

电子资源建设：从 COUNTER4 到 COUNTER5 的转换分析

任 武

(上海立信会计金融学院, 上海 201620)

摘要: [目的/意义] 研究从 COUNTER4 到 COUNTER5 的转换问题, 以为图书馆电子资源建设和相关评估工作提供参考。[方法/过程] 通过概念、属性、计量等方面的对比, 分析 COUNTER5 报告所实施的根本变化, 详细介绍了基于 COUNTER4 XML 的电子资源使用统计数据通过映射的方式整合到 COUNTER5 的数据库的方法, 包括数据的收集、数据库格式的整理、查询字段的提取等几个方面。[结果/结论] 对 COUNTER5 的特性及其与旧版本差异的比较, 有助于加强标准的理解和实施工作; 所提出的数据转换分析方法, 可实现对旧版本数据的简化处理, 有利于保障馆藏数据资源的连续性和可比较性, 帮助图书馆改善及提高电子资源服务水平。

关键词: 电子资源 使用统计 COUNTER5

分类号: G250

DOI: 10.31193/SSAP.J.ISSN.2096-6695.2021.01.08

0 引言

COUNTER 的全称为 Counting Online Usage of Networked Electronic Resources, 是由图书馆界、出版界和中间商共同发起的研究项目, 目的是提供一致的、可靠的和兼容的使用标准和解决方案。2012年4月推出的第四版《COUNTER 电子资源使用统计实施规范》(简称 COUNTER4) 整合和升级了前三版的内容, 涵盖了在线期刊、在线图书、数据库、参考著作及多媒体内容等电子资源的使用统计规范^[1]。为尽可能覆盖各类资源的使用统计, 报告数量从规范制定之初的7个增至 COUNTER4 报告的27个, 与格式类型相关的计量类型增至25种, 由此带来了复杂性增加、计量不一致、术语歧义以及定义模糊等多方面的问题^[2]。为此, 第五版《COUNTER 电子资源使用统计实施规范》(简称 COUNTER5) 于2017年7月正式颁布, 并于2019年1月起生效, 进行

[作者简介] 任武 (ORCID: 0000-0001-5070-3511), 博士, 研究方向为数字图书馆建设, E-mail: renwu@lixin.edu.cn。

了全面的体系调整,在实用性和复杂性之间寻求平衡,以确保所有发布者和内容商都能实现合规性。COUNTER5采取的措施包括引入视图概念,更新报告的已有元素,增加计量维度,使用更多可启用临时报告的过滤器等,达到报告整体简化和明晰的目的^[3];同时取消了报告格式的选择,自由组合计量与属性,从而提高了定制需求的灵活性。

根据 COUNTER 项目网站给出的时间表, COUNTER4 服务时间至 2021 年 1 月^[4], 实施规范要求内容提供商从 2019 年 1 月开始提供 COUNTER5 的报告, 并同时提供过渡阶段的 COUNTER4 报告。现在全球有一百多家内容提供商遵守 COUNTER 新标准的工作^[5]。但由于 COUNTER 版本的转换属于可选项, 从国外用户的使用反馈来看, 版本转换工作并未完全实施, 一些著名的供应商如 ScienceDirect、Highwire 等, 主要是提供本年度的报告以及前两年的 COUNTER4 格式的报告^[6]。主要原因有两个方面: 一是使用许可资源的组织需要时间适应新标准的要求, 甚至一些已经以新格式生成报告的组织还尚未通过 COUNTER 审核, 报告数据可能会发生事后调整, 而原始数据的处理过程无法回溯; 二是版本之间差异过大, 计量并不完全对应, 有的只是近似值。从资源建设的角度看, 利用统计数据帮助分析馆藏资源的使用情况, 有利于制定决策计划。如果统计单位不一致, 就可能造成同类型资源无法横向比较的问题。因此, 进行新旧版本的数据转换是非常必要的。

目前在国外, 主要的研究机构对 COUNTER4 到 COUNTER5 的研究不多, 且较少涉及底层 XML 格式的变动, 例如 Usus 持续关注 COUNTER 的使用^[7], 其主要集中在前四个版本的研究上, 对 COUNTER5 仅限于内容的介绍; Oliver Pesch 领导的 EBSCO 信息服务团队为 COUNTER5 创建了 SUSHI 收割工具^[8], COUNTER4 数据则以 CSV 格式引入到数据库, 并未统一化其数据格式。该团队提供了转换的相关工具, 但主要是版本 3 和版本 4 之间的模板转换^[9], 以及将 COUNTER4 XML 输出到 EXCEL 等方面的内容; 由 Melissa Belvadi 团队开发的最新的 COUNTER5 SUSHI 收割工具, 提供了不同版本在文本格式的替换, 但转换没有涉及较为底层的 XML 格式的变动, 使用功能主要是将收割的 JSON 数据用于生成 EXCEL 报告^[10]。

