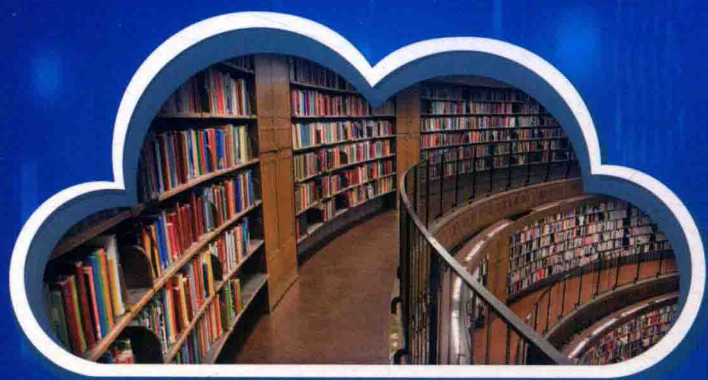




云计算时代 信息技术在图书馆中的应用研究

赵春燕◎著



中国石化出版社

[HTTP://WWW.SINOPEC-PRESS.COM](http://www.sinopec-press.com)

云计算时代 信息技术在图书馆中的应用研究

赵春燕 著

中国石化出版社

图书在版编目(CIP)数据

云计算时代信息技术在图书馆中的应用研究 / 赵春燕著.
—北京: 中国石化出版社, 2018. 8
ISBN 978-7-5114-5017-3

I. ①云… II. ①赵… III. ①云计算—应用—图书馆
工作—研究 IV. ①G250.72

中国版本图书馆 CIP 数据核字(2018)第 193482 号

未经本社书面授权, 本书任何部分不得被复制、抄袭, 或者以任何形式或任何方式传播。版权所有, 侵权必究。

中国石化出版社出版发行

地址: 北京市朝阳区吉市口路 9 号
邮编: 100022 电话: (010) 59964500
发行部电话: (010) 59964526
<http://www.sinopec-press.com>
E-mail: press@sinopec.com
北京艾普海德印刷有限公司印刷
全国各地新华书店经销

*

787 × 1092 毫米 16 开本 15.5 印张 261 千字
2018 年 8 月第 1 版 2018 年 8 月第 1 次印刷
定价: 68.00 元

目 录

第一章 云计算与图书馆	1
第一节 云计算研究	1
一、什么是云计算	1
二、云计算的历史	3
三、云计算特点	3
四、云计算架构	4
五、云计算技术	4
六、云计算的服务形式	5
第二节 图书馆研究与实践	6
一、图书馆的起源	6
二、图书馆的定义	7
三、数字图书馆概述	8
四、特色数字图书馆建设研究与实践	14
第二章 计算机网络技术研究与实践	16
第一节 计算机网络技术概述	16
一、计算机网络的定义	16
二、计算机网络网络发展史	17
三、计算机网络的性能	19
四、计算机网络的体系结构	23
五、计算机网络的组成分类	25
六、局域网技术	27

七、计算机网络	33
八、无线计算机网络	34
九、移动互联网技术	36
十、VPN 技术概述	37
第二节 图书馆无线网络建设案例	39
一、项目概述	39
二、无线网络建设方案	40
三、WLAN 各子系统详细设计	42
四、无线网络解决方案的技术特点	44
第三节 图书馆网络行为管理系统实践案例	58
一、项目背景	58
二、需求分析	59
三、上网行为管理的解决方案	60
四、项目成效	66
第四节 图书馆资源校外访问系统实践案例	66
一、图书馆的需求分析	67
二、技术解决方案	68
三、图书馆校外访问案例	69
第三章 移动图书馆技术研究与实践	70
一、移动图书馆概述	70
二、基于 AHP 的移动图书馆信息服务质量影响因素分析	71
三、移动图书馆建设实例	74
第四章 Web 技术研究与实践	79
一、Web 技术	79
二、Web Services 网站系统开发技术	84
三、语义网 (语义 Web)	89
四、图书馆网站建设常用技术与应用案例	91

