第三代图书馆服务平台中的数据安全研究*

赵想飞 常颖聪 黄闽 崔增辉
（河北师范大学图书馆，石家庄 050024）

摘要：随着云服务理念的深入和大数据、云计算、人工智能等新技术在图书馆领域的广泛应用，大数据收集及分析和云端部署为第三代图书馆服务平台的运行提供了数据基础，但同时增加了数据安全风险。本文从第三代图书馆服务平台的特点出发定性分析数据安全的风险因素，并针对数据采集、数据存储、数据访问与利用、数据共享与更新及数据安全评估各阶段分别提出相应的数据安全措施，以便更好地提升图书馆的管理和服务水平。

关键词：第三代图书馆服务平台；数据安全；数据采集；数据安全评估
中图分类号：G250.7；G252.0　DOI：10.3772/j.issn.1673-2286.2021.08.006
引文格式：赵想飞，常颖聪，黄闽，等. 第三代图书馆服务平台中的数据安全研究[J]. 数字图书馆论坛，2021（8）：39-44.

大数据、云计算及人工智能等新技术给图书馆带来了各种应用需求的机遇，要求图书馆以知识化和智慧化的形态来适应巨变形势下的信息环境。为了更好地履行职责，图书馆的业务管理系统需要不断地适应数字资源管理和读者服务需求的变化。在这种趋势的引领下，图书馆管理系统由最初以馆藏管理为中心、以OPAC为代表的第一代图书馆自动化系统，发展到“藏用一体”并使图书馆服务标准化的第二代图书馆集成管理系统（Integrated Library Management System，ILS），该系统虽然能够实现资源导航与发现、文献传递及新媒体服务等功能，但是难以有效地对数字资源进行集成管理，也无法按照知识信息流的规律去开展有效的知识服务，同时也存在跨平台信息孤岛等诸多问题，难以满足图书馆深度专业的知识化和智慧化发展需求。基于这一趋势，第三代图书馆服务平台（Library Service Platform，LSP）应运而生，这一概念由Breeding[1]在2012年的报告中提出，很快得到图书馆界的认同，并传播开来。他认为LSP应包括对印刷和电子资源的纸电合一的全媒体管理，支持业务和服务的全流程统一管理，具备全球知识库，以多租户的方式提供SaaS服务，通过APIs支持图书馆互联互通和系统扩展，是将现有集成管理系统、电子资源管理工具以及数字资产管理系统整合在一起的平台[2]。

除以上特点外，肖铮等[3]认为LSP的主要特征还具有云端化，借助云服务的高可用性、高扩展性和通用性，实现对图书馆各类数据的细致深入地分析，为图书馆的发展提供决策支持；张磊等[4]认为LSP通过可视化数据分析工具深度分析平台上的海量数据，建立大数据分析平台，帮助图书馆进行用户监测分析和服务优化决策；施晓华等[5]认为LSP是在公有云和私有云协同的混合架构下构建的，要充分考虑平台上的数据安全。可以说LSP不再仅仅是一个简单的系统，而是一个以“数据驱动”为目标的“互联网+云平台+大数据”的共赢的开放性学术生态系统，数据安全是平台建设要考虑的重要问题。

1　第三代图书馆服务平台的数据安全问题分析

数据安全是图书馆提升管理效率和服务质量的基础，是支撑图书馆应用LSP的前提。国际标准化组织（ISO）对计算机系统中数据安全[6]定义为：数据安全

* 本研究得到河北省高等学校人文社会科学研究项目“微服务架构下第三代图书馆服务平台数据安全研究”（编号：SQ201035）资助。

主要分为数据本身的安全、数据防护安全及数据存储安全三方面内容，并采用相应的技术和安全保护措施以保护数据不被偶然与恶意的原因遭到破坏、更改与泄露。2021年6月颁布的《数据安全法》[7]规定："数据安全，是指通过采取必要措施，保障数据得到有效保护和合法利用，同时明确了组织、个人在开展数据活动中对数据安全要承担相应的保护义务和责任"。保障图书馆的各类数据安全是图书馆向服务智能化转型要考虑的首要问题。

1.1 第三代图书馆服务平台应用过程中存在数据安全风险

LSP的特点决定了其在部署应用过程中必然会涉及大量的用户信息、隐私及知识产权问题。随着大数据中心和云平台成为网络攻击的主要目标，图书馆如何保障生产并存储在云平台上的数据安全是面临的重要问题。虽然图书馆应用LSP后会在信息处理与利用方面提高工作绩效，但是不可避免地会带来数据安全风险。因此，做好图书馆数据安全工作，对服务质量的提升至关重要。

