



大数据时代观下的教育质量监控（下）

● 徐 勇

大数据时代观：一种可以采纳的思维模式

依据现有传统思维模式实现上述的“全数据”评估，必然存在人力、物力和海量数据的困境。因此，我们必须转变思维模式，革新传统监控评估技术。大数据时代观就是一种可以采纳的思维模式。其关键是“样本=总体”。

在信息处理能力受限的时代，世界需要数据分析，却缺少用来分析所收集数据的工具，因此随机抽样应运而生，“抽样”的目的就是用最少的数据得到最多的信息，它可以被视为那个时代的产物。但是，当感应器、手机导航、网站点击等被动地收集了大量数据，而计算机可以轻易地对这些数据进行处理，我们可以获得海量数据的时候，其缺陷越来越难以忽视了。〔见维克托·迈尔-舍恩伯格（英）《大数据时代：生活、工作与思维的大变革》（中文版）.浙江人民出版社.2012版，以下简称“维克托文”〕例如用抽样的方法分析整个人口的情况，正确率可达97%。对于某些事物来说，3%的错误率是可以接受的。但是对于教育来说，3%的错误率是不能接受的。数据处理技术已经发生了翻天覆地的改变，我们的方法和思维也要跟上这种改变，要以每一个学生数据中心，合法地收集他们所有数据，用“样本=总体”为思路，设计评估，设计质量监控。

（一）大数据时代观点综述

进入2012年，大数据(big data)一词越来越多地被提及，人们用它来描述和定义信息爆炸时代产生的

海量大数据时代来临，并命名与之相关的技术发展与创新。大数据时代观下有许多成功案例，如2009年甲型H1N1流感爆发的时候，谷歌更有效、更及时的发布了“流感”指示标。令人惊人的是，谷歌的方法是建立在大数据的基础之上的，而不需要分发口腔试纸或联系医生。这是当今社会所独有的一种新型能力：以一种前所未有的方式，通过对海量数据进行分析，获得有巨大价值的产品和服务，或深刻的洞见。（见“维克托文”）

为了引入大数据时代的基本观点，我们对这段时间内与监控有联系论点做一个综述。

首先，“从硅谷到北京，大数据的话题正在被传播。”“我们的行为、位置，甚至身体生理数据等每一点变化都成为了可被记录和分析的数据。……如今，一个大规模生产、分享和应用数据的时代正在开启。”“互联网时代，尤其是社交网络、电子商务与移动通信把人类社会带入了一个以PB（1024TB）为单位的结构与非结构数据信息的新时代。”大数据也将改变我们生活中最重要的方面，因为它为我们的生活创造了前所未有的可量化的维度。大数据已经成为了新发明和新服务的源泉，而更多的改变正蓄势待发。

其次，让数据“发声”。“大数据”专家杰夫·乔纳斯（Jeff Jonas）提出要让数据“说话”。在数字化时代，数据处理变得更加容易、更加快速，人们能够在瞬间处理成千上万的数据。但当我们谈论能“说话”的数据时，我们指的远

远不止这些。

其三，“样本=总体”。当数据处理技术已经发生了翻天覆地的变化时，进行抽样分析在大数据时代就像在汽车时代骑马一样。（见“维克托文”）随机抽样取只是一条捷径，是在不可收集和分析全部数据的情况下的选择，它本身存在许多固有的缺陷。一旦抽样过程中存在任何偏见，分析结果就会相去甚远。探索以大数据为基础的克服抽样缺陷的解决方案，是提高数据利用效率的重要手段。联系到我国现代化所面临的种种问题以及教育、交通、医疗保健等各方面挑战，通过大数据这种创新方式来解决，意义就更大。（见“维克托文”）

其四，大数据的核心就是建立在海量数据的基础之上的1预测。有预测的监控，才有可能不断完善受监控的学校或区域。而这种“完善”是自主，内在的，随着监控系统接收到的数据越来越多，它们可以聪明到自动搜索最好的信号和模式，并监控系统也将自己改善自己。将目前单纯依靠人类判断力的监控模式，转变为监控在计算机系统帮助下，成为受监测的主体和监控实施者共同自主完善模式下的监控行为。

从上述大数据的观点来看，目前在教育质量监控评估上，过于依赖“抽样”数据，都比较重视各种终端信息的采集，但对于管理和教学过程中的信息及发展性的信息关注不够，教学质量信息采集时效性较差，是出现“两费两差一不足”，不利于对教学质量进行全过程管理，影响教学质量评估诊断、促进效用的及时发挥。



师道

二〇一四年第五期