在国内, 高校图书馆数字资源采购联盟 DRAA 门户于 2013 年即开始支持通过 SUSHI 协议自动获取 COUNTER 报告^[11]。杨巍等介绍了 COUNTER5 的新特征, 以及如何通过 SUSHI 将数据收割到相应的字段, 并重新设置数据库表的过程^[12]; 胡大琴等基于 COUNTER5 的草案版本分析了统计计量的变化情况^[13]。陈大庆等介绍了实现电子资源自动收割和多资源整合的综合性平台的方法^[14]。

在上述研究的基础上, 本文从 COUNTER XML 出发, 简化了数据分析和字段提取的流程; 利用工具及程序开发来实现数据的转化和 EXCEL 报表的生成; 在转换处理中融合了不同的数据格式, 从而有利于资源的整体处理和报告扩展。

1 COUNTER 实施规范的发展

COUNTER5 在概念上重新设计了结构模型, 按照概念、属性、计量三个方面, 组合不同的属性和计量指标。这种从复杂且缺乏变化的层次结构, 过渡到 COUNTER5 的多方面分类结构的处理方法是合理的。

1.1 COUNTER5 的概念变化

COUNTER5 标准确定了四个主报告 (Master Report)。其中, 平台主报告 (PR) 用于显示可以承载多个不同数据库的平台使用情况数据。数据库主报告 (DR) 用于显示数据库使用的数据, 如摘要和书目、全文和混合等。标题主报告 (TR) 用于显示与传统图书馆出版物类似的数据, 如在线期刊、在线图书、报告等。项目主报告 (IR) 则用于显示如文章、章节和注释等更多小对象的数据。各个主报告分别有相应的标准视图 (Standard View), 视图实质上是一组预定义的属性和过滤器, 如 DR 主报告对应 DR_D1/DR_D2, TR 对应 TR_J1(期刊)/TR_B1(图书) 等各视图, 这是版本设计上的一个显著变化。

1.2 COUNTER5 的属性变化

COUNTER5 修改了大部分 COUNTER4 XML 的属性, 如 counterElements4_1.xsd 框架元素 IdentifierType 中的 Online_ISSN、Print_ISSN、DOI 等项, 都被提取出来作为 COUNTER5 独立的属性, 在 AttributeType 和 DateType 中的元素则分别只保留 ArticleVersion 项和 PubDate 项; 另一部分则是新增的用于描述资源属性的计量, 如 Parent Item 与 Component Item 等项。COUNTER5 主要的属性包括数据类型 (Data_Type)、章节类型 (Section_Type)、访问类型 (Access_Type)、访问方法 (Access_Method) 和计量类型 (Metric_Type) 等五种。

其中, 数据类型包含文章 (Article)、图书 (Book)、平台 (Platform)、数据库 (Database)、数据集 (Dataset)、期刊 (Journal)、多媒体 (Multimedia) 及论文 (Thesis or Dissertation) 等元素; 章节类型包括文章 (Article)、图书 (Book)、章节 (Chapter or Section) 和其他 (Other) 等项; 访问类型包括受控 (Controlled)、金色开放获取 (OA_Gold) 等项; 访问方法有常规 (Regular) 和文本数据挖掘 (TDM) 两项; 数据类型从 COUNTER4 的 6 项扩充至 COUNTER5 的 13 项。

在 COUNTER5 主报告上, 用户通过设置视图过滤器和属性参数, 订制丰富的使用统计报告。对不同的主报告, 属性的选择范围有所不同。TR, DR 和 PR 主报告有五项共同属性, 分别是访问方法 (Access_Method)、平台 (Platform)、计量类型 (Metric_Type)、数据类型 (Data_Type) 和报告期间总量 (Reporting_Period_Total) 章节类型和访问类型等仅对 TR 和 IR 主报告视图生效, 而数据类型、访问方法和计量类型则对所有的报告视图都有效。

