

大数据第二专业建设

汪卫

复旦大学计算机科学技术学院

上海数据科学重点实验室

weiwang1@fudan.edu.cn

目录

- **大数据人才培养的现状**
- **对大数据第二专业培养目标的理解**
- **复旦大学大数据第二专业建设**

大数据人才培养的现状

- 大数据人才缺口大
 - 我国3-5年内大数据人才缺口150万
 - 目前大数据人口46万
 - 麦肯锡预测，到2018年，在“具有深入分析能力的人才”方面，美国可能面临着14万到19万的缺口，而“可以利用大数据分析来做出有效决策的经理和分析师”缺口则会达到150万
- 大数据人才的薪酬高
 - 据Information Week统计，BI领域的工资至少超过其它IT领域10%

2018 届互联网校招高薪清单

公司	岗位	年薪 (人民币)
谷歌中国	人工智能 (ssp)	56w
微软	算法工程师	51w
google	算法工程师	50w
腾讯	基础应用研究 (ssp)	45-50W
腾讯	腾讯云后台研发工程师	32.4w
大疆	算法工程师	30w-40w
海康	算法工程师	30w
华为	研发工程师 (硕士 ssp)	30w+
Works Applications	软件工程师	36w
网易游戏	游戏研发工程师	32.8w
网易游戏	算法工程师 (sp)	36w+
阿里巴巴	算法工程师	38.4
美团	研发工程师 (sp)	30w
大疆	后台开发工程师	30-40w
OPPO	AI 视觉 (sp)	31-32w
滴滴	算法工程师 (新锐)	30.4w- 37.5W
51 信用卡	Java 工程师 (sp)	30w
好未来	数据挖掘工程师	35w
百度	算法工程师	32W
今日头条	算法工程师 (博士 sp)	33w
28 所	论证设计师	32W+11W 安家费
京东	算法工程师 (ssp)	31w
亚马逊	软件工程师 SDE	31w 加股票
深信服	软件研发 (ssp)	30w

数据来源：校招薪资水公众号，转载请注明

大数据人才培养的现状

- 大数据人才培养模式
 - 高校已经开始大数据方向的人才培养
 - 很多大学已经开始“数据科学与大数据技术”专业的人才培养
 - 大量的学校开始相关专业的硕士和博士研究生的培养
 - 第二专业、交叉专业或能力提升计划受到各专业学生的欢迎
 - 国外高校大数据人才培养情况和国内大致同步
 - 企业在大数据人才培养方面
 - 阿里、中科曙光、慧科等在高校承建相关专业
 - 章鱼大数据、Talking data等企业提供了大数据实验平台
 - IBM、理想等众多企业通过举办竞赛、夏令营等方式丰富了大数据人才培养的实践环境
 - 新兴专业和大数据专业的差异发展成为新的课题
 - 人工智能、智能科学与技术