LSP在图书馆应用过程中，会产生海量数据，可分为以下两种类型。①业务管理数据，例如：读者数据（身份信息、行为数据等）；数字资源（各种数据库、电子书、特色库）的利用数据，同时还包括版权信息、使用范围、利用率等方面的数据；图书管理系统中的馆藏数据（入馆时间、分类、元数据、闭/开架位置等书目信息）。②服务数据，如参考咨询命题/项目、馆际互借、查引查新、专题、阅读推广活动等数据。

LSP会深入分析、挖掘数据，具有全面化、自动化、深度化及动态化等特征。第一，LSP尽可能全面、实时地采集读者的各种行为数据，除了借阅数据、入馆数据，还包括读者在馆的运动轨迹数据、信息检索数据、读者空间设施使用数据，甚至还有读者参与图书馆各种活动的数据，进而绘制出多维度的"读者画像"，而这些数据具有敏感性，需要对其访问过程、方式及后续使用进行合理限制，以保障图书馆数据安全；第二，图书馆对读者数据的采集程度变得非常细致，数据的形式不再局限于文本形式，也可以是图像、声音和视频；第三，LSP能够将OPAC系统、空间设施管理系统、门禁系统和数据库管理系统等实现互联互通，各种数据的集中管理和相关性分析能最大化地挖掘数据价值，但同时增加了数据安全的风险。

1.2 新技术安全风险

尽管LSP给图书馆应用环境和服务方式带来深刻变革，但是云计算的引入使图书馆大量数据上传至云服务器中，导致图书馆面临云服务信息安全的威胁，影响数据的安全性及保密性。云服务提供商提供的数据安全服务协议中只有概括性的规定，没有具体的数据安全方案，这也会给图书馆数据安全存储带来许多未知风险。区块链技术虽然能够提高数据存储的安全性，但是其自身缺乏体系化安全防护，其全量备份的机制也容易遭遇到存储瓶颈。

因此，在分析LSP本身安全风险的基础上，再结合相关技术的安全风险分析，可以看出，LSP平台下的数据安全保障策略非常必要，能够帮助图书馆制定数据安全制度、框架、政策及服务流程[8]，进而更好地促进图书馆在管理、服务、建设等方面的智慧化发展。

2 第三代图书馆服务平台的数据安全方案

国内很多学者针对图书馆数据安全工作进行了深入研究，试图寻求一些关于数据安全的解决方案。例如：周秀霞等[9]提出用Five Safety安全框架来规划图书馆数据的安全访问，对数据进行敏感度分级，提高图书馆数据安全访问水平；梁俊荣[10]从图书馆信息系统安全管理方面着手，设计了图书馆安全风险识别与管理系统用于提高图书馆数据的安全性；万映红等[11]对智慧图书馆个人数据安全存在的问题进行解析，提出了保护个人数据安全的解决方案；张娟等[12]提出了完善信息安全管理和加强馆际协作来应对图书馆安全风险的方案。

图书馆数据安全一直贯穿于数据的采集、存储、访问与利用、共享及更新等数据活动的全过程。但是，在不同阶段，图书馆对数据的应用与数据安全的保护期望不同，因而，不同阶段采用不同的措施能够更详细地制定数据安全方案（见图1）。