第五章 服务器技术研究与实践	95
一、服务器技术	95
二、服务器虚拟化技术	98
三、图书馆服务器虚拟化建设项目案例实践	102
第六章 存储技术研究与实践	110
一、DAS	110
二、SAN	110
三、NAS	112
四、云存储	112
五、存储技术研究热点	115
六、数字图书馆建设中的存储技术方案	116
第七章 图书馆数据中心建设研究与实践	118
一、项目背景	118
二、升级改造的目标	120
三、建设内容	120
四、各系统功能描述	121
五、各系统设计实施方案	127
第八章 云计算时代的智慧图书馆技术研究与实践	136
第一节 物联网技术研究及图书馆应用实践案例	136
一、物联网的发展历程	136
二、物联网的定义	138
三、物联网的分类	139
四、物联网的四大技术形态	140
五、物联网的支撑技术	142
六、RFID 技术	144
七、RFID 技术应用案例	148

第二节 大数据环境下数据挖掘技术及图书馆应用实践研究	158
一、大数据技术概述	158
二、数据挖掘技术	165
三、数据挖掘技术在高校图书馆的应用	184
四、Web 使用挖掘技术在图书馆网站建设中的案例研究	187
 第九章 图书馆服务领域的信息技术应用研究与实践	208
第一节 新媒体技术在图书馆信息服务中的应用	208
一、研究背景	208
二、新媒体技术及其与高校图书馆信息服务的关系	209
三、新媒体在高校图书馆信息服务中的应用	209
第二节 论文抄袭检测技术研究	213
一、研究背景	213
二、抄袭检测技术概述	214
三、抄袭检测技术的发展	214
四、抄袭检测系统的应用	215
五、抄袭检测技术存在的问题	217
第三节 基于 MOOC 的图书馆服务实践	217
一、什么是 MOOC	217
二、MOOC 在高职教育中应用研究	221
三、MOOC 背景下高职院校图书馆服务转型的探索与实践	224
 第十章 图书馆特色数字资源建设研究与实践	230
一、项目背景	230
二、项目基本情况	233
三、项目建设成果	237
四、项目建设意义	238
五、项目创新点	238
六、共享机制设计与实践	239
 参考文献	240

第一章 云计算与图书馆

第一节 云计算研究

近年来，云计算已经渗入到各个领域，虽然云计算这个词已经大家耳熟能详，可是究竟什么是云计算呢？

一、什么是云计算

云计算的出现是一种历史的必然。一方面，技术的飞速发展为云计算的实施扫清了障碍，光接入技术已经能够为用户提供百兆乃至更高的带宽，SOA、虚拟化等技术的发展已经为云计算的实施做好了技术储备。另一方面，市场正在快速推动云计算的落地，企业用户希望通过构建私有云来降低 IT 系统运营成本，设备商和服务供应商也希望通过提供云服务来赢取更多的利润，开拓更为广阔的市场。技术和市场在云计算这个切入点找到了统一的节拍，从而使得云计算的发展前景极为光明。毫无疑问，云计算是当下计算机行业最为时髦的术语之一，但是当具体谈到什么是云计算时，大家的观点往往互不相同。

虽然随着 Amazon.com 在 2006 发布其弹性云计算产品，“云计算”这一提法开始普及，但最开始提到“云计算”一词是在 1996 年的 Copaq 内部文档中。

云符号用来表示 1977 年的 ARPANET 和 1981 年的 CSNET 中的计算设备的网络——互联网的前身。云这个词被用作互联网的隐喻，用一个标准化的云状形状来表示电话示意图上的一个网络。通过这种简化，隐含的是，网络的端点如何连接的细节与理解图表的目的无关。

云计算是一种信息技术的范例，它能够无处不在地访问可配置的系统资源和更高级别的服务的共享池，这些服务通常可以在互联网上以最小的管理投入而快速提



供。云计算依赖于资源共享来实现一致性和规模经济，类似于公用事业。使用第三方的云服务使组织能够专注于他们的核心业务，而不是在计算机基础设施和维护上花费资源。云计算可以使公司避免或最小化前期的 IT 基础设施成本，使企业的应用程序运行的更快，提升可管理性和减少维护量，并且它使 IT 团队能够更快地调整资源以满足波动和不可预测的需求。云供应商通常使用“按需付费”模式。

