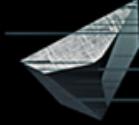


# 数据人才培养的若干思考

主讲人：陈娟（DataFountain创始人）

微信：chenjuanc



# 一张岗位薪酬表

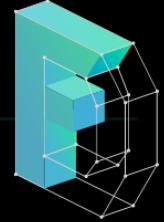


排名	职位	需求数/百万职位	平均薪资	2013-2016职位需求增长率
1	机器学习	58	\$134306	36%
2	数据科学家	360	\$129938	108%
3	计算机视觉工程师	20	\$127849	34%
4	开发运维工程师	731	\$123165	106%
5	云工程师	217	\$118878	67%
6	高级审计经理	53	\$118692	52%
7	渗透测试工程师	317	\$115557	52%
8	Oracle人力资本管理	44	\$113107	41%
9	全栈开发人员	641	\$110770	122%
10	Salesforce开发人员	230	\$108089	83%

数据来源：美国最大的招聘网站Indeed.com



# 数据人才的争夺战



AI时代的数据人才荒已经成为全行业乃至全社会最为关注的问题

政府：人才引入政策；

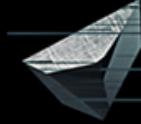
企业：共建课程、共建专业

高校：兴建数据科学相关专业

人才存量：5万+

人才需求：500万

数据来源：linkedin《全球AI领域人才报告》



# 数据人才培养面临的共同问题

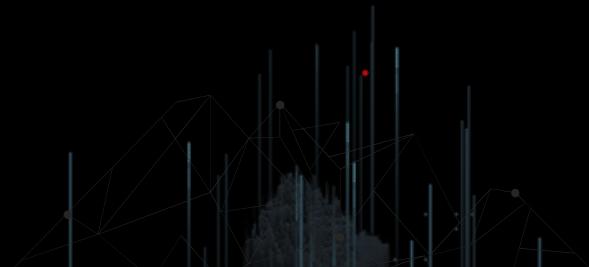


师资  
缺乏

数据  
缺乏

案例  
缺乏

教材  
缺乏





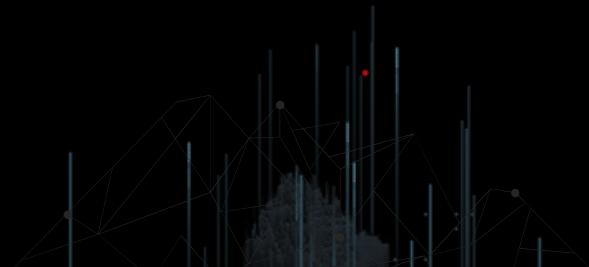
DF的发心

