

大数据下读者行为分析 ——以武昌分校图书馆为例

分享人：华中科技大学武昌分校王辉

1 大数据的发展与利用

2 武昌分校图书馆积累的大数据可以做什么？

3 分析读者行为的结果对我们工作的指导作用

4 我馆的大数据应用所做的工作及还可进一步开展的工作

大数据的发展与利用

什么是大数据？

大数据是指无法在一定时间内用传统数据库软件工具对其内容进行抓取、管理和处理的数据集合

1. Volume

数据量巨大

全球在2010年正式进入ZB时代，IDC预计到2020年，全球将总共拥有35ZB的数据量

3. value

沙里淘金，价值密度低

以视频为例，一部一小时的视频，在连续不间断监控过程中，可能有用的数据仅仅只有一两秒。如何通过强大的机器算法更迅速地完成数据的价值“提纯”是目前大数据汹涌背景下亟待解决的难题

2. Variety

结构化数据、半结构化数据和非结构化数据

如今的数据类型早已不是单一的文本形式，订单、日志、音频，能力提出了更高的要求

4. Velocity

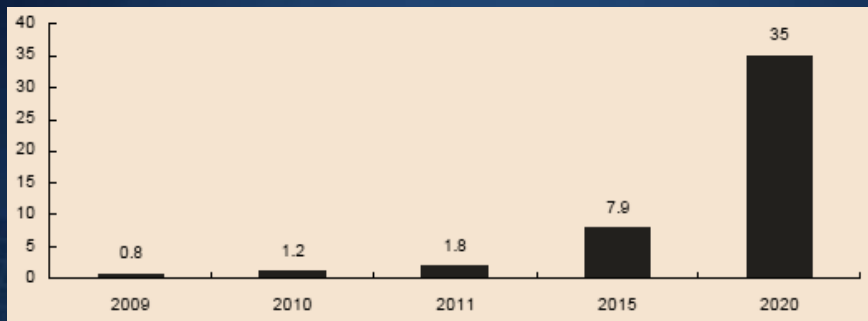
实时获取需要的信息

大数据区别于传统数据最显著的特征。如今已是ZB时代，在如此海量的数据面前，处理数据的效率就是生命

大数据的发展与利用

国内外大数据的发展与现状

数据量增加

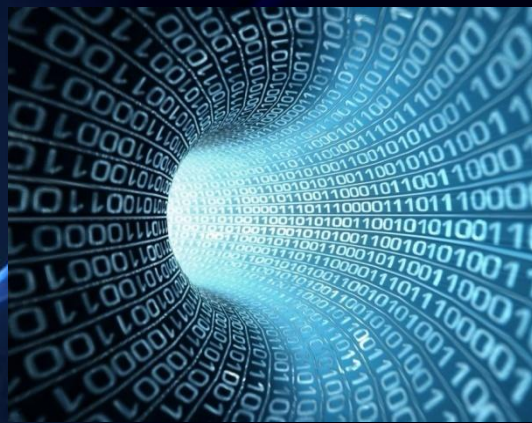


根据IDC 监测，人类产生的数据量正在呈指数级增长，大约每两年翻一番，这个速度在2020 年之前会继续保持下去。这意味着人类在最近两年产生的数据量相当于之前产生的全部数据量

TB ⇨ PB ⇨ EB ⇨ ZB

数据结构日趋复杂

大量新数据源的出现则导致了非结构化、半结构化数据爆发式的增长



大数据的发展与利用

国内外大数据的发展与现状

“大数据”这个概念早在上个世纪的1980年，著名未来学家阿尔文·托夫勒便在《第三次浪潮》一书中，就将“大数据”热情地赞颂为“第三次浪潮的华彩乐章”。2008年9月《科学》(Science)杂志发表了一篇文章“BigData: Science in the Petabyte Era”。“大数据”这个词开始被广泛传播。目前国内外的专家学者对大数据只是在数据规模上达成共识：“超大规模”表示的是GB级别的数据，“海量”表示的是TB级的数据，而“大数据”则是PB级别及其以上的数据。但对其的准确定义给出的结论不一。美国国家科学基金会(NSF)则将大数据定义为“由科学仪器、传感设备、互联网交易、电子邮件、音视频软件、网络点击流等多种数据源生成的大规模、多元化、复杂、长期的分布式数据集”。

