

大数据时代：  
运用EOI分析架构智能驱动企业运营和战略决策

Jan 2017

Michael Li, Head of Analytics



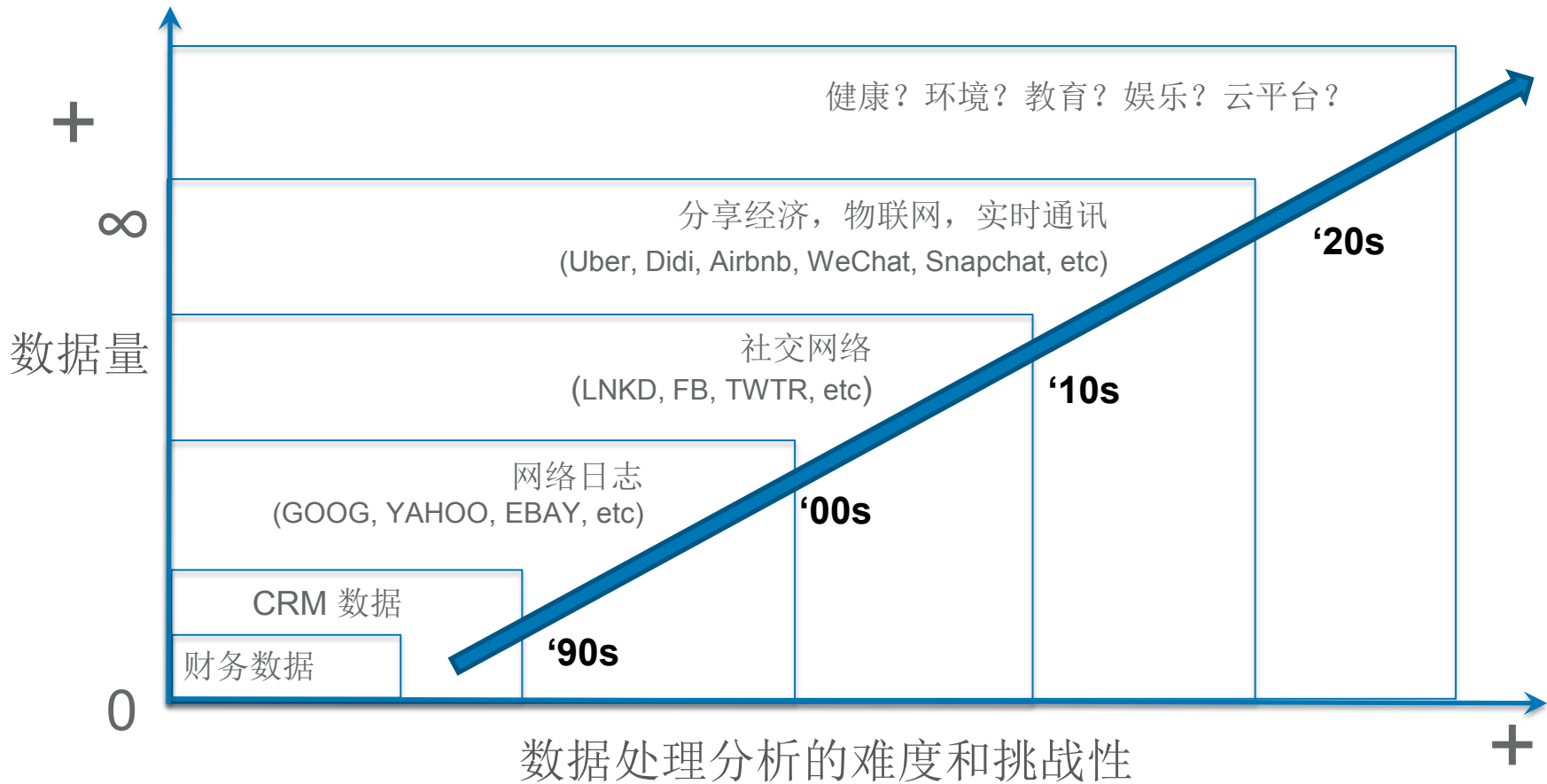


## Michael Li 李玥

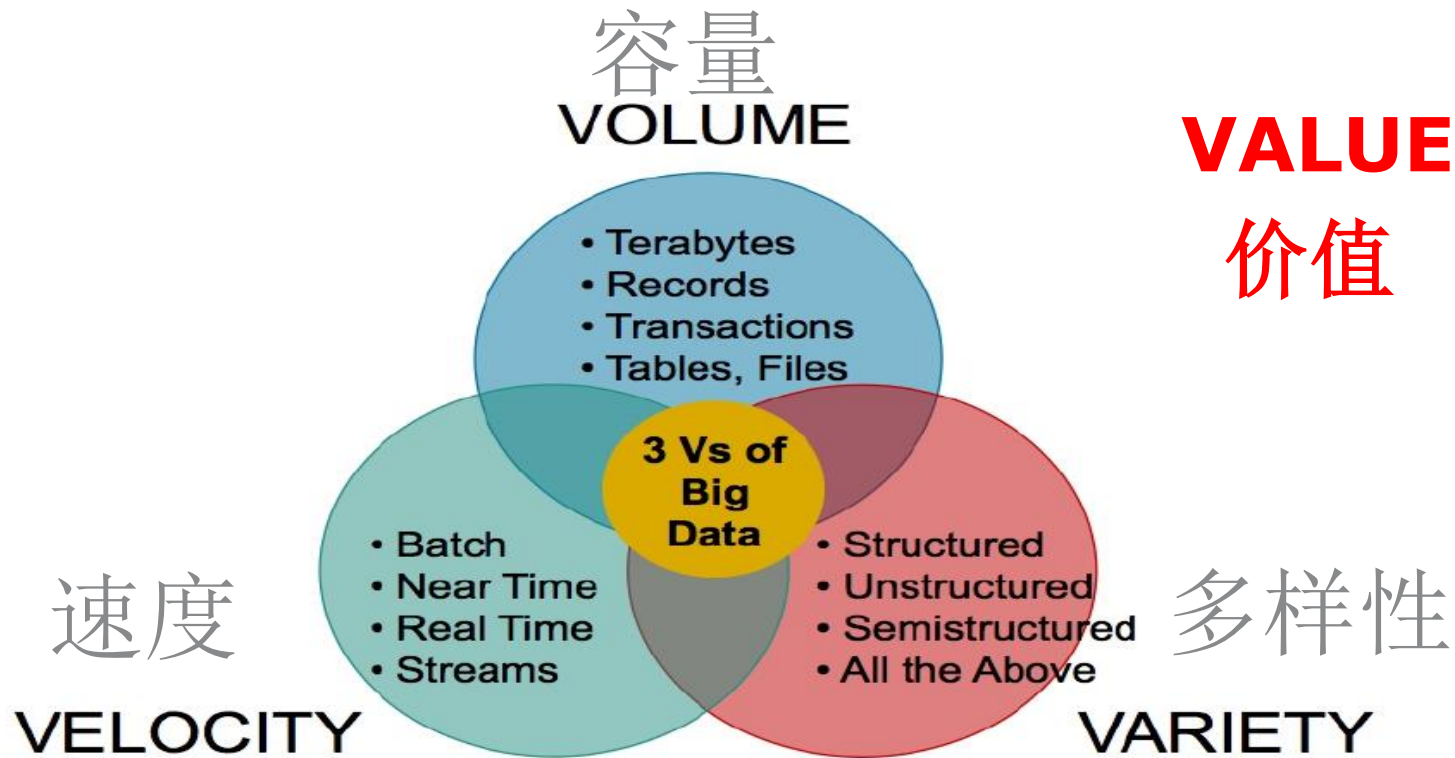
- LinkedIn(领英), 大数据分析和数据科学总负责人, *Head of Analytics (150+人团队)*
- Coursera, 在线教育初创公司, *Advisor (顾问)*
- GAGO 佳格, 环境科学大数据初创公司, *Advisor (顾问)*
- 曾供职于美国信用卡公司第一资本*Capital One*和电子商务平台*eBay*
- *Santa Clara University, Advisory board for M.S. Business Analytics (商业分析学位导师)*
- 毕业于 *Fudan University (复旦大学)* and *University of Illinois at Urbana-Champaign*
- *Teaching at Stanford and UC Berkeley*
- **Big Data Evangelist and Practitioner** 大数据宣传员和实践者