2.1 数据采集阶段

在数据采集阶段，LSP会尽可能地细致、全面并动态地采集数据，这些数据主要包括两种类型[13]：一是

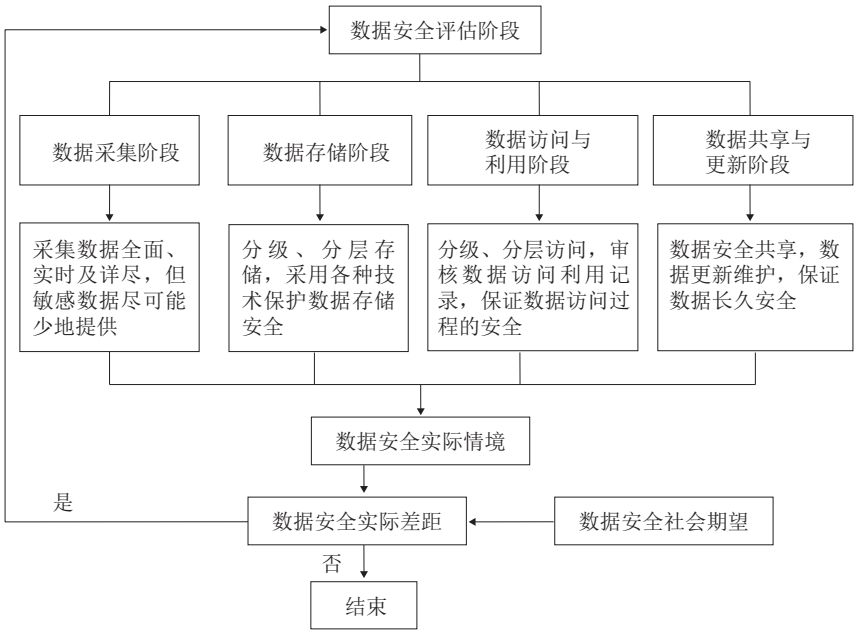

图1　数据安全方案

读者身份数据和行为数据，如读者身份信息、进馆数据以及座位预约数据；二是资源数据，主要包括文本数据、数值型数据、数据库记录、媒体文件、检索历史，或其他数据及信息。数据采集的目的是为图书馆后期的数据分析和利用提供更多的参考依据。

（1）制定相应的数据安全政策。随着网络安全事件的频发，读者对数据安全就更敏感，对信息保护的意愿就更加强烈。读者在披露如身份证号、学工号等个人敏感信息时变得越来越谨慎。据《2019全国网民网络安全感满意度调查统计报告》[13]显示：37.4%的网民认为个人信息泄露的比较多，网络安全态势仍然很严峻；50.15%的用户遭遇过信息泄露的问题，个人信息安全的状况仍需改善。为了保证各类数据安全，图书馆应该告知读者收集的内容、目的以及关于数据安全的政策和承诺。目前，国内大部分图书馆都没有明确地与读者就个人数据采集和安全达成协议[14]，也很少在图书馆网站上发布关于数据安全的管理规则。同时，云服务提供商通过相关技术很容易采集到图书馆的敏感数据，甚至拥有采集数据的超级权限，一旦管理失控，很容易带来巨大的数据安全风险。因此，图书馆应该出台相应的政策和措施来弥补这方面的漏洞。

（2）根据内容和目的确定不同的采集方式。第一，对于为满足图书馆优化资源、管理及服务需求而采集的凸显读者群体行为的数据，尽可能采用匿名化、模糊化的方法进行数据处理。如图书馆管理系统中的借阅量、资源使用量以及远程访问量、下载量的分析数据等都是按照时间轴来呈现借阅量、访问量和下载量等读者使用信息的[15]，通过数据脱敏技术对读者敏感数据进行模糊处理。第二，面向读者采集带有明显身份特征的数据时，如学工号、院系以及访问数据库的时间和IP地址等，图书馆应该根据不同的数据应用环境，制定不同的数据安全策略，提前做好数据安全规划。

2.2　数据存储阶段

基于LSP的云部署，其所采集的数据由本地存储转移到云平台存储，这意味着图书馆不再独自享有数据，而云服务提供商也拥有这些数据，图书馆的数据安全面临很大的威胁。虽然图书馆会尽可能地保证数据安全，但是在实际操作过程中会受到各种因素的影响，如数据存储设备大小、安全技术储备等。因此，现阶段可以采用以下三大措施来保障数据安全。第一，根据重要性及安全等级等指标，采用不同的方式进行分级、分布存储。涉及读者身份、行为、图书馆核心资产的数据（专题活动数据、特色库数据），需要本地存储。第二，可以将一些安全级别低的数据进行加密脱敏存储在公有云平台上，借此降低存储成本。例如，关于借阅量、数据库访问量、空间环境数据等表征读者群体行为的数据。第三，图书馆应参照国家发布的《信息安全技术网络安全等级保护基本要求》《信息安全技术网络安全等级

保护测评要求》等网络安全领域的国家标准[16]，引入Web防火墙软件、网络入侵检测报警系统及采用网络隔离等手段对主要业务系统进行分区分域和定级防护，建立一个“打防管控”的一体化网络安全防护体系来保障数据存储安全。

2.3 数据访问与利用阶段

数据访问与利用是图书馆、云服务提供商在管理和服务中对数据进行提取、分析，向读者推荐其感兴趣的主题，统计和识别读者并分析读者行为的过程，以便更深入地挖掘数据价值。在此阶段，图书馆需要做出以下3个方面的努力。

