
关于深入推进大数据引进、使用、创新、提升，促进人才培养效果转化的发言

贵阳职业技术学院 刘海波

2016 年 11 月 15 日

2016 年 7 月 4 日至 8 月 2 日,贵阳市委组织了为期 21 天的大数据人才一选一专题培训班。我有幸作为贵阳职业技术学院选派的代表,参加了培训。此次培训会规格之高史无前例,邀请到一批国内一流学者和活跃在大数据产业生态各个领域的许多专家前来授课,为我们提供了一个极为难得的契机,能够与国内大数据顶尖人才进行深入交流。近一个月的时间,让我收获颇丰。

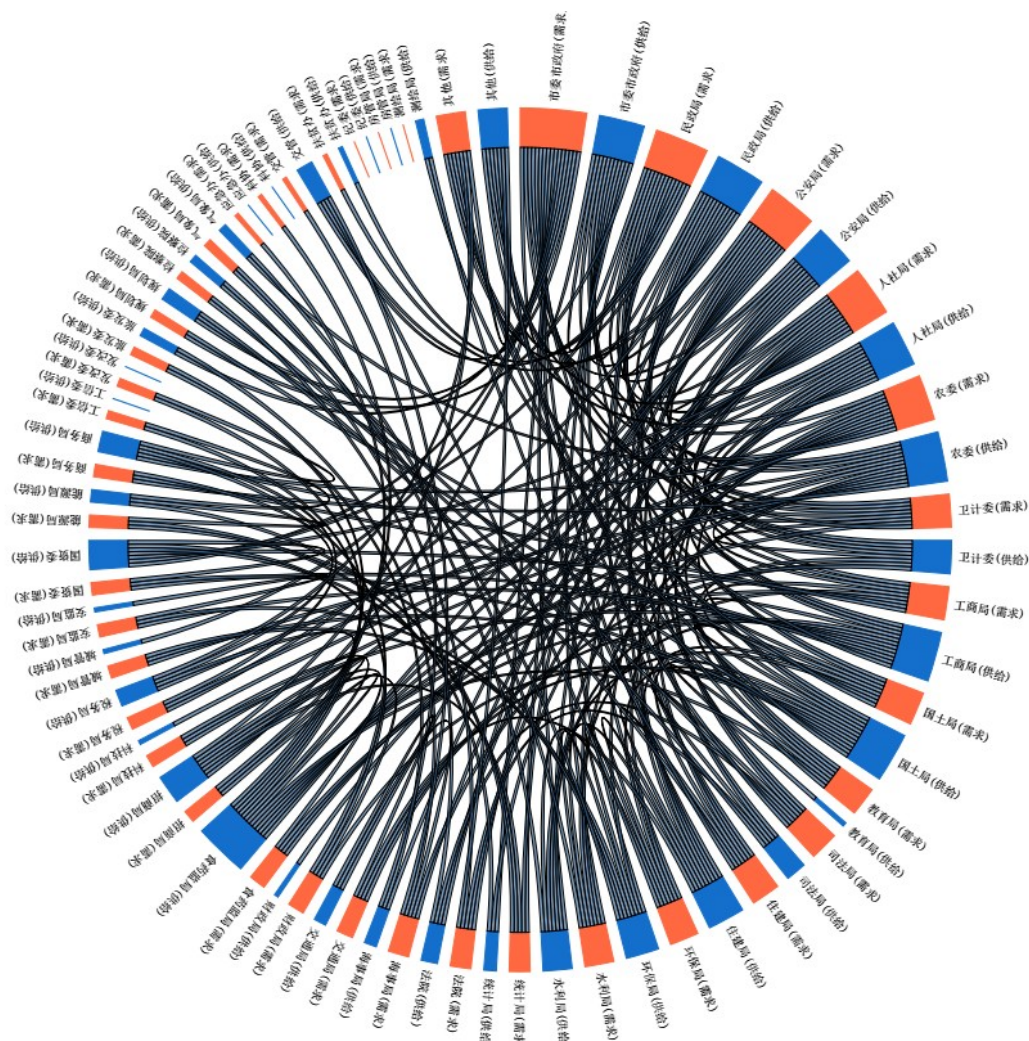
首先,此次培训让我能够从国家战略和贵阳市发展大数据的理论依据及方针政策的高度去理解贵阳市为何要发展大数据及怎样发展大数据。

其次,通过和全国专家的深入交流,让我对大数据概念的核心和外延有了较为深刻的认识。对大数据产业的业态分布、生态系统、支撑技术及应用领域有了全面的了解,并通过实际案例厘清了大数据从根本上改变传统产业发展规律及格局的逻辑脉络。

最后,作为贵阳市直属的教学单位的一线教师,我对于如何将学校教学和建设工作与市政府的发展战略及方针政策结合起来有了更清晰的认识;对于如何促进培训成果向大数据工作推进成效转变有了一定的体会。下面,我就结合自己所在单位的功能和特点,谈一下如何借助贵阳市发展大数据的历史契机,提升学校信息化水平及如何建设大数据人才培养专业这两个方面的问题。

实事求是地讲,贵州省高职高专院校的信息化水平和一线城市及沿海发达地区相比是严重落后的。那要如何才能借助大数据的战略发展契机实现“弯道取直”,后发赶超呢?我认为在这次培训中给我们作了精彩报告的清华大学数据科学院韩亦舜教授为我们布置的一个互动环节可以给我们以最好的启迪。韩教授在报告行将结束时,要求在场的每位培训学员写下自己的单位,自己的单位想从哪些参