# 到底什么是“大数据”？



# 大数据本身的三个基本技术维度：3Vs



• Chart from Philip Russom- Research Director: TDWI

# LinkedIn's BIG data



**450M+**  
Members

**35B+**

Quarterly page views

**8B+**

Endorsements

**8M+**

Active company profiles

**8M+**

Jobs

# LinkedIn's Economic Graph 经济图谱

经济图谱就是用数字化的方式呈现各种经济元素。有了这个经济图谱，我们就能更好的洞察经济发展的规律，让经济元素更有效的在行业之间、地区之间流动，创造更多更好的经济机会。



# 大数据分析目标 Goal of Big Data Analytics

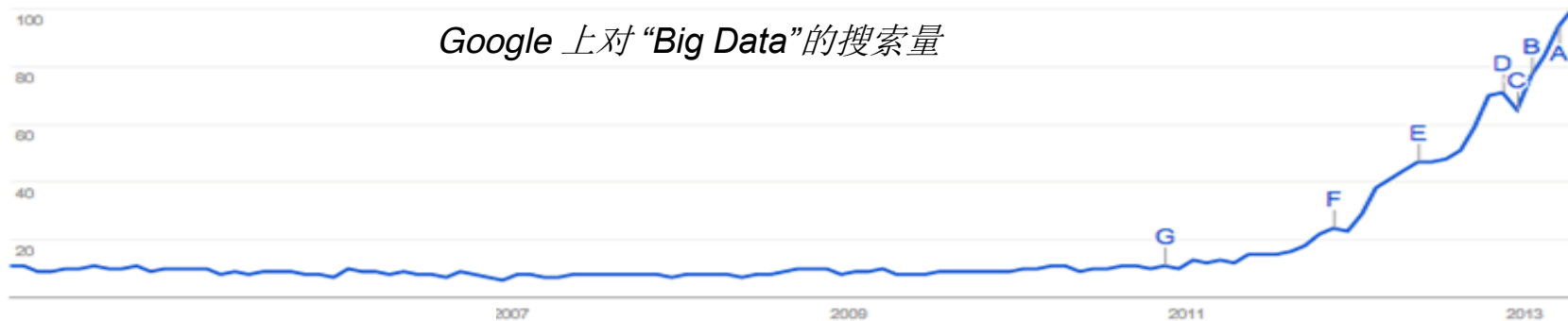


Enable high quality, agile, and impactful decision making for all professionals by providing **actionable insights** and **scalable analytics solutions** leveraging big data  
- **“Data Science with impact”**

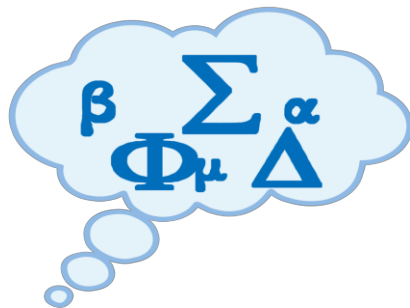
利用大数据为所有职场人员做出迅捷，高质，高效的决策提供有实践指导意义的洞察和可规模化的解决方案  
- **“创造商业价值的科学”**



# 分析师团队将技术和业务有机结合起来



技术部门



分析师



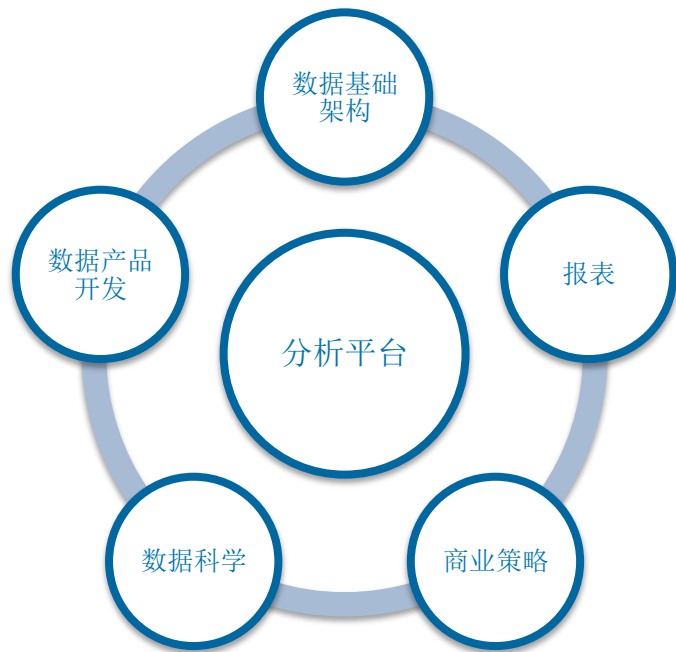
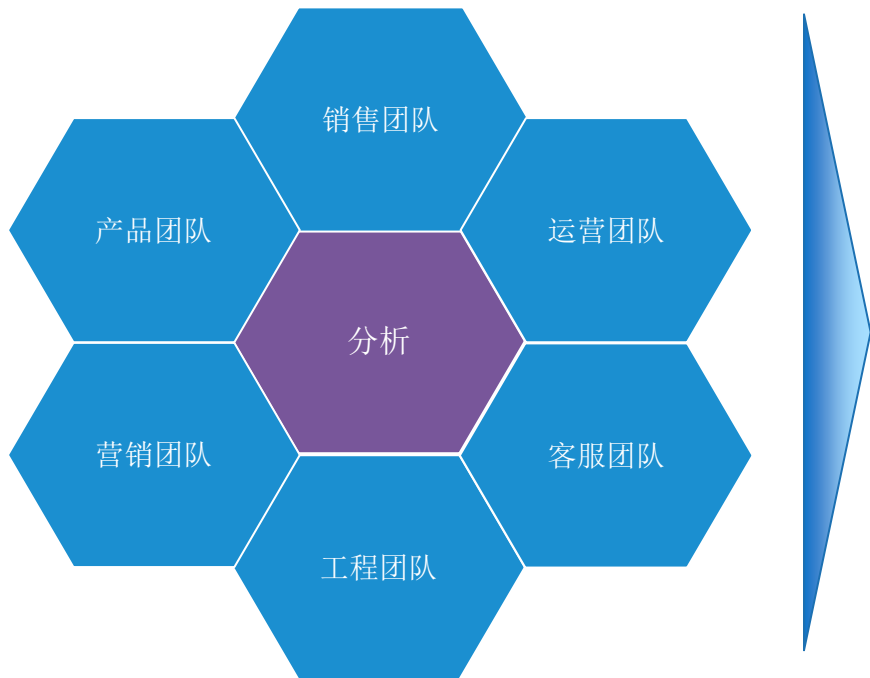
业务部门

