



华东师范大学

数据科学与工程专业

建设探索

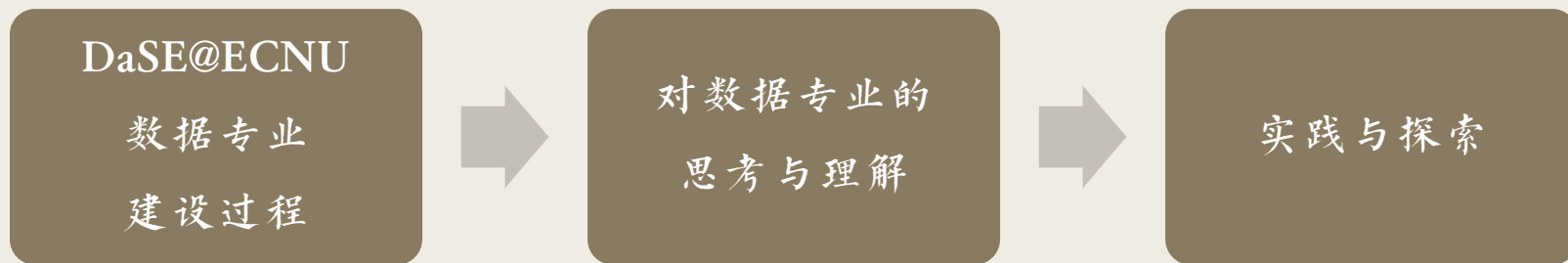
钱卫宁

华东师范大学 数据科学与工程学院

wnqian@dase.ecnu.edu.cn



提纲



探索过程

■ 2012.6 云计算与大数据研究中心

- 学术伙伴、企业赞助、访问学者
- **研究生培养**



■ 2013.9 数据科学与工程研究院

- 全国高校第一个大数据研究院
- 学科交叉，应用驱动创新
- 2014: **完善“数据科学与工程”二级学科**
- 着手“数据科学与工程”本科专业建设

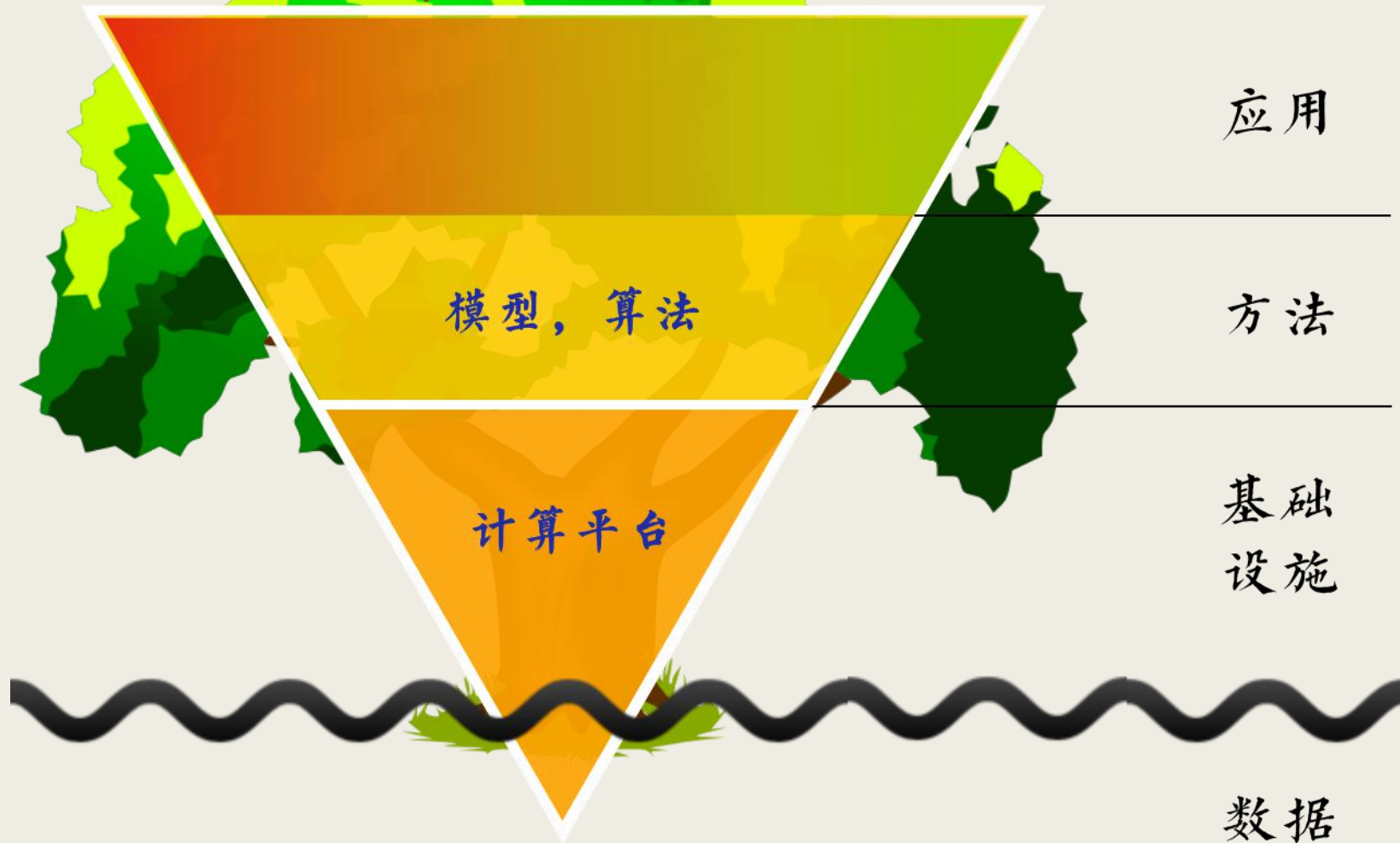


■ 2016.9 数据科学与工程学院

- 2017: **招收“数据科学与大数据技术”专业本科生**
- 2019.9: **从高中毕业生招收数据专业本科生**

数据科学与工程

科学 企业 社会 教育 健康 电子 搜索 物流 金融
研究 管理 治理 学习 医疗 商务 引擎 交通 投资



数据专业设计

避免成为简单的 CS+ 或者 STAT+ 专业

培养目标：造车人 | 高端网信人才

- 系统架构师、数据科学家、算法和系统工程师

知识结构：化学 | 生物融合的交叉学科

