

“数据科学与大数据技术” 专业 课程体系与教学环节探讨

对外经济贸易大学 曹淑艳 教授

主讲教师



曹淑艳，博士，对外经济贸易大学信息学院教授，信息化处副处长。自1985年起在对外经济贸易大学从事计算机教学及信息系统开发研究工作，本科计算机软件，管理科学与工程硕士、金融学博士。担任教育部大学计算机课程教指委委员、文科计算机教指分委副主任委员，全国高校计算机基础教育研究会文科专业委员会副主任委员、财经专业委员会副主任委员。主持及参与国家社科基金项目、国家自科基金项目、教育部人文社科一般项目、北京市社科重点项目、北京市社科基金项目、APEC国际组织研究项目、北京市教改联合项目等；先后四次获北京市优秀教学成果二等奖、主编的教材获北京市精品教材、国家“十一五”规划教材；参与多家出版社电子商务、信息管理、大数据专业、文科及财经类计算机基础系列教材编委会的教材编写组织工作；作为专业负责人成功申报首批“数据科学与大数据技术”本科专业。

主要内容

- 一、大数据专业认识
- 二、数据科学与大数据技术专业建设
- 三、数据科学与大数据技术专业课程体系探索
- 四、实践性教学环节和主要实验
- 五、数据科学与大数据技术专业就业领域
- 六、2015-2017专业申报情况回顾

1. 设置大数据专业的理由和基础

(1) 国家重视

- 政府重视大数据的战略意义和数据资源对社会经济发展的积极作用，国务院**2015**年出台了《促进大数据发展行动纲要》，并将“实施国家大数据战略，推进数据资源开放共享”纳入“十三五”期间规划和建设的重要目标。

1. 设置大数据专业的理由和基础

(2) 产业发展需要大数据人才

- 根据麦肯锡报告，仅美国市场2018年大数据人才和高级数据分析专家的人才缺口将高达19万；
- 美国企业还需要150万能够提出正确问题、运用大数据分析结果的大数据相关管理人才；
- 85%的500强企业已经或正在推出大数据项目，未来几年这些企业在数据分析上的投资将平均上涨36%；
中国大数据产业发展.....
- 目前国内大数据产业发展已初具规模，归纳提炼出了大数据产业核心要素，有待于形成业界共识。

(3) 大数据及其相关专业发展情况

- 国外:

美国Top50院校中的哈佛、斯坦福、卡内基梅隆、哥伦比亚大学、芝加哥大学、麻省理工学院等15所高校均开设了大数据相关专业;

- 国内:

2015年教育部批准三所高校设立本科专业;

2016年有32/38所申报本科专业;

2017年 ? /263所申报本科专业;

(3) 大数据及其相关专业发展情况

- 清华大学（多层次大数据人才培养：
硕士、能力提升项目、本科辅修专业）
- 中国人民大学（大数据科学与工程，
目录外二级学科硕博点）
- 复旦大学（多目标大数据人才培养：数据科学
本科第二专业、二级学科-博科硕专硕、相关院系大数
据课程-新媒体硕士）

北京航空航天大学、上海交通大学、厦门大学、天津大
学、浙江大学等也在研究生层次建立大数据专业方向。

(4) 高校陆续进行了开设大数据专业的基础工作

1. 师资团队建设与培训
2. 大数据专业建设小组
3. 实习实训等实践教学环境建设
4. 参与大数据学科竞赛
5. 与大数据企业和大数据产业协会联系紧密
6. 参与大数据人才培养论坛
7. 复合型人才就业前景展望

(5) 学科越来越交叉

- 现在各种学科都在往大数据的方向上发展。比如像数学、统计学，信息管理与信息系统专业，管理科学与工程等等，这些学科都在向大数据靠，所以大数据不好说归为某一学科（虽然目前在计算机门类中）。
- 就像电子商务一样，现在电子商务成为了管理门类里面独立的一级学科，但是要说电子商务属于技术还是属于管理，还是属于营销，或是其他方面，这取决于开设本专业的培养目标和办学定位。
- 大数据这个学科是一个融合，用来解决**领域问题**是非常明确的，就是通过数据的分析和挖掘来获得价值。