大数据与新工科教育

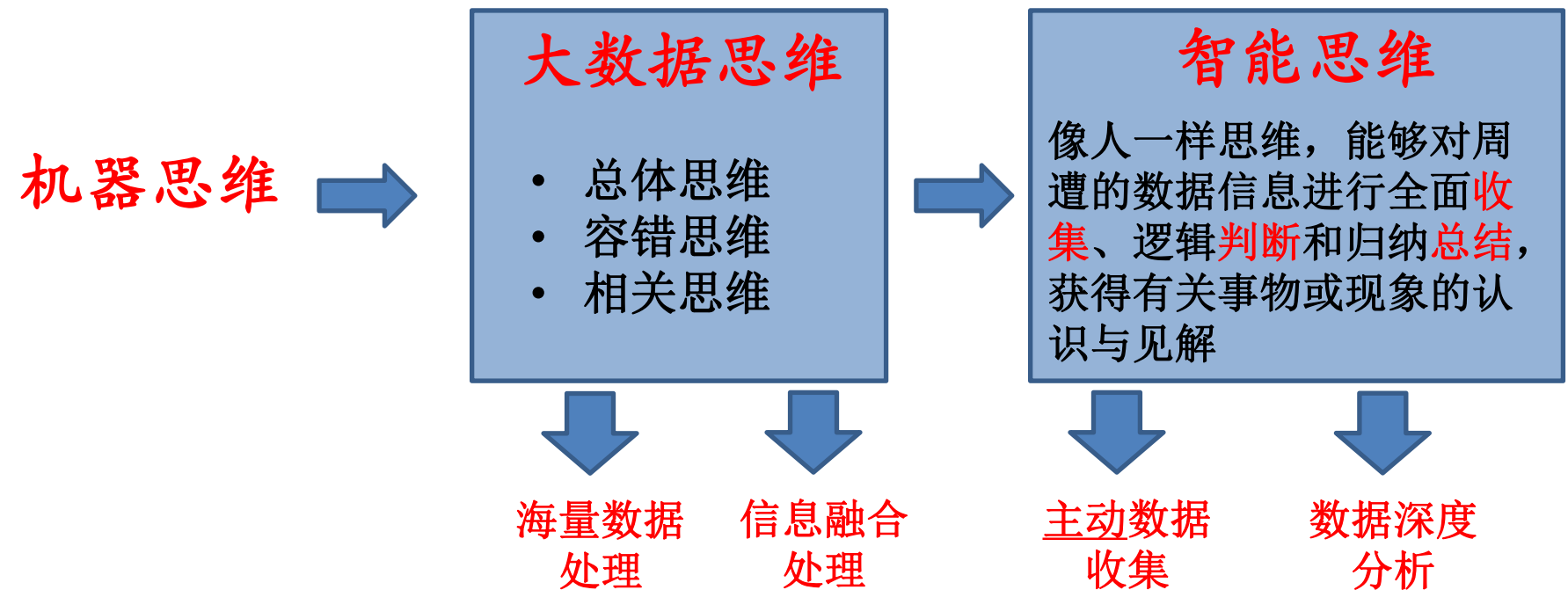
- 大数据是信息技术领域新的产业方向
- 具有天然的学科交叉的基因
 - 代表新的科学研究范式
 - 第四范式
 - 对其它学科有强大渗透力
- 创新创业的热土
 - 大数据是和市场紧密结合的方向
 - 大量的创新创业的项目都和大数据有关

目录

- 大数据人才培养的现状
- **对大数据第二专业培养目标的理解**
- 复旦大学大数据第二专业建设

数据思维是大数据人才的核心能力

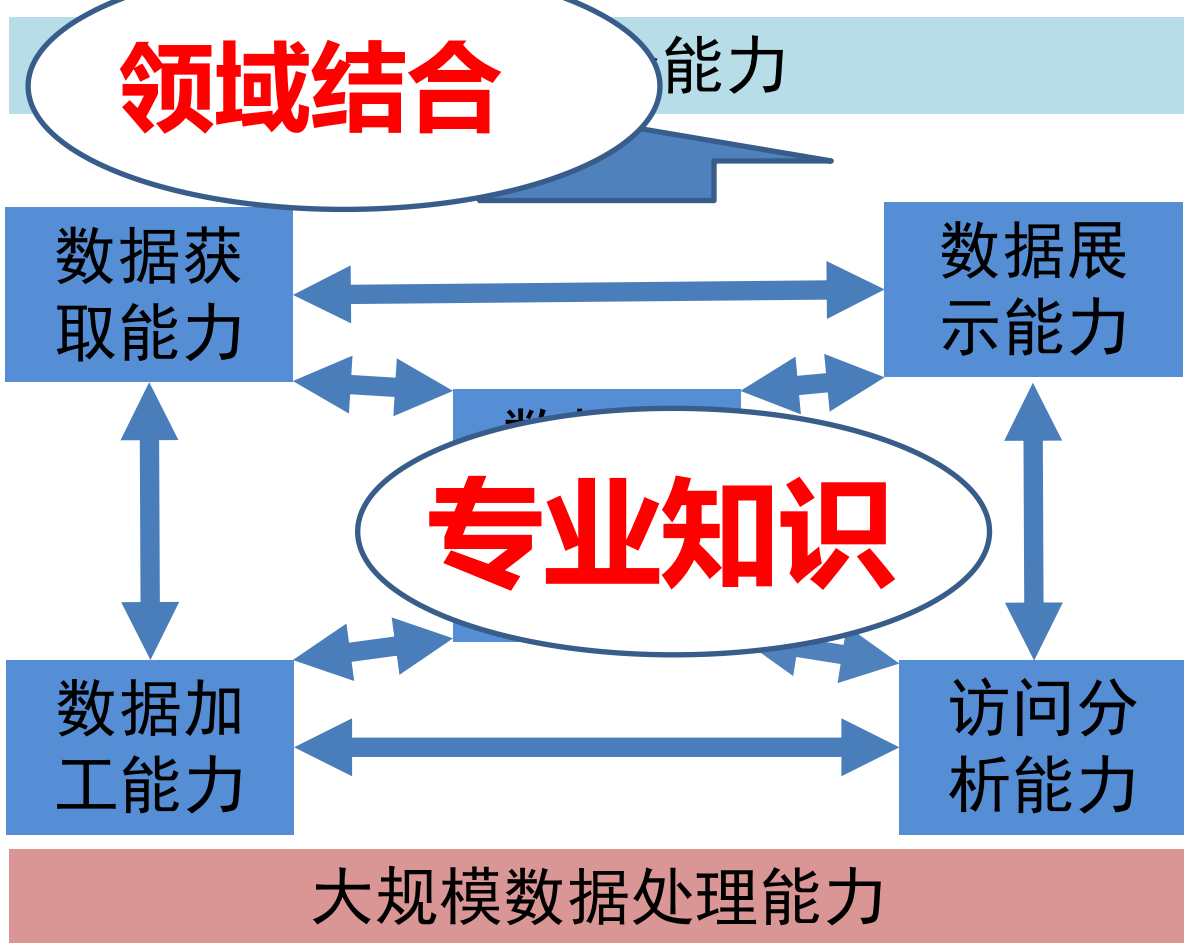
Data thinking is the generic mental pattern observed during the processes of picking a **subject** to start with, **identifying** its parts or components, **organizing** and **describing** them in an **informative fashion** that is relevant to what **motivated** and initiated the whole processes.



