

# ITStar大数据公开课 —

@潭州教育旗下大数据培训

# 老师介绍

- 15年以上的IT行业从业经历，清华大学计算机软件工程专业毕业，现任Oracle中国有限公司高级技术顾问；曾在BEA、甲骨文、摩托罗拉等世界500强公司担任高级软件架构师或咨询顾问等要职，精通大数据、数据库、中间件技术和Java技术。有多年的大数据培训经验。

- 在线课程涉及：

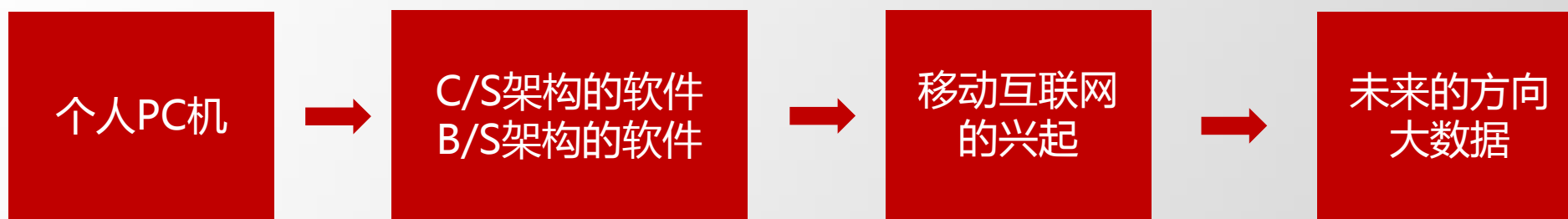
- 大数据、大数据实时处理
- Oracle数据库
- Weblogic中间件
- NoSQL数据库



# 过去的工作和现在的工作

- 目前就职于Oracle（甲骨文中国）有限公司
  - 在Oracle公司工作已经超过10年
- 职位：高级技术顾问
- 过去的工作
    - 03年，在校期间在IBM研究中心（IBM CRL）实习生
    - 亿阳信通
    - 摩托罗拉（中国）
    - 北电网络（中国）

# IT 软件行业的发展



任何时候：得数据者，得天下！！

# 大数据的就业前景

## Java开发工程师的薪资待遇

Java开发工程师 [酒仙桥] 10:19发布

15k-30k 经验3-5年 / 本科

Java 架构 阅读

自如网

移动互联网,O2O / 成熟型(不需要融资)

"15-17薪 带薪年假 房补 餐补 六险一金"



Java [西直门] 13:33发布

15k-25k 经验3-5年 / 本科

后端开发 Java MySQL

Udesk - 企业级智能客服...

企业服务,数据服务 / 成长型(B轮)

"B轮融资,团队强悍,待遇丰厚"



Java工程师 [中关村] 2017-04-21

15k-25k 经验5-10年 / 本科

高级 后端开发 分布式 J2EE SOA

拉勾网

企业服务,招聘 / 成熟型(C轮)

"技术挑战、工作居住证、美味果盘、运动健身"



Java [通州区] 1天前发布

20k-35k 经验3-5年 / 本科

Java web

京东商城

电子商务 / 上市公司

"扁平管理,领导好,美女多"



## Hadoop开发工程师的薪资待遇

Hadoop 研发工程师 [望京] 12:38发布

20k-40k 经验1-3年 / 本科

资深 高级 架构师 hadoop 数据治理

数美

企业服务,数据服务 / 成长型(A轮)

"源码研发,大牛如云,技术氛围,大佬领投"



Hadoop [新街口] 15:50发布

25k-40k 经验3-5年 / 本科

弹性工作 敏捷研发 股票期权 年底双薪

和创科技 (红圈营销)

移动互联网,企业服务 / 上市公司

"上市公司,极牛汇集,持续的股权激励政策!"