- 计算机 + 应用数学/统计 + 信息系统 (领域应用)

课程设置：实践驱动 | 螺旋式上升

- 加强应用数学训练，精简计算机课程，注重系统理念培养

教学特色：站在用户角度

- 注重应用，践行设计思维

矛盾与思考

- 对数据专业的**理解差异**
- “**硬专业**”与学生发展的矛盾
- **个性化**和**规范化**的矛盾
- **丰富的学习内容**与**有限的学分/课时**的矛盾
- **综合能力锻炼**与**客观评估**之间的矛盾
- **快速发展的技术**与**相对固定的培养计划**的矛盾
- **学生学习新技术的热情**与**循序渐进的学习规律**的矛盾
- **高要求**与**普遍的厌学**、“**水课**”环境的矛盾
- **数据专业**与**同类专业**间的竞争

入口与出口问题

入：招生

- 从“转专业”和“插班生”学生开始
- 形成文化后延伸至高中毕业生
- 充分沟通
 - 意愿
 - 知识
 - 能力
 - 习惯
 - 理解
 - 意志
- “菁英班”模式

出：就业与深造

- 希望学生能够进一步深造
 - 相关专业博士或学术型硕士
 - 专业工程硕士
- 目标行业
 - 金融
 - 互联网
 - 物流
 - 教育
 - 制造、电信、能源等
- 从事工作
 - 科研、数据分析、系统架构师/工程师