2. 大数据专业与相近专业的联系与区分分析

- 该专业所属**学科门类及专业类是计算机**，主要围绕与计算机相关专业的关系及区分度进行讨论：

计算机专业、统计学专业、信息管理与信息系统专业
与计算机专业的区别：

- 首先数据科学与大数据技术专业主要支撑技术来源于计算机；
- 其次是该专业具备前所未有的复合性特征；
- 最后本专业以大数据为核心研究对象，强调学生对**专业领域数据**的理解能力，深刻体现了技术为数据服务的思想———**复合型人才培养**

2. 大数据专业与相近专业的联系与区分分析

与统计学专业的区别：

- 统计学使用结构化数据；
- 本专业：强调“大数据”处理技术平台的要求
- 强调数据从获取、清洗、存储、处理和展示等各个环节与计算机技术的深度融合；
- 重视对非结构化和半结构化数据的处理
- 统计学工具是“大数据处理”环节需要的

第 6 期

计算机教育

30 2016 年 6 月 10 日

Computer Education

文章编号: 1672-5913(2016)06-0030-04

中图分类号: G642

大数据本科专业申报及认识

曹淑艳, 张 莉, 黄 浩

(对外经济贸易大学 信息学院, 北京 100029)

摘 要: 从产业发展、数据科学的学科特征、大数据专业与其他相关专业的不同等 3 个方面, 阐述增设大数据本科专业的合理性和必要性; 以对外经济贸易大学专业建设为例, 指出大数据专业人才应该能够围绕互联网平台上经济金融数据的商业价值进行挖掘并揭示数据间关系。

关键词: 数据科学; 大数据技术; 大数据科学与应用

DOI:10.16512/j.cnki.jsjy.2016.06.007

2017年12月10日

- 师资队伍建设
- 教学条件建设
- 课程体系建设
- 实践教学环节

师资队伍建设：16位教师，其中4位企业兼职

序号	姓名	性别	年龄	专业技术职务	第一学历毕业学校、专业、学位	最后学历毕业学校、专业、学位	现从事专业	拟任课程	专职/兼职	
授课教师表										
1	姓名	性别	年龄	拟任课程	专业技术职务	最后学历毕业学校、	最后学历毕业专业	最后学历毕业学位	研究领域	专职/兼职
2				程序设计、数据结构	教授					专职
3				大数据分析	教授					专职
				数据获取和清洗	副教授					专职

2017年12月10日

教学条件建设

其他办学条件情况表

申报专业副高及以上职称(在岗)人数	10	其中校外兼职人数		可用于该专业的教学实验设备(千元以上)	
可用于该专业的教学设备总价值(万元)					

主要设备

学校名称	设备名称	型号规格	数量	购入时间
	计算机(双屏)	lenovo M4500Q	70	2014年8月
	机架式服务器	IBM 3850 x6	3	2015年1月

三、数据科学与大数据专业课程体系探索

遵照产业定义的通用原则，产业各环节的参与实体应能够提供可交付的产品和服务，并形成上下游产业链供需关系。依据数据价值的提升路径和IT领域的产品布局。

大数据产业核心要素



1. 大数据产业的核心要素

- 数据资源部分负责原始数据的供给和交换，是数据资产作为生产要素的直接表现，根据数据来源的不同，可以细分为数据资源提供者和数据交易平台两种角色；
- 数据基础能力部分负责与数据生产加工相关的基础设施和技术要素供应，为数据加工和价值提升提供生产工具，主要包括数据存储、数据处理和数据库等多个角色；
- 数据分析和展示部分负责数据隐含价值的挖掘、数据关联分析和可视化分析与展现等，是智力要素在数据价值中的集中体现，包括传统意义上的BI、可视化和通用数据分析工具，以及面向非结构化数据提供的语音、图像等媒体识别服务；
- 数据应用部分根据数据分析和加工的结果，面向**电商、金融、交通、医疗**等细分行业提供精准营销、信用评估、出行引导、健康治疗等企业或公众服务。
- 来源：孙鑫（中国信息通信研究院政策与经济研究所）