（1）建立明确的数据访问与利用制度。数据安全事件大多数情况下是由人为造成的，例如，学生身份证号的泄露，导致学生信息被企业冒用虚报个税，引起法律纠纷[17]。图书馆需要从制度上制定一个数据访问与利用的可行规范来保护各类数据安全，形成数据安全把控机制。一方面，图书馆应规范工作人员访问与利用数据的行为，对相关人员进行定岗定责，本着“谁使用谁负责”的原则，根据岗位需求分配相应的权限，落实相关人员的主体责任，尤其是对数据安全人员录用及转岗等进行规范化管理[18]；另一方面，可以将动态机器人安全防御技术应用在图书馆数据安全中，变被动防御为主动防御，使数据安全自动化和工具化，可以全时段、全方位地保护数据安全，保障图书馆与读者、云服务提供商及数据商之间在数据访问与利用过程中的数据安全。

（2）审核数据访问和利用记录。在大批量访问、调取及分析图书馆核心数据时，一方面，图书馆应根据工作人员的权限执行相应的审核程序，对馆员的行为加以限制，同时对数据访问、利用的时间、责任人、用途及去向做好记录；另一方面，图书馆应就数据的访问与利用和云服务提供商签订相应的访问控制协议，如数据访问主体的身份验证、访问及利用数据时是否告知并获得图书馆数据授权书等。只有这样才能保证数据不被滥用和泄露。

（3）分级与分布访问数据。基于数据的分级与分布存储特性，图书馆和云服务提供商可以对数据进行分级与分布访问。第一，对于安全级别高且涉及读者隐私及图书馆核心财产的数据，只有获得权限的图书馆工作人员才能访问，这些数据对云服务提供商不开放；第二，对于表征图书馆群体性质的管理与服务方面的数据，可以供云服务提供商访问及利用，这样做可以最大限度地保障数据安全。

2.4 数据共享与更新阶段

图书馆资源、管理、服务方面的数据可以为图书馆的智能化和个性化服务提供数据支持，也可以为第三方机构提供改进和评估产品服务质量的数据依据。

因此，图书馆需要建立与读者、第三方机构数据共享机制，在数据共享过程中必须遵循数据合理使用原则，综合运用可靠的安全管理技术来保障数据安全。伴随着数据应用及读者信息的变化，数据的残余价值也会发生阶段性降低，因此为了降低数据运维成本，需要定期地对过期及冗余数据进行更新维护。

2.5 数据安全评估阶段

数据安全评估是从数据安全的角度出发，分析数据活动全过程的安全风险，使风险可视化、可控化，提升图书馆数据安全防护水平，实现数据安全的规范化和精细化管理。

在数据采集阶段，图书馆应关注采集环境、采集行为、采集传输及采集管理等方面的风险，并定期进行有针对性的安全评估，制定相应的数据安全评估指标。例如：采集人员的权限和角色是否明确；采集行为是否规范；采集后数据传输是否加密；采集的数据是否进行安全分级处理。

在数据存储阶段，需要关注数据存储环境安全、数据存储加密、数据存储空间分级分布、数据存储访问控制、数据容灾备份与恢复等安全问题，并针对这些安全风险定期评估。

在数据访问与利用阶段，需要评估以下风险：①云平台对人员认证和权限管理不当，导致非法用户越权访问数据；②云平台本身缺乏敏感数据发现与识别机制，导致敏感数据在分析后泄露[19]；③云平台的安全评估及审计手段缺失，导致无法有效监督用户访问行为，增大敏感数据泄露的风险，给数据所有权人造成巨大的经济损失和社会不良影响。

在数据共享与更新阶段，需要评估以下风险：①在数据共享过程中，云服务提供商直接将未加密或未脱敏的数据传输给第三方机构共享；②云服务提供商与第三方机构共享图书馆数据，是用于改进自身的服务或产品

还是有其他用途，图书馆和云平台需要在签订使用服务协议时应该明确用途；③在数据生命周期结束后，数据未被彻底更新维护，还存有敏感数据的残余介质。

LSP正在图书馆领域如火如荼地应用中，通过分析以上数据活动各阶段安全风险，明确了需要评估的内容及指标，希望逐步缩小数据安全风险的现实情况与社会期望之间的差距，对于构建全方位防御的数据安全保障体系起到一定的作用。