培养计划设计

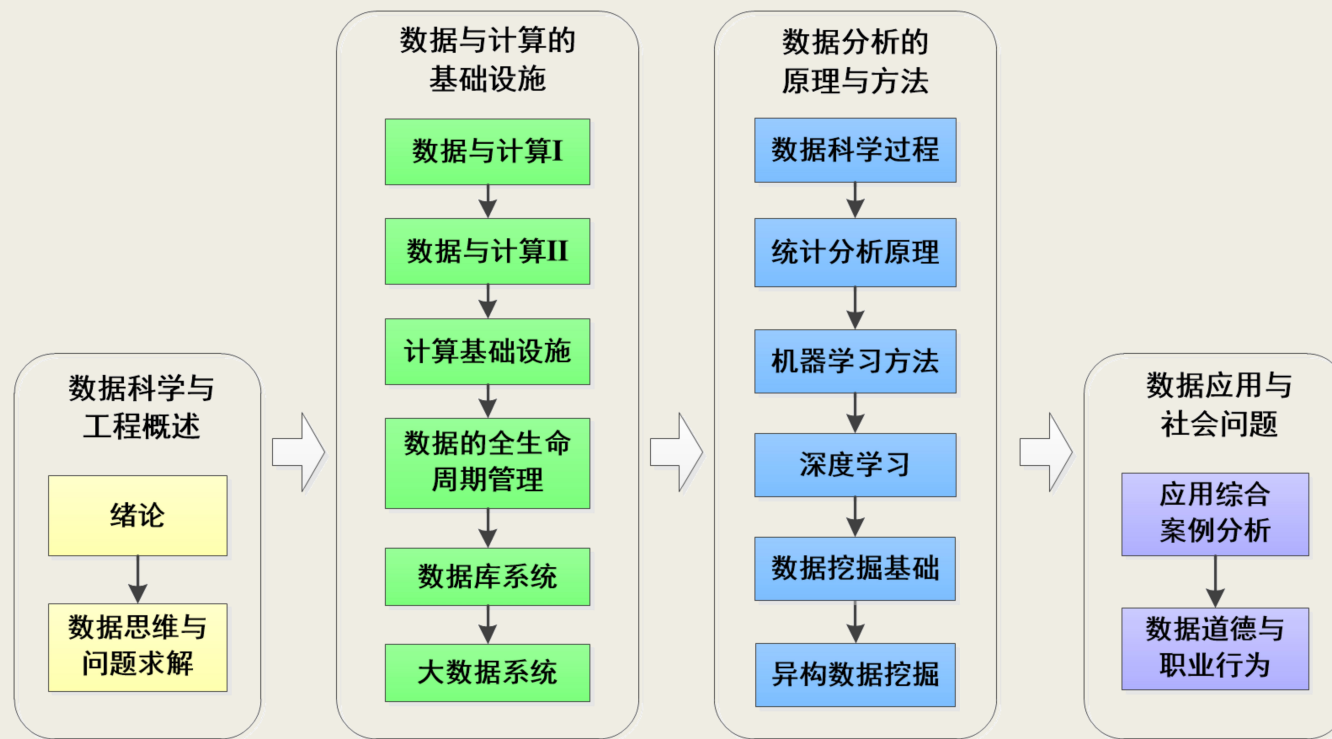
	第一学期	第二学期	一暑期	第三学期	第四学期	二暑期	第五学期	第六学期	第七学期	第六、七、八学期，三暑期
基础/ 导论	线性代数	高等数学(2)		DaSE 导论				AI导论		选修课： 深度学习； IR+NLP； CV+MM； 高级DBMS； 事务处理； 智慧城市； 社会计算； 区块链； 设计思维； 电子商务； 商业分析... 实习实践 毕业论文
	高等数学(1)			概率论						
	程序设计									
计算机理论		数据结构		算法设计与分析			DaSE 算法			
		离散数学								
系统				计算机系统	操作系统		数据管理系统	分布式编程模型与系统		
实践		Web编程	数据可视化		数据伦理		计算机网络原理与编程		软件工程	
数据科学					数理统计			机器学习		
					DaSE 数学					

数据科学与工程导论



嘉数汇

技术传播文明，践行同步世界，我是
MR. WANG，讲述信息文明时代的创意和故事。



数据科学与工程算法

- 衔接“数据结构/算法”课程与机器学习和分布式系统实现中所需的算法知识
- 六个模块
 - 基础：概率不等式、随机算法初步、hash等
 - 优化算法：运筹学算法初步
 - 矩阵算法：特征值计算、SVD/PCA、矩阵分解
 - 图算法：PageRank/HITS、图分割等
 - 统计计算：回归、抽样、EM算法等
 - 系统算法：并发数据结构、分布式一致性等