云计算一词用来同时描述一个系统平台或者一种类型的应用程序。一个云计算的平台按需进行动态地部署（provision）、配置（configuration）、重新配置（reconfigure）以及取消服务（deprovision）等。在云计算平台中的服务器可以是物理的服务器或者虚拟的服务器。高级的计算云通常包含一些其他的计算资源，例如存储区域网络（SAN），网络设备，防火墙以及其他安全设备等。云计算在描述应用方面，它描述了一种可以通过互联网 Internet 进行访问的可扩展的应用程序。云应用使用大规模的数据中心以及功能强劲的服务器来运行网络应用程序与网络服务。任何一个用户可以通过合适的互联网接入设备以及一个标准的浏览器就能够访问一个云计算应用程序。

云计算包含了两个方面的含义：一方面描述了基础设施，用来构造应用程序，其地位相当于 PC 机上的操作系统；另一方面描述了建立在这种基础设施之上的云计算应用。

在与网格计算的比较上，网格程序是将一个大任务分解成很多小任务并行运行在不同的集群以及服务器上，注重科学计算应用程序的运行。而云计算是一个具有更广泛含义的计算平台，能够支持非网格的应用，例如支持网络服务程序中的前台网络服务器、应用服务器、数据库服务器三层应用程序架构模式，以及支持当前 Web 2.0 模式的网络应用程序。云计算是能够提供动态资源池、虚拟化和高可用性的下一代计算平台。现有的云计算实现使用的技术体现了以下 3 个方面的特征：

（1）硬件基础设施架构在大规模的廉价服务器集群之上。与传统的性能强劲但价格昂贵的大型机不同，云计算的基础架构大量使用了廉价的服务器集群，特别是 x86 架构的服务器。节点之间的互连网络一般也使用普遍的千兆以太网。

（2）应用程序与底层服务协作开发，最大限度地利用资源。传统的应用程序建立在完善的基础结构，如操作系统之上，利用底层提供的服务来构造应用。而云计算为了更好地利用资源，采用了底层结构与上层应用共同设计的方法来完善应用程序的构建。

(3) 通过多个廉价服务器之间的冗余, 使用软件获得高可用性。由于使用了廉价的服务器集群, 节点的失效将不可避免, 并且会有节点同时失效的问题。为此, 在软件设计上需要考虑节点之间的容错问题, 使用冗余的节点获得高可用性。

通过这些技术手段, 云计算达到了两个分布式计算的重要目标: 可扩展性和高可用性。可扩展性表达了云计算能够无缝地扩展到大规模的集群之上, 甚至包含数千个节点同时处理。高可用性代表了云计算能够容忍节点的错误, 甚至有很大一部分节点发生失效也不会影响程序的正确运行。

二、云计算的历史

在 20 世纪 60 年代, 分时的初始概念通过 RJE (远程作业录入) 得到普及; 这个术语主要与诸如 IBM 和 DEC 这样的大型供应商相关联。用户向操作员提交作业到 IBM 主机去运行的“数据中心”模式占据主导地位。

在 20 世纪 90 年代, 以前主要提供专用点对点数据电路的电信公司开始提供具有类似服务质量的虚拟专用网络 (VPN) 服务, 可以更有效地使用整个网络带宽, 但成本较低。他们开始使用云符号来表示提供者负责和用户负责的分界点。云计算扩展了这个边界以覆盖所有的服务器和网络基础设施。随着计算机越来越普及, 科学家和技术人员探索了通过分时使更多的用户获得大规模计算能力的方法。

随着信息技术的飞速发展, 云计算成为了一种普遍认可的计算模式。近年来, 云计算的应用为用户提供可靠的各种资源服务, 它是一种一切皆为服务的模式, 为用户提高质量的服务另外, 云计算与互联网 + 的结合和云计算与大数据之间的应用, 使云计算得到新的发展和应用现阶段, 由于得到相关技术标准和各个服务层协议的支撑, 一些云计算产品的推出和应用, 为客户提供高品质的服务体验, 因此云计算具有广阔的发展前景。本文主要研究云计算的主要特点架构和相关技术, 重点阐述云计算的研究相关工作, 展望云计算研究中应解决的关键问题和研究方向。

三、云计算特点

云计算是在原有网格计算 (Grid Computing), 并行计算 (Parallel Computing) 和分布式计算 (Distributed Computing) 的基础上扩展而成的。通常把云计算定义为一种网络的增值服务, 即用户可以在任何地方、任何时间最大限度地使用相关服务,



