





慧辰资道 HCR 商业大数据洞察领先者







2017 中国大数据产业地图

HCR Business Consulting

研究分析 排名第一

2016 收入 3.02亿



深入认知消费者/个性化服务提升/提升企业经营 政府公共管理





汽车|快消|运营商|家电|TMT|通信数码|医药|零售|教育 政府管理|智慧城市|安监



























SIEMENS



































HCR 业务服务模式 数据科学技术 + 垂直 / 人民邮电出版社 POSTS & TELECOM PRESS



产品服务形态

多种形式

Adhoc 研究报告

标准化产品

专项 在线系统

多行业 专家团队

按行业线深入

行业标准产品 研究员洞察能力 行业大数 据产品

行业研究 模型

汽车

行业大数 据产品

行业研究 模型 快消

移动运营商

行业大数 据产品

行业研究 模型

IT/数码/家电

行业大数

据产品

行业研究

模型

行业大数 据产品

行业研究 模型

医药/旅游

行业大数 据产品

更多

行

行业研究 模型

教育/金融

数据科学 团队

数据建模

基础计算/挖掘

SPark/NLP/ML..

基础数据库

千亿/TB/异构/分布式

大数据处理

数据源整合

多种渠道/来源接入

数据清 洗融合

分析

机器 学习

言处理

深度 学习 常用商业 业务模型 用户标签画 像分析

移动互联 网数据

用户各类 行为数据

互联网/ 论坛数据 社交媒 体数据 企业业务 /CRM..

线下数据 (超市/交易..)

外部渠道

自主采集/搜索

众包

客户企业提供 (销售/CRM/客服) 合作方/行业 性



HCR 数据科学团队 领先的商业分析建模与 人民邮电出版社 POSTS & TELECOM PRESS



自有的面向深度洞察分析的大数据分析平台

分布式集群 / 千亿级数据处理能力 / 多级建模能力支持

分析 工具

统计性分析 SPSS/SAS/R

数据挖掘 SPSS Modeler/Weka 可视化分析 **TableAU**

在线可视化 (ECharts/D3)

工具/语言 (Java/Python)

业务 计算

数据

管理

机器学习 Spark MLib

Spark Sql

分布式计算 Spark

类Sql检索

NLP 处理 **API**

数据订阅 Kafka

ETL Kettle

业务 规则 引擎

流式实 时计算 Storm

Infobright/GreenPlum

结构化数据 HBase/Cassandra 结构化数据

非结构化数据 全文检索 **Elastic Search**

文件 系统

HDFS(Hadoop)

Linux Ext4 / NFS

专业的数据科学团队

清华/北邮/中科院专业博士领衔 统计/NLP/图挖掘/机器学习/深度学习

深入的行业业务理解/建模能力

国家高新技术企业(行业唯一)

数据研究 相关专利 2 申请中 数据分析 相关软著 13个







K12教育大数据分析应用





HCR 从商业研究角度 认知 K12教育大数据 | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L R | L



二三四线城市

用户群巨大 诉求强/上升通道 低成本需求

学校(教育局)体系

中国教育的主战场 影响/辐射能力最大 长于教学/交互 弱于分析

核心业务诉求

更好的升学能力 需要更强的学科知识能力

教育局/学校/教师/学生/家长

如何解 决核心 诉求

因材施教

识材:深入的学生分析 施教:个性化的教学

> 可汗学院 已经证明其可行性

HCR

精准的分析:深入认知每个学生特性 个性化提升:每个学生针对建议/策略

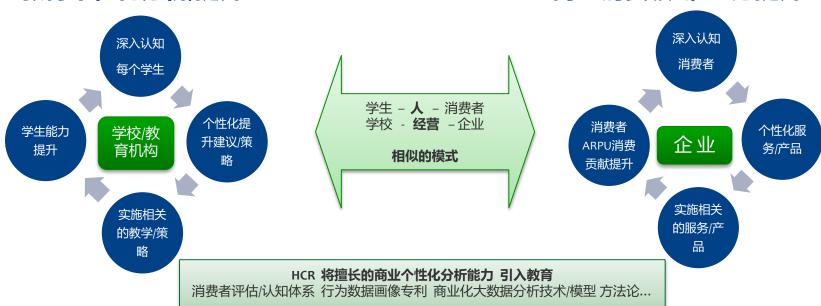
让学校/老师 专注于擅长的教学与交互 教师/家长:基于个性化建议更有效提升学生





教育个性化分析/提升

商业消费领域 业务提升



To B 帮助教育局/学校 学情分析/策略 提升教学能力

To C 帮助学生/老师/家长 精细认知学生的学习能力不足

指导学校/学生 如何有效提升 提供个性化教学/增强建议/策略



HCR 教育大数据分析服务 三驾马车







消费者 认知模 型

商业分析方法

研究方法论/模型

分析 方法

心理学/ 行为学

行为画 像专利

知识遗 忘函数..