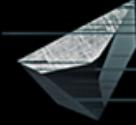


# 社会的 强劲需求

通过连续两年CCF大数据大赛的组织，我们感受到了来自学生端对于真实数据及实战经历的强烈需求，和来自企业多于人才和创新方案的强烈渴望……

数据+需求+悬赏=人才+方案+品牌  
众筹数据+众包方案

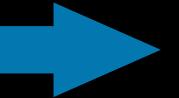




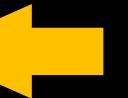
# 竞赛的概念与由来



古老智慧

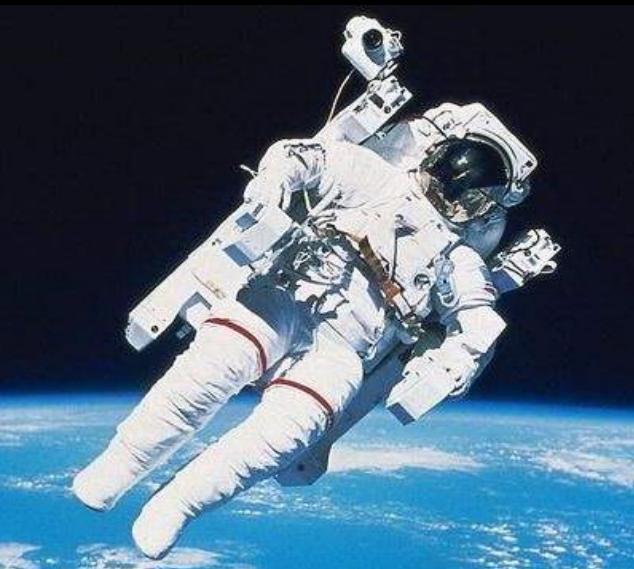


现代思维



悬赏

众包



曾经，英国议会悬赏2万英镑的“经度”大奖促使一个钟表匠发明了航海天文钟，完全改变了航海史和征服史。

拿破仑悬赏1.2万法郎征集储存食物的方法，一个商人之子发明了罐头。

近现代史上这样的悬赏还有很多，比如跨大西洋飞行，月球车、宇航员手套等。

而现代的众包思维完全改变了当代知识的生成和解决问题的方式，比如维基百科。



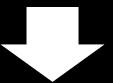
# 竞赛平台的人才培养逻辑



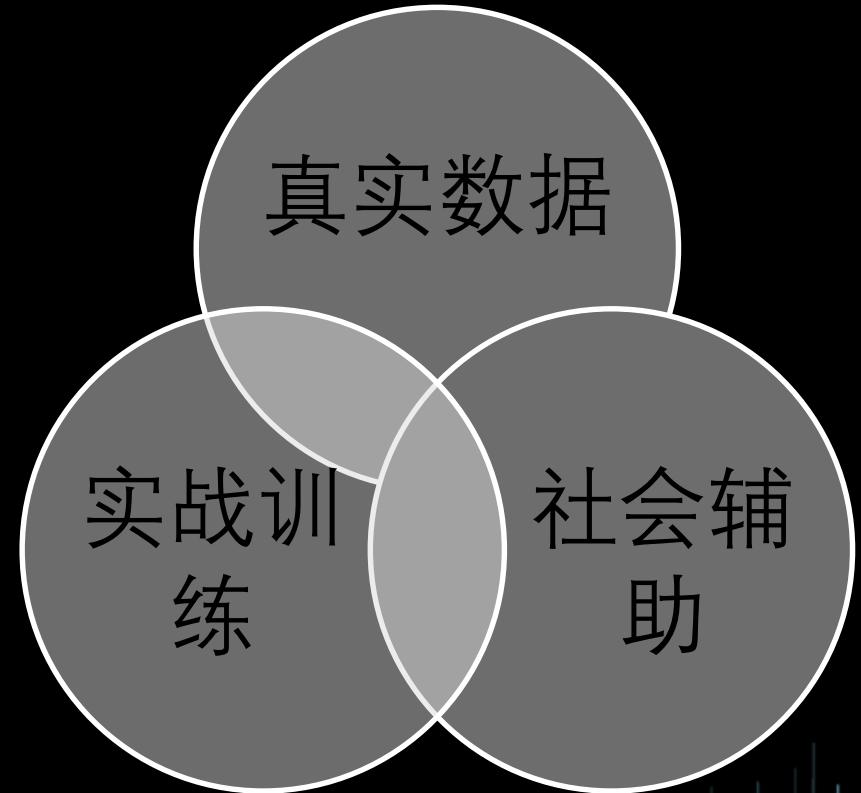
从实际问题出发

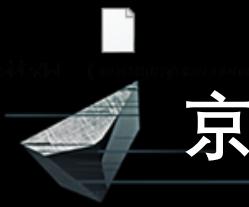


模型、算法和应用的深度理解



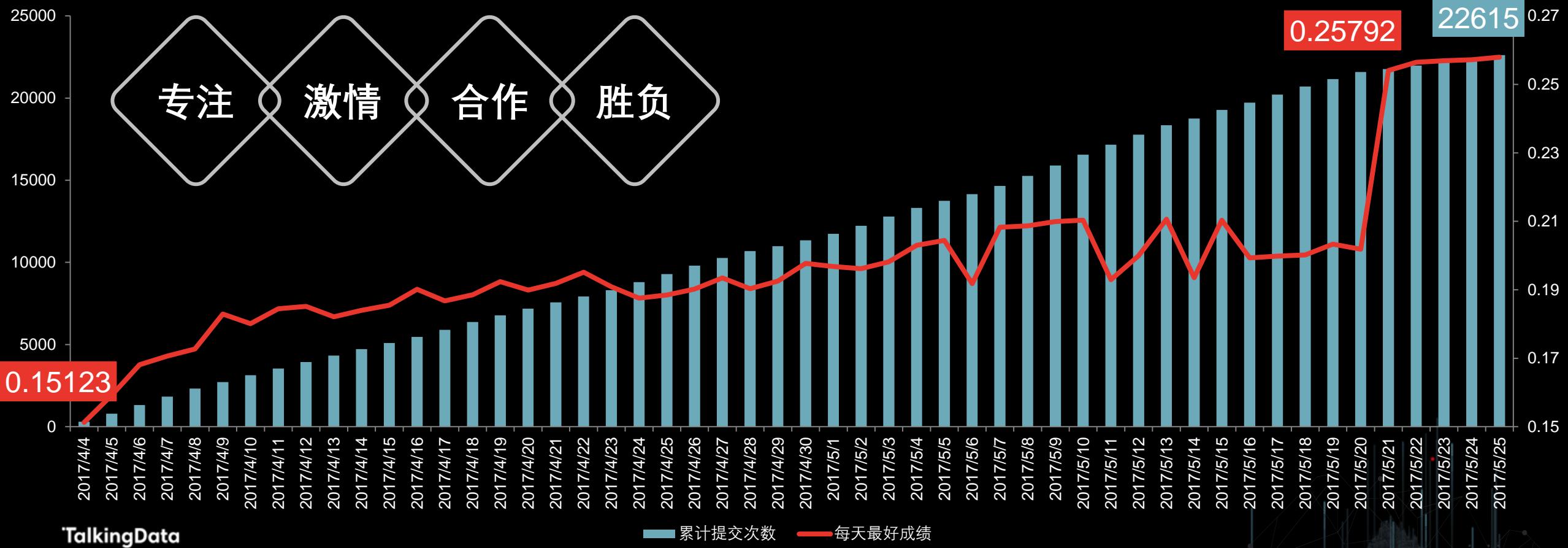
应用创新

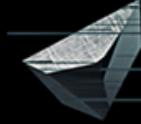




# 京东的JData算法挑战赛的真实案例

赛题背景：京东商城提供真实的用户、商品和行为数据（脱敏后）为基础，参赛者构建用户购买商品的预测模型，为精准营销提供高质量的目标群体。