大数据专业人才能力需求

- 具有黑客的头脑
 - 对数据有好奇心
 - 对商业有热情
- 是有影响的、有创意的

数据敏感



软能力

硬能力

大数据人才所应具备的素质

思维素质

跨专业学习

创新能力

专业素质

数据管理
(海量异质数据的组织管理)

数据加工
(数据质量的提升和语义提升)

数据分析
(各种数据分析模型的理解和运用)

专业能力

数学基础
(面向数据分析的数学基础)

编程基础
(利用各级编程语言实现模型和应用的开发)

系统基础
(先进的计算机系统的运用能力)

基础能力

大数据第二专业的培养对象

- 技术融合型
 - 将大数据技术作为专业的主要支撑能力，具有较好的数学和计算机基础
 - 如智能科学专业、电子、数学
 - 深入理解大数据和人工智能领域的主要技术，特别是对大数据分析技术有重要的需求
 - 培养特点
 - 掌握大数据管理和分析的主要技术和模型
 - 强调对大数据分析工具的应用
 - 理解大数据分析应用的建模

大数据第二专业的培养对象

- 模型融合型
 - 将大数据和人工智能技术作为专业的工具型支撑能力
 - 如经济、管理、医学、社会学、。。。
 - 深入理解大数据技术应用领域的**需求与主要模型**
 - 培养特点
 - 掌握业务需求、能够设计并实现分析模型，了解获取数据的途径
 - 强调对大数据分析工具的应用，强调对业务和分析需求的掌握
 - 了解大数据管理技术，不同程度的理解大数据分析模型

学生选择大数据第二专业的原因

- 为转行做准备
 - 转计算机或数据科学等相关行业
 - 转系转专业、读研、找工作
- 了解大数据技术，拓展知识面
 - 人文、社会科学、...
- 提升在大数据分析方面的能力
 - 经济、管理、医学、数学、物理、...

