

嗨， 咱们和大师聊聊

通识教育名家讲坛

杨士军 张海森 韩云松\主编

专辑
2

大教授走进小教室 小课堂融入大社会

处于人生**成长关键期**的学生，不能陷在课本里
来，一起聆听**大师们的智慧**
共同走上**博闻强识**之路

江苏凤凰科学技术出版社 | 凤凰顾合

嗨，咱们和 大师聊聊

——通识教育名家讲坛 专辑 2

主编：杨士军 张海森 韩云松

顾问：朱清一

编委（按姓氏笔画排序）：卜昱翎 王树生 代龙妹 乐 敏
朱超宇 向晋榜 刘 丹 刘 雪 刘 赞 刘子源 刘春燕
李 荔 李晓颖 吴蒙昊 张 隽 张 靓 张心圣 张卓君
张晓敏 陆卫忠 陆宏亮 陆剑毅 陆鹏娟 陈 虎 周孝芳
庞维成 胡婷婷 袁瑞茜 徐 城 奚高峰 曹 彦 程 芳
蔡英洁

图书在版编目 (CIP) 数据

嗨,咱们和大师聊聊:通识教育名家讲坛.专辑2/
杨士军,张海森,韩云松主编.--南京:江苏凤凰科学
技术出版社,2016.1

ISBN 978-7-5537-5296-9

I. ①嗨… II. ①杨… ②张… ③韩… III. ①社会科
学-文集 IV. ①C53

中国版本图书馆 CIP 数据核字 (2015) 第 212821 号

嗨,咱们和大师聊聊——通识教育名家讲坛 (专辑 2)

主 编 杨士军 张海森 韩云松

责任编辑 傅 昕

责任校对 郝慧华

责任监制 曹叶平 周雅婷

出版发行 凤凰出版传媒股份有限公司

江苏凤凰科学技术出版社

出版社地址 南京市湖南路 1 号 A 座, 邮编: 210009

出版社网址 <http://www.pspress.cn>

经 销 凤凰出版传媒股份有限公司

印 刷 无锡市证券印刷有限公司

开 本 787mm × 1092mm 1/16

印 张 15.75

字 数 233 000

版 次 2016 年 1 月第 1 版

印 次 2016 年 1 月第 1 次印刷

标准书号 ISBN 978-7-5537-5296-9

定 价 39.80 元

图书如有印装质量问题,可随时向我社出版科调换。

序

风雨兼程不止步

成长，是一种无需声张的厚实，是一种明亮但不刺眼的光辉。

2005年，复旦大学率先在国内高校中推行通识教育，这给复旦附属学校及时提了一个醒，那就是：教育不是百米短跑，而是一场接力赛。因此，所有的教育教学活动的开展均需要遵循学生身心发展的规律，一切按教育规律办事，于是想到了要进一步推进通识教育理念，何况，在复旦二附中实践通识教育观具有得天独厚的条件。我们对于在中学渗透通识教育的进一步解读就是“重德育体验，重创新实践，实行有宽度的教学，培养全面的可持续发展的人才”，相应的举措即“五个多”：读的书比别人多，听的讲座比别人多，参观访问比别人多，参加的社会实践比别人多，经历的创新活动比别人多。

于是，2005年秋季，我校开发并推出了“名师讲坛”这一资源型课程，到如今已经整整十个年头了。一门校本课程，能够坚持做10年，这就足以说明这门课程存在的必要性和其中的价值所在。

2013年，我校从前8年的60余讲名师讲坛学者讲座里精选了20讲，和江苏凤凰科学技术出版社合作，编辑出版了《听大师讲课——中学通识教育名师讲坛（专辑1）》（后更名为《嗨，咱们和大师聊聊——通识教育名家讲坛（专辑1）》），有效地辐射了课程资源，也得到了社会好评。时间又过去了两年多，而我们的专家讲座也接近100讲，其中又有大量对学生的通识教育和可持续发展产生助力的精彩讲座。为此，我校与江苏凤凰科学技术出版社再度携手，在专辑1的基础上推出了专辑2。