Michael Li, Head of Analytics





# 大数据分析团队与全公司各部门合作驱动业务增长



Michael Li, Head of Analytics



# 商业分析团队利用大数据手段来持续推动商业价值的方法 - EOI (助力、优化、创新)的分析架构

## Empower 助力

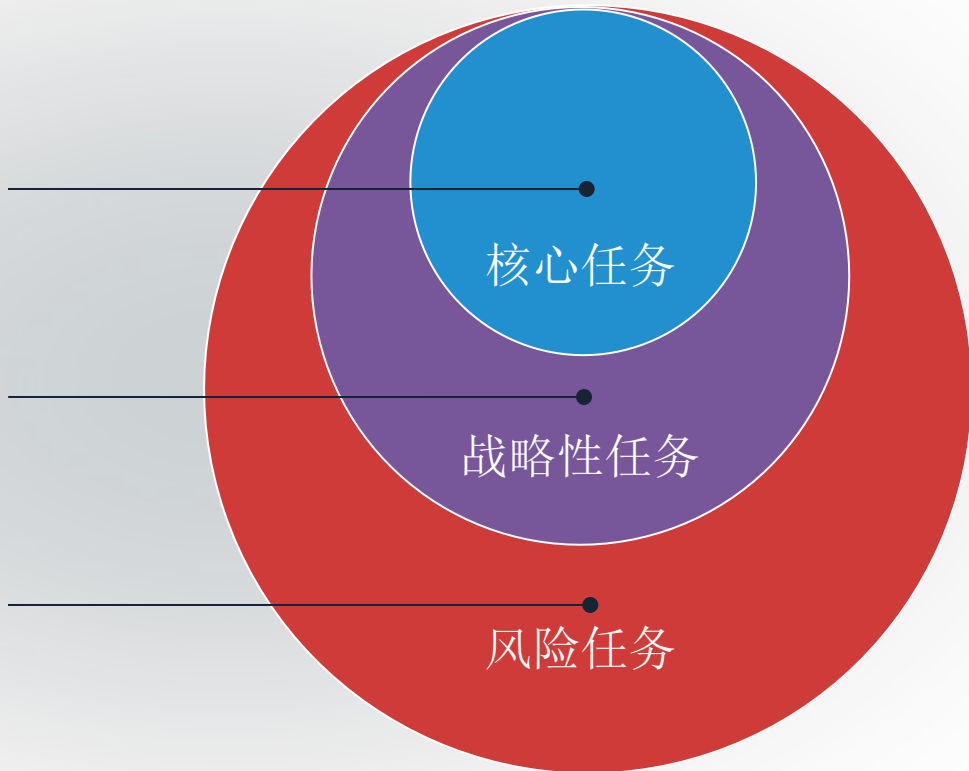
Empower business partners to have access to the data and insights they need when they need them

## Optimize 优化

Optimize business performance leveraging the powerful & unique LinkedIn data we have

## Innovate 创新

Innovate the way on how analytics can help our business grow leveraging both internal & external data





李明  
LinkedIn 首席分析师

## LinkedIn 的大数据新玩法

基于大数据技术的 EOI 分析架构已被证明有助于企业商业价值的最大化。

在信息发达的今天，我相信任何一个对科技略有兴趣的人，都不会对大数据这个词感到陌生。尽管大多数人并不一定知道它的确切含义，但仍然会为它可能为这个世界带来的变化而感到兴奋。这正是大数据的魅力之处。

作为一个在大数据领域出现前就已在相关领域工作的人（我的研究生课程是用全美最快的并行电脑来构建三维气候模型，以模拟大气物理过程和化学反应）；我很高兴地看到，大数据已渐为人知，并已展示出强大威力。与此同时，令人失望的是，大数据越来越多地被用于市场营销，而非用于创造对商业及人类社会真正有价值的东西。这也是我开始在 LinkedIn 上自称大数据高传者和执行者的原因。在最近参加的一些活动和会议上，我已开始用这样的称号介绍自己。

今天我想谈的是一个 EOI (Empower/Optimize/Innovate, 即助力、优化、创新) 的分析架构，这是 LinkedIn 的商业分析团队用大数据手段来持续推动商业价值的方法。下面我将详细阐述这一分析架构：

### E：助力信息分享

在该领域，业界公认的做法是，根据商业伙伴提出的问题予以专门分析，比如“去年、上月或上周我们赚了多少钱”、“导致核心业

务表现指数大幅下降的主要原因是什么”等。这可能是大多数人对分析一词的定义。的确，这种分析对业务发展很重要，因为它可以帮助管理者做出基于此类数据的决策，至少可以促进他们考虑使用数据来做决策。

很多分析团队在解决这类问题上花费了大量时间。但随着工作效率的不断提高，问题来了——分析师会对一遍遍重复类似的分析感到厌倦。避免厌倦的方法之一是，尽可能多地利用技术手段简化分析流程，并尽可能将分析步骤自动化，比如自动化数据清理、自动化数据格式转化等。这样他们就可以腾出时间去做更有意思的事情——发掘更多可洞察事物本质的结论并为商业伙伴提供相关决策参考。

一个典型案例是我们团队推出的一个名叫“Merlin”的内部分析网站。该网站是 LinkedIn 销售团队建立的，其功能能自动生成结论型信息，一键搜索后，团队成员可迅速将此类信息分享给客户。每天有数千名来自销售团队的工作人员在使用这个网站，完全自动化获取数据、指标、报告、图表等。由于对地面销售业务给予了巨大支持而目前者因此得到了可观的经济回报，该项目被公司管理层评选为 2011 年 LinkedIn 10 大最具改革力案例之一，并获得了公司国际销售部门授予的“影响力奖”。

### O：优化业务表现

该领域包括诸多高级分析工作，比如基于商业报道的深入分析、市场推广定位以及用户倾向模型的建立等，用于回答类似“如果我们这样做，会发

生什么”、“最好的结果是什么”等问题。尽管这类分析执行起来通常要花更多时间，但也会在业务上带来更多回报。更重要的是，因为几乎总是要从 E，即“助力”环节建立起来的知识基础起步，所以分析人员能够更好地理解数据本质，并可将其与实际业务需求有效结合。

业界经常出现的情况是，当一个分析团队想跳过“助力”环节直接进入 O，即“优化”环节时，往往会遭遇缺乏数据基础架构和基本商业知识的情况。最终必须回过头来夯实基础，然后才能进入下一步。一个关于“优化”的典型案例是，我们为 LinkedIn 高级销售业务建立的倾向模型。在这个模型中，我们利用用户身份、用户行为和社交网络数据来作为邮件营销行为区分人群属性。该模型已成为对 LinkedIn 最大线上业务在市场营销侧面的核心驱动力。

### I：探索创新模式

在硅谷，每个人都会因“创新”而兴奋。LinkedIn 的分析团队有很多创新。我们相信，衡量分析团队创新水准衡量的终极标准是其对公司核心业务影响力的大小。当我们评估一项创新或风险项目的发展潜力时，我们会关注其在未来 1-3 年的潜在商业影响，主要评价指标是它的营收、利润、用户粘性、访问量增长等。我们还要确保影响力巨大

的商业活动可以利用我们项目的研究成果，从而快速验证我们的分析解决方案在市场上是否可行，而非为创新而创新。

一个最新案例是，我们和公司市场共同建立的企业用户兴趣指数，对那些有可能成为 LinkedIn 的企业型客户，在预期上进行排序。该项创新的关键在于，结合加权后的公司内部个人层面分值，以及 B2B 销售流程中决策者的影响，可得出该企业转化为 LinkedIn 客户的可能性。该系统建立迄今，一直被我们的地面销售团队广泛采用，有效提高了客户转化率，对销售额和工作效率的提升均颇有助益。

自倡导这一分析架构以来，我经常被问到一个问题是，在合理的 EOI 架构下，资源配比情况是怎样的。事实上，根据分析团队工作进展和公司所处发展阶段，花费在 EO 上的分析资源也各有不同。就此可大致绘出一条高低起伏的配置曲线。此中关键在于，在 E、O、I 任何一个层面，你至少都要投入能够产生实际作用的资源，并设计一个你认为对于现阶段业务增长能够发挥最佳效果的配置比例。总体而言，基于我和行业内众多分析师同事的讨论结果，合理的资源分配比例一般是 E>O>I。对于年均业务增速在两位数以上的企业来说，尤其如此。■

——翻译/本刊记者 史晓丹

#### E人E本公司收购苹果飞虎牌

事实上，大型公司正在将大数据变成，不仅是 LinkedIn 将大数据用于社交平台，国内的“阿里系”的电商平台，在深耕网络商城的同时，深谙大数据之道，获取大量企业以及消费者的行为数据，通过建立精准推荐系统获得市场，创建小型企业的优势。

大企业将大数据应用到何处为止，那要看企业可在利用大数据做什么，比如苹果 iPhone 汽车业务，以在不到 20 人的团队发挥，收购了本田汽车，未来将推出汽车业务领域的全球人口统计数据，成为可能。事实上，在金融领域，许多小企业开发大数据业务的数据量并不少。