处理大规模计算问题。云计算主要有以下特点：

(1) 服务资源池化：通过虚拟化技术，将硬件软件等资源进行资源化，并根据用户需求进行动态分配资源，提高了资源的共享和群体的合作沟通能力。

(2) 弹性的计费模式，减少成本投入：用户根据需求来访问基础设施平台服务和软件服务等服务，按需分配资源而不必要自己搭建相关的硬件和软件平台，减少成本的投入，并解决了用户计算能力不足而不满足应用需求，或计算能力空闲而浪费相关资源，提高了使用效率。

(3) 可靠性：云服务在性能和质量上具备 QOS 保证，通过制定的协议和技术提供 QOS 保证，提高服务的可靠性和有效性

四、云计算架构

云计算架构可分为 4 层，自下到上有：

(1) 物理资源层，主要为基础设施设备，包含计算机存储设备数据库软件和网络设施等；

(2) 资源池层，资源池是通过虚拟技术实现对物理资源的访问和管理，主要包含计算资源池网络资源池存储资源池数据资源池软件资源池网络资源池等

(3) 核心中间件层，实现应用服务接口对资源池的访问，主要负责资源管理任务管理用户管理安全管理性能监控和计费管理等任务

(4) 云计算应用服务接口层，具有服务调度服务查询服务 workflow 和服务选择等功能从云计算服务的角度划分可分为三层，分别为基础设施即服务（Infrastructure as a Service, IaaS），即向云用户提供模拟完善的计算机存储器等基础设施平台即服务（Platform as a Service, PaaS），是在软件的研发平台提供给上层的开发用户软件即服务（Software as a Service, SaaS）是以租用的方式向用户提供软件服务，并按一定的计费方式使用从云计算的部署划分，可分为私有云社区云公共云和混合云。

五、云计算技术

1. 虚拟化技术

虚拟化是指计算元件不是在真实的基础上而是虚拟的基础上运行，是一种优化资源和简化管理的解决方案虚拟化技术适合在云计算平台中的应用，虚拟化的核心

解决了脱离硬件的依赖,提供统一的虚拟化界面,通过虚拟化技术,可以在一台服务器上运行多台虚拟机,从而达到服务器的优化和整合目的。虚拟化技术使用动态资源伸缩的手段降低云计算基础设施的使用成本,提高负载部署的灵活性比如当虚拟化数据中心需要维护和管理时,并不需要关闭虚拟机或关闭程序,只需要把虚拟机迁移到另一台服务器上因此,云计算在数据中心虚拟化过程中,具有在线迁移低开销管理服务器整合灵活性和高可用性等优势,为云计算在部署资源池层中提供新的发展思路

2. 中间件技术

中间件即是运行在两个层次之间的一种组件,是在操作系统和应用软件之间的软件层次,支持应用软件的开发运行部署和管理的支撑软件称为中间件中间件可以屏蔽硬件和操作系统之间的兼容问题,并能具有管理分布式系统的中节点间的通信节点资源和协调工作等功能,通过中间件技术,可将不同平台的计算机节点组成一个功能强大的计算机分布系统而云环境下的中间件,其主要功能是对云服务资源进行管理,主要包含用户管理任务管理安全管理,为云计算提供可靠的部署运行开发和应用提供高效支撑

3. 存储技术

在云计算中,存储技术通常和虚拟技术相互结合起来,通过对数据资源虚拟化,提高访问效率,目前数据存储技术有 HDFS (Hadoop Distributed File System, 开源) 和 Google 的 GFS (Google File System, 非开源),该技术具有高吞吐率、分布式和高速传输等优点,适合云计算中为大量用户提供云服务

六、云计算的服务形式

云计算包括以下几个层次的服务:基础设施即服务 (IaaS),平台即服务 (PaaS) 和软件即服务 (SaaS)。

1. IaaS 基础设施即服务

IaaS (Infrastructure as a Service): 基础设施即服务。消费者通过 Internet 可以从完善的计算机基础设施获得服务。例如:硬件服务器租用。IaaS 是云计算的基础。通过建立大规模数据中心, IaaS 为上层云计算服务提供海量硬件资源。同时,在虚拟化技术的支持下, IaaS 层可以实现硬件资源的按需配置,并提供个性化的基础设



