

业界动态

2016.06.07

韩岳岐

HANYUEQI@CNIC.CN

亚马逊正在蚕食这个世界

- ▶ 亚马逊提供了Hadoop、Spark、ElasticSearch，并不断上移堆栈，程序员发现将堆栈托付给Amazon更便宜、更快

- ▶ 机器的海洋

- ▶ 可租用你想要的不同时期的机器
 - ▶ 各种性能的机器

- ▶ 市场化运作机制

- ▶ 防误删等准硬件技术支持
 - ▶ 自动化创建、测试和部署
 - ▶ 现货市场出售容量，价格随市场波动

- ▶ 云创新

- ▶ Lambda，Amazon的无服务事件驱动计算服务
 - ▶ Elastic BeanStalk，模仿Google，自动弹性扩容和缩减集群，自动负载均衡，云管家
 - ▶ 几十个公共数据集

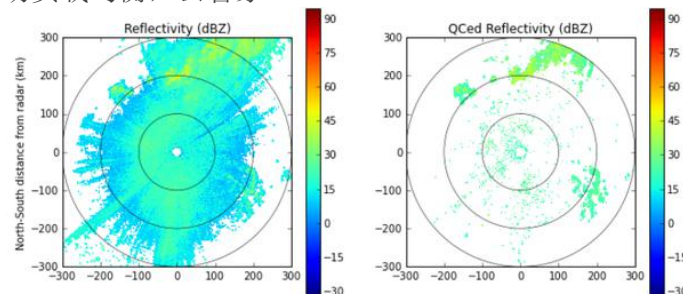
- ▶ 越飞越高

- ▶ 实例化功能集群，Spark、Hadoop、Presto、Pig、Hive
 - ▶ 更完整的机器学习工具，分析数据卖给用户
 - ▶ 为用户收集公共数据

The screenshot shows the 'Settings: Test1' page for an AWS CodeCommit repository. It includes sections for 'Repository name', 'Repository ID', 'Repository ARN', 'Repository description', and 'Default branch'. On the right, there are settings for 'Number of instances', 'Purchasing option' (Request Spot instances), 'Network', 'Subnet', 'Auto-assign Public IP', 'IAM role', 'Shutdown behavior', 'Enable termination protection', 'Monitoring', and 'Tenancy'.

资源拍卖

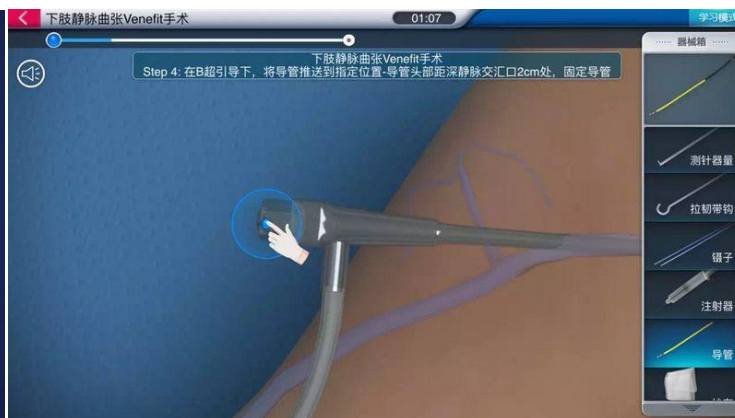
资源申请



气象数据

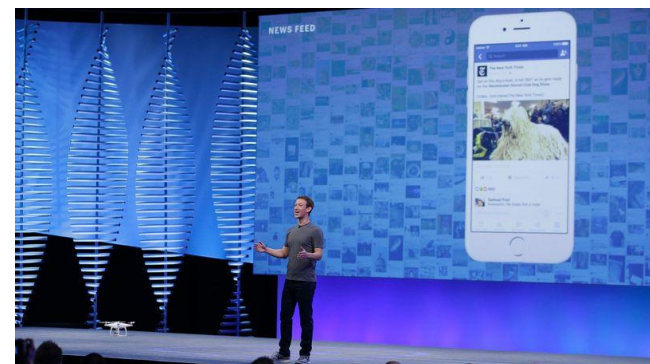
柳叶刀客利用VR“解剖” 医疗行业

- ▶ 面向青年外科医生、主打医疗+VR在线手术学习工具
- ▶ 获得千万元天使投资
- ▶ 36种手术，4个科室
- ▶ 手术（3D交互模拟）+VR（360全景视频直播和录播）