大数据带来的不仅只是商机，有人想大数据行业为未来世界提供资源，更有人觉得掌握了大数据的人可以登上全球一统的宝座，但现实如何，大数据时代已经到来。

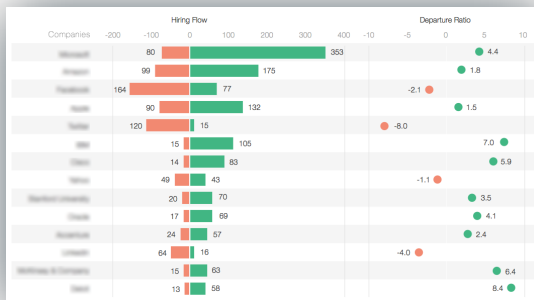


E人E本 商务平板电脑

# 商业分析实例：在EOI的架构下创造商业价值

## 1. 助力

利用交互性的数据应用给职场人员建立数据通道

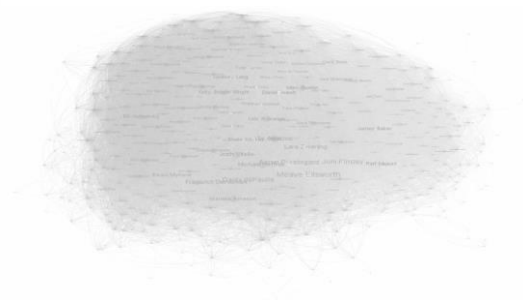


### 人才流动看板

动态可视化工具帮助挖掘商业洞察

## 2. 优化

精准营销：通过分析和倾向模型精准定位优化营销策略



### 用户倾向预测模型(B2C)

识别正确分块市场，在最好的时间宣传最适合的产品

## 3. 创新

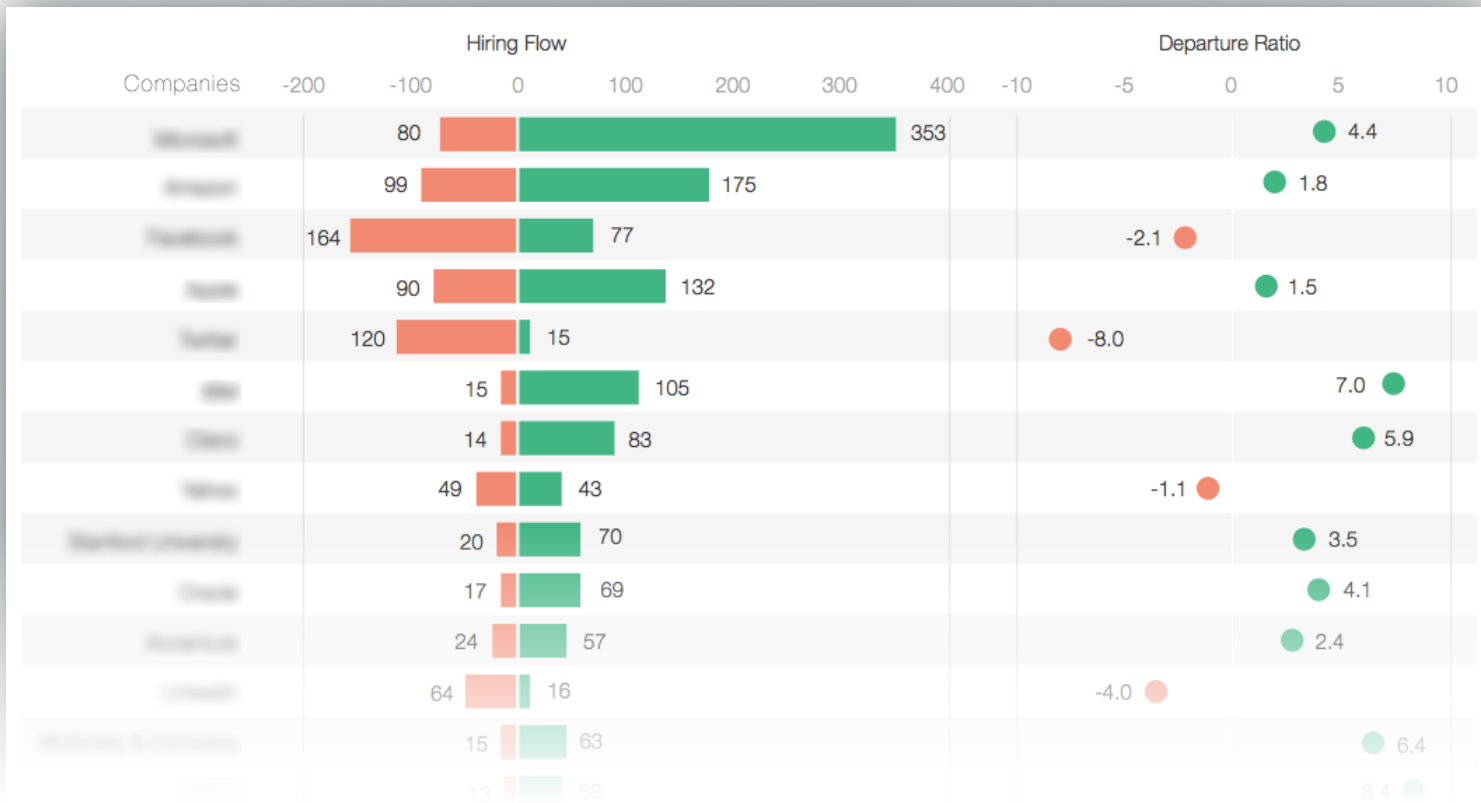
用商业分析的创新将营销战略带到新的高度



### 大客户兴趣指数(B2B)

商业大客户（企业及团体）对相关产品的兴趣度的倾向模型

# 人才流动画板帮助你发现公司在人才争夺中的战况



# 用职场大数据来衡量雇主品牌\*

450M+ 职场人员

45B+ 与公司的互动

\* We cross-referenced our data with thousands of survey responses to pinpoint the specific activities that best indicate familiarity and interest in working for a company: **connecting with employees, viewing employee profiles, visiting Company and Career Pages, and following companies.**

<http://blog.linkedin.com/2012/10/11/indemand-employers-2012/>

The screenshot shows a LinkedIn blog post from Michael Li, dated October 11, 2012. The post title is "Ranking the 100 Most InDemand Employers Using LinkedIn Data [INFOGRAPHIC]". It has 4,050 shares. The main text describes how LinkedIn used its massive professional data set to identify the most attractive companies to potential candidates. It mentions that the rankings were unveiled at the Talent Connect event in Las Vegas. Below the text is an infographic titled "LinkedIn's Most InDemand Employers". The infographic includes a world map showing the most in-demand employers globally and a list of 15 companies ranked by geography and function. The companies listed are: 1. Google, 2. Apple, 3. Microsoft, 4. Facebook, 5. The Walt Disney Company, 6. Pepsico, 7. P&G, 8. McKim & Company, 9. The Coca-Cola Company, 10. Nike, 11. salesforce.com, 12. Twitter, 13. The UPS Store, 14. NESSE, 15. BCG. There is also a badge for "Most InDemand Employers 2012".

