改进中信所的科技报告管理平台系统的各类功能，避免一些资源质量形式上的错误发生。比如：对一些可选项的要素（如“备注”“报告密级”“合作单位”等）可添加“有/无”选项。没内容时选“无”，避免出现字段项空值而造成统计偏差。

（2）加强对科技报告撰写人、报告提交人（如学术秘书等）、部门审核者等关键管理岗位人

员相关质量规范的培训，以便统一和规范科技报告元数据项的填写内容。

（3）改进科技报告资源管理流程，细化各种质量规范和约束规则，既要从软件检验工具入手加强科技报告资源的形式质量，还要指派专人从内容审核入手加强对内容的质量控制，提高科技报告的“含金量”。

（4）制定和细化科技报告资源的质量评估细则，可通过科技报告分类来建立不同权重的评估指标体系，细化评分标准，使评估具有可量化、可操作性。

5 结论与展望

本文以中信所内部科技报告资源的元数据项进行案例分析，尝试探索元数据质量评估方法在科技报告资源建设过程中可行性。从结果分析来看，这一方法在一定程度上可以帮助发现在科技报告资源建设中存在的一些规范性、专业性的质量问题，以便改正，从而提高科技报告资源质量。

但是，由于缺乏明确的元数据质量规范标准，本次研究尚有一些不足。本文只是就现有科技报告资源中的描述性元数据项进行统计分析，尚没有收集、整理科技报告产生过程的各个环节的管理元数据项内容数据，故无法就过程环节中质量问题进行统计分析。另外，关键元数据项在评估质量维度中的权重分配，应首先通过用户和专家调查法来调查分析确定，但由于众多用户和专家对科技报告资源关键元数据项的认识不一致，也缺乏相关国家标准参考，故本次调查数据处理暂对各个元数据项权重采取均分做法，这需要在后续研究中结合科技报告产生全过程确定科技报告项目中关键元数据项，以确定不同元数据项对不同质量维度的质量影响程度不同。

参考文献

- [1] 张新民.国家科技报告服务系统构建中相关问题的探讨[J].中国科技资源导刊,2014,46(1):9-13,27.
- [2] 中央政府门户网站.关于深化科技体制改革加快

- 国家创新体系建设的意见 [EB/OL].(2012-09-23) [2015-3-15]. http://www.gov.cn/jrzq/2012-09/23/content_2231494.htm.
- [3] 中国政府网.关于加快建立国家科技报告制度指导意见的通知[EB/OL]. (2014-08-31)[2015-3-15].http://www.gov.cn/zhengce/content/2014-09/10/content_9071.htm.
- [4] 曾建勋. 基层科技报告体系建设研究[J].情报学报,2014(8):800-806.
- [5] 曾建勋.科技报告技术标准体系研究[J].情报学报,2013,32(5):459-465.
- [6] 贺德方,胡红亮,周杰,等.中国科技报告体系的建设模式研究[J].情报学报,2009,28(6):803-808.
- [7] THOMAS R BRUCE, DIANE I Hillman. The continuum of metadata quality: Defining, expressing, exploiting[M]//Diane I Hillmann. Elaine L Weatbrooks. Metadata in Practice. Chicago:American Library Association, 2004:238-256.
- [8] 黄莺,李建阳.元数据质量评估方法及模型研究[J].图书馆学研究,2013(12):52-56.
- [9] 曹月珍,马建玲.国内外元数据质量控制的研究进展与发展趋势[J].图书与情报,2013(6):101-103.
- [10] QIN Jian, MARCIA Lei Zeng. Metadata[M]. New York: NealShuman Publisher,Inc,2004:247-249.
- [11] 黄莺,李建阳,元数据质量评估研究现状剖析[J].中国电子商务,2013(4):164-165.
- [12] GB/T 30535-2014科技报告元数据规范[S].北京:中国标准出版社,2014.
- [13] TOSAKA Yuji, PARK Jungran. Metadata quality control in digital repositories and collections: Criteria, semantics, and mechanisms[J]. Cataloging & Classification Quarterly,2010(48):696-715.

(上接第41页)

作的深入开展,科技报告工作必须适应科技计划体系以及科技计划管理模式的变化和需求,实现科技报告和科技计划管理全过程的有机统一,将建立山东省科技报告工作“一盘棋”的格局必须摆上日程。

5 结语

科技报告制度建设是一项系统工程。鉴于其复杂性、艰巨性和长期性,不可能一蹴而就,也不可能一劳永逸。目前,科技报告制度已陆续在国内多数省份启动建立。由于各地普遍存在基础薄弱、能力不足、经验欠缺等困难,省级科技报告制度建设面临巨大挑战。山东省科技报告工作实践表明,科技报告制度建设机遇大于挑战。不仅要结合自身实际夯实工作基础,更要注重顶层设计,超前谋划,建立科学的协同和推进工作机制。我们相信,科技报告制度在省级层面必将取得突破性进展,建立全国统一的科技报告制度指日可待。

参考文献

- [1] 袁清昌,张玉华,乔振.从美国科技报告体系看科技报告制度的建设发展[J].中国科技资源导刊,2015,34(5):40-44.
- [2] 张爱霞,沈玉兰.美国政府科技报告体系建设现状分析[J].情报学报,2007,26(4):496-502.
- [3] 张新民.国家科技报告制度建设试点工作的回顾与展望[J].中国科技资源导刊,2015,34(5):1-7.
- [4] 贺德方.中国科技报告制度的建设方略[J].情报学报,2013(5):452-458.
- [5] 袁清昌,马文哲.山东近千份科技报告面向公众开放[N].科技日报,2015-04-06(3).
- [6] 贺德方,胡洪亮,周杰.中国科技报告体系的建设模式研究[J].情报学报,2009,28(6):803-808.
- [7] 中华人民共和国促进科技成果转化法(2015年修订)[EB/OL].(2015-08-29)[2015-09-20]. http://www.most.gov.cn/fggw/fl/201512/t20151203_122619.htm.
- [8] 国办发〔2014〕43号.关于加快建立国家科技报告制度的指导意见[EB/OL].(2014-09-10)[2015-09-20]. http://www.gov.cn/zhengce/content/2014-09/10/content_9071.htm.
- [9] 袁清昌,姜媛,高巍.“科技报告”“科技报告制度”和“科技报体系”概念辨析[J].中国科技资源导刊,2015,34(3):84-87.