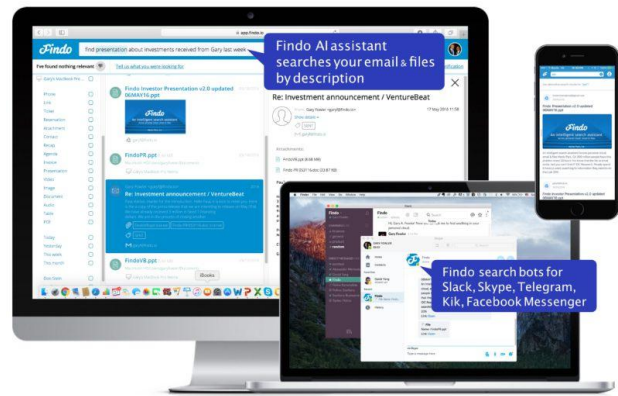
Facebook发布人工智能引擎 Deep Text

- ▶ 像用户一样理解Facebook发布内容与聊天信息，挖掘用户感兴趣信息，背后隐藏的是想成为一个强大搜索引擎的雄心
- ▶ 对海量信息进行分类
- ▶ 基于神经网络接近于人类识别精确度理解数千条内容，并识别20余种语言。
- ▶ 分析销售情况给出销售建议，分析朋友出行情况，智能化连接信息孤岛
- ▶ 无人机连接世界，治愈所有疾病，登月计划。。



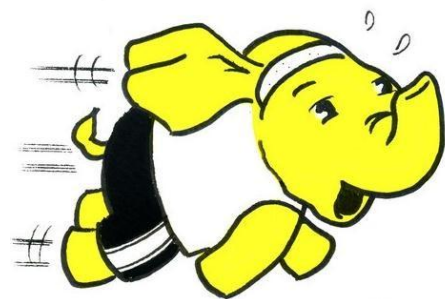
Findo获300万美元种子轮融资

- ▶ 利用AI和自然语言查询帮你搜到所需要的任何资料
- ▶ 通过自然语言查询和聊天机器人的方式找到深藏在文档、文件、邮件附件和通讯软件等众多地方的有用信息
- ▶ 帮助用户搜索储存在本地的文档资料，同时也能搜索到Facebook、Dropbox、Google Drive、Evernote、Box、Gmail、Exchange 和 Outlook 等这些云端账户里的资料文档。如果记不清具体的文档资料标题里的关键词，用户甚至可以通过记忆中一些模糊的细节信息进行搜索
- ▶ iOS 和网页端的智能助手 Findo.io
- ▶ 关键词+自然语言处理+语义分析



比Spark快10倍的Hadoop3.0有哪些实用新特性？

- ▶ Mapreduce，内存+io+硬盘
- ▶ hdfs就近计算原则，本地块内存化，计算，IO，共享内存计算
- ▶ 重大特性
 - ▶ Hadoop Common,精简内核、Classpath独立机制、Shell脚本重构
 - ▶ HDFS支持数据擦除编码，支持多NameNode部署
 - ▶ MapReduce，Tasknative、内存参数自动推断
 - ▶ Yarn，cgroup内存与硬盘隔离、curator选举、container resize



阿里投资大数据电商搜索引擎 Twiggle公司

- ▶ Twiggle，电子商务搜索引擎，以色列公司
- ▶ 成立于2014年，解决大数据、搜索及电子商务领域问题，拥有数据科学、人工智能知识背景
- ▶ 更好地利用数据去理解消费者真正需求，例如用户搜索的笔记本究竟是看重笔记本的哪种特性
- ▶ 阿里人工智能布局
 - ▶ 小AI，通过神经网络、社会计算、情绪感知进行本质洞察和实时预测，应用交通管理、能源管理、天气预测等应用
 - ▶ 淘宝小蜜，训练型人工智能，商品推荐