# 找出硅谷最具潜力的初创公司

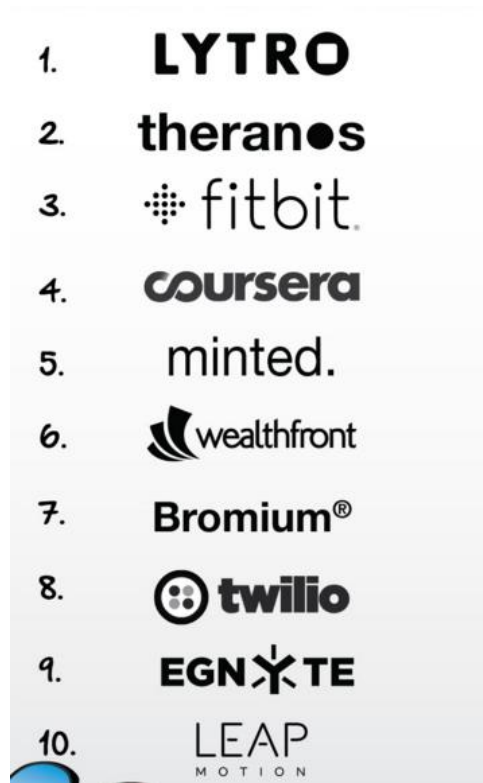
## 2012



## 2013



## 2014



# 2015年最in雇主

年度  
大奖



华为技术有限公司  
2015年最in雇主年度大奖

全球领先的信息与通信解决方案供应商，最早引入社交招聘的中国企业之一，雇主品牌典范，唯一曾上榜全球最受欢迎雇主的中国企业。

阿里巴巴   
Alibaba.com

阿里巴巴（中国）有限公司  
2015年最in雇主

从最初的B2B业务到现在的多元化，人才管理根据业务变化迅速调整，培训每一位管理者成为雇主品牌大使，实现精准社交招聘。



华晨宝马汽车有限公司  
2015年最in雇主

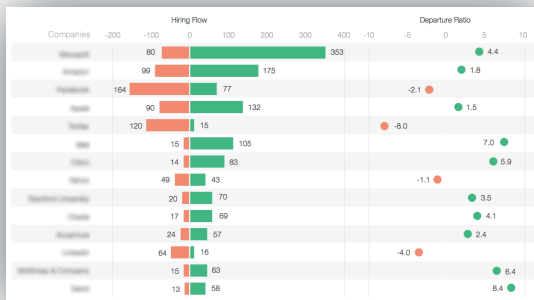
国际高档汽车品牌本土化的榜样，不仅做到了商业成功，也将员工视作最重要的资源，积极推广社交招聘，雇主品牌指数位居行业前列。



# 商业分析实例：在EOI的架构下创造商业价值

## 1. 助力

利用交互性的数据应用给职场人员建立数据通道

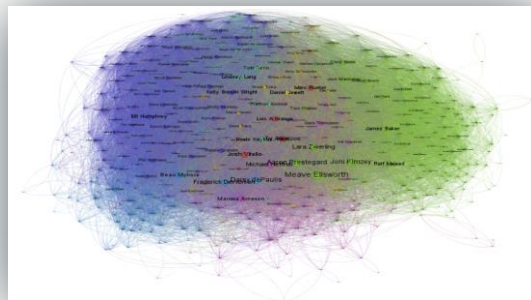


### 人才流动动画板

动态可视化工具帮助挖掘商业洞察

## 2. 优化

智能销售：通过倾向模型预测用户订单额度，优化销售策略



### Size Of Prize 客户消费潜力

利用职场社交数据预测客户消费潜力

## 3. 创新

用商业分析的创新将营销战略带到新的高度



### 大客户兴趣指数(B2B)

商业大客户（企业及团体）对相关产品的兴趣度的倾向模型

# 数据驱动B2B销售和营销团队

哪个公司? (SOP)

联系谁?

怎样联系?

哪个销售人员?

什么内容?





# Size of Prize (SOP):

What is Size of Prize (SOP)?

- **Size of Prize intends to estimate how much an account could spend with LinkedIn on Talent Solutions products (based upon the behavior of current customers.)**

What is **NOT** Size of Prize (SOP)?

- How much an **account currently spends** with LI
- **A prediction of how easy it will be** for an account to spend this amount (will discuss adapting with similar measure)

# 领英数据的三种主要类型

## 用户个体



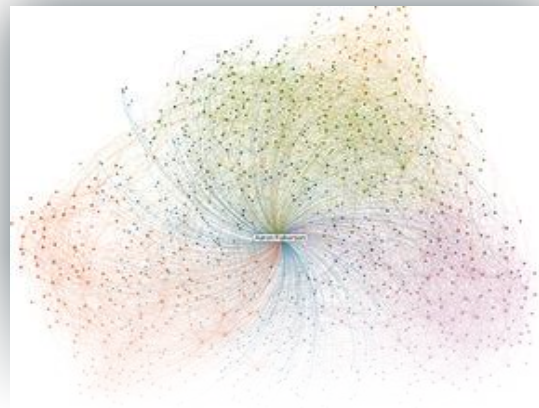
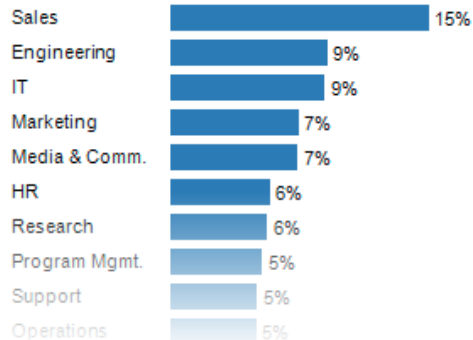
## 用户行为



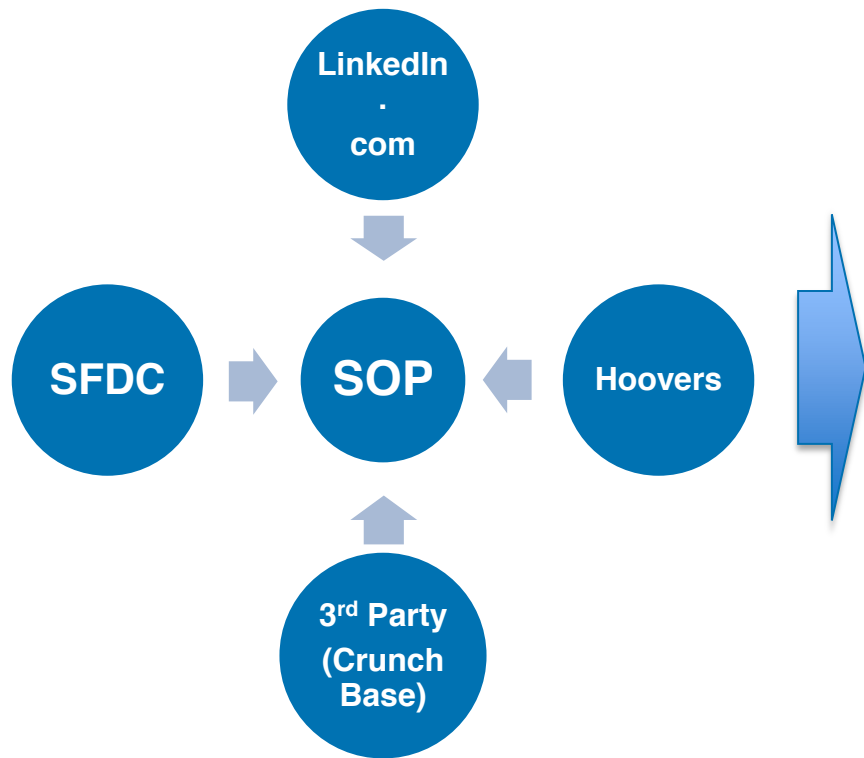
## 用户网络



Visitors By Function (Top 15 Functions)



# 根据行业知识整合外部数据



## Thousands of Variables Selections

- # Hires weighted by country and recency
- # Recruiter
- Recruiter to hire ratio
- Industry
- Company size
- Recruiter penetration rate
- Country
- Average salary of posted jobs
- Hiring accelerating rate
- . . .



# 找到预测性和实用性最高的指标

## SOP driven by 3 LI-based factors

### 1 # of Hires on LI

- On LinkedIn, the # of employees who joined a company from a different organization in a given time period
  - E.g., Jonathan Lister leaves Google to join LinkedIn in July 2010 = a new hire for LinkedIn for July 2010

### 2 # of Recruiters on LI

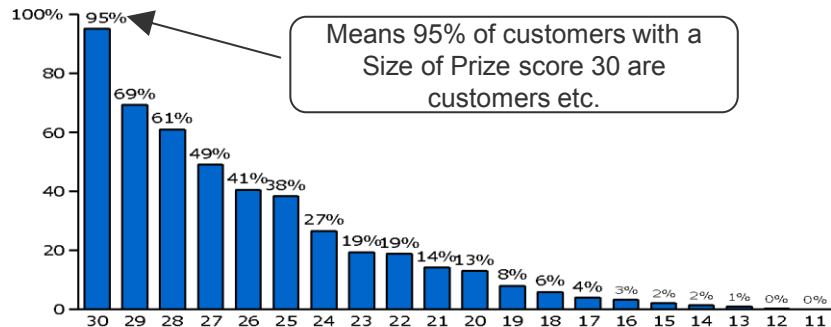
- On LinkedIn, how many employees serve a recruiting-related role?
  - Coding titles for words such as “recruiting,” “talent” etc.