1.3 COUNTER5 的计量变化

COUNTER4 具有与格式类型相关的 25 种计量类型, 显得非常混乱和复杂。因此 COUNTER5 进行了更新, 主要增加了请求 (Requests) 和调查 (Investigations) 两个关键概念, 与 Access_Type、Data_Type 等属性的组合, 提供细粒度查询分析。根据内容和对各种对象的适用性原则, COUNTER5 组合后形成了 12 种计量类型。

其中, Investigations 是 COUNTER5 的新增项, 用于记录用户对内容项执行任何操作活动, 包括浏览摘要 (View abstract)、连接到链接解析器 (Link to Link Resolver)、浏览文章概要 (View article preview) 和浏览参考文献 (View cited reference) 等。该计量的引入是为了便于统计用户对某段内容以及下载内容感兴趣的部分。Requests 是 Investigations 的子集, 用于描述全文 HTML 浏览或 PDF 下载的术语, 范围仅限于资源内容本身。Total_Item 与 COUNTER4 的计量类型中的 ft_total 相似, 而 Unique_Item 则在各个平台之间具有更高的可比性。在先前版本中, 全文 HTML 视

图和一篇文章的PDF下载作为两次记录，而在COUNTER5中，使用Unique_Item作为唯一计数的请求或调查，如果用户活动在同一会话中，则全文HTML视图和PDF下载将仅计为一次。

检索(Searches)主要针对平台和数据库对象。对平台上所有内容检索采用平台检索(Searches_Platform)，而对数据库，检索方法则分三种：常规检索(Searches_Regular)表示有目的地访问数据库；联合检索(Searches_Federated)用于在多个数据库中的联合检索；自动检索(Searches_Automated)表示使用一个界面中的单个索引的发现服务。与旧版本不同，COUNTER5进行了细分：Searches_Regular和Searches_Platform两项对应于COUNTER4的“常规检索”(Regular Searches)，Searches_Federated和Searches_Automated则对应于COUNTER4的“联合与自动检索”(Searches-federated and automated)一项。

COUNTER5中列出了等效的版本度量比较类型^[15]，本文归纳为计量转换对照关系(Metric Transition Equivalent Relationships)，如表1所示。

表1 新旧版本报告和计量类型对照关系

COUNTER4 报告	COUNTER5 标准视图	COUNTER4 计量类型	对应的 COUNTER5 计量类型
PR1	PR_P1	常规检索 (Regular Searches) 联合与自动检索 (Searches-federated and automated) 结果点击 (Result Clicks) 记录视图 (Record Views)	检索平台 (Searches_Platform) 自动检索 (Searches_Automated) 总访问量 (Total_Item_Investigations) 总项目请求量 (Total_Item_Requests)
DR1	DR_D1	常规检索 (Regular Searches) 联合与自动检索 (Searches-federated and automated) 结果点击或记录视图 (Result Clicks or Record Views)	常规检索 (Searches_Regular) 自动检索 (Searches_Automated) 总访问量 (Total_Item_Investigations)
JR1, JR1GOA JR5, JR1a BR2	TR_J1, JR_J3 TR_J4 TR_B3	总访问数 (ft_total) Pdf 访问数 (ft_pdf) Html 访问数 (ft_html)	总项目请求量 (Total_Item_Requests) 唯一项目请求量 (Unique_Item_Requests) 忽略 (just ignore this metric)
JR2, DR2 BR3	TR_J2, DR_D2 TR_B2	超量 (limit exceeded) 访问拒绝 (Access_Denied)	超量 (Limit_Exceeded) 未授权 (No_License)
BR1	TR_B1 (Excluding GOA)	总访问数 (ft_total) Pdf 访问数 (ft_pdf) Html 访问数 (ft_html)	唯一标题请求量 (Unique_Title_Requests) 唯一项目请求量 (Unique_Item_Requests) 忽略 (just ignore this metric)

1.4 COUNTER XML 与 JSON 的格式比较

COUNTER4 是利用 SUSHI 协议传输数据，以 XML 的格式返回到服务器。带有 SUSHI 名称的模式一般是 XML 格式的 COUNTER 报告的检索封装，实际上 COUNTER XML 报告和 SUSHI 协议是可以分开使用的。XML 的内容通过其 XSD 定义来保证严格性。COUNTER XML 支持报告的完整范围，包括期刊、数据库、图书和多媒体等项目使用情况的报告。在 XML 中必须包含 ReportItem、ItemDataType 等项，XML 的具体数据存储在 Instance 中。