施服务。典型的 IaaS 平台有 Amazon EC2、Eucalyptus 和东南大学云计算平台等。

2. PaaS 平台即服务

PaaS (Platform as a Service): 平台即服务。PaaS 实际上是指将软件研发的平台作为一种服务。典型的 PaaS 平台有: Google App Engine、Hadoop 和 Microsoft Azure。这些平台都基于海量数据处理技术搭建, 且各具代表性。

3. SaaS 软件即服务

SaaS (Software as a Service): 软件即服务。它是一种通过 Internet 提供软件的模式, 用户无需购买软件, 而是向提供商租用基于 Web 的软件, 来管理企业经营活动。典型的 SaaS 应用包括 Google Apps、Salesforce CRM 等。

第二节 图书馆研究与实践

一、图书馆的起源

图书馆的英文词是 library, 这一词源于拉丁语 librarium, 原义为藏书之所。图书馆的历史可追溯到公元前 3000 年以前, 当时的美索不达米亚已有保存泥板文献的图书馆。此外, 古代埃及、中国和希腊等人类文明的发源地也出现了图书馆。古希腊、罗马时期的一些图书馆如亚历山大图书馆等, 藏书已有相当规模。中世纪欧洲图书馆的发展受到宗教势力的束缚, 逐渐衰落, 印刷术的推广和文艺复兴推动了西方图书馆事业的发展。18 ~ 19 世纪, 西方的国家图书馆开始出现并获得了较大发展, 19 世纪中叶出现了由政府举办的向社会开放的公共图书馆。

我国的图书馆历史悠久。只是起初并不称作“图书馆”, 而是称为“府”“阁”“观”“台”“殿”“院”“堂”“斋”“楼”等。如西周的盟府、两汉的石渠阁、东观和兰台、隋朝的观文殿、宋朝的崇文院、明代的澹生堂、清朝收藏四库全书的“七阁”等。中国古代藏书楼主要为官府所设。宋代以后私家藏书楼迅速发展。此外, 寺观和书院藏书也很普遍。这些藏书楼的服务范围相当狭窄, 但对保存文化典籍起了一定作用。直到 20 世纪初, 中国才出现近代意义的公共图书馆。

20 世纪, 图书馆类型和数量不断增多, 向社会开放的范围不断扩大, 提供的服务更趋多样化。现代科学文献的大量增长, 使图书馆馆藏规模和结构都发生了很大

变化,文献整理和加工更加深入细致,各类图书馆特别是专门图书馆大量开展情报服务。随着现代科学技术在图书馆的广泛应用,图书馆管理已逐步走向自动化、现代化和数字化。

现代图书馆是信息时代的产物,它已由单纯的收集、整理、利用文献的比较封闭的系统,发展到以传递文献为主的、全面开放的信息系统。计算机技术、存储技术和数据通信技术在图书馆工作中的广泛应用,以及这三者的相互结合,有力地改变着图书馆工作的面貌。随着云计算技术的应用,图书馆正向着智慧图书馆方向发展。

二、图书馆的定义

国外学者们给图书馆的定义有很多种,从图书馆的馆藏到图书馆的职能等不角度描述了什么是图书馆。美国图书馆学家巴特勒认为:“图书馆是将人类记忆的东西移植于现在人们的意识之中的一个社会装置”。美国图书馆学家谢拉提出:图书馆是这样的一个社会机关,它用书面记录的形式积累知识,并通过馆员将知识传递给团体和个人,进行书面交流。因此,图书馆是社会中文化交流体系的一个重要机关。德国图书馆家卡尔施泰特则认为:图书是客观精神的容器,图书馆是把客观精神传递给个人的场所。客观精神的作用就在于它能够建立和维持各种社会形象。因此,图书馆就是在维持社会形象中,使世代结合的纽带。实际上,客观精神主要指的是人类创造的文化。在这个意义上,图书馆就是使文化的创造和继承成为可能的社会机构。《英国百科全书》的解释:图书馆的意思是很多书收藏在一起,这是为了阅读、研究或参考用的。《苏联大百科全书》的解释:图书馆是组织社会利用出版物的文化教育和科学辅助机关。图书馆系统地从事搜集、保藏、宣传和向读者借阅出版物,以及进行图书情报工作。