和专辑1相比，专辑2在讲座编辑体例上做了一些调整。首先是优化了“导读”环节。“导读”以简洁明了的语言，向读者介绍了每个讲座的主要内容和应该特别关注的环节与知识。其次是把整个讲座用小标题的方式分成若干个段落，让读者在阅读的时候能够不因讲座篇幅过长而产生阅读的疲劳感，实效性也得以提升。第三，与专辑1相比，专辑2在每个讲座中配上了若干与讲座有关的照片、示意图等，增强可读性。第四，改进了“拓展延伸”环节，在每个讲座里，向读者推荐1~3个与讲座相关的拓展知识，使讲座的外延进一步拓展，也便于师生基于讲座进行必要的探究实践活动，培育学生的创新精神。

在本书出版之际，我们要特别感谢专家学者们的大力支持，其中有不少学者来自于复旦大学及周边的高等院校，还有不少学者本身就是我们的学生家长，是因为他们的无私奉献和鼎力相助，每月一至两讲精彩纷呈的讲座才能如期顺利进行。我们还要感谢上海市杨浦区教育局和杨浦区教师进修学院在政策、资金、精神、技术等诸多方面的支持，尤其是信息中心的技术支持，使讲座能及时、顺利地以网络直播的方式让区内的其他初中兄弟学校的学生同步观看，讲座的受益面更广了，那么我们的工作也就更有价值了。此外，我校作为复旦大学高等教育研究所的实践基地，每年都有研究生在我校实习，他们也参与了将讲座视频翻成文字实录稿的艰苦而细致严谨的工作，在这里我们一并致谢。

践行“给学生更宽广的教育”是我们一以贯之的办学理念，促进学生的全面可持续发展是我们教育工作者的不懈追求。风雨兼程不止步，“名师讲坛”这一能对学生的思考力、想象力和创造力发展起到正面推动作用的举措，我们必将坚持作为。为了持续推进并借助更高、更大的平台去拓展孩子们的视野，我们期待着有更多的专家学者来为中学生做各类讲座，我们也期待在日后的若干年里，会陆续看到《嗨，咱们和大师聊聊》专辑3、4、5……的问世，那将是我们学校的幸事，更是我们学生的福祉。