与 COUNTER XML 不同，COUNTER5 SUSHI 采用 RESTful 接口的方式返回 JSON 格式的报告。JSON 报告分为表头和与正文相关的报告项两个部分。具体使用数据和计量类型显示

在 Performance 节点下的 Instance 结构中, 对于计数为零的实例项、没有 Instance 的项及没有 Performance 项的 Report_Items 都将被省略。与 XML 一样, 在 COUNTER5 JSON 中也不得包含报告总数。COUNTER5 托管 Swaggerhub API 文档是一个全面的参考版本, 其中包含整个 COUNTER_SUSHI API 的详细说明。Swagger 文档包含了 PR/DR/TR 等 JSON 格式的示例^[16]。

从结构上看, COUNTER4 XML 和 COUNTER5 JSON 都包含了根节点, 呈现相似的树状结构, 易于实现数据的转换。由于在 COUNTER4 和 COUNTER5 版本的各对应节点中, 存在缺失和名称差异, 所以采用数据映射和生成报表比较的处理方式。

2 COUNTER 数据转换分析

美国国家信息标准组织 (National Information Standards Organization, 简称 NISO) 网站提供了用于 COUNTER4 报告的 SUSHI Starters 客户端测试服务^[17]。在该服务页面选择需要下载的报告名称和起止时间, 即获取 COUNTER4 的 XML 格式数据。本实验下载了从 2016 年至 2019 年的 JR1/JR1GOA/JR5 等项报告的数据, 以年为单位分别编号保存至文件目录, 对照建立起相应的数据库表。

2.1 XML 与 JSON 转换成数据库表

COUNTER4 的 XML 数据在转换时, 其各节点的具体数据都需要保留。COUNTER5 的 JSON 则只需要保留其节点名称及关联即可。实验首先采用编辑工具 Altova XMLSpy, 该工具拥有直观的用户界面和种类丰富的 XML 编辑视图和选项, 可以直接将 XML 和 JSON 转换成关系数据库主从表的关联的格式, 保留除了主键、外键的所有字段。COUNTER4 XML 中没有的字段, 例如 YOP, ACCESS_TYPE 等, 则通过 SQL 语句将值填充到数据库表的对应项中。

2.2 COUNTER4 各组数据映射到 COUNTER5 的关系表

实验采用工具 Alt ova MapForce 实现数据映射, 该工具通过映射数据源的方法, 实现多种文件之间的转化。本实验将不同年份和格式的 COUNTER4 的数据表, 通过字段关联以追加的方式转换到对应的 COUNTER5 的数据库表中, 如图 1 所示:

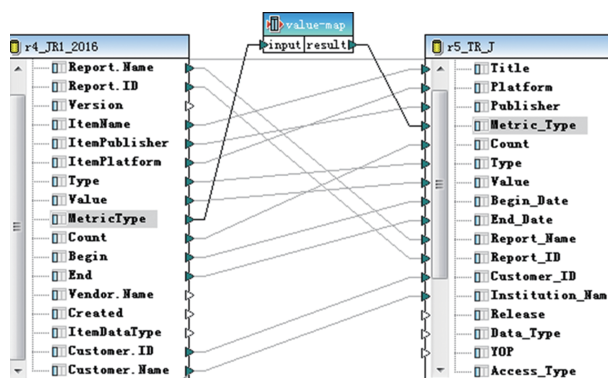


图 1 COUNTER 数据库期刊字段映射

其中不同版本的计量类型字符串转换是在转换工具的 value-map 控件中设置的, 分别在该控件的输入和结果栏中填写本文表 1 中的对应计量值, 对于无法进行转换的选项, 如从 COUNTER4 转换来的 ft_html, ft_ps 等无效项, 则选择 Otherwise 并在结果栏中输入“NULL”, 以便在后续处理步骤中删除。

2.3 生成报告的程序实现界面

将 COUNTER4 XML 数据转换成 COUNTER5 格式后, 下一步将转换的数据以 COUNTER5 的标准 EXCEL 报表输出。为此, 实验下载了 COUNTER 的模板文件并保存在设定的位置, 并利用程序开发了使用平台界面, 生成报告的主要步骤如下:

// 根据查询记录集的大小确定要输出的单元格的选择范围

Do Until recordset.EOF: oApp.Rows (row).Select:

oApp.Selection.Insert:recordset.MoveNext: Loop: recordset.MoveFirst

// 根据选择范围, 将查询记录集数据循环输入至 EXCEL 单元格

Do Until recordset.EOF: For i = 1 to FieldNum

oApp.Cells (InitRow, i).Value = rst.Fields (i - 1)

