中国教育财政

怀仁怀朴　唯真唯实

北京大学中国教育财政科学研究所

2015 年第 8-2 期（总第 107 期）　　2015 年 6 月 30 日

编者按：本文为和君集团高级投资经理高步云在 2015 年 5 月 5-6 日“教育、技术、融资与政策——ICT 和 EMIS 的应用及其政策影响”研讨会上的发言。

教育领域的大数据应用

高步云*

一、大数据的本质

大数据的经典定义可以归纳为 4 个 V：海量的数据规模（volume）、快速的数据流转和动态的数据体系（velocity）、多样的数据类型（variety）和巨大的数据价值（value）。整个大数据的处理又分为获取、整理、储存、分析、呈现五个环节。数据的活性、粒度、是否在线也是考查大数据的指标。有些数据粒度不够，数据不够精细，也无法成为大数据。从大数据处理的五个环节来看，对数据的处理很大程度上是要整理掉那些无用数据和一些活性不大的数据。原有的教育体系也沉淀了很多的数据，但这些数据根本没有在线，也就不能用大数据方式和大数据思维去处理。未来世界是万物互联，物联网加上大数据技术，很可能到最后就可以帮助我们实现按需分配，这是未来科技发展的终极目标。

* 作者简介：高步云，和君集团高级投资经理。

二、教育领域大数据的本质

对教育领域来说，大数据到底是什么样的？加拿大阿萨巴斯卡大学教育技术领域的资深专家乔治•西蒙斯比较了数据挖掘和大数据，他认为最大的区别可能就是后者的自适应。在教育领域，大数据从学术的效应来说产生了两个学科，一是学习分析学，一是自适应学习。二者是一脉相承的。也就是说，后台分析出哪块需要加强，考学练配套的部分也就自动加强，即要因材施教。

三、数据维度教育领域大数据的分类

从数据维度展开，教育领域大数据分为四类：教育管理（行政）数据、教育行为数据、教育情景及环境数据、教育情景中的其他数据。

1. 教育管理（行政）数据。目前有很多的数据都是教育管理数据，这些数据因为面临一些问题，不能形成为大数据：（1）数据孤岛，数据之间没有互联互通，是完全割裂的；（2）数据泥潭，数据活性特别差，缺少更新和流动；（3）数据荒漠，即根本不知道发生了什么，如在学生安全方面，不知道学生行为和安全到底有什么隐患。除此之外，还存在教育行政主管部门对学生数据的统计往往都集中在简单的生理数据上的问题。在管理的手段跟被管理者之间有激烈冲突的情况下，教育行政数据该如何采集、如何整理，以及如何应用，还不是特别清晰。

2. 教育行为数据。因为在线教育或者电子化教育形式还不够普及，目前已有的教育数据里，教育行为数据是最少的。在课堂上到底发生了什么，以前原始课堂里是没法记录的。但随着未来信息化技术不断进步，以及设备成本不断降低，在课堂上捕获学生数据将成为可能，如慕课平台已经实现记录使用者的行为数据。

3. 教育情景及环境数据。记录教室里到底发生了什么，不能靠人，而应该靠一套智慧系统。北师大黄荣怀教授提出了一套智慧教室的模型，包括五个维度：内容呈现、环境管理、资源获取、及时互动和情景感知。通过智能传感技术、物联网技术以及一系列体感和可穿戴设备，可以把在教室发生了什么记录下来。

4. 教育情景中的其他数据。以学生为中心展开其他非教育本体的相关数据都属于大数据范畴，往往可以从另外一个角度反映出教育的情况。如把大学生常用的所有 APP 根据其定位进行重新分类：沟通、社交、BBS、电商、教育、展示、生活若干维度，进行一个综合的统计。学生在网络上的行为数据可以在一定程度上反映出学生的学习水平以及他们的认知水平。

四、教育情境维度教育领域大数据的分类

1. 幼教领域。幼教领域最核心的数据是家长与孩子的互动数据，而这种数据产生必须要有一个场景化的东西，以这种数据为延伸的教育机构及其提供的服务，是比较符合这个产品方向的。在这个领域，纯线上也有像宝宝树这样做得非常成功的，但如果没有一个线下实体，是很难提供服务的，所以更适合 O2O 的商业模式。在线下有了一定基础之后，把线下数据提出来放到线上，然后以线上数据作为核心资产，再做商业延展。

2. K12 领域。在中国，K12 领域最优秀的数据提供就是以考为纲的学习助理，抛开考试一定是脱离实际的。该领域家长最核心的两个需求，一是提分，二是升学。所以“考学练”，把考试放在第一即以高考为导向。如创数教育，在中国三四线甚至以下城市的县域中学开自习室，把数据平台接进去，然后让这些县域学生有机会通过教育设备（如平板电脑、远程教学设备等）享受到更平等的教育。具体做法就是以考试为中心，将高考数学题进行情景还原，学生在平台上做题，系统给出自适应的针对性的诊断，根据诊断推送相应的资源，同时就留下了以考试为中心的学习数据。

3. 高等教育领域。高等教育领域的数据应该是自主学习、自主生活、全面发展的数据。例如，彩校园，武汉的一个企业，针对民办学校，通过 O2O 模式把学生饮食起居数据全都收到平台上，然后做有针对性的推送。

4. 职业及成人教育领域。该领域强调重度垂直，以职业发展和认证为双核。以这两个场景为核心展开的数据，都是非常接地气的。例如：搜名律，是国内现在最大的一个律师平台；和君商学，是重度地为高校精英群体打造商学的生态圈。

五、总结

首先，大数据不是万灵药，而是人生平常、吃饭穿衣。其次，一切不以场景为基础谈大数据都是纸上谈兵，只有在一定场景才有数据有效性。再次，以数据属性维度展开的产业整合是未来发展的方向。以前的产业并购是以产业维度上下有整合作为一个发展方向，现在互联网公司之间的跨界整合特别多。第四，三大力量改变中国教育：（1）消费升级，家庭会有更多的钱投入到教育领域；（2）制度创新，现在整个教育体制都在进行改革，很多政策的红利就会释放出来；（3）技术革命，尤其是慕课这样的先进技术，还有未来虚拟现实等等一系技术。

“教育为本，政策为势，技术为魂，金融为器”。教育，是一个最根本的命题；政策，是大的战略，是未来的方向；技术，代表着创新的方向，是所有问题的魂魄；金融，是一个手段。

上期回顾

2015 年第 8-1 期（总第 107 期）

通过教育管理信息系统（EMIS）强化教育系统的国际经验

摘要：世界银行资深教育专家 Husein Abdul-Hamid 结合世界银行的实践, 介绍了什么是EMIS，如何系统地考查、对比不同国家的EMIS，发达经济体的EMIS发展情况以及EMIS的实践，并就大数据能否强化教育系统提出了几点意见。

《中国教育财政》由北京大学中国教育财政科学研究所主办；旨在反映本所最新的学术科研活动；相关内容仅体现作者本人观点，并不必然代表本所的立场。

文章内容仅供参考，如需转载须事先征得本研究所同意。

本期印发：1800 份

下载网址：http://ciefr.pku.edu.cn

主办单位：北京大学中国教育财政科学研究所

责任编辑：毕建宏　黄春寒

电　　话：010-6275-9700

地　　址：北京市海淀区颐和园路 5 号

北京大学教育学院楼 413 室（100871）

电子信箱：newspaper@ciefr.pku.edu.cn

传　　真：010-6275-6183

微信公众号：中国教育财政