大数据专业与人工智能专业的差别

- 相同点
 - 数据驱动的分析是专业的核心能力
 - 分析模型的掌握是共同的需要
 - 与应用相结合是培养的目标
- 差异点
 - 大数据专业更强调
 - 数据的获取过程
 - 管理模式
 - 处理平台
 - 人工智能技术更强调
 - 系统的整体架构
 - 与生物领域的结合
 - 学习模型的核心算法

目录

- 大数据人才培养的现状
- 对大数据第二专业培养目标的理解
- 复旦大学大数据第二专业建设

復旦大學大數據人才培養的情況

- 大數據學科支撐
 - 上海市數據科學重點實驗室
 - 國內第一個省部級大數據方向重點實驗室
 - 由多個院系的老師構成
 - 具有良好的研究基礎
 - 大規模數據管理
 - 數據挖掘
 - 機器學習
 - 多媒體信息處理
 - 語義計算
 - 大數據應用
 - 。 。 。

复旦大学大数据相关专业建设基本情况

- 计算机学院数据科学人才培养
 - 本科
 - 数据科学第二专业（2015年）
 - 研究生
 - 数据科学二级学科专业（2014年）
 - 博士
 - 科学硕士
 - 专业硕士（全日制与非全日制）
 - 相关院系的大数据课程
 - 新媒体专业

復旦大學大數據相關專業建設基本情況

- 大數據學院
 - 本科
 - 數據科學與大數據技術第一專業
 - 研究生
 - 數據科學專業碩士
- 信息學院（結合計算機學院、類腦研究院、微電子學院）
 - 智能科學與技術（新工科項目）
 - 智能信息處理方向