Next I: rst.MoveNext: InitRow = InitRow + 1: Loop

// 按照 COUNTER5 模板的设定, 除表头外初始行设为 15, 调用报告生成函数

Call tr_excel (15, strSet, comboYOP.Value, ComboReportID.Value, strWhere (), 10)

MsgBox (“COUNTER5 EXCEL 报表已生成”)

2.4 TR_J1 与 JR1/JR1GOA 比较

COUNTER5 的 TR_J1 报告与 COUNTER4 的 JR1 报告相比, TR_J1 仅限于 Controlled 使用量, 因此在转换中, 需要排除 OA_Gold 的使用量。等式为: TR_J1 使用量 = JR1 使用量 - JR1GOA^[18]。在实验中根据设定的模板分别生成了 TR_J1, JR1 和 JR1GOA 的 EXCEL 报告 (见图 2)。核对各组数据均符合上述等式, 证明了所实施的转换方法的正确性。从 COUNTER5 报告格式看, 与 COUNTER4 报告的差异较大, 例如没有期刊总数这一行, 没有 HTML 及 PDF 的期间总数值, 而是用计量类型的“唯一项目请求”近似值代替。

Report Name	Journal Requests (Excluding OA_Gold)									
Report ID	TR_J1									
Release	5									
Title	Publis	Platform	Print ISSN	Online ISSN	Metric_Type	Reporting_Period	Jan-2017	Feb-2017	Mar-2017	
ACM Computing Surveys (CSUR)	ACM	ACM Digital Library	0360-0300	1557-7341	Total_Item_Requests	294	8	14	8	
ACM Computing Surveys (CSUR)	ACM	ACM Digital Library	0360-0300	1557-7341	Unique_Item_Requests	292	8	14	8	
ACM Inroads	ACM	ACM Digital Library	2153-2184	2153-2192	Total_Item_Requests	28	4	8	0	
ACM Inroads	ACM	ACM Digital Library	2153-2184	2153-2192	Unique_Item_Requests	26	4	8	0	
ACM Proceedings	ACM	ACM Digital Library		2168-4081	Total_Item_Requests	3274	260	214	265	
ACM Proceedings	ACM	ACM Digital Library		2168-4081	Unique_Item_Requests	3339	270	226	264	
Journal Report 1 (R4)	Number of Successful Full-Text Article Requests by Month and Journal									
<Customer>										
Journal	Publis	Platform	Print ISSN	Online ISSN	Reporting_Period	Reporting_Period	Reporting_Period	Jan-2017	Feb-2017	Mar-2017
Total for all journals	ACM	ACM Digital Library			7628	469	7154	1056	1376	1194
ACM Communications in	ACM	ACM Digital Library	1932-2240		7	0	7	0	0	0
ACM Computing Surveys	ACM	ACM Digital Library	0360-0300	1557-7341	296	0	296	16	28	16
ACM Inroads	ACM	ACM Digital Library	2153-2184	2153-2192	28	2	26	8	16	0
ACM Letters on Program	ACM	ACM Digital Library	1057-4514	1557-7384	2	0	2	0	0	0
ACM Proceedings	ACM	ACM Digital Library		2168-4081	3343	1	3339	540	452	530
ACM SIGACCESS Acce	ACM	ACM Digital Library	1558-2337	1558-1187	8	0	8	4	0	0
Journal Report 1 GOA (R4)	Number of Successful Gold Open Access Full-Text Article Requests by Month and Journal									
<Customer>										
Journal	Publisher	Platform	Print ISSN	Online ISSN	Reporting_Period	Reporting_Period	Reporting_Period	Jan-2017	Feb-2017	Mar-2017
ACM Computing Surveys (CSUR)	ACM	ACM Digital Library	0360-0300	1557-7341	118	24	39	16	56	
ACM Computing Surveys (CSUR)	ACM	ACM Digital Library	0360-0300	1557-7341	2	0	4	0	0	
ACM Proceedings	ACM	ACM Digital Library		2168-4081	69	0	0	10	12	
ACM Transactions on Comput	ACM	ACM Digital Library	0734-2071	1557-7333	4	0	4	0	4	