Hadoop研发工程师 [亚运村] 11:10发布

20k-35k 经验3-5年 / 本科

软件 hadoop 大数据 Java python

京东

移动互联网,电子商务 / 上市公司

"源码研发,大牛汇聚,各种福利,上班购物"



Hadoop开发 [西坝河] 13:35发布

15k-30k 经验1-3年 / 本科

高级 hadoop 数据挖掘 ETL 数据仓库 BI

大鱼自助游

旅游 / 成长型(B轮)

"工程师文化,两次调薪,B轮融资,期权激励"



# 需要具备的基础

- 会一定 Java SE编程基础
- 有一定的网络基础，起码知道客户端和服务器的概念
- 有一定 Linux基础
- （最好）了解一些关系型数据库，比如：Oracle、MySQL等等

# 大数据能做什么？

数据爆炸式的增长：（每秒.....）

## 什么是数据？

得数据者，  
得天下

13000+个 iPhone 应用下载	Skype上 37万+分钟的 语音通话	Twitter上发 布98000+新 微博
上传6600张 新照片到 flickr	发出1.68亿+ 条Email	Facebook上 更新69.5万+ 条新状态
YouTube上 上传600+新 视频	淘宝光棍节 10680+个新 订单	12306出票 1840+张





# 什么是大数据？

- 思考：

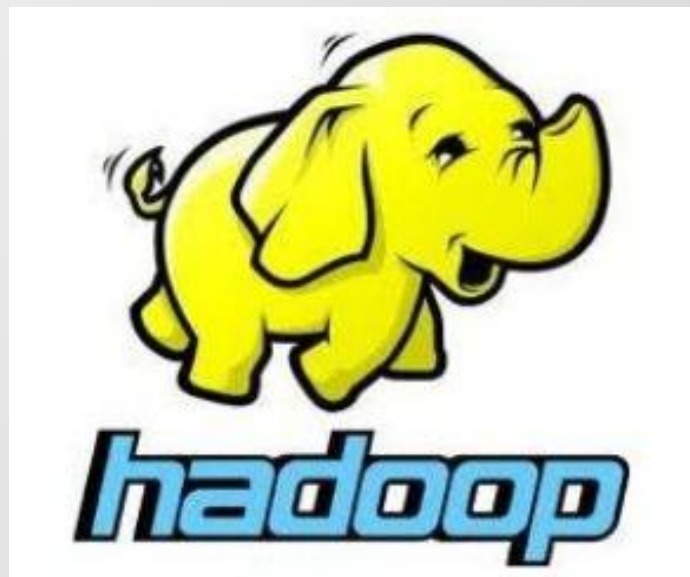
- 如何进行商品推荐？
- 如何进行天气预报？

- 大数据的本质：

- 数据的存储
- 数据的计算

- 五个特性（五个V）：

- Volume（大量）、Velocity（高速）、Variety（多样）、Value（价值）、Veracity（真实性）





# 睿思博考大数据业务部 2.0 产品架构图



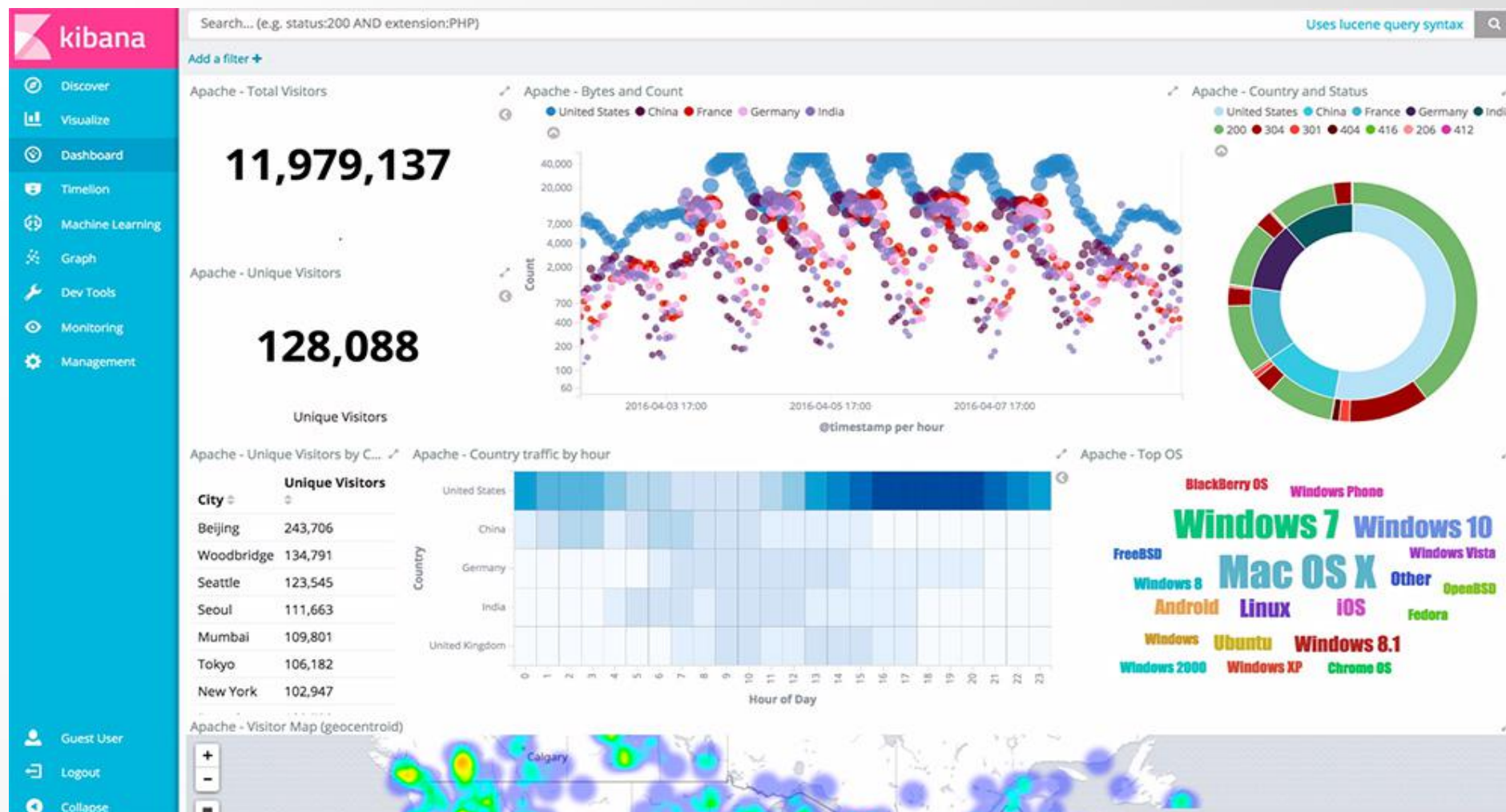
## 课程特点及课程内容

- 理论+实践
- 强大的课程体系

# 项目实战课程一：大数据网站离线日志分析系统

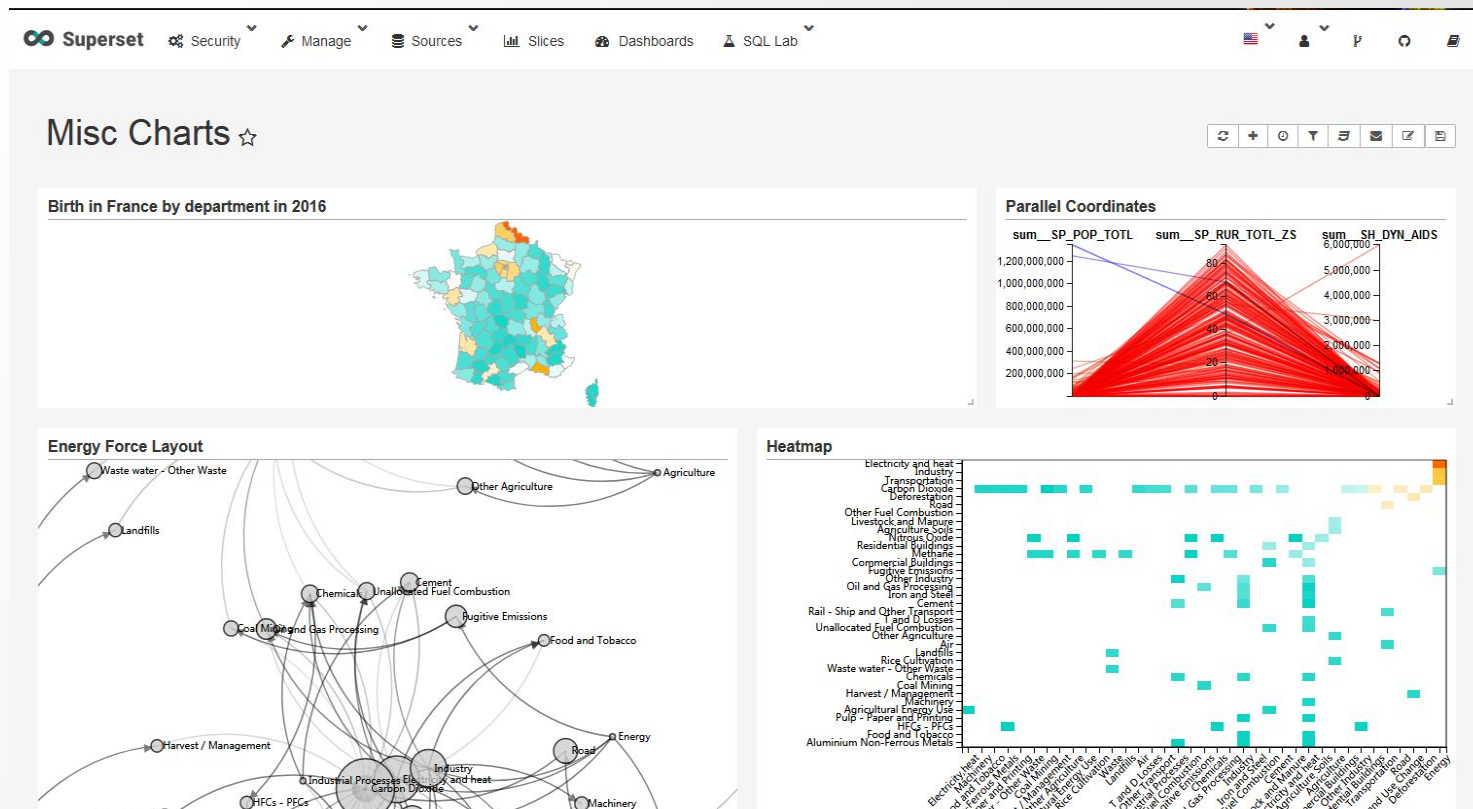
- Web日志分析概述Web日志由Web服务器产生，可能是Nginx，Apache，Tomcat等。从Web日志中，我们可以获取网站每类页面的PV值（PageView，页面访问量）、独立IP数；稍微复杂一些的，可以计算得出用户所检索的关键词排行榜、用户停留时间最高的页面等；更复杂的，构建广告点击模型、分析用户行为特征等等。
- 使用哪些技术：Hadoop Spark Hive Hdfs Elkstack Flume HUE OOZIE  
Kafka

# 效果图



# 项目实战课程二：互联网+在线电商平台行业分析系统

- 使用哪些技术：Kafka SparkStreaming Redis Hue Superset
- 效果图



# 问题与回答

## •问题1

- 人工智能是要基于大数据吗？大数据和机器学习之间的关系？

## •问题2

正式课上课时间和包括了项目实战在内的课程总课时有多少

## •问题3

- 如果大家还有其他问题，可以联系ITSTAR的老师，如果问题比较多，我们会在录制视频给大家讲。