数据科学与工程数学基础

- 衔接以高等数学、线性代数、离散数学、概率论为基础的 CS 专业数学类课程与机器学习、人工智能算法开发与研究所需的数据知识
- 两大部分
 - 数据的表示和数学模型
 - 向量、矩阵；矩阵分解与降维；最小二乘与回归
 - 贝叶斯决策与分类；信息论基础
 - 数值优化模型和方法
 - 一阶方法；二阶方法
 - 理论基础：凸性，最优性，对偶性等

数据管理系统

- 打破 OldSQL, NoSQL, NewSQL 界限
- 实践驱动, 理解数据和数据管理的本质
- 极大的课程量 (6学分, 7课时/周 x 18周)
- 课程条理
 - 应用中的数据
 - 数据管理 (文档与关系型数据库)
 - 使用与运维
 - 扩展
 - 实现
 - 设计与理念

分布式编程模型与系统

- 衔接 OS, DB 和研究生以及实践工作中所要面对的各类数据系统使用和设计、实现需求
- 平衡理论和实践
- 五类系统
 - 批处理系统
 - 流计算系统
 - 批流融合系统
 - 图数据处理系统
 - 机器学习系统
- 系统基础知识
 - 分布式文件系统, 元数据管理, 资源管理

数据伦理

- 研讨类课程：无定论
- “一节课可以讲完的内容拉长到一个学期讨论”
- 案例教学：
 - 讨论数据的权利属性、资产属性、货币属性
 - 讨论算法对于社会的影响
 - 讨论数据和算法的社会属性和技术发展的双向作用

人工智能导论

- 迎接智能系统时代
 - 给所有人的人工智能课程
- 主讲：
史维学 (Vishal Sikka) 博士
华东师范大学 荣誉顾问教授
Hang Ten Systems 创始人 & CEO
前 SAP CTO, Infosys CEO
- 6 x 90min

Preparing for the Age of Intelligent Systems

A Class on AI for Students of All Ages



史维学博士
(Vishal Sikka, Ph.D.)
Hang Ten Systems 创始人 & CEO
华东师范大学 荣誉顾问教授

10.31	8:30 - 10:00	Topic 01: AI breadth and history
10.31	10:10 - 11:40	Topic 02: AI techniques
11.01	8:30 - 10:00	Topic 03: AI systems
11.01	10:10 - 11:40	Topic 04: Opportunities and road ahead
Topic 05 - Topic 06: TBD		

课程内容

There has been a recent surge of interest and investment into the applications of AI in many domains. While this resurgent interest is new, many of the core concepts underlying these applications are many years, indeed many decades, old. Regardless, as a disruptive and transformational technology, AI has the potential to impact every aspect of human existence in a profound manner - with both positive and negative consequences. This is an review class on AI for students of all ages and levels of expertise with AI. It will discuss AI's opportunities, boundaries - and the road ahead. It will cover the theoretical underpinnings behind many of the AI concepts being applied today - but also include developing an understanding on how to build systems with AI. It is designed to be of value to beginning students, advanced students and even teachers and educators in computer science and all other domains where there could be relevant applications of AI.

主讲简介

Vishal Sikka completed his PhD in Artificial Intelligence from Stanford University, and was an Executive Vice Chairman of Infosys and former Indian American CEO and MD of Infosys. He is also the honorary professor of ECNU. Prior to joining Infosys, Sikka was a member of the Executive Board and the Global Managing Board of SAP AG, leading all SAP products and innovation globally. In his 12 years at SAP, Sikka led SAP's product portfolio including the breakthrough in-memory platform, SAP HANA, and all their applications, cloud and technology solutions. He is credited with accelerating SAP's development processes, transforming its innovation culture and leading several successful product co-creation initiatives with clients. He is also the creator of the concept of 'timeless software, which represents the renewal of products without disruption to customer environments.