全球每秒钟发送 **2.9 百万封** 电子邮件，一分钟读一篇的话，足够一个人昼夜不息的读**5.5 年**……
每天会有 **2.88 万个** 小时的视频上传到Youtube，足够一个人昼夜不息的观看**3.3 年**……
推特上每天发布 **5 千万条** 消息，假设**10 秒钟**浏览一条信息，这些消息足够一个人昼夜不息的浏览**16 年**……
每天亚马逊上将产生 **6.3 百万笔** 订单……
每个月网民在Facebook 上要花费**7 千亿分钟**，被移动互联网使用者发送和接收的数据高达**1.3EB**……
Google 上每天需要处理**24PB** 的数据……


武昌分校图书馆积累的大数据可以做什么？

目前我馆做到的

读者行为分析

武昌分校图书馆积累的大数据可以做什么？

读者行为有哪些？

- 
- (1)、进馆刷卡，
 - (2)、运用座位系统进行选座在馆内进行普通自习，
 - (3)、利用馆内的实名制座位进行考研考证备考，
 - (4)、馆内浏览电子资源（可细分为各类数据库），
 - (5)、远程访问图书馆的电子资源，
 - (6)、馆内阅读图书，
 - (7)、馆内阅读期刊，
 - (8)、进馆借书，
 - (9)、进馆续借图书，
 - (10)、通过手机或者web页面续借图书，
 - (11)、进馆缴纳超期罚款，
 - (12)、进馆还书，
 - (13)、进馆借爱心伞，
 - (14)、进馆还爱心伞。

武昌分校图书馆积累的大数据可以做什么？

怎么分析读者行为？

- 1、分析内阅数据，借出文献来知道读者阅读行为习惯的
- 2、分析读者资源使用习惯及偏向。
- 3、分析读者进出馆次数，时间段来知道到馆行为习惯
- 4、分析实名制座位申请使用行为习惯
- 5、分析读者对爱心伞的借用行为习惯
- 6、分析读者续借的行为习惯
- 7、分析读者阅读需求趋势

武昌分校图书馆积累的大数据可以做什么？

怎么样挖掘与分析读者相关数据？

1、我馆数据存储方式

我馆图书自动化管理系统数据采用oracle数据库来进行存储，数据的稳定性，安全性有一定的保障性，另外也方便大量取出进行分析，从而提出对于我馆有价值的数据。

2、关联规则挖掘及灰关联分析法

关联规则挖掘的基本原理：

给定交易数据集 T ， $T=\{t_1, t_2, \dots, t_m\}$ ，其中 $t_i (1 \leq i \leq m)$ 是每笔交易的数据记录，每条记录 t_i 由若干个数据项组成。设 X 、 Y 是任一交易记录 t_i 中可能出现的数据项。若在一个交易记录中既含有 x ，又含有 Y ，则称 x 昔 Y 在此交易中成立，即存在着数据项 x 对数据项 Y 的关联。亦可表示为包括数据项 x 的交易记录中具有包括数据项 Y 的趋势。

武昌分校图书馆积累的大数据可以做什么？

怎么样挖掘与分析读者相关数据？

在数据挖掘的过程中，通常依据下述两个技术指标对满足要求的关联规则进行筛选

(1)、支持度(support, 用S表示)

(2)、可信度(confidence, 用C表示)

* (3) 增益度指标 (lift) 。增益度通过将关联规则x号Y的可信度C与任意情况下出现数据项Y的可能性(即包含数据项Y的记录在交易集合T中所占的百分比)相比较，来修剪挖掘出来的关联规则，以避免产生错误的结果。

灰关联分析法：

灰关联分析方法实际上是将图书管理系统视为“既含有已知信息，又含有未知的、非确知的信息”的灰色系统，通过不同时段读者行为对馆藏数量、新书量、借阅方式等的影响之间的隐含关系进行细化分析，从而发现各影响因素对图书馆管理的影响程度。

武昌分校图书馆积累的大数据可以做什么？

怎么样挖掘与分析读者相关数据？

3、具体我馆的做法：

读者的相关信息在我馆存储的数据中可以通过“读者证号”来唯一确定读者信息，图书文献在我馆可以通过馆藏的“索取号”进行唯一确定，图书馆的座位可通过座位管理系统中的“座位号”来唯一确定，爱心伞可通过“伞编号”来唯一确定。