### 3 Industry

- Which industries have highest LI usage?

## SOP algorithm is highly predictive

% of accounts with given Size of the Prize score that are LCS customers



# cust.:

# total co's:

Avg. seats/acct.:



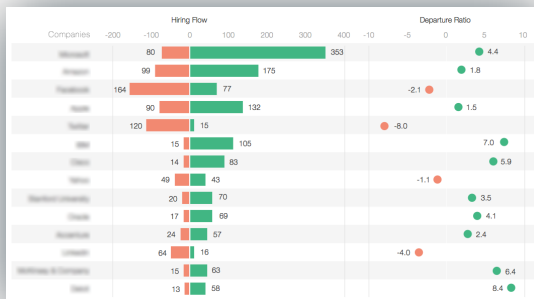
“Size of the Prize” approach use to:

- Allocating right accounts to right sales reps
- For each rep, prioritizing accounts, targeting spend

# 商业分析实例：在EOI的架构下创造商业价值

## 1. 助力

利用交互性的数据应用给职场人员建立数据通道

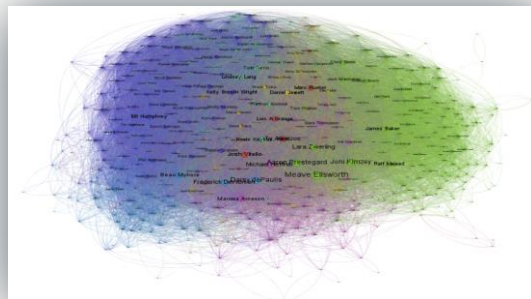


### 人才流动画板

动态可视化工具帮助挖掘商业洞察

## 2. 优化

智能销售：通过倾向模型预测用户订单额度，优化销售策略



### Size Of Prize

利用职场社交数据预测客户消费潜力

## 3. 创新

用商业分析的创新将营销战略带到新的高度



### 大客户兴趣指数(B2B)

商业大客户（企业及团体）对相关产品的兴趣度的倾向模型



LOW

NO  
SMOKING

LOW

HIGH

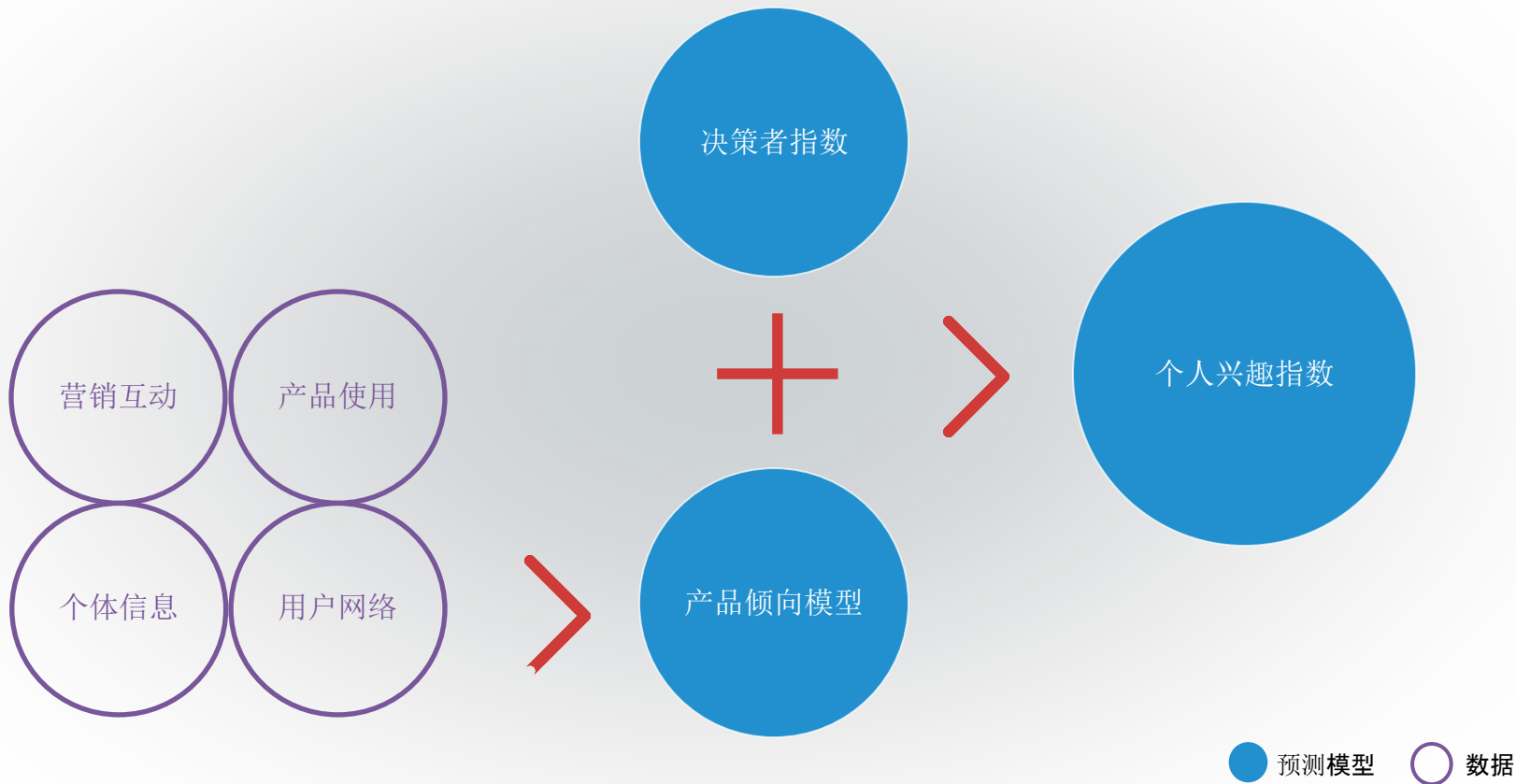
MEDIUM

决策者

决策者在B2B商业模式里起着非常重要的作用



# 根据决策者指数和产品倾向模型建造个人兴趣指数





LOW

NO  
SMOKING