面向人的综合分析体系

基于商业 结合教育业务特点 可通过数据量化分析/可实施 ..





数据科学 大数据 + AI

建模分析技术





HCR 个性化分析提升服务 流程





教育大数据采集

教学资源数据/教学数据 学生属性与行为数据...

系统题库 校本题库 学生背景数据/在校行为数据 教师授课数据 考试评价数据(连续)... 持续 迭代

2 分析决策

【面向教育局/学校】 大数据学情分析/决策

跨校联考分析 上线分布评估 临界生锁定与个性化提升策略 走班教学选科建议

教师教学能力分析

持续 迭代

3 提升建议/策略

【面向学生端】 个性化提升建议/服务 反馈 闭环

学科知识点弱点提升 个性化补练方案 个人学情分析/加强

学生个性化试题/强化作业 学校选科/培训/个性化在线课堂

自动标注题 目的知识点

知识点之间的 自动匹配个性 化练习题目 关联关系

模拟教师出题 思路出题/组卷 根据现有成绩 预测未来成绩

智能选科

学科知识点体系

自然语言处理/图像处理

机器学习

深度学习

个性化推荐



HCR TOB服务场景举例: 帮教育局/学校 钊 人民邮电出版社 POSTS & TELECOM PRESS







临界生: 预期在分数线区域 波动的 大量学生群体

上线结果 影响最大 学校/家长 重点关注

教育行业现有模式

教育局统一组织学校 开会研讨:

历史+近期表现



预估分数线

本地区终模/高 考预期分数线

各学 校

粗粒度的经验模型 取决于决策者的经验/能力 耗时耗力/不准确

按经验筛选

老师依据该预测线 根据经验感觉筛选



临界生





精确到每个学生



LR + 深度学习RNN(连续序列)+分布异常估算



举例:某市重点中学2017 高考临界生』 人民邮电出版社 POSTS & TELECOM PRESS





最大潜力提升科目 以及擅长该内容教学的老师



学校根据建议 进行针对性突击辅导/潜力科目专项提升



ToC服务场景: 面向学生 个性化深入 MACHIBLE AND A STREET OF THE TOTAL TOTA











薄弱学科



根据数据 自动分析





圆的切线方程..





推荐 (自动)

从学科细化到学科内知识点 (三级)

薄弱学科 -> 薄弱知识点

基于学生的薄弱学科,通过学科知识 点能力分析模型(基于AI模拟学科老 师的相关分析能力),自动发现学生 更深层/更具体的知识点缺陷



针对薄弱知识点强化

- ✓ 个性化巩固练习
- ✓ 小班专讲(针对薄弱考点)
- ✓ 走班补差/在线课堂



HCR 我们 通过 大数据+人工智能 学习模》 ARMed H版社 POSTS & TELECOM PRESS





中学题 库数据

人工 智能 建模

教师 学科 认知 分析 模型

教育大数据资源

- 1学生考试/答题记录
 - 2 学生背景信息
 - 3 教师授课数据



自然语言处理/图像识 别/深度学习(RNN)

每学生 学科能力画 像

个性

化

推荐

细化到每知 识点(缺陷 /问题)

个性化策略

反馈强化学习

自动组卷 个性化作业 知识点强化专 项题目

学生答题

答题 结果 识别

题库大数据 题目/教师组卷习惯)

Step1:构造每科的学科知识分析模 型,可分析题目考核知识点和难度

Step2: 基于教学/考试行为数据, 生成学生的学科能力深度画像

Step3: 基于画像,个性化推荐最适 合学生提升的策略/试题/组卷







中学题库 大数据

230万 中学学科题 目/标注数据

自然

语言

理解

图像 识别



学科知 识点体 系



教师组卷 习惯

30万试卷

人类学习函数

学科教学规律

学科与知识点 分析认知模型 数学 初中/高中 深度 机 学习 器学习 优 语文 化 高中/初中 **RNN LSTM**

> 细化到三级知识点 独立建模

错误结果 反馈/优化

题目自动分析

初中/高中 数学题目



如 某试题如下:

①与圆只有一个公共点的直 线是圆的切线;②与圆心的 距离等于半径的直线是圆的 切线;③垂直于圆的半径的 直线是圆的切线 ...

所考核学科知识点:

高中数学

结

- -> 平面解析几何
- -> 圆的切线方程的求法



HCR □ 举例: ToC 学生提升服务平台(家长 (□ 人民邮电出版社 POSTS & TELECOM PRESS





通过学校学情大数据,及时 诊断学生近期学习薄弱点



对应强化策略 提升 将缺陷/漏洞及时填补









World Class Research from China