若读者对图书馆资源有相关需求，则会产生相应的行为，则上述相关唯一字段会产生关联，这样我馆数据记录中就相应存储了其行为，要对其行为进行分析也还得一定量的数据作为分析因素，所以我馆通过处理了大量相关数据进行分析。

武昌分校图书馆积累的大数据可以做什么？

1 分析内阅数据，借出文献了解读者阅读行为习惯



武昌分校图书馆积累的大数据可以做什么？

1

分析内阅数据，借出文献了解读者阅读行为习惯



内阅分析：可以从左边的表中看出某段时间内我馆F、K、TN、TM、TP、TB、TH、TG类图书的内阅情况，其中排在前三比较受读者喜爱的图书类型分别为：TM、K、TB类，从侧面反映了我校专业设置的一个倾向等现象，也反映了我校学生对历史、地理类书籍的喜爱程度，我校学生比较注重人文修养的提高。

武昌分校图书馆积累的大数据可以做什么？

1

分析内阅数据，借出文献了解读者阅读行为习惯

通过灰关联分析法来分析，一个专业的学生在一段时间内的借阅的专业书籍来分析其是否存在借阅其它专业书籍的趋势。以2012级本科“计算机科学与技术”专业2014年9月到2014年11月借阅情况为例。

借书人	R01A	R01B	R021	R024	R02T	LOGO	LOGA1	LOGT	LOGG	LOGI	TITLE	
38 陈鹏飞	...	18086506877	... 计算机科学与技术 ... 学生	001	17949093	20140909	LT	...	60439	A0951787	... 数码单反摄影从新手到高手
39 肖世勇	...	18086506942	... 计算机科学与技术 ... 学生	001	17949738	20140909	LT	...	60481	A0410231	... Windows 9X/Me/NT/2000/XP/2003 DOS命令行网络服务与管理
40 王启	...	18086506933	... 计算机科学与技术 ... 学生	001	17958549	20140911	LT	...	60477	A0925852	... 《网管世界》2012超值精华本
41 王启	...	18086506933	... 计算机科学与技术 ... 学生	001	17958551	20140911	LT	...	60477	A1119525	... 不要等到毕业以后
42 王启	...	18086506933	... 计算机科学与技术 ... 学生	001	17958552	20140911	LT	...	60477	A0934664	... 重遇未知的自己：爱上生命中的不完美
43 王启	...	18086506933	... 计算机科学与技术 ... 学生	001	17958554	20140911	LT	...	60477	A0794875	... Google知道你多少秘密
44 王启	...	18086506933	... 计算机科学与技术 ... 学生	001	17958555	20140911	LT	...	60477	A0856071	... Windows 7应用技巧
45 王启	...	18086506933	... 计算机科学与技术 ... 学生	001	17958556	20140911	LT	...	60477	A0973721	... 网络硬件设备入门与实战
46 王启	...	18086506933	... 计算机科学与技术 ... 学生	001	17958558	20140911	LT	...	60477	A0973721	... 网络硬件设备入门与实战
47 尹谦	...	18086506950	... 计算机科学与技术 ... 学生	001	17965781	20140913	LT	...	60489	A0974411	... C语言与程序设计
48 程伟	...	18086506879	... 计算机科学与技术 ... 学生	001	17965982	20140913	LT	...	60452	A0367774	... Java编程思想
49 王萍欢	...	18086506932	... 计算机科学与技术 ... 学生	001	17973318	20140916	LT	...	60431	A0882567	... 实用汉英分类生活词典
50 王萍欢	...	18086506932	... 计算机科学与技术 ... 学生	001	17973319	20140916	LT	...	60431	A0999652	... 学校里学不到的机关生存经验
51 王萍欢	...	18086506932	... 计算机科学与技术 ... 学生	001	17973320	20140916	LT	...	60431	A1102724	... 女人要读点婚恋心理学
52 陈祥	...	18086506878	... 计算机科学与技术 ... 学生	001	17981148	20140919	LT	...	60492	A0858470	... 计算机组成原理联考辅导教程：2012版
53 余燕	...	18086506952	... 计算机科学与技术 ... 学生	001	17992399	20140923	LT	...	60451	A0843985	... 51单片机应用开发实战手册
54 余燕	...	18086506952	... 计算机科学与技术 ... 学生	001	17992400	20140923	LT	...	60451	A1076398	... Java编程手记：从实践中学习Java
55 余燕	...	18086506952	... 计算机科学与技术 ... 学生	001	17992448	20140923	LT	...	60451	A0909104	... 无线电台订本：2011年