图 2 生成的 TR_J1/JR1/JR1GOA 的 EXCEL 报表比较

2.5 TR_J3 与 JR1/JR1GOA 比较

COUNTER5 的 TR_J3 报告包含了“Controlled”和“OA_Gold”两部分内容。TR_J3 的总使用量等于 COUNTER4 的 JR1 的使用量。由于按照访问类型区分 TR_J1 与 TR_J1 GOA 的整合, COUNTER5 在报告上比 COUNTER4 多出 Access_Type 部分。

2.6 TR_J4 与 JR5 报告比较

COUNTER5 的 TR_J4 报告与 COUNTER4 的 JR5 报告都包含年度汇总部分, 在格式上具有较高的可比性。但 TR_J4 仅限于 Controlled 使用量, 并不包括 OA_Gold 访问量, 而 JR5 则包含所有部分的内容, 所以两者的数字并不匹配, 且 JR5 报告不包含按每月的统计。TR_J4 报告的转换更多的是来自 JR1/JR1GOA 部分, 并具有 JR5 所不具备的功能, 比如按月统计期刊的下载次数等。

2.7 COUNTER 报告的扩展

基于已经建立的数据库表, 用户可以设计出具有所需属性和计量类型任意组合的报告, 例如查询图书馆的按出版年限使用的免费和有偿获取期刊文章的报表情况, 即形成按出版年份划分的使用情况(“Controlled”和“OA_Gold”)的数据透视表。在开发界面上, 选择访问类型、计量类型、出版年和总数后形成的透视图, 其中用实线表示 Controlled 期刊, 对应左边的主纵坐标; 用虚线表示 OA_Gold 期刊, 对应右边的次纵坐标(见图 3)。可以看出, 图书馆近年来在所有期刊中使用 Controlled 获取文章的数量略有减缓, 而 OA_Gold 获取的请求数量正在增加, 表明金色开放访问或将成为电子资源建设的一种趋势。

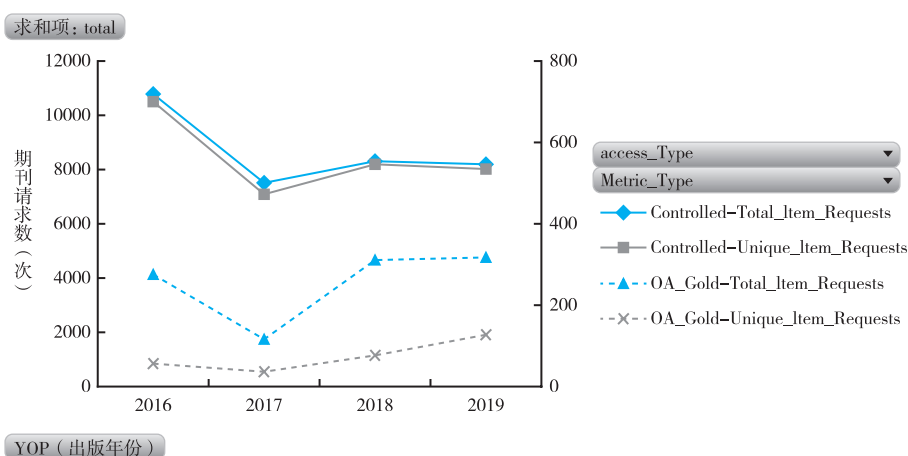


图 3 按出版年份划分的期刊请求数比较

除获得所有期刊的使用情况外, 用户还可以比较某一平台下各个数据库商的期刊下载情况, 而这些数据在相应的标准报表中并不直观。例如查询 ACM Digital Library 平台下的各年度 PDF 的下载排名情况, 通过在查询字段中选择出版商、出版年及计量类型的“Unique_Item_Requests”获得图形表示。从图 4 可见, 包含“OA_Gold”在内的 ACM 平台下的各出版商的 PDF 下载量是逐年递增的。通过结合数据库的年度采购金额除以其年度总下载量以计算其单篇平均成本, 进而与其他同类数据库比较以帮助采购评估分析。