LOW

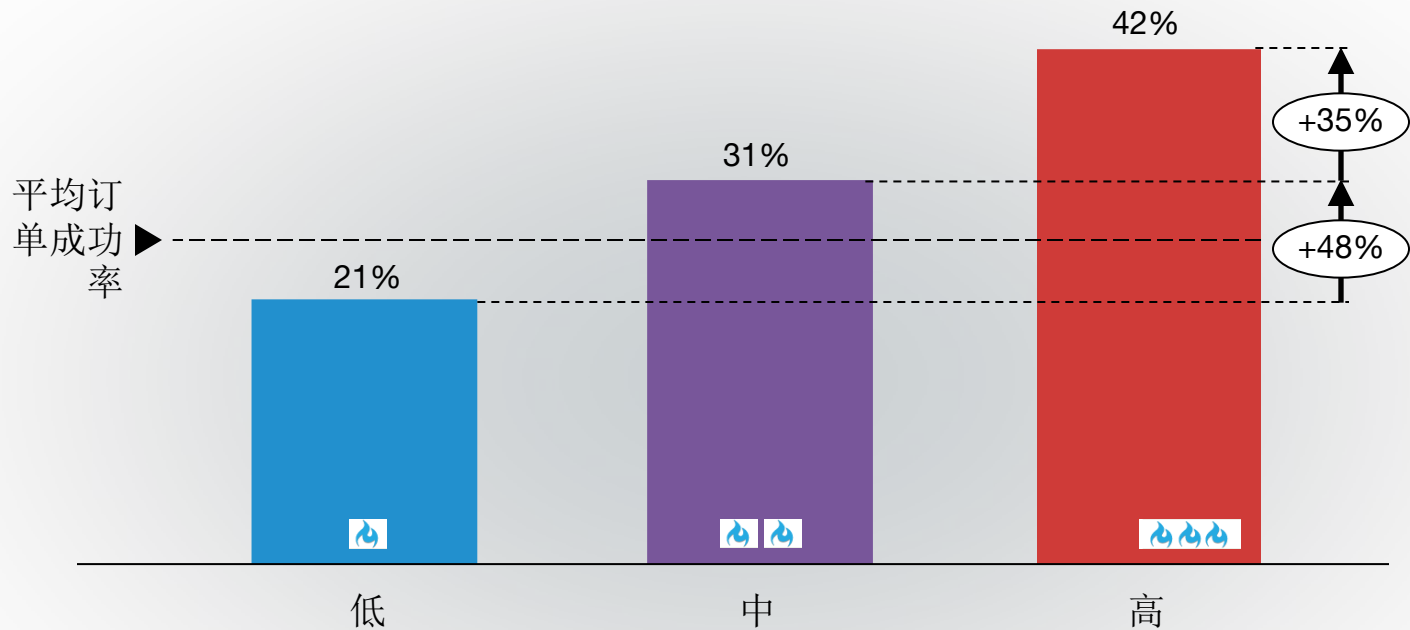
HIGH

MEDIUM

大客户兴趣指数

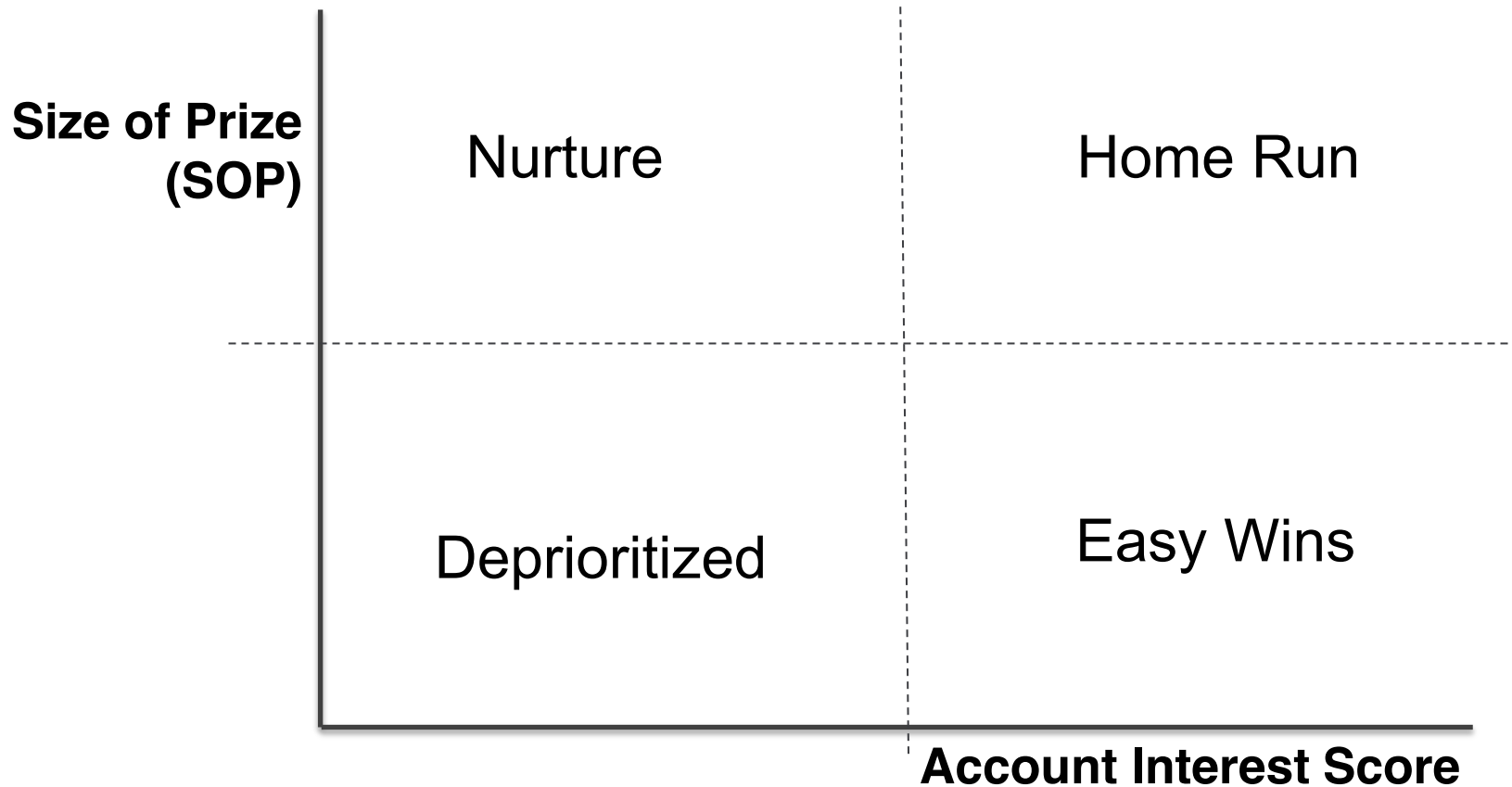
从个人兴趣指数整合到大客户兴趣指数

# 较高的大客户兴趣指数带来更高的交易效率和成功率



在不同大客户兴趣指数上的订单成功率

# SOP 和客户兴趣指数共同驱动销售业绩提升



# 冰山下的真正秘密



salesforce.com

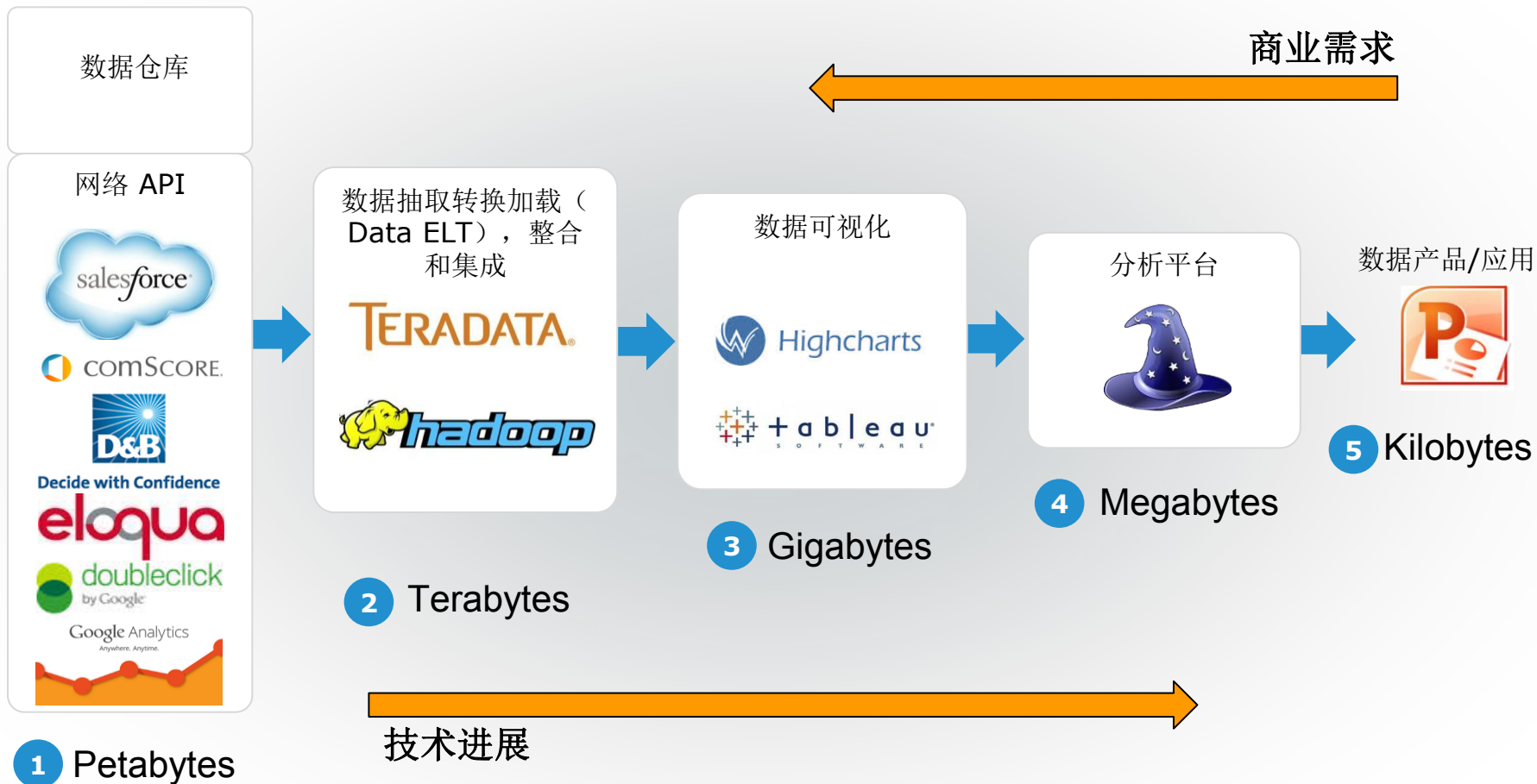


ELOQUA

TERADATA ORACLE



# 科技是实现可规模化大数据分析的基石



# 数据不是金钱，是资本

- 公司建立初期应以商业模式和用户体验为中心，数据需要积累到一定规模会开始对公司运营提供直接帮助。公司模式越成熟，数据的作用越巨大。
- 数据很难直接变现（不要卖数据），但是公司业务中积累和拥有的独特数据会是商业竞争的长期资本，以及长期战略的决定性因素。
  - 例子：LinkedIn's Economic Graph
- 每一轮创新大潮归根结底都是新型数据的产生和积累过程，最终胜出的公司成为某类数据的唯一拥有者并且以这种数据为基础建立起独裁者一样的商业模式
  - 例子：Google, Facebook, Uber
- 从长远的角度讲，数据的拥有权和使用权将是企业最核心的竞争力并且决定企业长期发展的高度





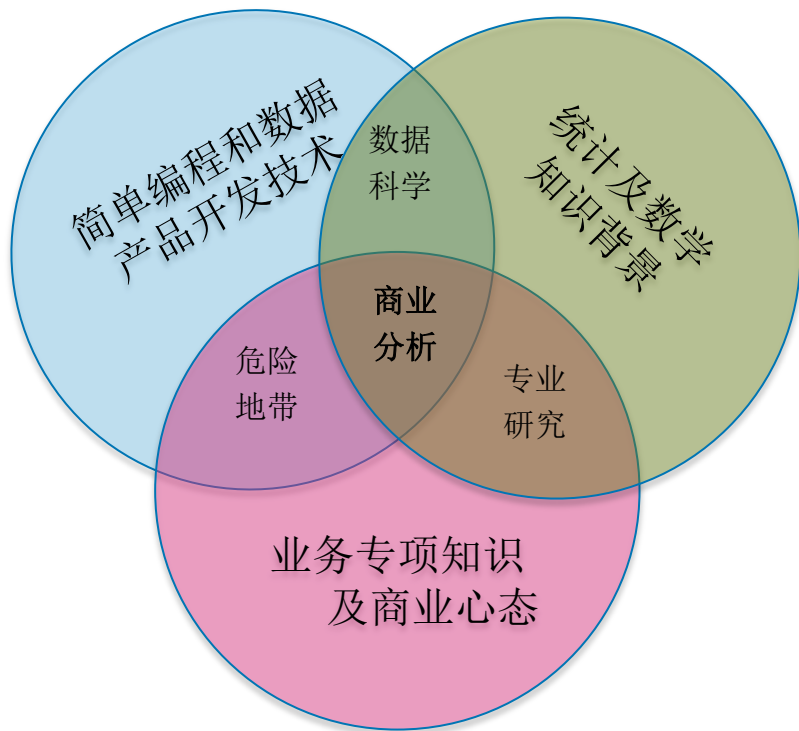
*“What is the most important thing for a team or a company in the 21st century?”*

**Talents is our No.1 priority!**

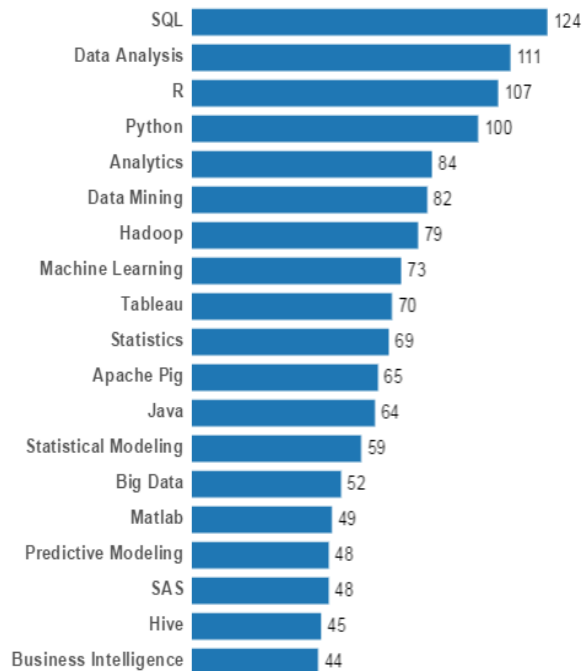