在我国,20世纪30年代以来一些图书馆学者相继对图书馆的定义进行了研讨。

1958年卢震京在《图书馆学辞典》中对图书馆定义做了如下解释:图书馆系根据其特定需要,搜集一切或一些人类文化在科学、技术、艺术及文学个方面所创造的精华记载,用科学的经济的方法,整理保存,以便广大人民使用,并进而帮助其接受马列主义为完成社会主义建设所必需的知识的文化中心。

1985吴慰慈等在《图书馆学概论》中提出:“图书馆是搜集、整理、保管和利用书刊资料,为一定社会的政治、经济服务的文化教育机构”。这个定义反映了20



世纪 90 年代以前人们对图书馆的认识,它是对传统图书馆本质的概括。这个定义可以回答有关传统图书馆的四个问题:一是图书馆的工作程序——对书刊资料进行搜集、整理、保管和利用;二是图书馆的工作对象——文献资料;三是图书馆活动的目的——为一定社会的政治、经济服务;四是图书馆的性质——文化教育机构。

1988 年黄宗忠在《图书馆学导论》中认为:“图书馆是对信息、知识的物资载体进行收集、加工、整理、积聚、存贮、选择、控制、转化和传递、提供给一定社会读者使用的信息系统。简言之,图书馆是文献信息的存贮与传递中心”。

2002 年吴慰慈、董焱在《图书馆学概论》中提出:“图书馆是社会记忆(通常表现为书面记录信息)的外存和选择传递机制。换句话说,图书馆是社会知识、信息、文化的记忆装置、扩散装置。”

近年来,随着图书馆空间服务、阅读推广、文化传播等职能在图书馆中发挥了越来越大的作用,数字图书馆建设也在信息技术的助推下,不断向智慧图书馆方向发展,图书馆的馆藏与服务也在不断转变,相信图书馆的定义也将不断随之变化。

三、数字图书馆概述

图书馆作为人类智慧的宝库,历来受到社会和政府的重视,纵观图书馆的发展史,图书馆的每一次进步无不反映着社会生产力的提高和文明的进步。随着计算机技术、通信和网络技术、高密度存贮技术以及多媒体技术的飞速发展,给图书馆的工作带来了深刻的变化,极大地改变了文献信息的处理方式、存贮方式和服务方式,提高了工作效率和质量,其职能也在不断扩展和加强,图书馆已从单纯的藏书、目录中心,逐步转变为人类科技文化信息中心、咨询服务中心和科学教育中心。图书馆已经发展到了实现真正意义上的信息资源共享为目标的新领域——数字图书馆。

数字图书馆(Digital Library)是现代高新科学技术和文献知识信息以及传统历史文化完美结合的体现。它改变了传统图书馆的静态书本式文献服务特征,实现了多媒体存取、远程网络传输、智能化检索、跨库无缝链接、创造出超时空信息服务的新境界。目前,数字图书馆已成为评价一个国家信息基础设施水平的重要标志,也是世界各国在高科技领域展开较量的焦点之一。由于数字图书馆与 Internet/Intranet 的紧密结合,深受广大读者的青睐,已成为图书馆信息化建设达标的标志之一。

1. 什么是数字图书馆

对于数字图书馆有不同的理解和认识:从图书馆学的角度来看,数字图书馆是

图书馆自动化向更高层次的发展；从数据库和信息检索的角度来看，数字图书馆是各种大型数据库，特别是全文数据库的大联合；而从计算机和通信技术的角度来看，数字图书馆是超文本和网格技术的特殊应用。随着对数字图书馆研究和建设的不断发展，对数字图书馆的认识也在不断加深，数字图书馆的定义也在不断深化中，目前关于数字图书馆的定义主要有以下几种：

定义之一：数字图书馆是组织数字化信息及其技术进入图书馆并提供有效服务，几乎图书馆的所有载体信息均能以数字化的形式获得，读者通过网络组织访问外界数字化图书馆和文献信息数据库系统。图书馆和网络系统连接，让人们很方便地共享资源。