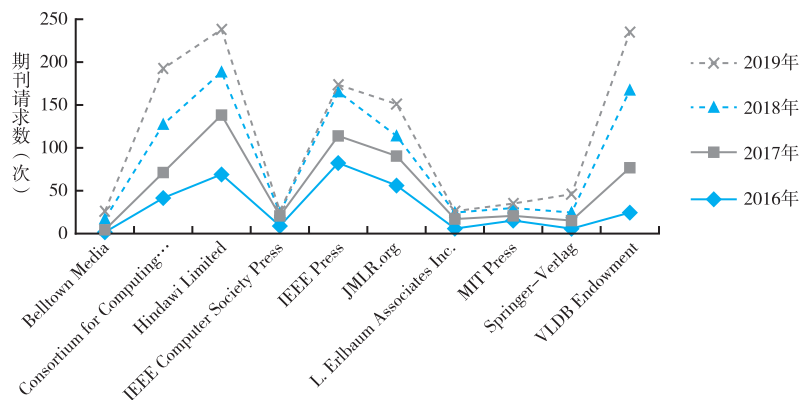


图4 2016 ~ 2019年ACM平台的各数据库商下载数比较

3 讨论

3.1 COUNTER 报告转换的意义

转换的意义有两方面：一方面由于 COUNTER4 将不再应用，这表明今后的 COUNTER 报告都将以 COUNTER5 的标准格式出现，但以往的 COUNTER4 数据报告仍然具有参考价值。另一方面由于在报告字段和计量指标方面差别较大，而原始数据处理应用无法回溯，会造成数据的比较无法实现，因此需要进行 COUNTER 版本数据的转换。

3.2 COUNTER 报告转换的不足

如前所述，COUNTER4 XML 数据和报告格式是分离的，例如，COUNTER4 期刊 JR1 报告中只包含 HTML 和 PDF 计数，XML 中包含每月的细目分类，而 COUNTER5 保持了格式和报告之间的一致性。因此版本间的报告和计量单位都差别较大，部分计量并不完全对应，有很多只是近似值，例如 COUNTER5 的新的计量单位 Unique_Item_Requests，使用情况是独立于所交付的格式进行计数的，其数量大于 COUNTER4 中看到的 PDF 数量和 HTML 数量，在数值上近似采用 ft_pdf 值，而 ft_html 则在 COUNTER5 中忽略。

3.3 COUNTER5 报告的优势

数据转换后，COUNTER5 报告格式的好处是，比较和合并的灵活性更高，并且由于 COUNTER5 所有报告都将采用相同的结构，使得报告处理过程变得简化。另一个好处是分析报告变得容易。例如 COUNTER4 的报告中，如果要计算已付费内容的每次使用成本，则必须检索 JR1 和 JR1GOA 两个报告，并且必须从 JR1 的总数中减去 JR1GOA 的总数。而转换成 COUNTER5 TR_J1 的标准视图后，TR_J1 不包括 OA_Gold，这样可以更容易对受控期刊的内容进行使用费用的计算。在本文转换中，通过添加 SQL 语句到映射工具完成， $UPDATE (SELECT * FROM tr_j1 WHERE Report_ID = 'JR1') AS a, (SELECT * FROM tr_j1 WHERE Report_ID = 'JR1 GOA') AS b SET a.count = a.count - b.count WHERE (a.Metric_Type = b.Metric_Type) AND (a.Value = b.Value) AND (a.Begin_Date = b.Begin_Date)$ 。

对于混合期刊, 其粒度级别处于文章级别。因此本文生成 COUNTER5 TR_J3 视图报告, 其访问类型包含用于 Controlled 文章和 OA_Gold 文章两个部分。而 COUNTER5 TR_J4 可以用于查询各个年度付费的总请求数和唯一请求数, 以及各个月份的使用量情况; 而在 COUNTER4 JR5 部分, 则不能够分辨付费使用量的情况, 且不能够查看不同月份的用量。

3.4 COUNTER5 报告的扩充

COUNTER5 报告领域将扩大到包括研究数据报告, 包括将 COUNTER 统计报告和图书馆使用成本明细表项结合, 创建相对简单的每次使用费用分析等方面。因此, 可以在报告订制中补充成本字段, 从成本明细工作表中提取成本数据; 与以 KBART 文件的形式提供的机构的馆藏数据结合^[19], 在两个报告上使用相同标识符或 ISSN/ISBN, 以促进报告准确匹配, 从而更精确地表示使用情况和持有量。

4 结论

COUNTER 标准实际上已成为大多数图书馆、出版商和数据库商所共同遵循的电子资源行业准则。为了适合需求变化和具备可扩展性, COUNTER5 对以前版本进行了根本的修订, 表现为其概念及计量等更加抽象, 这也为版本数据的转换工作带来了困难。本文从 COUNTER4 XML 出发, 简化了字段提取的分析流程, 实现了预定的转换功能。从实验效果上看, 还有待进一步总结和完善。仅此抛砖引玉, 以期为图书馆电子资源建设和相关评估工作提供参考。