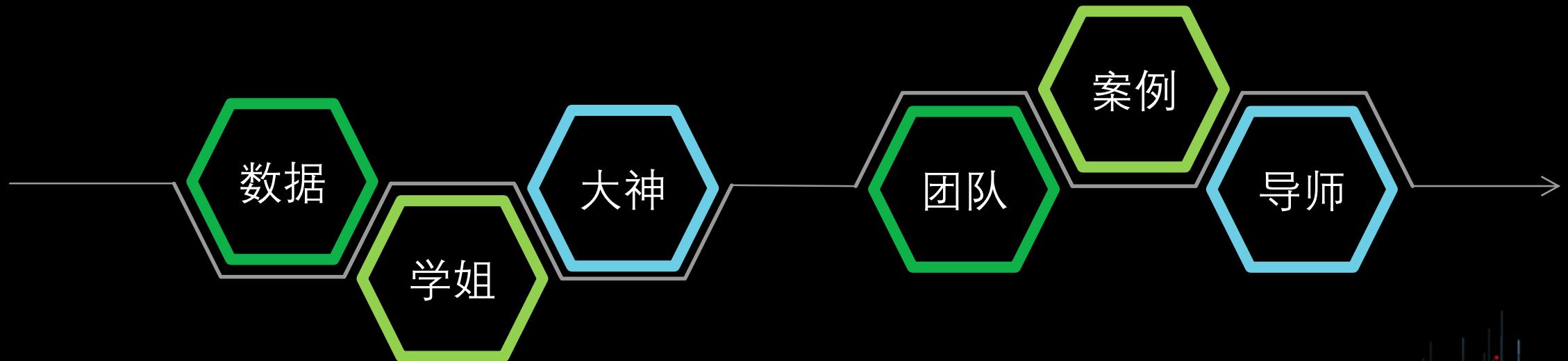




# 三个月的成长历程



一群志同道合的数据爱好者的化学反应，从此我们有个圈叫竞赛圈





# 竞赛圈有故事的男同学



销售-电脑运维-图文店老板-网管

炒股亏损的2万元，让他加入数据圈，希望算法来预测股票走势，挽回2万元的损失

组了一个20人的兴趣群，和一群“教授”学习，但是股票还是越炒越亏

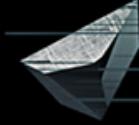
认识了海淀吴彦祖进行竞赛的圈子，从Xgboost、到R，到hadoop，渐渐开始从竞赛榜上找到了自己的排名，也找到了未来的方向

2016年CCF大赛上国家电网赛题的第二名  
入职国网信通产业集团技术部

浪叫兽说，志同道合的人走到一起是及其享受的，竞赛两年我有了家庭，有了战友，我很满足.....



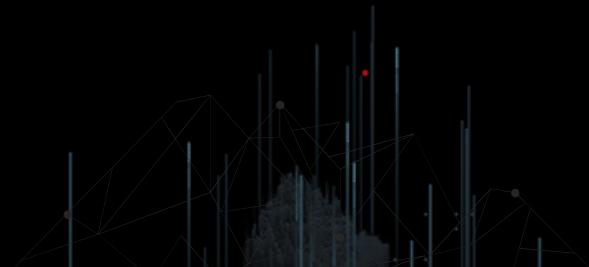
浪叫兽的故事



# 他们的故事



- 陈夏明，第二届CCF大赛一等奖获得者，上海交大，坐标上海，大数据领域创业者，C轮融资阶段
- 张学武，第二届CCF大赛三等奖获得者，北京大学，坐标深圳，机器人领域创业者，A轮融资，队友去国外攻读博士学位，研究自动驾驶技术，另一个队友去了华为，做大数据和人工智能的研发；
- .....
- 尹宁，阿里（北京）
- 吴加海，阿里（北京）
- 吴昌，网易（北京）
- 魏逸，百度（上海）
- 宋乐怡，摩根（上海）



期待共同成长与合作

易观智库

明略数据  
MININGLAMP

邀请更多志同道合的伙伴一起，助力数据人才的培养

• 用户分布：平台已有用户近**3万**人，分别来自**1000+**所高校，**51**所国家级科研院所，**700+**企业；

• DF盟友：目前在我们平台上举办**55**场竞赛，而大部分企业都愿意将数据和方案进行开源。

# THANKS