3　推动第三代图书馆服务平台数据安全的建议

3.1　健全数据安全的法律法规

现阶段，我国对于数据安全保护的法律制度仍然不够健全。近年来我国先后出台了《网络安全法》《公共图书馆法》《数据安全法》等法律，落实了数据安全的主体责任，明确了数据安全对于国家安全的重要性，但是在如何保障图书馆等公益性单位的数据安全方面却描述得不详细。为了保证LSP在图书馆的顺利运行，图书馆应制定符合本馆的数据安全规定，避免将来因数据泄露或数据产权而产生不必要的纠纷。相关的数据安全管理规定应涉及以下3个方面：①图书馆应当明确数据活动各阶段的范围、用途，数据安全所采用的技术及手段，数据公开的格式等，在网站上发布数据安全的管理规定；②对于图书馆数据的调用及访问，要制定相关的身份认证、权限、职责审查机制，避免权责不明确而导致数据泄露；③须和云服务提供商等第三方签订数据安全协议，明确双方在维护数据安全方面的权利和义务。

3.2　应用新技术加强数据安全保障

图书馆数据安全的很多措施和设想的付诸实施都需要相关技术的支撑。图书馆需要关注的相关技术有3个。①区块链技术。它是将数据区块以链的形式顺序串联在一起的数据结构，彼此相邻的两个数据区块存在关联，在其中的一个数据区块数据不修改的情况下，其余数据区块的数据几乎无法篡改，能够提高图书馆数据存储安全和可信度。国内学者对区块链技术在数据隐私保护[20-22]、数据安全领域[23]、用户画像数据安全[24]等方面都有深入的研究，因此，区块链技术在图书馆数据安全制度的建立方面会大有作为。②网络安全态势感知技术。加强数据安全传输防御措施，对可能遭遇的数据安全传输风险进行提前预判与监测，布置安全防御手段，提高数据安全的保障水平。③云存储技术。为了保证数据存储的安全，各种云平台及大数据的技术安全框架都在不断升级，通过加密、脱敏等手段，加强数据在云平台上访问、存储等方面的安全性，明确图书馆对数据的所有权。图书馆应当经常关注新技术的发展趋势，探讨新技术在图书馆数据安全方面应用的各种可能性，这是LSP将来应用过程中必须要重视的工作。

3.3　提升图书馆的数据管理水平

目前，LSP越来越受到广泛重视，图书馆的发展也正朝着数据化、信息化、云端化的方向迅速迈进，但同时数据也时刻面临安全威胁。因此，图书馆应该从以下3个方面提升数据管理水平：①图书馆应该根据本馆的具体情况制定数据安全管理制度，对数据活动的各阶段做出相应的程序规定[25]，改进馆内安全管理规则，降低数据被破坏的概率；②加强专业化人员的培训，提升其处理和利用数据的安全知识技能，并根据业务岗位的需要进行定岗定责，遵循最小授权原则分配工作权限，让馆员树立数据安全意识；③做好网络安全等级保护工作，完善安全物理环境、安全通信网络、安全区域边界及安全计算环境等方面的工作，提升网络及业务系统的安全应用防护能力、隐患发现能力及应急处置能力，切实保障图书馆的数据安全。

4　结语

LSP是基于云平台及大数据技术架构的开源信息业务系统，其安全稳定运行是一件极其重要的工作，如何保障数据安全是关乎图书馆发展和服务升级的重要因素。本文从数据活动的各个阶段，分析了LSP运行过程中所面临的数据安全威胁，提出构建数据安全方案，在《网络安全法》实施及《数据安全法》发布的当下，期待能够促进图书馆数据保护措施的制定和实施，对开展数据安全工作具有一定指导和实践意义。

参考文献

[1]　BREEDING M. Perceptions 2012：An Inter-national Survey

of Library Automation [EB/OL]. [2021-05-22]. https://librarytechnology.org/perceptions/2012/.

[2] 钱国富. 下一代图书馆服务平台的研究与发展 [J]. 图书馆论坛，2019，39（5）：62-66.

[3] 肖铮，林俊伟. 用微服务构架下一代图书馆服务平台：以FOLIO为例 [J]. 图书馆杂志，2018，37（11）：63-69.

[4] 张磊，贺晨芝，赵亮. 面向数据与知识服务的第三代图书馆服务平台 [J]. 国家图书馆学刊，2018，27（6）：40-47.

[5] 施晓华，王昕，徐璟，等. 新一代智慧图书馆服务平台的发展现状与特征研究 [J]. 大学图书馆学报，2019，37（2）：49-54.

[6] 数据安全 [EB/OL]. [2021-05-12]. https://baike.so.com/doc/6144889-6358066.html.