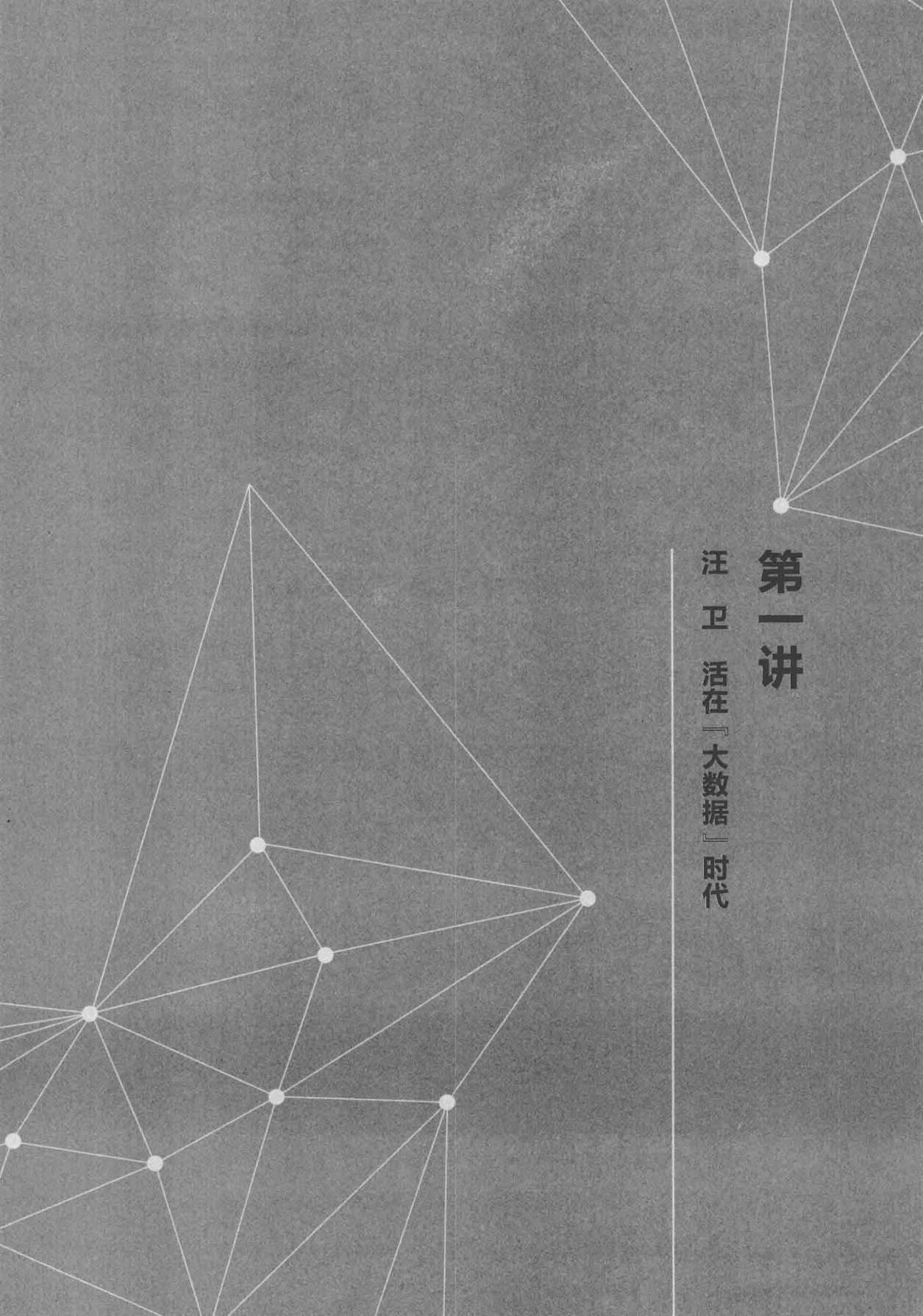
杨士军

2015年6月18日

目录

CONTENTS

第一讲	汪 卫	活在“大数据”时代	01
第二讲	李大维	“玩”出来的创新——创客文化和创客空间	17
第三讲	娄永琪	设计漫谈	29
第四讲	孔爱国	华尔街之门——货币与金融常识	49
第五讲	陆大江	基于“云健康”的生活方式——远程监控	63
第六讲	庄松林	光的神奇世界	75
第七讲	柴之芳	辐射是把双刃剑	91
第八讲	金亚秋	仰望星空——电磁波信息科学与空间遥感技术	109
第九讲	施蕴渝	奇妙的生物大分子机器	125
第十讲	金 力	解读写在基因中的人类历史	137
第十一讲	江 明	化学创造另一个世界	153
第十二讲	麻生明	与我们息息相关的化学	165
第十三讲	沈 愈	2013年上海夏季极端高温及成因	181
第十四讲	戴慎志	灾难面前如何自救——城市灾害与安全防灾	195
第十五讲	段 枫	从《哈利·波特》看英语和英语文化	211
第十六讲	杨东援	上海交通的昨天、今天、明天	229



第一讲

汪卫 活在『大数据』时代

专家简介

汪卫：复旦大学计算机科学技术学院教授、副院长，中国计算机学会数据库专业委员会委员，上海市计算机学会理事，数据库专业委员会副主任，ICDM、SIAM DM、CIKM等重要国际学术会议的程序委员。初中和高



中均就读于山东师范大学附属中学，1988年考入山东大学计算机系攻读本科学位，1992年考入复旦大学计算机系攻读硕士学位，并于1998年在复旦大学获得博士学位，毕业后留校任教。

长期从事数据库与数据挖掘领域的研究和开发工作。在SIGKDD、SIGMOD、VLDB、WWW、ICDE、IEEE TKDE、JIIS等国际权威的学术会议上发表论文近30余篇，被引用1000余次（Google Scholar）。作为主要参与者获得上海市科技进步一等奖一项、二等奖两项、三等奖一项，并获高等教育国家级教学成果二等奖一次。