2. 新专业, 人才培养方案、课程体系难以成熟、需探索

- 目前难以有完整的课程体系。可以围绕核心要素，根据各个学校特色进行课程体系设计；
- 目前的阶段就像早期的信息管理和电子商务专业，需要一个时间过程、探索, 才会有较为成熟的课程体系。
- 自**2015**年首次设立以来，高校进行了探索，如：

3. 围绕大数据产业核心要素 **四层次梳理主要课程**

- 数据资源提供者和数据交易平台+数据存储、数据处理和数据库（主要体现在计算机软件技术）
 - 信息资源管理、Web技术、软件体系结构
 - 计算机应用基础、C（C++、C#、Java等）程序设计
 - 算法与数据结构、数据库（NoSQL）
 - **※ R语言/Python语言程序设计※**
- 数据分析和展示：体现在BI、统计、数据挖掘、机器学习、可视化分析与展示、通用数据分析工具等，为非结构化数据提供文本挖掘、语音与图像等媒体识别服务；
- 主要课程大致为：

数据分析和展示主要课程：

- 分布式系统原理、人工智能基础、
- 应用统计（**统计学**）、多元统计分析
- **机器学习**、数据挖掘、
- 可视化数据分析、可视化数据展示、
- 数据建模与统计分析（R语言）
- 自然语言处理、社会网络分析
- 云计算及其应用、领域大数据分析。

数据应用部分（领域知识及其应用）：

- 结合高校特色进行设计,如UIBE申报本专业时分为如下三个方向:
 1. 贸易金融大数据:
 2. 网络营销大数据:
 3. 电子商务大数据:
- 围绕上述三个方向设置经贸、金融、营销、电子商务等相关领域的基础课程，使学生具备领域基础知识；

专业人才培养方案的课程体系应以大数据产业的四个层次为指引，系统地培养学生掌握大数据应用中的各种典型问题的解决办法，提升学生解决实际问题的能力。

具体框架为：

- 公共基础课+专业基础课（大一、大二）
- 专业基础课+专业选修课+领域选修课（大二、大三）
- 专业选修课+领域选修课（大三、大四）

核心课程示例：略去计算机+专业领域部分.....

序号	课程名称	周学时	序号	课程名称	课程周学时
1	多元统计分析	2	8	数据可视化	2
2	统计分析软件与应用	2	9	模式识别与机器学习	2
3	统计学习理论基础	2	10	云计算与数据中心	2
4	人工智能/商务智能	2	11	大数据分析软件	2
5	数据科学导论	2	12	NoSQL数据库及应用	2
6	Python程序设计	2	13	Hadoop与MapReduce	2
7	数据挖掘	2	14	。 。 。 。 。	