【参考文献】

- [1] 李洪. 新版 COUNTER 的特征及未来发展 [J]. 中国图书馆学报, 2012(6): 29-37.
- [2] PESCH O. COUNTER: Looking Ahead to Release 5 of the COUNTER Code of Practice [J]. The Serials Librarian, 2016, 71(2): 83-90.
- [3] 陈越, 都平平, 江艳霞. COUNTER 5: 关键变化及应用展望 [J]. 图书馆论坛, 2020, 40(1): 132-139.
- [4] Transition Timeline and Transition Options [EB/OL]. [2021-3-25]. <https://www.projectcounter.org/transition-timeline-transition-options/>.
- [5] Organisations release 5 compliant and working toward compliance [EB/OL]. [2021-3-25]. <https://www.projectcounter.org/about/register/>.
- [6] Usage Report Issues [EB/OL]. [2021-3-25]. <https://www.projectcounter.org/usage-report-issues/>.
- [7] The community website for issues with library usage reports [EB/OL]. [2021-3-25]. <https://www.projectcounter.org/usus-website-where-you-will-find-the-usage-report-issues-the-usus-community-has-discussed-and-resolved-so-far/>.
- [8] COUNTER R5 SUSHI HARVESTER FOR MULTI-SITE LIBRARIES [EB/OL]. [2021-3-25]. <https://www.projectcounter.org/counter-harvester-tools/>.
- [9] Osterman A, Pesch O, Schmidt K. Introduction to Usus, a Community Website on Library Usage, and a Discussion about COUNTER 4 [J]. The Serials Librarian, 2016, 70(1-4): 211-216.
- [10] COUNTER-5-Report-Tool [EB/OL]. [2021-3-25]. <https://github.com/CS-4820-Library-Project/COUNTER-5-Report-Tool>.
- [11] 高校图书馆数字资源采购联盟 [EB/OL]. [2021-3-25]. <http://www.libconsortia.edu.cn/index.action>.

[12] 杨巍,叶仁杰,吴元业,等. COUNTER Release 5 的新特征及其应用研究 [J]. 大学图书馆学报, 2020, 38(1): 18-25+41.

[13] 胡大琴. 中文数字资源使用统计数据的调查研究 [J]. 图书情报工作, 2017, 61 (15): 104-110.

[14] 陈大庆,叶兰,杨巍,崔晓松. 电子资源使用统计平台 USSER 的设计与实现 [J]. 图书情报工作, 2015, 59(1):106-112.

[15] Transitioning from COUNTER R4 to R5 [EB/OL]. [2021-3-25]. <https://www.projectcounter.org/code-of-practice-five-sections/13-transitioning-previous-releases-new-reporting-services/#fourtofive>.

[16] The SUSHI examples on the SwaggerHub website [EB/OL]. [2021-3-25]. https://app.swaggerhub.com/apis/COUNTER/counter-sushi_5_0_api/1.0.0

[17] SUSHI Starters Client for COUNTER Release 4 - Quick Online Test [EB/OL]. [2021-3-25]. <https://www.niso.org/standards-committees/sushi/sushi-tools>

[18] BARBERS I. Using COUNTER Release 5 Usage Reports for strategic decision making [C]. Proceedings of the UKSG 42nd Annual Conference Telford, 8-10 April 2019

[19] RESCH O. E-Resource Standards You Should Know About [J]. The Serials Librarian, 2011, 61(2), 215-230.

E-Resources Construction: Analysis of Transition from COUNTER4 to COUNTER5

REN Wu

(Shanghai Lixin University of Accounting and Finance, Shanghai 201620, China)

Abstract: [**Purpose/significance**] This paper studies the construction of COUNTER5 electronic resource statistical data through data transition from the old version. [**Method/process**] Through the comparison of concepts, attributes and metrics, it analyzes the fundamental changes implemented in the COUNTER5 report and introduces in detail the method of integrating electronic resource usage statistics based on COUNTER4 XML into the COUNTER5 database through mapping, including aspects such as data collection, database formatting, query field extraction, etc. [**Result/conclusion**] (1) The comparison of the characteristics of COUNTER5 and the old version will help to strengthen the understanding and implementation of the COUNTER5 standards. (2) The proposed transition analysis approach can simplify the processing steps of the old version data, which is conducive to the promotion of continuity and comparability of library data resources, as well as to help libraries improve their electronic resource services.

Keywords: Electronic resources; Usage statistics; COUNTER5

(本文责编: 孔青青)