演讲关键词 | “大数据” “大数据”时代 “大数据”思维智能

导读

最近，我们常常能在新闻里、报纸上看到“大数据”（Big Data）这个词，有人说我们已经进入了一个“大数据”时代。可是，“大数据”究竟是什么呢？只是大量数据的集合吗？它离我们的生活有多远呢？其实，“大数据”虽然有很强的专业技术性，不容小觑，但是它并不神秘，就在我们的生活中。卫生部门预测并预防疾病、商家投放广告、警察指挥交通等都得到了“大数据”的帮助。但更重要的是，“大数据”为我们提供了一种思维方式，让我们生活中的智能活动变得简单。你是不是已经迫不及待地想要知道更多“大数据”的秘密了呢？那就让我们一起跟随汪教授进入“大数据”的世界吧！

嗨，
咱们
和
大师
聊聊

活在“大数据”时代

一、“大数据”在哪里

首先，让我们通过一些案例来看看“大数据”存在于哪些地方。

1. 医疗卫生领域中的“大数据”

一位谷歌工程师在《自然》杂志上发表了一篇文章，预测美国会在什么时候爆发季节性流感，也就是我们所说的感冒。感冒在美国被认为是一种很严重的疾病，由于某些体质的原因，感冒在美国引发的死亡率较高，因此美国人对于流感的预防非常重视。但是，以前美国的卫生组织在预测何时会发生流感时，只能先去当地的医院了解数据，再判断是否有流感大规模爆发和蔓延的趋势，以此来作相应的防控工作。现在，这位谷歌工程师通过在谷歌的搜索引擎里对比美国人检索最频繁的5000万条词汇后发现，当某个地区流感较为严重时，该地区的居民搜索和“流感”相关联的词的比例就会非常大。由此，他经过了一定的数据分析后得出了一个预测方法，预测流感爆发的准确性高达97%，几乎某个地区一爆发流感马上就能发现。而以往的疾控中心在了解相关情况时往往要1~2周时间才能得出分析结果。谷歌工程师只需通过谷歌搜索引擎数据的分析，就能很快了解到流感的蔓延趋势，这就是“大数据”应用

的一个非常好的例子。

2. 经济领域中的“大数据”

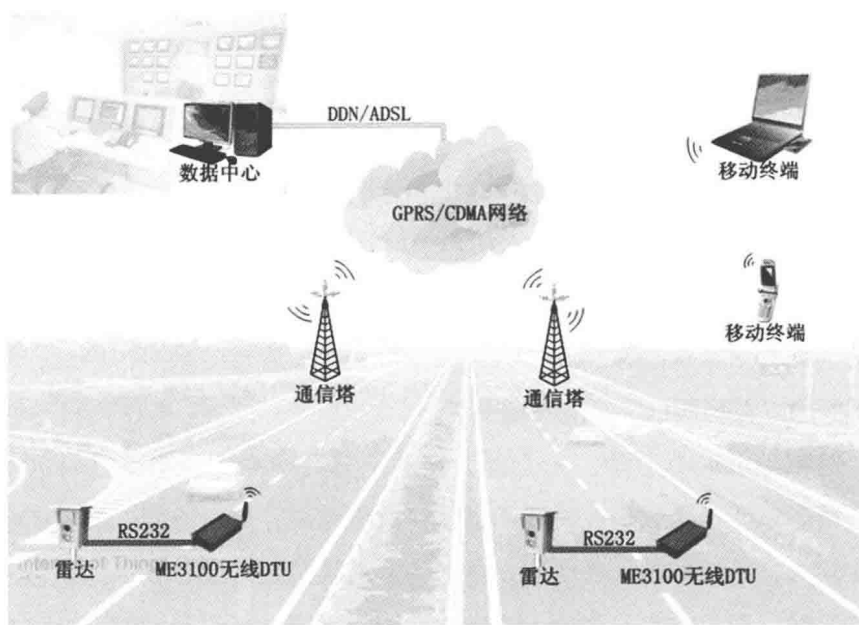
在中国，淘宝公司除了提供网上购物服务外，还曾做过一件伟大的事情：2008年，淘宝发现在其阿里巴巴的平台上，买家的询盘数^①急剧下降。通常买家在购买商品时，总会多找几家商铺询盘，经过比较后找到最为合算的。而询盘数量的减少就意味着最后成交数量肯定会减少，由此，淘宝就预测出“经济危机”可能快到来了。