安全大数据公司微步在线 (ThreatBook)获3500万A轮融资

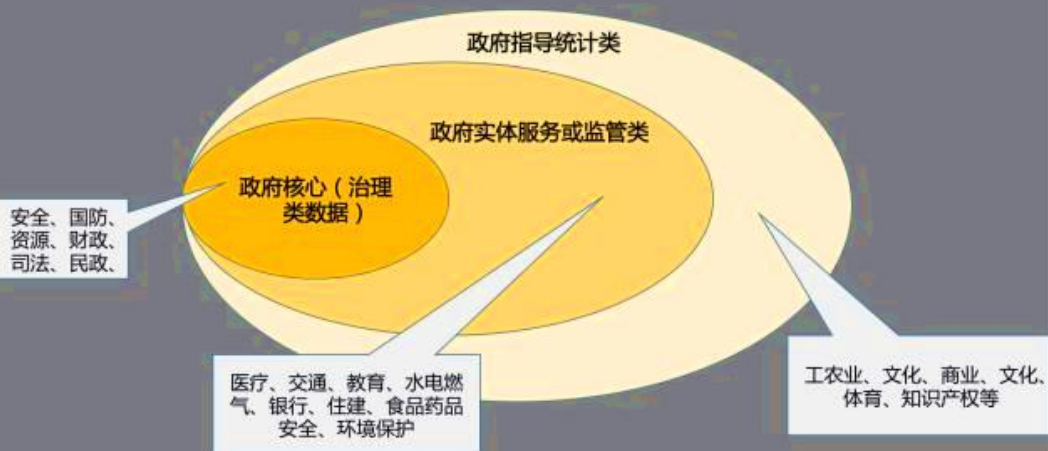
- ▶ 2015年6月成立于北京，国内首个以安全情报威胁服务为中心的互联网公司
- ▶ “聚焦威胁，情报驱动”，专注于提升威胁分析和情报能力
- ▶ 生产环境安全、企业内部安全，根据行业特点提供业务安全解决方案
- ▶ XcodeGhost事件溯源，Darkhotel对国内APT供给行为溯源

2分钟，全面了解政府大数据门类及应用！

► 价值

- 使政府治理与决策更精细化、科学化
- 促进公共服务能力与水平的全面提升
- 引发民生大数据开发应用的新高潮

- 2018年前构建政府大数据统一开放门户
- 政府数据的级别分类



政府大数据的应用领域



新加坡大数据公司Lynx Analytics拿到1000万美元投资

- ▶ 利用图理论汇聚信息，精准刻画消费者画像，帮助企业进行产品推广
- ▶ 营业收入每年50-60%增长，曾拒绝多家公司投资与并购



Stack Overflow网站：30台服务器支撑日均1.7亿PV？

- ▶ 互联网、负载均衡、web层、服务层、缓存、推送、搜索、数据库
- ▶ 4 台 Microsoft SQL Server 服务器（其中 2 台使用了新的硬件）
- ▶ 11 台 IIS Web 服务器（新的硬件）
- ▶ 2 台 Redis 服务器（新的硬件）
- ▶ 3 台标签引擎服务器（其中 2 台使用了新的硬件）
- ▶ 3 台 Elasticsearch 服务器（同上）
- ▶ 4 台 HAProxy 负载均衡服务器（添加了 2 台，用于支持 CloudFlare）
- ▶ 2 台网络设备（Nexus 5596 核心 + 2232TM Fabric Extender，升级到 10Gbps 带宽）
- ▶ 2 台 Fortinet 800C 防火墙（取代了 Cisco 5525-X ASAs）
- ▶ 2 台 Cisco ASR-1001 路由器（取代了 Cisco 3945 路由器）
- ▶ 2 台 Cisco ASR-1001-x 路由器

Stack Overflow网站：30台服务器支撑日均1.7亿PV？