基础资源建设与社会资源引入

■ 固定机位学生机房

- 实验课教室
- 实践场所
- 文化建设场所



■ 服务器

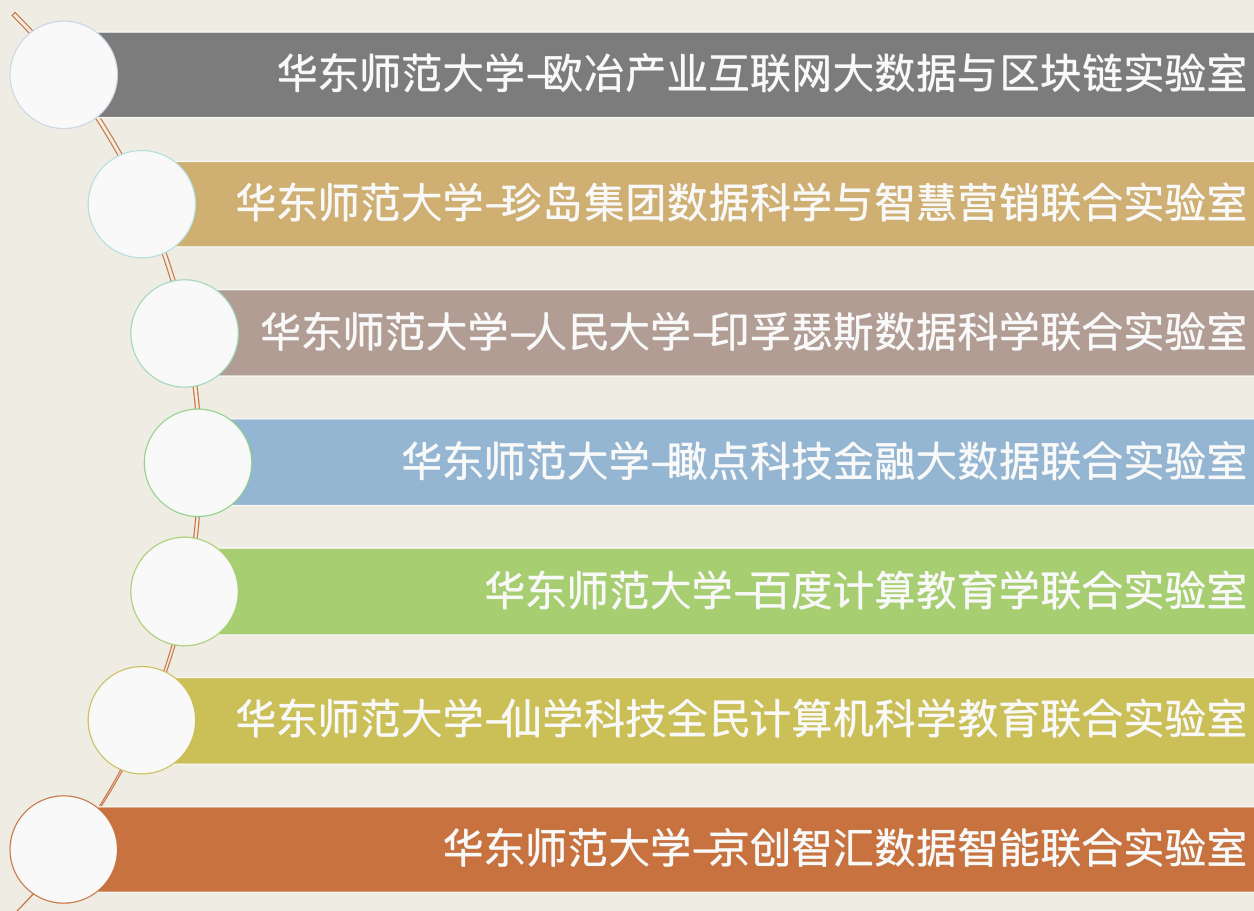
■ 云计算资源捐赠

- 学生学习过程数字痕迹
- 计算和存储资源租用



■ 企业冠名奖学金

校企联合实验室



小结

- 数据专业是**互联网环境下的新 IT 专业**
- 在教学过程中逐渐明晰其**内涵与外延**
- 学生培养要**兼顾**深造、科研、工程、学生发展的要求
- **“新工科”**建设需要打破传统理工科人才培养的方式和限制
- **“应用驱动” + “开放办学”**

谢谢！