这种预测很有道理。一方面，当时淘宝在电商市场份额占有90%以上，所以它可以有效地反映我国零售行业的情况；另一方面，淘宝的覆盖面比较大，不仅中国人在使用，世界各地的人都在使用。我去阿里巴巴参观的时候，看见它有一张很大的世界地图，地图上标注并实时反映哪个地方的人买了一件什么样的商品。我国统计局也经常发布一些指数，但是和淘宝一对比就会发现原来的统计方式非常落后。原先统计局会在全国各地安排1万~2万名市场调研人员，去各个商场观察消费者的消费情况，现在有了淘宝之后，他们发现不需要安排那么多的人员，可以直接使用淘宝的数据进行分析。所以，有了“大数据”之后，我国一些诸如统计的工作都可以用新的方式来解决。

3. 互联网营销中的“大数据”

我们非常熟悉的“移动互联网”也与“大数据”有着非常大的关系，特别是互联网营销，比如广告的精准投放。实际上，我们在网上购物、即时通讯聊天时，都或多或少地在互联网上留下了一些信息，一些互联网企业就针对我们所留下的信息进行数据分析，然后进行精准的广告投放。

^① “询盘数”是指“询盘”的数量。“询盘”是买方为了购买或卖方为了销售货物而向对方提出有关交易条件的询问，内容可以是询问价格或交易条件等。——王君. 浅析外贸询盘的判定处理及常见问题[J]. 外贸业务探讨, 2014(6): 56.



移动互联网

二、“大数据”如何发挥作用

那么，“大数据”是如何发挥作用，并给一些行业带来利益的呢？

1. 通过收集数据、分析数据，“大数据”能提供全面的信息

电商网站对于我们每个人在诸如社交网站、电子商务网站上的行为都有着很全面的信息积累，他们收集数据之后，就可以给每一个人“画像”。以腾讯为例，用户会使用QQ聊天、浏览腾讯新闻、玩腾讯游戏等，在这一过程中，腾讯就把用户所有的信息都收集下来，随后分析每一个用户的情况。首先，腾讯会分析用户是哪个区域的，比如是在上海、北京、国内其他地区或者国外。其次，腾讯可以判别用户的身份，比如学生、公司职员、政府职员等。如果学生登录到一个社交网站，那么他所接触的社交成员大部分也是学生，因此可以通过社交网络中的身份特征来确定用户的身份属性。腾讯还可以判断用户的爱好。如果用户在浏览新闻时，更多关注体育，那么他的爱好可能就是体育方面的；如

果更多关注音乐、文化等，那么也基本可以确定他的爱好。诸如此类，都可以通过收集用户在互联网上行为的数据来进行分析。对于腾讯来说，这是他们内部的一些数据，但对于每一个用户来说，腾讯有一千多种关于该用户的描述信息，包括身份、性别、爱好等等。因而从某种角度来说，真正了解你的可能不是你自己，而是腾讯。

2. 通过数据收集与人工智能的结合，“大数据”能有效协助警务工作

“大数据”另外一个较成功的案例是智能交通。平时我们走在马路上的时候，会发现上方安装有很多摄像头，尤其是在交通信号灯附近，这些摄像头起着非常大的作用。因为现在计算机智能技术已经比较成熟，当拍到一辆汽车的照片之后，摄像头可以很清楚地识别出汽车的型号、牌照等信息。有了这些信息以后，电子警察可以监控这辆车的运动轨迹，从而了解它是否有违章停车、超速等违法行为。另外，有了这些数据之后，也可以用来指导交通，比如计算出一段时间内、在某一段道路有多少辆车经过，来确认该段道路是否拥堵。再比如发生在济南的一个案例，有一个人抢劫了一家金店并驾车逃逸，监控中心发现其车辆号牌被遮挡，于是警方通过智能系统观察以往经常和这辆车在一起的车辆，发现了另一辆没有遮挡号牌的车，于是通过该车车主的信息最终抓获了抢劫金店的犯罪嫌疑人。从中我们可以看到，这样一些交通数据可以提供很多帮助。

三、“大数据”从哪里来

那么，“大数据”又是如何产生的呢？

其实自从有了计算机以后，人类社会就在不断积累各种各样的数据，当数据积累到某一个特定的时刻，它就发生了一种变革，形成了“大数据”的概念。过去我们常说电脑的内存和硬盘是多少MB，现在则已经升到了GB，而真正的“大数据”，其规模要比我们日常所知道的单位要大很