小结：

- 数据资源和数据基础层依托于传统的计算机技术
- 数据分析和展示层在统计学支撑下与IT结合有所创新
- 数据应用层则体现领域综合与办学特色

强调：校企合作、产学研结合

实践性教学环节按照知识体系分为三个层面：

□ 技术、平台类实验：

- 1、数据库及非结构化数据处理
- 2、程序设计、数据结构与算法
- 3、云计算平台的建设与开发。。。

□ 数据分析及展示类实验：

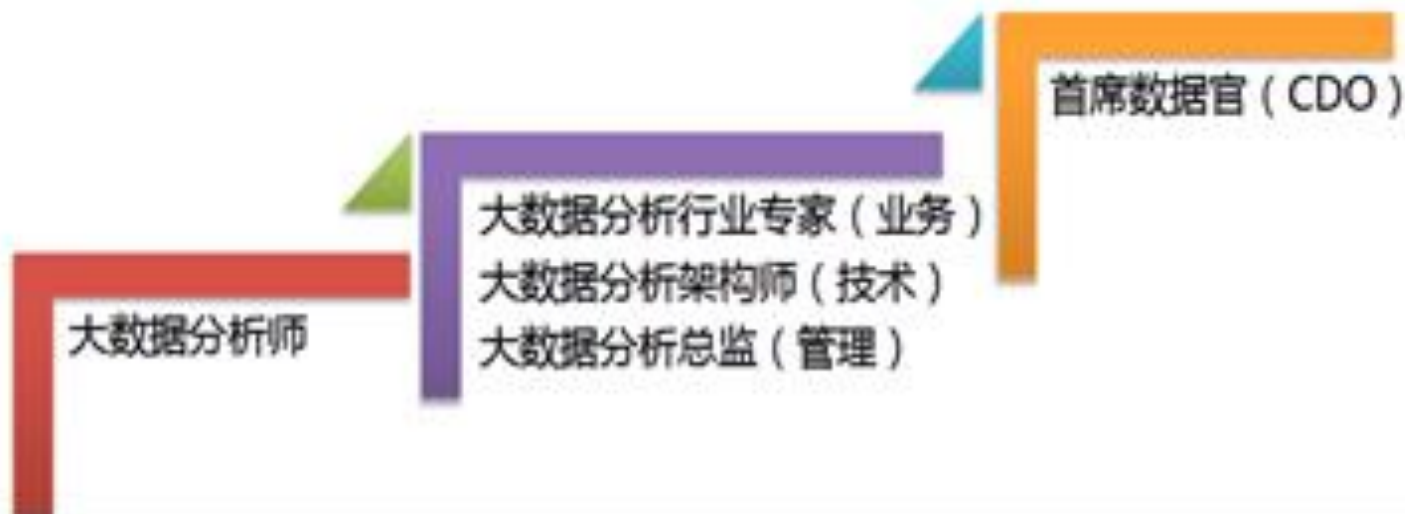
- 1、结构化数据统计分析
- 2、机器学习（商务智能）
- 3、数据挖掘与大数据分析处理技术
- 4、大数据可视化实验。。。

□ 综合应用类实验：金融、营销、物流大数据综合实验 （与相关大数据公司合作，联合培养）

人才就业领域及职业生涯发展----来源:亿赞普大数据工程师--两会大数据

三个方向就业：

1. 贸易金融方向 大数据分析师：供应链融资公司、P2P信贷征信平台、商业银行。
2. 网络营销方向 大数据分析师：互联网广告、O2O营销公司、大型网络媒体。
3. 电子商务方向 大数据分析师：电商公司、第三方支付公司、电商物流公司。



大数据分析引爆央视两会报道

2014年03月17日17:02 来源: [人民网-传媒频道](#)

<http://media.people.com.cn/n/2014/0317/c40606-24658421.html>

<http://tv.sohu.com/20140304/n395994317.shtml>

2015-2016地域分布

- **2015:** 北京大学 理学
对外经济贸易大学 工学
中南大学 工学
- **2016:** 申报**38**所，批复**32**所
工学**29**所，理学**3**所

北京	5	安徽	1
贵州	5	福建	1
上海	4	广西	1
广东	3	河南	1
四川	3	湖北	1
云南	3	湖南	1
山西	2	吉林	1
		宁夏	1
		浙江	1
		重庆	1

2017申报：工学190所，理学73所

- 中国有23个省、5个自治区、4个直辖市2个特别行政区共34个省级行政区。其中29个参与申报

地域统计：

北京	12	安徽	14	黑龙江	7
贵州	8	福建	11	江西	7
上海	7	广西	7	辽宁	13
广东	13	河南	20	天津	3
四川	6	湖北	11	新疆	1
云南	9	湖南	6	江苏	13
山西	12	吉林	7	内蒙古	10
甘肃	5	宁夏	2	山东	15
海南	3	浙江	10	陕西	9
河北	14	重庆	8	总计	263

结语

2017年申报结果期待中。。。

- **申报新专业，机遇与挑战**
- **建设新专业，任重而道远**

欢迎指正

BDTC 2017 中国大数据技术大会
Big Data Technology Conference 2017