56 余燕	...	18086506952	... 计算机科学与技术 ... 学生	001	17992505	20140923	ZQGP	...	60451	GF909104	... 无线电台订本：2011年
57 肖世勇	...	18086506942	... 计算机科学与技术 ... 学生	001	17992326	20140923	LT	...	60481	A0947540	... 巧学巧用Flash CS6制作动画
58 余燕	...	18086506952	... 计算机科学与技术 ... 学生	001	17992402	20140923	LT	...	60451	A0686827	... Java语言程序设计百问百例
59 赵保康	...	18086506958	... 计算机科学与技术 ... 学生	001	18006774	20140925	LT	...	60432	A0659486	... 马云精彩语录
60 钱志祥	...	18086506922	... 计算机科学与技术 ... 学生	001	18006775	20140925	LT	...	60434	A0880497	... 任正非正传：指挥挥之队驰骋沙场的商界思想家传奇
61 余燕	...	18086506952	... 计算机科学与技术 ... 学生	001	18014867	20140927	LT	...	60451	A0536425	... 全国计算机等级考试实用应试教程：二级公共基础知识
62 余燕	...	18086506952	... 计算机科学与技术 ... 学生	001	18014869	20140927	LT	...	60451	A1083233	... 全国计算机等级考试二级教程：Visual Basic语言程序设计：2013年版
63 肖世勇	...	18086506942	... 计算机科学与技术 ... 学生	001	18015348	20140927	LT	...	60481	A1089341	... 黑客攻防技术宝典：Web实战篇
64 沈齐红	...	18086506925	... 计算机科学与技术 ... 学生	001	18013739	20140927	LT	...	60496	A0975963	... C程序设计伴侣：帮你更好地理解解谭浩强老师的那本书以及更多!
65 尹谦	...	18086506950	... 计算机科学与技术 ... 学生	001	18027197	20140930	LT	...	60489	A1063240	... 全国计算机等级考试二级教程：网络技术：2013年版
66 程伟	...	18086506879	... 计算机科学与技术 ... 学生	001	18028916	20141003	Web	...	60452	A0367774	... Java编程思想
67 余燕	...	18086506952	... 计算机科学与技术 ... 学生	001	18030853	20141005	ZQGP	...	60451	GF909104	... 无线电台订本：2011年
68 余燕	...	18086506952	... 计算机科学与技术 ... 学生	001	18030855	20141005	LT	...	60451	A0909104	... 无线电台订本：2011年
69 余燕	...	18086506952	... 计算机科学与技术 ... 学生	001	18033876	20141007	LT	...	60451	A0904788	... 2012年全国计算机等级考试系列辅导用书：二级Visual Basic(含公共基...
70 肖世勇	...	18086506942	... 计算机科学与技术 ... 学生	001	18042760	20141008	LT	...	60481	A1101232	... 编译器构造
71 肖世勇	...	18086506942	... 计算机科学与技术 ... 学生	001	18042761	20141008	LT	...	60481	A0986119	... 新编计算机组成原理习题与解析
72 肖世勇	...	18086506942	... 计算机科学与技术 ... 学生	001	18037944	20141008	LT	...	60481	A0706033	... Red Hat Enterprise Linux系统管理
73 肖世勇	...	18086506942	... 计算机科学与技术 ... 学生	001	18037946	20141008	LT	...	60481	A0912509	... Linux网络服务配置详解
74 李胜林	...	18086506902	... 计算机科学与技术 ... 学生	001	18043805	20141009	LT	...	60464	A1083275	... 全国计算机等级考试二级教程：Java语言程序设计：2013年版
75 王萍欢	...	18086506932	... 计算机科学与技术 ... 学生	001	18051505	20141010	LT	...	60431	A0882567	... 实用汉英分类生活词典
31 程伟	...	18086506879	... 计算机科学与技术 ... 学生	001	17949739	20140909	LT	...	60452	A0514567	... 燕各款词头成案例解析
38 陈鹏飞	...	18086506877	... 计算机科学与技术 ... 学生	001	17949093	20140909	LT	...	60439	A0951787	... 数码单反摄影从新手到高手