多，已经从GB跨越过TB、PB、EB，来到了ZB，也就是GB的 2^{40} 倍这样的规模。

“大数据”确实能在我们的现实生活中起到很多作用，最著名的例子就是预测2012年的美国大选。对于哪一个州、哪一个党派会胜选的预测准确率非常高，使得人们相信“大数据”的价值。

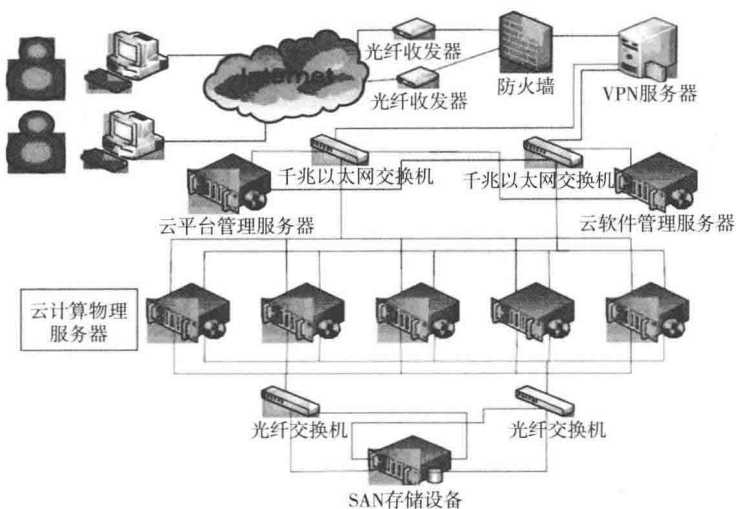
1. “大数据”到底是怎样形成的

实际上，自从有了计算机，研究者就一直在摸索如何让计算机智能化，如何让计算机去分析一些数据，而随着数据的分析，我们可以获得更多的知识。所以，一开始我们有了概率统计的分析、数据挖掘的分析，到今天就有了“大数据”的分析。

2. “大数据”概念的形成并不是孤立的

“大数据”与其他很多技术有关系，比如云计算^①，它使得“大数据”走进了民间。普通人家一般都有一台计算机，最多三四台。但是“大数据”里的数据规模非常大，一般的计算机里仅能放1TB的数据而已，而所谓的“大数据”一般要达到1000TB，那么如何来处理那么庞大的数据呢？云计算就起到了这样一个作用。在云计算方面出现了许多企业，它们的工作就是租赁机器。这些企业本身拥有上千台机器，通过某种技术将其中的一些机器租赁出去，但不是通常意义上的被搬走，而是以技术手段租赁。用户可以使用这些机器，甚至可以一次性租用1000台，当然也可以一次只租用1小时。这种技术的出现就成为了分析“大数据”非常有利的一个方面。当用户在完成一项任务的时候，不需要去买1000台机器，而只需临时租用，省去了很大一部分成本，因此，在云计算的基础上，不论有多大的数据量，都可以有条件进行处理。

① 云计算将IT相关的能力以服务的方式提供给用户，允许用户在不了解提供服务的技术、没有相关知识以及设备操作能力的情况下，通过Internet获取需要的服务。比如在医学领域，就可以通过云计算进行DNA信息分析等。——陈全，邓倩妮：云计算及其关键技术[J]. 计算机应用，2009（9）：25-63.



云平台典型物理架构

3. 我们对“大数据”概念的认识还要再深入

实际上，认为数据量多就是“大数据”，这种认识还是比较浅显的。为什么到今天我们才提出这个概念，而不是三年前或五年前？原因是：虽然以前也有各种各样的系统在收集数据，但能力比较有限，很难得到关于一个人各方面的数据，而如今可以很方便地得到关于一个人各方面的信息。比如复旦大学通过学生的校园卡就可以了解到学生的很多信息，包括吃饭的信息（如一个星期有多少天在食堂吃饭、每顿饭花费多少钱），以及作息的信息（如何时进出宿舍、图书馆等）。因此，通过校园卡的数据，就可以了解一个学生一天内的行为。学生是懒惰还是勤奋，都可以通过数据分析出来。通过教务处的数据比较，还可以了解学生成绩的升降情况。另外，校园里有许多摄像头，借助人脸识别技术，可以了解某个学生的校园行走路线。当学生上网时，根据校园网也可以知道他经常关注哪一类网站。