# 团队背景和技能



Skills

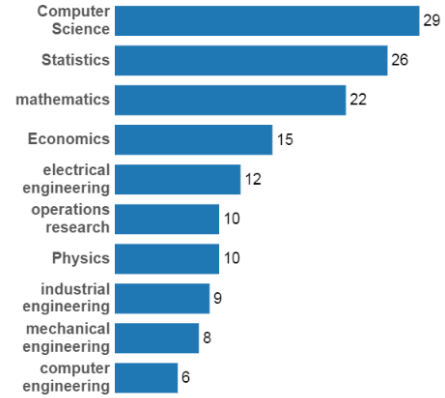


# Fun Facts

## Top Schools



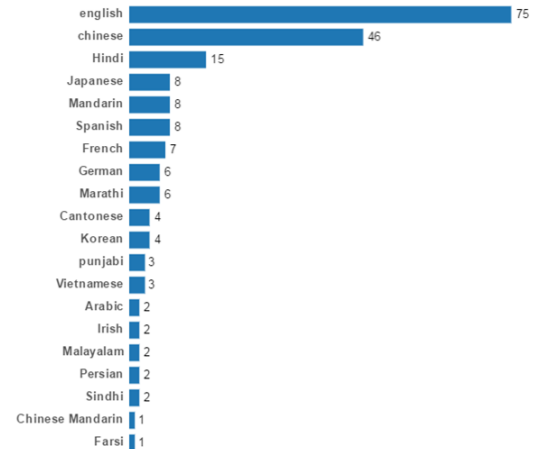
## Field of Education



## Previous Companies



## Languages



# Appendix



# 佳格

## 农业数据云平台

AGRICULTURAL DATA CLOUD PLATFORM

农业综合管理信息系统

智慧农业信息云平台

THE WISDOM OF AGRICULTURAL INFORMATION CLOUD PLATFORM

### 农作物类型分析



监测农作物种植面积

农作物适宜分析与规则

农作物生长周期估计

### 农作物长势监测



光合作用效率监测

卫星和无人机搭载的多光谱成像仪能够精确对农作物进行观测



自然灾害预测与评估

综合气象数据与卫星监测对农作物快速准确的监测



作物病虫害预警与评估

通过卫星监测作物，对作物的状况有实时的预警



土壤水分墒情精细估计

卫星遥感数据对土壤的水分进行实时观测，为灌溉提供指导

### 农作物产量估计

- 农作物产量远期预估
- 农作物产量近期预报
- 农作物产量收获评估

### 农业综合管理信息系统



种植面积

出苗情况

土地等级

长势监控

产权信息

收获管理

### 农业深度分析仪表盘系统



智慧农业是农业生产的高级阶段，是集新兴的互联网、移动互联网、云计算和物联网技术为一体，依托部署在农业生产现场的各种传感节点（环境温度、土壤水分、二氧化碳、图像等）和无线通信网络实现农业生产环境的智能感知、智能预警、智能决策、智能分析、专家在线指导，为农业生产提供精准化种植、可视化管理、智能化决策。