加培训的其他单位那里获取哪方面的数据，而自己的单位又能给其他单位提供什么有价值的数 据。这个小小的“课后作业”看似不起眼，却在大约一两周后变成一幅让所有人震撼的市级单位数据供需关联图：



这虽然仅是一个小范围的采样，至多不过是展示一下概念，没有多大的实用价值，但它背后体现出的大数据思路却是需要我们注意的。透过数据可视化技术的处理，我们能够发现传统数据处理方法和技术手段所容易忽视的特征，揭露数据供需关系背后的社会及产业发展需求，从而让数据在交叉引用和融合中产生出新的价值。

同理，高职高专院校借助大数据思维提升信息化水平也要从内部挖潜和寻求对外联系两个方面着手来发掘数据价值。过去，由于各个部门或者各个项目长期各自为政的信息化建设现状，导致各部门各项目的数据自成体系，“老死不相往来”，形成一个个的“信息孤岛”。俗话说“流水不腐”，这些数据洼地由于缺乏

有效的数据流动，已经是一滩死水，数据的日常维护不但不能成为“资产”，反而成为各个部门的“负担”。只有先建立起一种有效的机制，促进各部门，各项目之间数据的流通和融合，聚沙成塔，从孤立数据形成“块数据”直至“大数据”，才有可能从根本上改变现在“数据成为负资产”的尴尬现状。而如何建立起这种促进数据有效流通的机制，涉及到制度、技术、资金及发展规划等多个层面多种因素的制约，限于篇幅也无法进行详尽的讨论，故在此只能遗憾地以抛砖引玉代替。

从对外寻求联系的角度看，高职高专院校的信息化建设也不能只局限在从自身内部挖潜，同样要本着数据即资产的理念，一如韩教授精心制作的贵阳市级单位数据供需关联图所示，参与到市一级的数据交换大循环中来。只有充分参与外部交换，内部的数据才有评价的准绳和参考的依据，也才能对外体现出价值。

最后，我想谈谈开设大数据产业相关专业学科的高职高专如何进行大数据人才培养的问题。自市政府启动大数据发展战略以来，我校也在积极筹建大数据相关专业甚至大数据学院。当然这个过程是极为审慎的，我们不能搞大数据“大跃进”而强行上马相关的项目，违背学科建设规律只会给国家和社会造成巨大的浪费。那么应该如何进行相关学科的建设呢？我的思考如下：

在今年 8 月的学习培训会后心得体会总结中¹，我详细地梳理了大数据三种业态相关专业的人才培养类型和特点。现总结如下：

大数据人才从总体上分为两大类，一类是面向内容的大数据人才，以数据科学家、算法工程师和系统软件架构师及软件工程师为主；另一类是面向形式的云计算/雾计算人才，以应用系统架构师、前后端软件开发工程师、运维工程师、网络工程师和物联网应用人才为主。这两类人才有着完全不同的属性，前者科学属性更强，属学术型人才，培养周期长，难度大，只适合有深厚学术传统和学科优势的院校展开培养，对于这些类型的人才需求，只能从引进转化着手，是一个长期而没有立竿见影成效的工作。另一方面，不具备相关专业开设资格的院校也

¹ 详见大数据人才“一选一”专题培训总结文稿《大数据时代高校建设思考》，作者刘海波，2016 年 8 月

需要做好分工协同，配合围绕这些大数据内容型人才的培养做好相关专业人才培养的衔接工作使得高职高专的人才培养能够与这些大数据高端人才的培养相得益彰。

下面说一下后者。后者的工程属性更强，属于应用型人才，培养周期同样很长，而且必须脱离学院环境在企业生产实践中进行培养。而信息科学类高职院校与 IT 企业对专业人才的职业技能供需对接、职位对接及专业技能人才的职业素养对接是长期积弱的短板所在。这些都决定了这类人才的培养往往难度更大，对高职高专院校提出了巨大的挑战。我认为当前发展贵州大数据产业，本地高职高专院校应该学习国内优秀科技企业的互联网思维，进行聚焦突破，把资源优化配置到上述应用型人才的培养上来。而在专业设置上严格界定大数据内容与形式的方向差别，将项目落地实施过程中所需要的大量工程型、技术型人才培养作为突破口发力，才能有效地对接市场需求，服务于大数据产业在贵州的落地。

最后，我不得不提到的是，大数据人才的培养必须拥抱开源技术。

开源软件技术及社区经过几十年的发展，借助古老而神圣的 C 语言及其开发环境，以及源远流长，桃李天下的 UNIX 操作系统，成为当今大数据/云计算时代当之无愧的中流砥柱，在开源的沃土上，GNU/Linux、Android、Apache Hadoop、Spark、NodeJS、Docker 等改变云计算面貌的革命性技术枝繁叶茂。开源文化“代言人”Eric Raymond 将开源软件运动形容成“大集市”，并且与传统的“几代人前赴后继按照宏伟蓝图修筑大教堂”的传统“闭源”软件开发商业模式相比也丝毫不落下风。开源软件运动是信息科学发展史上的奇迹，拥抱开源为我们“后发赶超”提供了千载难逢的历史契机，抓住这个机会，我们便能够突破国外“软件航母”群的封锁，实现大数据各个业态核心知识产权自主发展的目标。

大数据是人类文明史上的“第四次浪潮”，将会深刻地改变人类社会的全貌。我们要以前所未有的决心投入到贵州省发展大数据产业的事业中来，才能实现社会经济发展的“弯道取直”，赶超经济发达地区。