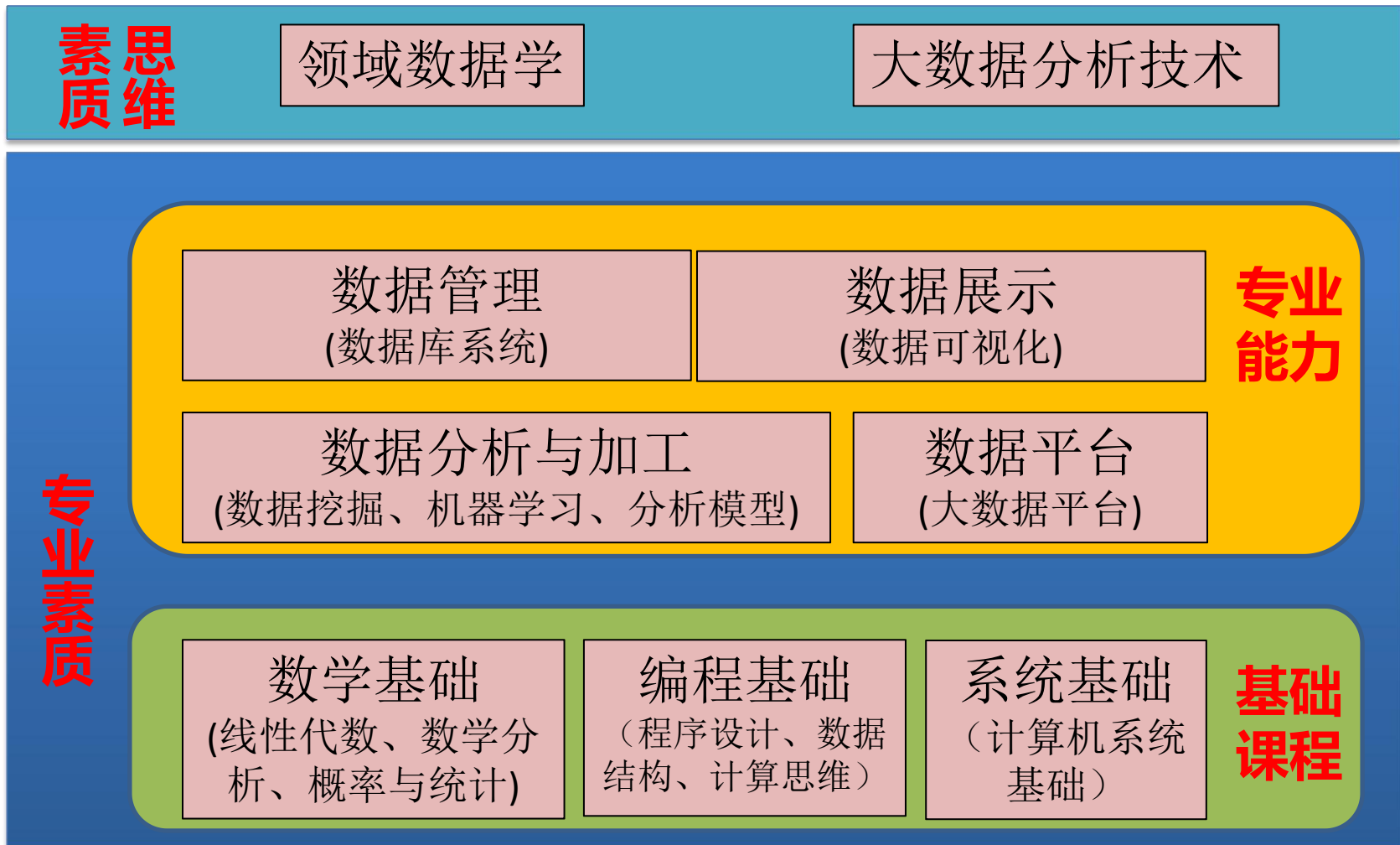
数据科学第二专业

- 目标
 - 掌握大数据的思维方式
 - 利用数据解决问题
 - 具有交叉学科思维
 - 跨学科的课程设计
 - 多学科背景的班级组织
 - 具有大数据分析的能力
 - 良好的数学基础
 - 基本的编程能力
 - 理解大数据分析平台技术
 - 理解数据分析与挖掘的原理和技术
- 教学计划规定的41 学分
 - 专业必修课26学分
 - 专业课程学习
 - 毕业设计，组队完成项目（学科交叉、学校和企业导师）
 - 选修专业课15 学分（模块化领域方向）
 - 数学、新闻、经济和管理、生物和医学等各个学科的基础课程
 - 学生需要修满至少一个专业的课程
- 成绩合格者，将颁发《数据科学》第二专业证书

本科第二专业

- 学生情况
 - 几乎覆盖复旦大学所有的专业
 - 文理科比例约35:65
 - 学生的学习积极性比较高
- 当前面临的主要问题
 - 学生背景的**差异性**
 - 由于教室规模的问题，限制了学生的规模
 - 最大的教室可容纳140人
 - 上海东北片高校联盟

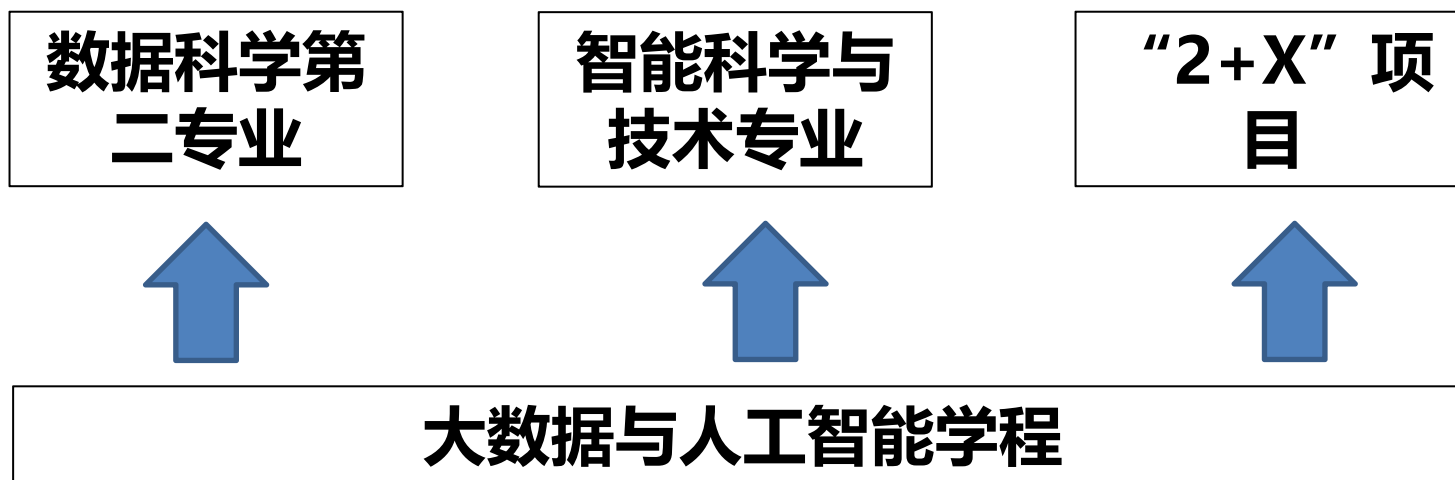
数据科学第二专业课程体系



智能科学与技术专业

- 目标
 - **系统**掌握智能系统构建的核心理论与技术
 - 具有较强的**综合**动手能力
- 由信息学院、计算机学院、类脑研究院、微电子学院共建
- 涵盖四个方向
 - 智能系统
 - 智能数据处理
 - 智能芯片
 - 类脑智能