武昌分校图书馆积累的大数据可以做什么？

1

分析内阅数据，借出文献了解读者阅读行为习惯

操作日期	操作时间	操作类型	使用者	设备	馆藏地点	读者证号	书目记录号	文献条码号
2014-11-28	08:16:38	流通借出	钟小利	zx1	中区分馆	17702733229	306759	A0873908
2014-11-28	08:21:13	流通借出	赵燕琼	zyq	自主学习中心	18086506095	183771	A0488892
2014-11-28	08:21:37	流通借出	赵燕琼	zyq	自主学习中心	18086506095	406406	A1084060
2014-11-28	08:21:41	流通借出	赵燕琼	zyq	自主学习中心	18086506095	360983	A0951229
2014-11-28	08:21:43	流通借出	赵燕琼	zyq	自主学习中心	18086506095	310008	A0848259
2014-11-28	08:22:00	流通借出	赵燕琼	zyq	自主学习中心	18086506031	413668	A1107359
2014-11-28	08:22:03	流通借出	赵燕琼	zyq	自主学习中心	18086506031	309868	A0848221
2014-11-28	08:22:05	流通借出	赵燕琼	zyq	自主学习中心	18086506031	413666	A1107398
2014-11-28	08:22:07	流通借出	赵燕琼	zyq	自主学习中心	18086506031	395169	A1065843
2014-11-28	08:22:09	流通借出	赵燕琼	zyq	自主学习中心	18086506031	395169	A1065842
2014-11-28	08:50:07	流通借出	钟小利	zx1	中区分馆	18120584645	371339	A0995805
2014-11-28	08:54:24	流通借出	钟小利	zx1	中区分馆	0408186	415218	A1107784
2014-11-28	08:54:26	流通借出	钟小利	zx1	中区分馆	0408186	300683	A0817641
2014-11-28	08:54:27	流通借出	钟小利	zx1	中区分馆	0408186	438445	A1146722
2014-11-28	08:54:28	流通借出	钟小利	zx1	中区分馆	0408186	412335	A1107798
2014-11-28	08:54:30	流通借出	钟小利	zx1	中区分馆	0408186	185410	A0506197
2014-11-28	08:54:31	流通借出	钟小利	zx1	中区分馆	0408186	290798	A0878280
2014-11-28	09:16:10	流通借出	钟小利	zx1	中区光盘库	18086507083	318400	GP872090
2014-11-28	09:23:21	流通借出	钟小利	zx1	中区分馆	18062045350	353155	A0965941
2014-11-28	09:34:34	流通借出	钟小利	zx1	中区分馆	17702737232	240828	A0709476

外借情况分析：又或者可以根据提出上一学年的某个学生的借阅书籍的清单，来分析该生可能是哪个专业的学生，该生偏向于哪一类的书籍，为该生做出一些阅读习惯方面的预估，并对该生进行相关类型书籍做一个个性化的推送。也可在学生毕业的时候对该生整个大学时代所借阅的图书做一个清单，并附上一些有针对性的总结。