[7] 数据安全法 [EB/OL]. [2021-06-11]. https://www.sohu.com/a/471563814_260616.

[8] 谢珍，陆溯. 智慧图书馆视域下用户数据应用与隐私保护平衡研究 [J]. 国家图书馆学刊，2020，29（2）：49-59.

[9] 周秀霞，刘万国，隋会民，等. Five Safes安全框架及其对我国图书馆领域敏感数据安全访问的启示 [J]. 情报理论与实践，2020，43（3）：85-89.

[10] 梁俊荣. 基于大数据决策的图书馆信息系统安全分析与管理研究 [J]. 图书馆理论与实践，2017（3）：93-98.

[11] 万映红，张泸月，万莉. 基于大数据应用的智慧图书馆个人数据保护研究 [J]. 图书馆学研究，2018（3）：31-34.

[12] 张娟，李仪. 云计算下图书馆读者个人信息的安全风险及应对 [J]. 情报理论与实践，2017，40（5）：39-43，49.

[13] 公安部：2019全国网民网络安全感满意度调查统计报告 [EB/OL]. [2021-05-11]. http://www.199it.com/archives/944906.html.

[14] 黄国彬，郑霞，王婷. 云服务协议引发的信息安全风险及图情机构的应对措施 [J]. 图书情报工作，2020，64（12）：38-47.

[15] 李仪，张娟. 面向图书馆读者的个人信息权利设置研究——以关怀读者人格为目的 [J]. 图书馆论坛，2015，35（6）：76-81.

[16] 陆康. 高校图书馆数字资源统计系统建设研究 [J]. 现代情报，2015，35（9）：140-145.

[17] 50余名学生个人信息泄露 被企业冒用虚报个税 [EB/OL]. [2021-04-07]. https://www.chinanews.com/sh/2020/02-03/9076654.shtml.

[18] 王丹，孙洋，谢辉，等. 基于网络安全等级保护2.0的农业科研单位网络安全体系研究——以中国农业科学院为例 [J]. 农业图书情报学报，2020，32（12）：97-103.

[19] 张辉. 基于网络安全等级保护2.0的高校网络安全体系研究 [J]. 网络安全技术与应用，2020（2）：83-84.

[20] 佟鑫，任望，冯运波. 大数据平台安全风险分析与评估方法 [J]. 保密科学技术，2018（2）：6-14.

[21] 祝烈煌，高峰，沈蒙，等. 区块链隐私保护研究综述 [J]. 计算机研究与发展，2017，54（10）：2170-2186.

[22] 柳林子，赵力. 区块链技术下图书馆读者个人信息保护研究 [J]. 图书馆工作与研究，2019（5）：96-101.

[23] 刘明达，陈左宁，拾以娟，等. 区块链在数据安全领域的研究进展 [J]. 计算机学报，2021，44（1）：1-27.

[24] 刘海鸥，姚苏梅，黄文娜，等. 移动图书馆用户画像大数据应用的困境与对策——基于区块链理念 [J]. 图书馆学研究，2019（23）：26-33.

[25] 王维秋，刘春丽，邱宇红，等. 智慧图书馆大数据安全问题研究 [J]. 图书馆学刊，2020，42（8）：91-94，106.

作者简介

赵想飞，男，1981年生，硕士，馆员，研究方向：图书馆信息资源管理、信息安全，E-mail：xiangfei1401@126.com。
常颖聪，女，1990年生，硕士，馆员，研究方向：阅读推广、读者服务。
黄闽，男，1973年生，馆员，研究方向：图书馆网络管理。
崔增辉，男，1974年生，馆员，研究方向：图书馆机房管理。

Data Security Research on the Third Generation Library Service Platform

ZHAO XiangFei CHANG YingCong HUANG Min CUI ZengHui
(Hebei Normal University Library, Shijiazhuang 050024, China)

Abstract: With the deepening of the cloud service concept and new technology such as big data, cloud computing, artificial intelligence is widely used in the library, the library service platform is implemented by many more libraries. Big data collection, data analysis and cloud deployment provide a data basis of deploy operation for the platform, but meantime increases risk of data security, The article qualitative analyzes risk factors for data security from characters of the platform, and then proposes some measures respectively in each stage of data collection, data storage, data access and application, data sharing and updating, data security assessment, in order to promote greatly the level of management and service on library.

Keywords: Library Service Platform; Data Security; Data Collection; Data Security Assessment

（收稿日期：2021-06-12）