大数据课程体系的组织规划



大数据与人工智能学程

- 背景

- 探索高校计算机基础教育的模式
 - 从单纯的技能培养到思维方式的培养
- 应对新工科背景下的计算机教育的需要
 - 在培养方案上推动跨学科培养
 - 学程作为学分的构成单位
 - 将计算机的最新技术发展融入到其它专业的教学过程，支撑“半专业”计算机人才的培养

面向学程的教学组织

- 大数据与人工智能学程
 - 设立一序列课程，帮助学生建立大数据与人工智能方面的知识体系
 - 课程设计原则
 - 基础性
 - 掌握的基本编程技能
 - 前沿性
 - 课程涵盖数据科学、机器学习等领域最新知识
 - 实践性
 - 通过1-2个综合课程实践，提高学生的动手能力

面向学程的教学组织

- 大数据与人工智能学程
 - 课程体系设计
 - 基础类
 - 程序设计、数据结构与算法设计、离散数学
 - 核心知识类
 - 人工智能、机器学习、图像处理与机器视觉、现代计算机与网络系统、数据库与数据挖掘、数学建模
 - 实践类
 - 大数据与人工智能实践、领域数据学与实践
 - 每门课4学分
 - 课程内容融合
 - 增加研讨性内容
 - 增加实践性内容

谢谢!

新工科下大数据人才培养的特点

- 新工科为大数据人才的培养提供了新的空间
 - 大数据专业的建设
 - 大数据与相关的人工智能可以在新设专业建设中发挥作用
 - 智能科学
 - 大数据可以在传统学科改造中发挥作用
 - 新媒体技术、医疗信息学、。。。
- 新工科也为大数据人才培养提出了新的要求
 - 聚焦核心能力
 - 加大与产业的结合

新工科是当前我国高教界的重要话题

- 国家对新工科建设给予了高度的重视
 - 充分发挥了产学研各方的力量
 - 设立了大量的教改项目
- 新工科触发了我们对原有教学方式的反思

对新工科的理解

- 复旦共识
 - 我国高等工程教育改革发展已经站在新的历史起点
 - 支撑服务以**新技术、新业态、新产业、新模式**为特点的新经济蓬勃发展
 - 世界高等工程教育面临新机遇、新挑战
 - 第四次工业革命正以指数级速度展开
 - 工科优势高校要对工程科技创新和产业创新发挥主体作用
 - 面向当前和未来产业发展急需，主动优化学科专业布局，促进现有工科的**交叉复合**、工科与其他学科的**交叉融合**，积极发展新兴工科，拓展工科专业的内涵和建设重点，构建创新价值链，打造工程学科专业的升级版
 - 综合性高校要对催生新技术和孕育新产业发挥引领作用
 - 以引领未来新技术和新产业发展为目标，推动应用理科向工科延伸，推动学科**交叉融合**和**跨界整合**，产生新的技术，培育新的工科领域

对新工科的理解

- 天大行动

- 探索建立工科发展新范式

- 总结技术范式、科学范式、工程范式经验的基础上，探索建立新工科范式
- 继承与创新、交叉与融合、协同与共享

- 问产业需求建专业，构建工科专业新结构

- 推动现有工科交叉复合、工科与其他学科交叉融合、应用理科向工科延伸，孕育形成新兴交叉学科专业

- 问技术发展改内容，更新工程人才知识体系

- 探索综合性课程、问题导向课程、交叉学科研讨课程，提高课程兴趣度、学业挑战度

- 问学生志趣变方法，创新工程教育方式与手段

- 加大学生选择空间，方便学生跨专业跨校学习
- 完善新工科人才“创意-创新-创业”教育体系

对新工科的理解

- 北京指南
 - 建设一批新型高水平理工科大学
 - 建设一批多主体共建共管的产业化学院
 - 建设一批**产业急需**的新兴工科专业
 - 建设一批体现产业和技术**最新发展**的新课程
 - 建设一批集教育、培训、研发于一体的实践平台
 - 培养一批工程实践能力强的高水平专业教师
 - 建设一批跨学科的新技术研发平台
 - 建设一批直接面向当地产业的技术创新服务平台
 - 形成一批可推广的新工科建设改革成果