武昌分校图书馆积累的大数据可以做什么？

2

分析读者资源使用习惯及偏向

以网上电子资源读者使用情况为

序号	数据库名称	7	8	9	10	11	13	14	15	16	16			
1	图书馆主数据库	(应试考试)考研数学(通关基础)	926	193	274	399	619					2411	#DIV/0!	
		(应试考试)考研政治	938	174	200	477	720					2509	#DIV/0!	
		(应用外语)英语语法	210	184	264	557	629					1844	#DIV/0!	
		(应用外语)英语词汇	435	119	374	538	745					2211	#DIV/0!	
		(应用外语)新概念第二册同步练习课	392	231	421	628	700					2372	#DIV/0!	
2	图书馆资源数据库	(求职辅导)大学生求职心理辅导	183	119	274	475	620					1671	#DIV/0!	
		(求职辅导)项目管理	140	98	213	174	283					906	#DIV/0!	
		(赠送课程)视频学习课程-时时直播课	900	63	394	573	592					2522	#DIV/0!	
3	本馆已购数据库	(赠送课程)视频学习课程-英语直播课	367	44	384	234	453					1432	#DIV/0!	
		(赠送课程)视频学习课程-大师讲堂	720	184	274	489	592					2259	#DIV/0!	
		博图外文图书数据库(2011年10月10日开始试用,2012年3月签订合同购买)												
		阅读次数	87	1184	505	69	284						2129	#DIV/0!
		下载次数	5	210	132	26	17						390	#DIV/0!
4	读秀学术数据库	天方有声数字图书馆(2012年3月23日开始安装本地试用)于2012年12月3日签订购买合同												
		访问量	72	120	120	97	48						457	#DIV/0!
		中国知网(2012年5月30日—2012年10月31日、12月31日试用)												
5	书生之家	登录次数	2408	6147	26216	35558	31036					101365	#DIV/0!	
		检索次数	186868	496073	317324	323225	396834					1720324	#DIV/0!	
		浏览次数	6020	10660	8274	11692	12645					49291	#DIV/0!	
		下载次数	7361	11494	31255	43670	36587					130387	#DIV/0!	
6	方正Apabi	超星数字图书馆(2014年6月1日正式购买)												
		总访问量(次)					5854	3119	2401	4853	5621		21848	#DIV/0!
		期刊论文下载量(次)					2876	492	592	1844	1772		7576	#DIV/0!
		中文期刊文献传递统计(篇)					170	22	45	108	149		494	#DIV/0!
		中文学位论文传递统计(篇)					116	148	22	98	48		432	#DIV/0!
		中文会议论文传递统计(篇)					4	0	0	5	0		9	#DIV/0!
6	超星移动图书馆(2014年6月1日正式购买)	外文期刊传递统计(篇)					103	22	12	12	3		152	#DIV/0!
		超星移动图书馆(2014年6月1日正式购买)												
		总访问量(次)					263747	119178	108414	502230	679296		1679374	#DIV/0!
6	超星移动图书馆(2014年6月1日正式购买)	(应试考试)考研英语宗题精讲	1201	204	438	473	601						2917	#DIV/0!

武昌分校图书馆积累的大数据可以做什么？

2

分析读者资源使用习惯及偏向

分析：总体的访问量上升了，偏学术性数据库、多媒体学习库的访问呈下降状态，大众化中文报刊库即博看数据库、天方有声数字图书馆和外文EBSCO数据库呈上升状态。原因如下：

- (1) 学术性数据库：2013年本馆新引进“中国知网”权威期刊库，从表可知其访问量相当高，“读秀、超星、书生、维普”等库的期刊需求访问量被分流了一部分；另外，如“读秀学术搜索”本学年访问量下降，虽然总体是下降，但纵向分析该库各月的访问量，可以发现靠近读书月活动的月份，其访问量明显上升，而且今年读书月活动有以读秀为主题的专项活动，其访问量就比去年四月还高2000多次，说明读者对该库的认知度还需要加强。
- (2) “博看期刊库”收录3000多种畅销期刊，得到读者的青睐；
- (3) “EBSCO外文期刊库”是本馆唯一一个正式订购外文库，对学生完成专业论文的有很大帮助，越来越多读者的喜爱；
- (4) “新东方多媒体学习库”的访问总体呈现下降，但其中的“应试类考试”模块是呈现上升状态。库的使用有关，从一定程度上分流了读者访问量；
- (5) 其他存在下降的数据库，常丰富全面，2013年申请试用数据库有22个，读者使用的选择面扩大，会导致已购数据库访问量的下降。

武昌分校图书馆积累的大数据可以做什么？

3 分析读者进出馆次数，时间段来知道到馆行为习惯

分析：可以通过我馆的进馆刷卡记录数可以知道某段时间内读者进馆的人数，某个读者的进馆次数，从而来预算出一天中哪个时间段到馆的人流量最大，或者通过几个时间段预计出每月中哪几天可能人流量大，从而来调整我们馆员的管理的方式，方法。

武昌分校图书馆积累的大数据可以做什么？

4

分析实名制座位申请使用行为习惯

分析：可以通过系统内对实名制座位的申请记录来知道，每次实名制座位的开放座位数，读者申请到实名制座位的用途，可以知道我校学生每年考研，考各类认证的人数，也可通过提取前几批开放的实名制座位数量来更准确的预留下一批的实名制座位数量，不至于出现预留不够的情况。

另外还可根据申请了实名制座位的学生的入馆记录进行判断，若该生一段时间内连续未能到馆，系统中可以对这类学生发出相应提示，这样也方便了我们的图书管理员对实名制座位的管理。