正是由于今天我们采集、收集数据的能力比过去提升了许多，所以可以形成对某一事物很全面的数据覆盖。有了这些数据，我们对一个事物的了解就可以更深入，有了积累“大数据”的条件，“大数据”也就应运而生了。

四、“大数据”带来了什么

1. “大数据”带来了思维的进步

“盲人摸象”的故事大家都知道。由于每个盲人摸到了大象的不同位置，他们对于大象的认识也是不一样的。在“大数据”时代来看这个故事，其实这也不是一件坏事。七个盲人摸象，摸到的是七个不同的位置，如果他们摸完之后讨论交流一下，或许就可以得出一个关于大象的完整印象了。所以在“大数据”里面，我们每个人看到的仅仅是一小块，或者是很片面的一块，如果我们每个人把自己认识的那部分分享出来，也许就可以得到对一个事物非常完整的认识，这就是“大数据”思维中很重要的一点。

2. “大数据”带来了智能的进步

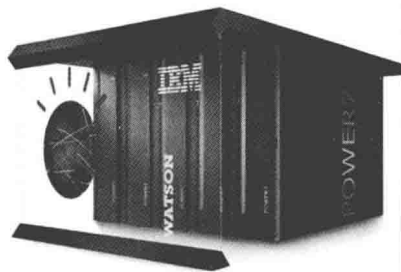
（1）“大数据”让智能数据分析更便捷

“大数据”思维和计算机智能有很大的关系。智能化一直是计算机研究的目标，智能=数据+推理，它一部分是拥有数据，一部分是拥有推理能力。就如学生的学习，首先是学会各种各样的知识，其次是学会思维的方式，人工智能也是如此，首先要掌握一定的数据，其次要在数据上推理，形成知识。过去我们没有现在那么多的数据，就只能使用采样的数据，采样数据只是整个数据中的一小部分。为了解决这个问题，计算机的研究者就需要尝试设计很多种模型，使数据尽可能完整。今天有了“大数据”之后，就可以在所有的数据上来做分析，模型就不需要像原来那么复杂了。从人工智能角度来看，这就是“大数据”带来的巨大进步。

（2）“大数据”让人工智能更“聪明”

2011年，IBM与美国德克萨斯大学联合研制的超级电脑——沃森（Watson）在美国最受欢迎的智力竞猜电视节目《危险边缘》中击败了

该节目历史上最成功的两位选手，成为节目的新王者。沃森的反应相当快，它在很短的时间内能正确回答问题的原因是：它存储有2亿页各种各样的文档资料，包括新闻、电视剧本资料、词典、文选和世界图书百科全书等。它能在不到3秒的时间里根据问题，基于“大数据”的资料进行搜索并推理。从沃森的成功来看，可能我们造不出具有多种情感的“人”，但我们可以造出一个百科全书式的“人”。虽然现在的计算机还不具备人类思考、创新的能力，但是它却可以具备很多知识，这也证明了人工智能方面的巨大提升。



超级电脑——沃森



电视节目《危险边缘》现场

（3）“大数据”让我们生活中的许多智能活动变得简单

以翻译为例，以前机器翻译被认为是极其复杂的，因为每个人都有不同的说话风格，不同人群和地区的发音也各不相同。如果需要做分析，就要把每个人说话的模式或语法记录下来，然后在这个基础上进行翻译。但是有了“大数据”之后，这种情况得到了改善。

谷歌和微软最早开始制作基于“大数据”的机器翻译，在它们的搜索引擎中有超过一百万个单词的双语文本语料库，即中文和英文之间的对照，它们也可以找到超过10亿个单词的语种语料库。随后它们对这两个语料库进行分析，观察中文和英文语法结构之间的对照关系和对照模式，比如主谓宾的顺序变化，以此为原则进行翻译。如果我们将“复旦大学”的英文介绍放入谷歌翻译进行中文的转换，语法结构和语义是基本正确的，而且最突出的是，它能够将复旦大学的党委书记和校长的名字以准确的文