对新工科的理解

- 新工科研究和实践围绕工程教育改革的新理念、新结构、新模式、新质量、新体系开展
 - 工程教育的新理念：分析研究新工科的内涵、特征、规律和发展趋势等，提出工程教育改革创新的理念和思路。
 - 学科专业的新结构：面向新经济发展需要、面向未来、面向世界，开展新兴工科专业的研究与探索，对**传统工科专业进行更新升级**等
 - 人才培养的新模式：开展深化**产教融合**、校企合作的体制机制和人才培养模式改革研究和实践
 - 教育教学的新质量：研究制订新兴工科专业教学质量标准，开展多维度的教育教学质量评价等。
 - 分类发展的新体系：提出推进工程教育办出特色和水平的宏观政策、组织体系和运行机制等。

对新工科的理解

- 如何在人才培养中贯穿新工科的思想
 - 如何实现学科的交叉与深度融合的目标
 - 交叉性课程
 - 力度不太够
 - 交叉专业
 - 要做到适配性组合，避免0.5+0.5的问题
 - 用能力的拓展代替知识的拓展
 - 产学研紧密结合的课程内容设计
 - “创意-创新-创业”教育体系
 - 通过具有真实产业背景的项目，帮助学生体验创新到创业的过程

目录

- 对新工科的理解
- **大数据人才培养的现状**
- 对大数据人才培养目标的理解
- 复旦大学大数据第二专业建设

目录

- 对新工科的理解
- 大数据人才培养的现状
- 新工科背景下对大数据人才培养目标的理解
- 复旦大学大数据第二专业建设

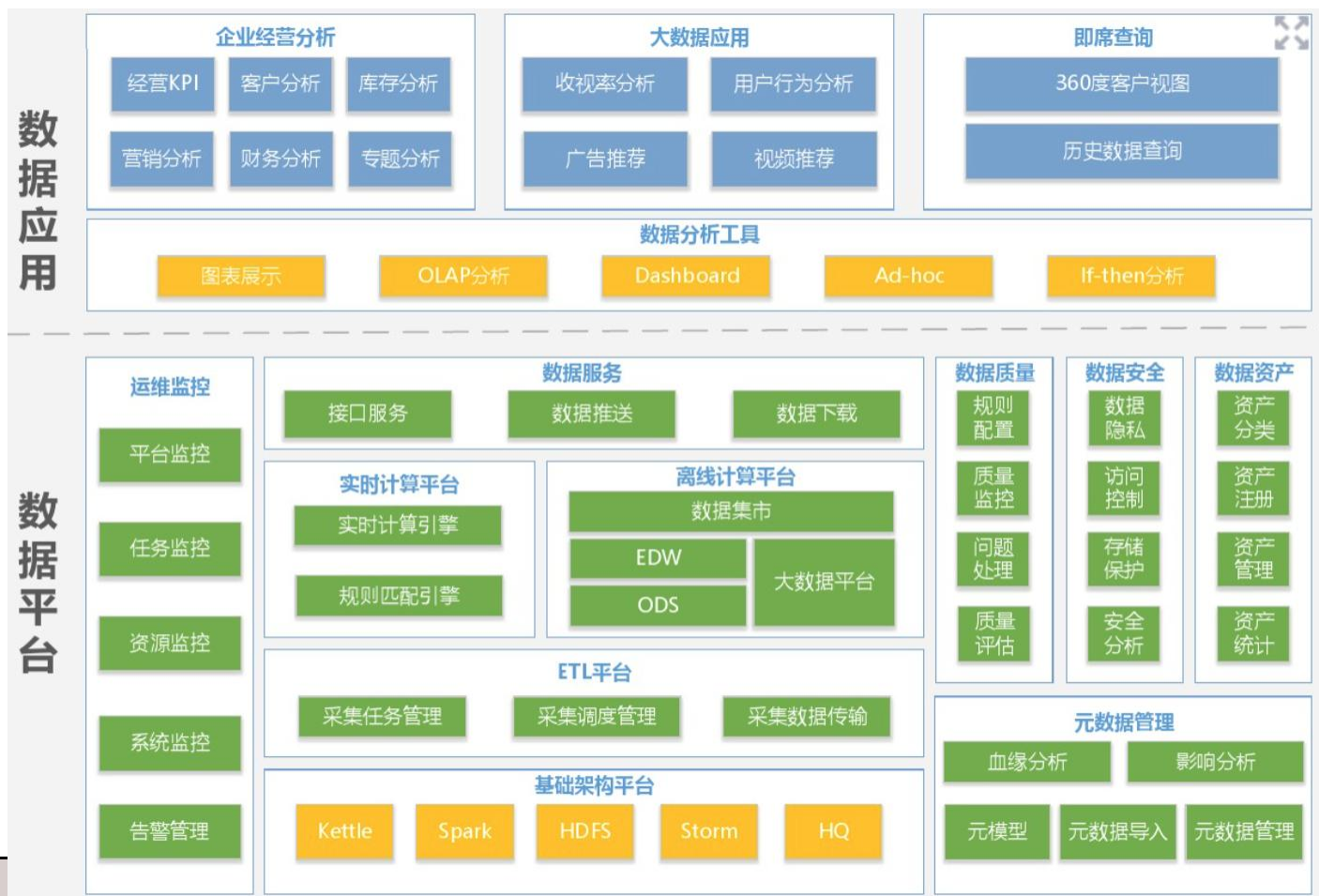
大数据架构看大数据人才素质

• TRS大数据系统



大数据应用看大数据人才素质

• 大数据系统架构



大数据应用看大数据人才素质

• 教育大数据



目录

- 对新工科的理解
- 大数据人才培养的现状
- 对大数据人才培养目标的理解
- 复旦大学大数据第二专业建设

谢谢！