武昌分校图书馆积累的大数据可以做什么？

5

分析读者对爱心伞的借用行为习惯

分析：可以通过系统内读者的历年借伞记录可以知道我馆爱心伞的使用情况如何，来确定是否有必要增加或者减少爱心伞的数量来满足读者需求，另外也可以通过爱心伞每年的借用时间段来分析当年的哪个时间段可能会下雨，以便于提前预备爱心伞供读者使用。

另外也可把借阅超时，爱心伞借用在规定时间内未还回，图书馆座位选座逾期未使用等记录相结合做一个该生在我校图书系统内的一个信用指标。

武昌分校图书馆积累的大数据可以做什么？

6

分析读者续借的行为习惯

分析：可以通过系统读者的续借记录知道有多少读者使用的是到图书馆前台进行续借，有多少读者使用的是web端或者手机移动端进行续借的。

通过续借方式来分析哪些读者更喜欢使用网络来访问图书馆的资源，发掘这些读者，在我馆有相关的网络方面的讲座，新的电子资源开始试用时，可以进行有针对性的推送。

武昌分校图书馆积累的大数据可以做什么？

7

分析读者阅读需求趋势

分析：可分析每月通过图书馆主页下的图书荐购提交到系统后台的数据，以及各院系老师直接反馈给图书馆管理员要求购买的书籍信息，来分析读者对于书籍的一个需求方向，了解到读者的一个阅读需求趋势。

另外在我馆有新书到来时可根据这些荐购记录，根据不同的读者需求可做定制性的推送。

分析读者行为的结果对我们工作的指导作用

1

可以通过自行开发的内阅扫描系统来统计出各类期刊，图书的内阅数量，内阅率，可以看出这段时间内读者馆内阅读喜好于某类期刊，图书，可以根据这个分析结果可以做些排架的相应调整，也可为接下来图书，期刊的配置的数量应该往哪方面倾斜。

2

可以通过我们的图书管理系统提取出有哪些图书，期刊被读者外借过，哪些书的借阅率较高，可以根据这些频率的估算我们可以估算出馆藏副本配置多少算合理，这样的话就不用所有的图书和期刊的副本量按照一个标准来，可以将副本量更合理化。

3

也可根据内阅率和外借率的分类来进行有针对性的新书展示推荐，也可以为我馆每年读书月活动提供一些参考信息，让我们的工作的效率更高。

分析读者行为的结果对我们工作的指导作用

4

(1) 合理发展馆藏文献资源。要结合本学院的专业设置，有计划地购书，切实保障重点学科的图书，引进高质量的通识性人文图书和数理化基础教育图书，为素质教育和创造性人才培养提供优质精神食粮，同时兼顾新兴学科、边缘学科、交叉学科文献需求，使经费和藏书得到合理的使用。

(2) 科学布局馆藏文献。以我馆为例，在充分分析读者图书外借、内阅情况下，充分掌握了读者需求量小而长期滞留在书架上的一批馆藏图书信息，结合本馆馆舍情况，规划建设三线书库，即把那些数量多但长期没人使用或副本多、内容陈旧的图书，从各个阅览室清理出来集中管理，使图书布局更加科学。

(3) 根据读者阅读喜好，可以定期展示相关的热门书籍推荐给学生阅读，从而来提升我馆的服务工作质量，从而吸引更多的读者利用图书馆。

我馆的大数据应用所做的工作及还可进一步展的工作

1

图书，期刊信息数据的增加，收集，存储，分析。

2

读者进馆，借书，还书，实名制座位申请，爱心伞借还等记录的收集，存储，分析。

我馆的大数据应用所做的工作及还可进一步开展的工作

3

接下来可进一步开展的工作：

(1)、将读者与图书馆产生的相关的数据进行进一步的深度挖掘，得出更可靠的，对于我们工作有价值的结果。

(2)、大数据时代我们图书馆的数据也在日益增长，数据的安全性，稳定性的保障以及它的存储方式也将是我们研究的一个方向。

The background is a deep blue gradient with a subtle grid pattern. It features several glowing, translucent spheres of varying sizes, some with bright highlights. There are also abstract, glowing blue shapes, including a large, curved ring-like structure on the left and a flowing, ribbon-like shape on the right. The overall aesthetic is clean, modern, and high-tech.

谢 谢!