定义之二：数字图书馆是一个大系统，它是分布式的、大规模的和有组织的数据库和知识库，用户和用户团体可以对系统内的数据库和知识库进行一致性的访问，获得自己所需要的最终信息。

定义之三：数字图书馆强调了它存储的信息是用电磁介质，按二进制编码的方法加以存储和管理，是把原先用纸张存储的信息转变为用计算机和多媒体技术存储和处理的信息。

定义之四：数字图书馆即是一个数字化的信息系统，它将分散于各种载体、不同地理位置的信息资源以数字化方式储存、以网络化方式互相连接，提供即时利用，实现资源共享。

定义之五：利用当今先进的数字化技术，通过诸如国际互联网计算机网络，使人数众多且又处在不同的地理位置的用户能够方便地利用大量的分散在不同储存处的电子物品的全部内容。

通俗地说，数字图书馆是用数字技术存储和组织信息，并通过计算机网络查询和检索信息的一种现代化的信息系统。数字图书馆所依附的是数字化技术，是一个数字化的信息系统，它把分散于不同载体、不同地理位置的信息资源以数字化的方式存贮、以网络他方式相连接，提供及时利用，实现资源共享。

数字图书馆是用数字技术处理和存储各种图文并茂文献的图书馆，实质上是一种多媒体制作的分布式信息系统。它把各种不同载体、不同地理位置的信息资源用数字技术存贮，以便于跨越区域、面向对象的网络查询和传播。它涉及信息资源加工、存储、检索、传输和利用的全过程。通俗地说，数字图书馆就是虚拟的、没有围墙的图书馆，是基于网络环境下共建共享的可扩展的知识网络系统，是超大规模



的、分布式的、便于使用的、没有时空限制的、可以实现跨库无缝链接与智能检索的知识中心。

数字图书馆是一门全新的科学技术，也是一项全新的社会事业。数字图书馆是一种拥有多种媒体内容的数字化信息资源，能够为用户提供方便、快捷、高水平的信息化服务机制。数字图书馆自从出现以来，在资源更新速度、信息共享范围、拉近与读者的距离、数据的随机性和全开放性等方面都显示出巨大的优势，其重要特色是基于 Internet 的网络化信息检索、信息加工和信息服务。

“数字图书馆” (Digital Library, 简称 DL), “Library” 这个词，在英文中它有两个解释，一是“图书馆”，二是“库”。“Digital Library” 的英文本意更强调的是“库”，而不是“图书馆”。现在关于“数字图书馆”的概念，很容易产生认识上的误区：认为数字图书馆就是将现有的图书馆数字化，这恐怕有点“望文生义”，失于简单、片面。“Digital Library” 是一个内涵很丰富的概念，其“解”并不唯一，数字信息馆、数字信息库、数字图书馆等都是 DL 的可能解释。

计算机技术、通信技术、网络技术、多媒体技术等新技术的飞速发展对当代图书馆的各个方面都产生了极大影响。其中关于数字图书馆的理论与实践研究是其中最热点的问题之一。1993 年，在德国的埃森召开了首届国际电子图书馆会议，1994 年，在美国得克萨斯又召开了国际数字图书馆会议。美国计算机协会 (ACM) 和美国信息科学学会 (ASIS) 及其他一些著名学会、协会的会刊都出版了与数字图书馆有关的专辑。1996 年 3 月，美国计算机协会信息检索专业组 (ACM SIGIR)、美国电气与电子工程师学会 (IEEE)、美国信息科学学会 (ASIS) 等几大学术组织在 Bethesda 召开了规模很大的首届 ACM 数字图书馆国际会议。

数字图书馆不是图书馆实体：它对应于各种公共信息管理与传播的现实社会活动，表现为一种新型信息资源组织和信息传播服务。它借鉴图书馆的资源组织模式、借助计算机网络通讯等高新技术，以普遍存取人类知识为目标，创造性地运用知识分类和精准检索手段，有效地进行信息整序，使人们获取信息消费不受空间限制，很大程度上也不受时间限制见。

2. 数字图书馆的由来

1945 年 7 月，当时担任美国科学研究和开发办公室主任的 Vannevar Bush 在《大西洋月刊》上首次谈到了数字图书馆的观点。二十年以后，美国学者 Licklider