相关课程

- 数据库与数据挖掘
 - 主要培养学生的数据管理与处理能力
 - 主要教学内容
 - SQL查询语言
 - 关系、图数据模型和数据库模式设计
 - 数据仓库
 - 数据挖掘过程与预处理
 - 主要数据挖掘模型
- 数据库与数据挖掘
 - 主要教学内容
 - 数据安全和隐私保护
 - NoSQL数据库
 - 主要实验内容
 - 数据库和数据仓库模式设计
 - SQL查询
 - 数据挖掘

相关课程

- 现代计算机网络与系统技术
 - 主要教学内容
 - 计算机系统
 - 计算机系统结构、网络、操作系统
 - 分布式系统
 - 分布式系统的基础理论
 - Map / Reduce计算框架与Hadoop、Spark
 - 流计算平台
 - NoSQL数据库系统
 - 大图数据管理系统
 - ...

相关课程

- 领域数据学与实践
 - 针对深度学习、企业征信、天文、工业 / 物联网、生物 / 医疗、脑科学等领域邀请了微软亚洲研究院的邓力、国科大张兵、田英杰、复旦大学冯剑锋、西澳大学吴晨等专家进行教学
 - 每学期会邀请不同的专家来复旦授课，通过视频拍摄可以积累了大量的视频
 - 学生在课堂上进行专题研讨
 - 结合领域问题完成数据分析实践
 - 主要集中在建模方面

领域数据学与实践



相关课程

- 大数据分析技术

- 目标

- 通过一个综合性实验，帮助学生建立大数据分析的整体概念和能力

- 主要作法

- 和企业建立紧密合作
 - 企业提出有第一线的问题和数据
 - Talking data、人人行、...
 - 充分利用各种大数据竞赛
 - 学生分组合作，利用一学期完成一个项目
 - WFI定位、手机APP分析、互联网金融数据分析、课程知识图谱、...

- 效果

- 培养学生的数据敏感性和对数据整个分析过程的掌握
 - 提升学生综合利用数据管理和分析工具的能力

谢谢！